Method for Shortening Dimensionality of Qualitative Attribute Space

Alexey B. Petrovsky[a,b,c,d]

[a] *Federal Research Center "Computer Science and Control" Russian Academy of Sciences, Moscow, Russia*
[b] *V.G. Shukhov Belgorod State Technological University, Belgorod, Russia*
[c] *Belgorod State National Research University, Belgorod, Russia*
[d] *Volgograd State Technical University, Volgograd, Russia*

Abstract

The paper presents a new method SOCRATES (ShOrtening CRiteria and ATtributES) for reducing the dimensionality of attribute space. In the method, a lot of initial numerical and/or verbal characteristics of objects are aggregated into a single integral index or several composite indicators with small scales of qualitative estimates. Aggregation of indicators includes various methods for a transformation of attributes and their scales. Multi-attribute objects are represented as multisets of object properties. Reducing the dimensionality of attribute space allows us to simplify the solution of applied problems, in particular, problems of multiple criteria choice, and explain the obtained results. An illustrative example is given.

Keywords 1

Attribute space, dimensionality reduction, attributes' aggregation, composite indicator, integral index, multi-attribute object, multiset, multiple criteria choice

1. Introduction

The tasks of a strategic and unique choice, in which there are very few objects, and the number of features characterizing their properties is large (tens or hundreds), are among the most difficult. Examples of such objects are places for the construction of an airport or power plant, routes of a gas or oil pipeline, schemes of a transportation network, configurations of a complex technical system, and the like. For decision-makers (DMs) and experts in real situations, it is very difficult to select the best object, to rank or classify objects that are described by numerous attributes, because, as a rule, many objects are formally not comparable in their characteristics. Additional difficulties arise in the case of ill-structured problems combining quantitative and qualitative dependencies, the modeling of which is either impossible in principle or very hard. The known decision-making methods [2-7] are poorly suitable for solving problems of multi-criteria selection of large dimension because they require a lot of labor and time to obtain and process large amounts of data about objects, a DM preferences and/or expert knowledge.

The following approaches are possible that facilitate the choice in a large space of attributes and reduce information loss: the use of psychologically correct operations for obtaining information from decision-makers and experts; the reduction of dimension of attribute space. It has been experimentally established that it is easier for a person, due to the peculiarities of his physical memory, to operate with small amounts of data, to compare objects according to a small number of indicators. To do this, it is enough to describe objects with three-seven indicators. At the same time, a person makes fewer mistakes when indicators have not numerical, but verbal scales [4, 5]. Shortening the dimension of attribute space by reducing the number of variables simplifies the solution of problems of individual

Russian Advances in Artificial Intelligence: selected contributions to the Russian Conference on Artificial intelligence (RCAI 2020), October 10-16, 2020, Moscow, Russia
EMAIL: pab@isa.ru
ORCID: 0000-0002-5071-0161

and group multi-criteria choice. Almost all the applied methods for shortening the space dimension work with numerical attributes [1, 2, 11]. Procedures for shortening the dimension of spaces of verbal attributes are described in [10].

This work describes a new method SOCRATES (ShOrtening CRiteria and ATtributES), in which a large number of initial characteristics of objects are aggregated into a single integral index or several indicators with small scales of verbal estimates. Representation of multi-attributes objects by multisets and aggregation of attributes can significantly diminish the complexity of solving the original problem of multi-criteria choice and explain reasonably the results.

2. Representation and comparison of multi-attribute objects

A multiset or a set with repetitions is a convenient mathematical model for representing and comparing objects, which are defined by many numerical and/or verbal attributes and are presented in several exemplars (versions, copies) that differ in the values of their characteristics [7-10]. This model allows us to take into account simultaneously heterogeneous attributes, possible combinations of attribute values, and the presence of different exemplars of objects.

Let objects $O_1,...,O_q$ exist in single copies and are described by attributes $K_1,\ldots,K_n$ with numerical and/or verbal rating scales. If each of the attributes $K_1,\ldots,K_n$ has the same rating scale $X = \{x^1,...,x^h\}$, then we associate the object O_p, $p = 1,\ldots,q$ with a multiset of estimates

$$\boldsymbol{A}_p = \{k_{Ap}(x^1)\circ x^1,...,k_{Ap}(x^h)\circ x^h\} \tag{1}$$

over the set $X = \{x^1,...,x^h\}$ of scale gradations. Here, the value of the multiplicity function $k_{Ap}(x^e)$ shows how many times the estimate $x^e \in X$, $e = 1,\ldots,h$ is present in the description of the object O_p.

If each attribute K_i has its own rating scale $X_i = \{x_i^1,\ldots,x_i^{h_i}\}$, $i = 1,\ldots,n$, we introduce a single general scale (hyperscale) of attributes – the set $X = X_1 \cup ... \cup X_n = \{x_1^1,\ldots,x_1^{h_1};\ldots; x_n^1,\ldots,x_n^{h_n}\}$, which consists of n groups of attributes and combines all the estimate gradations on the scales of all attributes. Then the object O_p will correspond to a multiset of estimates

$$\boldsymbol{A}_p = \{k_{Ap}(x_1^1)\circ x_1^1,\ldots,k_{Ap}(x_1^{h_1})\circ x_1^{h_1};\ldots; k_{Ap}(x_n^1)\circ x_n^1,\ldots,k_{Ap}(x_n^{h_n})\circ x_n^{h_n}\} \tag{2}$$

Here, the value of the multiplicity function $k_{Ap}(x_i^{e_i})$ shows how many times the estimate $x_i^{e_i} \in X_i$, $e_i = 1,\ldots,h_i$ by the attribute K_i is present in the description of the object O_p.

The expression (2) is easy to write in the "usual" form (1), if in the set $X = \{x_1^1,\ldots,x_1^{h_1};\ldots; x_n^1,\ldots,x_n^{h_n}\}$ make the change of variables: $x_1^1 = x^1,\ldots, x_1^{h_1} = x^{h_1}$, $x_2^1 = x^{h_1+1},\ldots, x_2^{h_2} = x^{h_1+h_2},\ldots, x_n^{h_n} = x^h$, $h = h_1+...+h_n$. Despite the seemingly cumbersome presentation of multi-attribute objects by multisets, such recording forms are very convenient when performing operations under objects, since calculations are performed parallel and simultaneously for all elements of all multisets.

The situation is complicated when the object O_p is present in several copies $O_p^{<s>}$, $p = 1,\ldots,q$, $s = 1,\ldots,t$, which differ in the values of the attributes $K_1,\ldots,K_n$. Different versions of the description of the object O_p arise, for example, when the object is evaluated by t experts according to many criteria $K_1,\ldots,K_n$, or the characteristics of the object are calculated t times by several methods $K_1,\ldots,K_n$, or measured t times by several tools $K_1,\ldots,K_n$.

A variety of operations on multisets provides the ability to group multi-attribute objects in different ways [9, 10]. A group of objects can be formed by defining the multiset $\boldsymbol{J}$ that represents the group by the sum $\boldsymbol{J} = \sum_s \boldsymbol{A}_s$, $k_{\boldsymbol{J}}(x^e) = \sum_s k_{As}(x^e)$, by the union $\boldsymbol{J} = \bigcup_s \boldsymbol{A}_s$, $k_{\boldsymbol{J}}(x^e) = \max_s k_{As}(x^e)$, by the intersection $\boldsymbol{J} = \bigcap_s \boldsymbol{A}_s$, $k_{\boldsymbol{J}}(x^e) = \min_s k_{As}(x^e)$ of multisets $\boldsymbol{A}_s$ that describe the grouping objects, or by one of the linear combinations of operations on multisets $\boldsymbol{A}_s$ such as $\boldsymbol{J} = \sum_s c_s\boldsymbol{A}_s$, $\boldsymbol{J} = \bigcup_s c_s\boldsymbol{A}_s$, $\boldsymbol{J} = \bigcap_s c_s\boldsymbol{A}_s$, $c_s > 0$ is an integer. When we add multisets, all properties (all values of all attributes), which are available to individual objects in the group, are aggregated. When we combine or intersect multisets, the best properties (maximum values of all attributes) or, correspondingly, the worst properties (minimum values of all attributes) of the grouping objects are strengthened.

If there are several versions of the object O_p, then this object is represented by a group of all its copies $O_p^{<s>}$, $p = 1,\ldots,q$, $s = 1,\ldots,t$. We associate the object O_p with the multiset $\boldsymbol{A}_p = \{k_{Ap}(x^1)\circ x^1,...,$

$k_{Ap}(x^h)\circ x^h\}$ of the form (1), (2), and the version $O_p^{<s>}$ with the multiset $\boldsymbol{A}_p^{<s>} = \{k_{Ap}^{<s>}(x^1)\circ x^1,..., k_{Ap}^{<s>}(x^h)\circ x^h\}$ over the set of estimates $X = \{x^1,...,x^h\}$ or $X = \{x_1^1,\ldots,x_1^{h_1};\ldots; x_n^1,\ldots,x_n^{h_n}\}$. We will form the multiset $\boldsymbol{A}_p$ as a weighted sum of multisets describing versions of the object:

$$\boldsymbol{A}_p = c^{<1>}\boldsymbol{A}_p^{<1>} + \ldots + c^{<t>}\boldsymbol{A}_p^{<t>},$$

where the multiplicity function of the multiset $\boldsymbol{A}_p$ is calculated by the rule $k_{Ap}(x^e) = \sum_s c^{<s>}k_{Ap}^{<s>}(x^e)$, and the coefficient $c^{<s>}$ characterizes the significance (expert competence, measurement accuracy) of the exemplar $O_p^{<s>}$.

3. Reduction of attribute scales

The reduction of dimension of the object description is a decrease of number of the indicators, that characterize the properties, state or functioning of objects, using some transformations of the initial data, during which the set of initial attributes $K_1,\ldots,K_n$ is aggregated into smaller sets of new intermediate attributes $L_1,\ldots,L_m,\ldots$ and final attributes $N_1,\ldots,N_l$. Transformations of characteristics can be formally written as

$$K_1,\ldots,K_n \to L_1,\ldots,L_m \to\ldots \to N_1,\ldots, N_l. \quad (3)$$

The initial attribute K_i has the scale $X_i = \{x_i^1,\ldots,x_i^{h_i}\}$, $i = 1,\ldots,n$, the intermediate attribute L_j has the scale $Y_j = \{y_j^1,\ldots,y_j^{g_j}\}$, $j = 1,\ldots,m$, the final attribute N_k has the scale $Z_k = \{z_k^1,\ldots,z_k^{f_k}\}$, $k = 1,\ldots,l$, $l < m < n$. Decreasing the dimension of attribute space is an informal multi-stage procedure based on the knowledge, experience and intuition of a DM/expert, who forms the rules for the conversion of attributes, establishes the structure, number, dimension and sense of new indicators.

In cases where multi-attribute objects are represented by vectors/tuples, the task (3) of reducing the dimension of attribute space has the form:

$$X_1\times\ldots\times X_n \to Y_1\times\ldots\times Y_m \to\ldots\to Z_1\times\ldots\times Z_l. \quad (4)$$

Then the dimension of the corresponding attribute space is defined as the power of the direct product of numerical or verbal gradations of attribute scales that are components of vectors/tuples. In the book [10], the problem (4) is considered as a multi-criteria classification problem, where the sets of estimates of the initial attributes are the classified objects, and the grades of the composite indicator scale are the classes of solutions [5, 7].

In cases where multi-attribute objects are represented by multisets, the task (3) of reducing the dimension of attribute space has the form:

$$X_1\cup\ldots\cup X_n \to Y_1\cup\ldots\cup Y_m \to\ldots\to Z_1\cup\ldots\cup Z_l. \quad (5)$$

Then the dimension of the corresponding attribute space is defined as the power of the hyperscale – the union of numerical or verbal gradations of attribute scales that are elements of multisets. The method SOCRATES (ShOrtening CRiteria and ATtributES) allows us to reduce the dimension of the description of multi-attribute objects that are presented in several different versions and are defined by multisets of numerical and/or verbal characteristics. The method uses two principal ways of data transformations: a reduction of attribute scale and an aggregation of attributes.

The reduction of attribute scale is a relatively simple transformation of attribute space and is aimed at reducing the number of gradations on the attribute scale. For this, several values of some object characteristic are combined into one new gradation of the same characteristic. The transition from the original characteristic scales to scales with the reduced number of grades is the transformation (5) into the form

$$X_1\cup\ldots\cup X_n \to Q_1\cup\ldots\cup Q_n, \quad (6)$$

where $X_i = \{x_i^1,\ldots,x_i^{h_i}\}$ is the original scale, $Q_i = \{q_i^1,\ldots,q_i^{d_i}\}$ is the shortened scale of the attribute K_i, $|Q_i| = d_i < h_i = |X_i|$, $i = 1,\ldots,n$. When forming (6) the shortened scales of attributes, it is desirable that they consist of a small number (2-4) of gradations, which have certain content for a DM/expert.

The representation of multi-attribute objects is transformed as follows. In the attribute space $K_1,\dots,K_n$, let the object O_p, $p = 1,\dots,q$ be defined by the multiset $\boldsymbol{A}_p$ (2) over the set $X_1\cup\dots\cup X_n$ of gradations of the original scales. Considering the properties of operations on multisets [9, 10], we rewrite expression (2) as multiset sums:

$$\boldsymbol{A}_p = \boldsymbol{A}_{p1} +\dots+ \boldsymbol{A}_{pn} = \{k_{Ap}(x_1^1)\circ x_1^1,\dots,k_{Ap}(x_1^{h_1})\circ x_1^{h_1}\} +\dots+ \{k_{Ap}(x_n^1)\circ x_n^1,\dots,k_{Ap}(x_n^{h_n})\circ x_n^{h_n}\} = \\ = \sum_{e_1=1}^{h_1}\{k_{Ap}(x_1^{e_1})\circ x_1^{e_1}\} +\dots+ \sum_{e_n=1}^{h_n}\{k_{Ap}(x_n^{e_n})\circ x_n^{e_n}\}. \quad (7)$$

When reducing the attribute scale, several gradations $x_i^{e_a}$, $x_i^{e_b},\dots,$ $x_i^{e_c}$ of the original scale $X_i = \{x_i^1,\dots,x_i^{h_i}\}$ of the attribute K_i are combined into a single gradation $q_i^{o_i}$ of the shortened scale $Q_i = \{q_i^1,\dots,q_i^{d_i}\}$. In the reduced space of attributes $K_1,\dots,K_n$ with rating scales $Q_1,\dots,Q_n$, the object O_p will correspond to a multiset

$$\boldsymbol{B}_p = \{k_{Bp}(q_1^1)\circ q_1^1,\dots,k_{Bp}(q_1^{d_1})\circ q_1^{d_1};\dots; k_{Bp}(q_n^1)\circ q_n^1,\dots,k_{Bp}(q_n^{d_n})\circ q_n^{d_n}\} \quad (8)$$

over the set $Q_1\cup\dots\cup Q_n$ of gradations of the shortened scales. The multiset $\boldsymbol{B}_p$ (8) can also be written in the equivalent form:

$$\boldsymbol{B}_p = \boldsymbol{B}_{p1} +\dots+ \boldsymbol{B}_{pn} = \{k_{Bp}(q_1^1)\circ q_1^1,\dots,k_{Bp}(q_1^{d_1})\circ q_1^{d_1}\} +\dots+ \{k_{Bp}(q_n^1)\circ q_n^1,\dots,k_{Bp}(q_n^{d_n})\circ q_n^{d_n}\} = \\ = \sum_{o_1=1}^{d_1}\{k_{Bp}(q_1^{o_1})\circ q_1^{o_1}\} +\dots+ \sum_{o_n=1}^{d_n}\{k_{Bp}(q_n^{o_n})\circ q_n^{o_n}\}. \quad (9)$$

The multiplicity of the element $q_i^{o_i}$, $o_i = 1,\dots,d_i$ of the multiset $\boldsymbol{B}_p$ (8) or (9), which corresponds to the gradation $q_i^{o_i}$ of the shortened scale $Q_i = \{q_i^1,\dots,q_i^{d_i}\}$, is determined by the rule:

$$k_{Bp}(q_i^{o_i}) = k_{Ap}(x_i^{e_a}) + k_{Ap}(x_i^{e_b}) +\dots+ k_{Ap}(x_i^{e_c}), \quad (10)$$

where the multiplicities of the elements $x_i^{e_a}$, $x_i^{e_b},\dots,$ $x_i^{e_c}$ of the multiset $\boldsymbol{A}_p$ (2) or (7), which correspond to the combined gradations of the original scale $X_i = \{x_i^1,\dots,x_i^{h_i}\}$ of the attribute K_i, are summarized.

4. Aggregation of attributes

The aggregation of attributes is a more complex transformation of attribute space and is aimed at reducing the number of attributes. For this, several attributes L_a, $L_b,\dots,L_c$ are combined into a single new attribute N_k, which we will call a composite indicator or a composite criterion. The aggregation of several attributes into a composite indicator is the transformation (5) that takes the form

$$Y_a\cup Y_b\cup\dots\cup Y_c \to Z_k, \quad (11)$$

where $Y_j = \{y_j^1,\dots,y_j^{g_j}\}$ is the scale of the original attribute L_j, $j = a,b,\dots, c$, $Z_k = \{z_k^1,\dots,z_k^{f_k}\}$ is the scale of the composite indicator N_k, $k = 1,\dots,l$, $|Z_k| = f_k < g_a + g_b +\dots+ g_c = |Y_a\cup Y_b\cup\dots\cup Y_c|$.

Sets of composite indicators and their scales can be formed by different methods (11) of granulation. This allows us to represent each gradation of the composite indicator scale as a combination of estimate gradations of initial attributes. It is recommended to combine 2-4 original attributes in a composite indicator with a small scale, including 2-4 gradations. In practical tasks, it is convenient to form the scales of the combined attributes and the composite indicator so that they have the same number of gradations, that is $g_a = g_b =\dots= g_c = f_k = d$, and each gradation of the scale of the composite indicator consists of similar gradations of the combined attribute scales.

The representation of multi-attribute objects during the attributes' aggregation is transformed as follows. In the space of original attributes $L_1,\dots,L_m$, let the object O_p, $p = 1,\dots,q$ be defined by a multiset

$$\boldsymbol{I}_p = \{k_{Ip}(y_1^1)\circ y_1^1,\dots,k_{Ip}(y_1^d)\circ y_1^d;\dots; k_{Ip}(y_m^1)\circ y_m^1,\dots,k_{Ip}(y_m^d)\circ y_m^d\} \quad (12)$$

over the set $Y_1\cup\dots\cup Y_m$ of scale gradations, where all scales $Y_j = \{y_j^1,\dots,y_j^d\}$, $j = 1,\dots,m$ have the same number d of gradations. Taking into account that the order of elements in a multiset is inconsequential, we rewrite expression (12) into the form of multiset sums:

$$\boldsymbol{I}_p = \boldsymbol{I}_{p1} +\dots+ \boldsymbol{I}_{pd} = \{k_{Ip}(y_1^1)\circ y_1^1,\dots,k_{Ip}(y_m^1)\circ y_m^1\} +\dots+ \{k_{Ip}(y_1^d)\circ y_1^d,\dots,k_{Ip}(y_m^d)\circ y_m^d\} =$$

$$= \sum_{j=1}^{m} \{ k_{Ip}(y_j^1) \circ y_j^1 \} + \ldots + \sum_{j=1}^{m} \{ k_{Ip}(y_j^d) \circ y_j^d \}. \quad (13)$$

In the reduced space of composite indicators $N_1,\ldots,N_l$, the object O_p will correspond to a multiset

$$\boldsymbol{J}_p = \{k_{Jp}(z_1^1) \circ z_1^1,\ldots,k_{Jp}(z_1^d) \circ z_1^d;\ldots; k_{Jp}(z_l^1) \circ z_l^1,\ldots,k_{Jp}(z_l^d) \circ z_l^d\} \quad (14)$$

over the set $Z_1 \cup \ldots \cup Z_l$ of scale gradations, where all scales $Z_k = \{z_k^1,\ldots,z_k^d\}$, $k = 1,\ldots,l$ have the same number d of gradations. The multiset $\boldsymbol{J}_p$ (14) can also be rewritten in the equivalent form:

$$\boldsymbol{J}_p = \boldsymbol{J}_{p1} + \ldots + \boldsymbol{J}_{pd} = \{k_{Jp}(z_1^1) \circ z_1^1,\ldots,k_{Jp}(z_l^1) \circ z_l^1\} + \ldots + \{k_{Jp}(z_1^d) \circ z_1^d,\ldots,k_{Jp}(z_l^d) \circ z_l^d\} =$$
$$= \sum_{k=1}^{l} \{ k_{Jp}(z_k^1) \circ z_k^1 \} + \ldots + \sum_{k=1}^{l} \{ k_{Jp}(z_k^d) \circ z_k^d \}. \quad (15)$$

The multiplicity of the element z_k^e, $e = 1,\ldots,d$ of the multiset $\boldsymbol{J}_p$ (14) or (15), which corresponds to the gradation z_k^e of the scale Z_k of the composite indicator N_k, is determined by the rule:

$$k_{Ip}(z_k^e) = k_{Ip}(y_a^e) + k_{Ip}(y_b^e) + \ldots + k_{Ip}(y_c^e), \quad (16)$$

where the multiplicities of the elements $y_a^e, y_b^e,\ldots, y_c^e$ of the multiset $\boldsymbol{I}_p$ (12) or (13), which correspond to the gradations $y_a^e, y_b^e,\ldots, y_c^e$ of the scales $Y_a, Y_b,\ldots, Y^e$ of the combined attributes $L_a, L_b,\ldots, L_c$.

An aggregation of attributes is carried out in stages, step by step. At each stage, it is determined which initial attributes should be combined into composite indicators, and which should be considered as independent final indicators. Verbal scales of composite indicators characterize desirable new properties of the compared objects and have a specific semantic content for a decision maker/expert. Consistently combining the attributes, a DM/expert designs acceptable intermediate and final indicators. The tree of attribute aggregation is built from the uniformed blocks that are selected by a DM/expert, and, in fact, is a form of semantic interpretation and granulation of a DM preferences and/or expert knowledge.

In practical situations, it is recommended to design several different schemes of combining indicators that include the procedures for reducing the attribute scale and aggregating attributes. Thus, the influence of each particular scheme is reduced and the validity of the results is increased. Depending on the specifics of the practical problem being solved, the last level of the attribute aggregation tree may consist of several final indicators that implement the idea of multi-criteria choice, or it may be a single integral index that implements the idea of holistic choice [7].

5. Illustrative example

Consider an illustrative example taken from [9, 10]. There are ten objects $O_1,\ldots,O_{10}$, which are described by eight attributes $K_1,\ldots,K_8$ with five-point rating scales $X_i = \{x_i^1, x_i^2, x_i^3, x_i^4, x_i^5\}$, $i = 1,\ldots,8$. Let the objects be pupils, and the attributes be estimates in studied subjects: K_1, Mathematics; K_2, Physics; K_3, Chemistry; K_4, Biology; K_5, Geography; K_6, History; K_7, Literature; K_8, Foreign language. Graduations of rating scales can be numerical or verbal and mean: x_i^1 is 1/very bad; x_i^2 is 2/bad; x_i^3 is 3/satisfactory; x_i^4 is 4/good; x_i^5 is 5/excellent.

Suppose that estimates are given twice a year for each half-year (semester). Then each object is present in two versions (copies) that differ from each other. Present the semi-annual estimates of the pupil O_p, $p = 1,\ldots,10$ (two versions $O_p^{<1>}$, $O_p^{<2>}$ of the object O_p) by the multisets $\boldsymbol{A}_p^{<1>}$, $\boldsymbol{A}_p^{<2>}$ of the form (2) over the set $X = X_1 \cup \ldots \cup X_8$ of gradations of the attribute scales $K_1,\ldots,K_8$. Let us define the annual estimates of the pupil O_p by the multiset $\boldsymbol{A}_p$, which we form as the weighted sum of the multisets describing the versions of the object: $\boldsymbol{A}_p = c^{<1>}\boldsymbol{A}_p^{<1>} + c^{<2>}\boldsymbol{A}_p^{<2>}$. Assuming that the semi-annual estimates are equally significant: $c^{<1>} = c^{<2>} = 1$, we obtain a multiset

$$\boldsymbol{A}_p = \{k_{Ap}(x_1^1) \circ x_1^1,\ldots,k_{Ap}(x_1^5) \circ x_1^5;\ldots; k_{Ap}(x_8^1) \circ x_8^1,\ldots,k_{Ap}(x_8^5) \circ x_8^5\}, \quad (17)$$

the element multiplicities of which form the rows of the matrix H ‘Object–Attributes’ (Table 1).

For example, the annual estimates of the pupil O_1 correspond to a multiset

$$\boldsymbol{A}_1 = \{0 \circ x_1^1, 0 \circ x_1^2, 0 \circ x_1^3, 1 \circ x_1^4, 1 \circ x_1^5; 0 \circ x_2^1, 0 \circ x_2^2, 0 \circ x_2^3, 0 \circ x_2^4, 2 \circ x_2^5;$$
$$0 \circ x_3^1, 0 \circ x_3^2, 0 \circ x_3^3, 1 \circ x_3^4, 1 \circ x_3^5; 0 \circ x_4^1, 0 \circ x_4^2, 0 \circ x_4^3, 0 \circ x_4^4, 2 \circ x_4^5;$$

$0 \circ x_5^1, 0 \circ x_5^2, 0 \circ x_5^3, 2 \circ x_5^4, 0 \circ x_5^5; 0 \circ x_6^1, 0 \circ x_6^2, 0 \circ x_6^3, 1 \circ x_6^4, 1 \circ x_6^5;$
$0 \circ x_7^1, 0 \circ x_7^2, 0 \circ x_7^3, 2 \circ x_7^4, 0 \circ x_7^5; 0 \circ x_8^1, 0 \circ x_8^2, 0 \circ x_8^3, 0 \circ x_8^4, 2 \circ x_8^5\}.$

This form of recording shows that, over the year, the pupil O_1 received one estimate "good" and one estimate "excellent" in mathematics, chemistry, history; two estimates "excellent" in physics, biology, geography, literature, and a foreign language. The pupil O_1 did not receive other estimates.

Table 1
Matrix H 'Objects–Attributes' (initial scales of attributes)

$O \backslash X$	x_1^1	x_1^2	x_1^3	x_1^4	x_1^5	x_2^1	x_2^2	x_2^3	x_2^4	x_2^5	x_1^1	x_3^2	x_3^3	x_3^4	x_3^5	x_4^1	x_4^2	x_4^3	x_4^4	x_4^5
$\boldsymbol{A}_1$	0	0	0	1	1	0	0	0	0	2	0	0	0	1	1	0	0	0	0	2
$\boldsymbol{A}_2$	0	0	1	1	0	1	1	0	0	0	1	1	0	0	0	2	0	0	0	0
$\boldsymbol{A}_3$	2	0	0	0	0	1	1	0	0	0	0	0	2	0	0	2	0	0	0	0
$\boldsymbol{A}_4$	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	0	0	0	1	1
$\boldsymbol{A}_5$	0	0	0	1	1	0	0	0	1	1	0	0	1	1	0	0	0	0	2	0
$\boldsymbol{A}_6$	0	0	0	1	1	0	0	0	0	2	0	0	0	2	0	0	0	0	2	0
$\boldsymbol{A}_7$	0	0	1	1	0	1	1	0	0	0	1	1	0	0	0	0	0	1	1	0
$\boldsymbol{A}_8$	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	1	1	0	0
$\boldsymbol{A}_9$	0	0	1	1	0	0	1	1	0	0	0	1	1	0	0	1	1	0	0	0
$\boldsymbol{A}_{10}$	0	0	1	0	1	0	0	0	1	1	0	0	1	1	0	0	0	0	1	1

$O \backslash X$	x_5^1	x_5^2	x_5^3	x_5^4	x_5^5	x_6^1	x_6^2	x_6^3	x_6^4	x_6^5	x_7^1	x_7^2	x_7^3	x_7^4	x_7^5	x_8^1	x_8^2	x_8^3	x_8^4	x_8^5
$\boldsymbol{A}_1$	0	0	0	2	0	0	0	0	1	1	0	0	0	2	0	0	0	0	0	2
$\boldsymbol{A}_2$	0	0	1	1	0	0	1	1	0	0	0	1	1	0	0	0	2	0	0	0
$\boldsymbol{A}_3$	0	0	0	1	1	1	1	0	0	0	2	0	0	0	0	0	0	1	1	0
$\boldsymbol{A}_4$	0	0	0	2	0	0	0	0	0	2	0	0	1	1	0	0	0	0	1	1
$\boldsymbol{A}_5$	0	0	0	2	0	0	0	0	1	1	0	0	0	1	1	0	0	0	2	0
$\boldsymbol{A}_6$	0	0	0	2	0	0	0	0	1	1	0	0	0	0	2	0	0	0	1	1
$\boldsymbol{A}_7$	0	1	1	0	0	0	0	1	1	0	1	1	0	0	0	0	1	1	0	0
$\boldsymbol{A}_8$	0	0	1	1	0	0	0	0	1	1	0	0	0	1	1	0	0	1	1	0
$\boldsymbol{A}_9$	0	1	1	0	0	0	0	2	0	0	0	1	1	0	0	0	0	2	0	0
$\boldsymbol{A}_{10}$	0	1	1	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0	2	0

The dimension of the attribute space is equal to $|X| = 5 \cdot 8 = 40$. The total number of object representations by elements of multisets (possible estimates in all studied subjects) is equal to the cardinality of a multiset $\boldsymbol{A}_p$ (17): $\text{card}\boldsymbol{A}_p = \sum_{x_i \in X} k_{Ap}(x_i^{e_i}) = 16$. Objects and multisets are generally incomparable.

Replace the initial five-point rating scales $X_i = \{x_i^1, x_i^2, x_i^3, x_i^4, x_i^5\}$ with the shortened three-point rating scales $Q_i = \{q_i^0, q_i^1, q_i^2\}$. Here q_i^0 is 0/high mark, including the estimates x_i^5 – 5/excellent and x_i^4 – 4/good; q_i^1 is 1/middle mark corresponding to the estimate x_i^3 – 3/satisfactory; q_i^2 is 2/low mark, including the estimates x_i^2 – 2/bad and x_i^1 – 1/very bad. If the initial gradations were ordered by preference, for example, $x_i^5 \succ x_i^4 \succ x_i^3 \succ x^2 \succ x_i^1$, then the new gradations will also be ordered as $q_i^0 \succ q_i^1 \succ q_i^2$.

When transiting from the five-point rating scales X_i to the three-point rating scales Q_i, $i = 1,\ldots,8$, the object O_p will correspond to a multiset

$$\boldsymbol{B}_p = \{k_{\boldsymbol{B}p}(q_1^0) \circ q_1^0, k_{\boldsymbol{B}p}(q_1^1) \circ q_1^1, k_{\boldsymbol{B}p}(q_1^2) \circ q_1^2; \ldots; k_{\boldsymbol{B}p}(q_8^0) \circ q_8^0, k_{\boldsymbol{B}p}(q_8^1) \circ q_8^1, k_{\boldsymbol{B}p}(q_8^2) \circ q_8^2\} \quad (18)$$

over the set $Q = Q_1 \cup \ldots \cup Q_8$ of shortened gradations of the attribute scales $K_1,\ldots,K_8$. The multiplicities of elements of the multiset $\boldsymbol{B}_p$ (18) form the rows of the matrix H_0 'Object–Attributes' (Table 2), which is a reduced matrix H (Table 1), and are determined by the rules (10):

$$k_{\boldsymbol{B}p}(q_i^0) = k_{Ap}(x_i^5) + k_{Ap}(x_i^4),\ k_{\boldsymbol{B}p}(q_i^1) = k_{Ap}(x_i^3),\ k_{\boldsymbol{B}p}(q_i^2) = k_{Ap}(x_i^2) + k_{Ap}(x_i^1).$$

In particular, the object O_1 is defined by a multiset

$$\boldsymbol{B}_1 = \{2\circ q_1^0, 0\circ q_1^1, 0\circ q_1^2; 2\circ q_2^0, 0\circ q_2^1, 0\circ q_2^2; 2\circ q_3^0, 0\circ q_3^1, 0\circ q_3^2; 2\circ q_4^0, 0\circ q_4^1, 0\circ q_4^2; \\ 2\circ q_5^0, 0\circ q_5^1, 0\circ q_5^2; 2\circ q_6^0, 0\circ q_6^1, 0\circ q_6^2; 2\circ q_7^0, 0\circ q_7^1, 0\circ q_7^2; 2\circ q_8^0, 0\circ q_8^1, 0\circ q_8^2\}.$$

We show that, over a year, the pupil O_1 received two high marks (estimates “excellent” and “good”) in all studied subjects: mathematics, physics, chemistry, biology, geography, history, literature, and a foreign language.

Table 2

Matrix H_0 ‘Objects–Attributes’ (shortened scales of attributes)

$O\backslash Q$	q_1^0	q_1^1	q_1^2	q_2^0	q_2^1	q_2^2	q_1^0	q_3^1	q_3^2	q_4^0	q_4^1	q_4^2	q_5^0	q_5^1	q_5^2	q_6^0	q_6^1	q_6^2	q_7^0	q_7^1	q_7^2	q_8^0	q_8^1	q_8^2
$\boldsymbol{B}_1$	2	0	0	2	0	0	2	0	0	2	0	0	2	0	0	2	0	0	2	0	0	2	0	0
$\boldsymbol{B}_2$	1	1	0	0	0	2	0	0	2	0	0	2	1	1	0	0	1	1	0	1	1	0	0	2
$\boldsymbol{B}_3$	0	0	2	0	0	2	0	2	0	0	0	2	2	0	0	0	0	2	0	0	2	1	1	0
$\boldsymbol{B}_4$	2	0	0	1	1	0	0	1	1	2	0	0	2	0	0	2	0	0	1	1	0	2	0	0
$\boldsymbol{B}_5$	2	0	0	2	0	0	1	1	0	2	0	0	2	0	0	2	0	0	2	0	0	2	0	0
$\boldsymbol{B}_6$	2	0	0	2	0	0	2	0	0	2	0	0	2	0	0	2	0	0	2	0	0	2	0	0
$\boldsymbol{B}_7$	1	1	0	0	0	2	0	0	2	1	1	0	0	1	1	1	1	0	0	0	2	0	1	1
$\boldsymbol{B}_8$	2	0	0	2	0	0	2	0	0	0	1	1	1	1	0	2	0	0	2	0	0	1	1	0
$\boldsymbol{B}_9$	1	1	0	0	1	1	0	1	1	0	0	2	0	1	1	0	2	0	0	1	1	0	2	0
$\boldsymbol{B}_{10}$	1	1	0	2	0	0	1	1	0	2	0	0	0	1	1	2	0	0	1	0	1	2	0	0

The dimension of the reduced attribute space is equal to $|Q| = 3{\cdot}8 = 24$, and the total number of object representations is expressed by the cardinality of a multiset $\boldsymbol{B}_p$ (18): $\mathrm{card}\boldsymbol{B}_p = \sum_{q_i\in Q} k_{\boldsymbol{B}}(q_i^{e_i}) = 16$. When the attribute scales are reduced, the dimension of the converted space decreases, and the total number of estimates in the studied subjects does not change. Objects and multisets are still largely incomparable. But a work with them is becoming easier.

We shall consider the transition from the original scales X_i to the shortened scales q_i as the zero scheme for aggregating attributes. To represent objects in reduced attribute spaces, we shall build other collections of indicators with different schemes of aggregating characteristics (Fig. 1). For simplicity, we assume that the scale of any new attribute has three gradations of estimates, as well as the scale Q_i. Each gradation of the scale of a composite indicator includes combinations of the same type of gradations on the scales of the initial attributes.

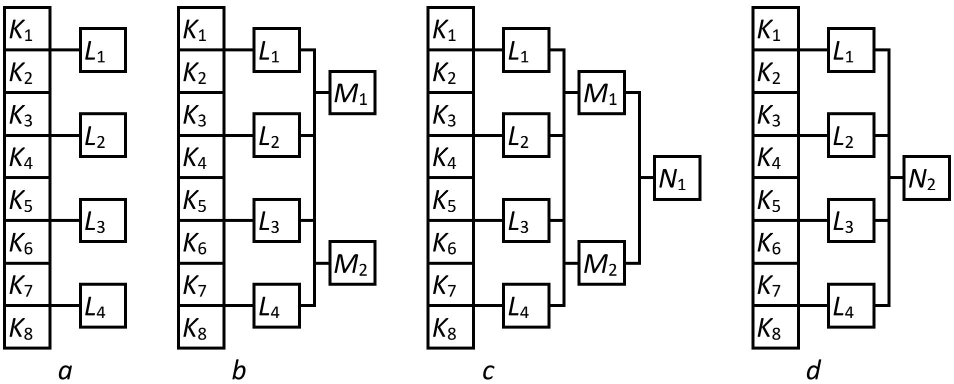

Figure 1: Aggregation of initial characteristics into composite indicators:
(*a*) the first scheme, (*b*) the second scheme, (*c*) the third scheme, (*d*) the fourth scheme.

According to the first aggregation scheme (Fig. 1, *a*), all the initial attributes $K_1,\ldots,K_8$ with scales $Q_i = \{q_i^0, q_i^1, q_i^2\}$ are combined into composite indicators, which are considered as final. The attributes K_1, Mathematics, and K_2, Physics form a composite indicator L_1, Physical-Mathematical disciplines: $L_1 = (K_1, K_2)$. The attributes K_3, Chemistry, and K_4, Biology form a composite indicator L_2, Chemical-Biological disciplines: $L_2 = (K_3, K_4)$. The attributes K_5, Geography, and K_6, History form a composite indicator L_3, Socio-Historical disciplines: $L_3 = (K_5, K_6)$. The attributes K_7, Literature, and K_8, Foreign language form a composite indicator L_4, Philological disciplines: $L_4 = (K_7, K_8)$. The composite indicators $L_1,\ldots,L_4$ have scales $Y_j = \{y_j^0, y_j^1, y_j^2\}$, $j = 1,2,3,4$ with the following verbal gradations: y_j^0 is 0/high mark, including estimates q_a^0, q_c^0; y_j^1 is 1/middle mark, including estimates q_a^1, q_c^1; y_j^2 is 2/low mark, including the estimates q_a^2, q_c^2. Here $a = 1$, $c = 2$ for $j = 1$; $a = 3$, $c = 4$ for $j = 2$; $a = 5$, $c = 6$ for $j = 3$; $a = 7$, $c = 8$ for $j = 4$.

Each object O_p is represented by a multiset

$$\boldsymbol{C}_p = \{k_{\boldsymbol{C}p}(y_1^0)\circ y_1^0, k_{\boldsymbol{C}p}(y_1^1)\circ y_1^1, k_{\boldsymbol{C}p}(y_1^2)\circ y_1^2;\ldots; k_{\boldsymbol{C}p}(y_4^0)\circ y_4^0, k_{\boldsymbol{C}p}(y_4^1)\circ y_4^1, k_{\boldsymbol{C}p}(y_4^2)\circ y_4^2\} \quad (19)$$

over the set $Y = Y_1 \cup \ldots \cup Y_4$ of gradations of the indicator scales $L_1,\ldots,L_4$. The multiplicities of elements of the multiset $\boldsymbol{C}_p$ (19) form the rows of the matrix H_1 'Object–Attributes' (Table 3) and are determined by rule (16) for forming the scales of composite indicators $L_1,\ldots,L_4$ from the scales of attributes $K_1,\ldots,K_8$. In particular, the object O_1 is defined by a multiset

$$\boldsymbol{C}_1 = \{4\circ y_1^0, 0\circ y_1^1, 0\circ y_1^2; 4\circ y_2^0, 0\circ y_2^1, 0\circ y_2^2; 4\circ y_3^0, 0\circ y_3^1, 0\circ y_3^2; 4\circ y_4^0, 0\circ y_4^1, 0\circ y_4^2\}.$$

We show that, over a year, the pupil O_1 received four high marks in physical-mathematical, chemical-biological, socio-historical and philological disciplines.

Table 3

Matrix H_1 'Objects–Attributes' (the first aggregation scheme)

$O \backslash Y$	y_1^0	y_1^1	y_1^2	y_2^0	y_2^1	y_2^2	y_1^0	y_3^1	y_3^2	y_4^0	y_4^1	y_4^2	$l(O_p)$	$s(O_p)$	$p(O_p)$	$b(O_p)$
$\boldsymbol{C}_1$	4	0	0	4	0	0	4	0	0	4	0	0	0,000	48	1-2	25,5
$\boldsymbol{C}_2$	1	1	2	0	0	4	1	2	1	0	1	3	0,700	24	9	1
$\boldsymbol{C}_3$	0	0	4	0	2	2	2	0	2	1	1	2	0,684	25	8	4
$\boldsymbol{C}_4$	3	1	0	2	1	1	4	0	0	3	1	0	0,211	43	4-5	16,5
$\boldsymbol{C}_5$	4	0	0	3	1	0	4	0	0	4	0	0	0,059	47	3	21
$\boldsymbol{C}_6$	4	0	0	4	0	0	4	0	0	4	0	0	0,000	48	1-2	25,5
$\boldsymbol{C}_7$	1	1	2	1	1	2	1	2	1	0	1	3	0,565	27	7	7,5
$\boldsymbol{C}_8$	4	0	0	2	1	1	3	1	0	3	1	0	0,211	43	4-5	16,5
$\boldsymbol{C}_9$	1	2	1	0	1	3	0	3	1	0	3	1	0,556	27	10	5,5
$\boldsymbol{C}_{10}$	3	1	0	3	1	0	2	1	1	3	0	1	0,253	41	6	12

According to the second aggregation scheme (Fig. 1, *b*), the first stage is the same as in the first scheme. In the next stage, the indicators L_1, Physical-Mathematical disciplines and L_2, Chemical-Biological disciplines form a composite indicator M_1, Natural disciplines: $M_1 = (L_1, L_2)$. The indicators L_3, Socio-Historical disciplines and L_4, Philological disciplines form a composite indicator M_2, Humanitarian disciplines: $M_2 = (L_3, L_4)$. The composite indicators M_1, M_2 are considered as final, which have scales $U_r = \{u_r^0, u_r^1, u_r^2\}$, $r = 1,2$ with the following verbal gradations: u_r^0 is 0/high mark, including estimates y_b^0, y_d^0; u_r^1 is 1/middle mark, including estimates y_b^1, y_d^1; u_r^2 is2/low mark, including the estimates y_b^2, y_d^2. Here $b = 1$, $d = 2$ for $r = 1$; $b = 3$, $d = 4$ for $r = 2$.

Each object O_p is represented by a multiset

$$\boldsymbol{D}_p = \{k_{\boldsymbol{D}p}(u_1^0)\circ u_1^0, k_{\boldsymbol{D}p}(u_1^1)\circ u_1^1, k_{\boldsymbol{D}p}(u_1^2)\circ u_1^2; k_{\boldsymbol{D}p}(u_2^0)\circ u_2^0, k_{\boldsymbol{D}p}(u_2^1)\circ u_2^1, k_{\boldsymbol{D}p}(u_2^2)\circ u_2^2\} \quad (20)$$

over the set $U = U_1 \cup U_2$ of gradations of the indicator scales M_1, M_2. The multiplicities of elements of the multiset $\boldsymbol{D}_p$ (20) form the rows of the matrix H_2 'Object–Attributes' (Table 4) and are determined by rule (16) for forming the scales of composite indicators M_1, M_2 from the scales of indicators $L_1,\ldots,L_4$. In particular, the object O_1 is defined by a multiset $\boldsymbol{D}_1 = \{8\circ u_1^0, 0\circ u_1^1, 0\circ u_1^2; 8\circ u_2^0, 0\circ u_2^1, 0\circ u_2^2\}$. This shows that, over a year, the pupil O_1 received eight high marks in natural and humanitarian disciplines.

According to the third aggregation scheme (Fig. 1, *c*), the first and second stages are the same as in the second scheme. In the next stage, the indicators M_1, Natural disciplines and M_2, Humanitarian disciplines form a final integral index N_1, Academic score: $N_1 = (M_1, M_2)$, which have a scale

$Z_1 = \{z_1^0, z_1^1, z_1^2\}$ with the following verbal gradations: z_1^0 is 0/high mark, including estimates u_1^0, u_2^0; z_1^1 is 1/middle mark, including estimates u_1^1, u_2^1; z_1^2 is 2/low mark, including the estimates u_1^2, u_2^2.

Each object O_p is represented by a multiset

$$\boldsymbol{E}_p = \{k_{\boldsymbol{E}p}(z_1^0)\circ z_1^0,\ k_{\boldsymbol{E}p}(z_1^1)\circ z_1^1,\ k_{\boldsymbol{E}p}(z_1^2)\circ z_1^2\} \quad (21)$$

over the set Z_1 of gradations of the indicator scale N_1. The multiplicities of elements of the multiset $\boldsymbol{E}_p$ (21) form the rows of the matrix H_3 'Object–Attributes' (Table 4) and are determined by rule (16) for forming the scale of composite indicator N_1 from the scales of indicators M_1, M_2. In particular, the object O_1 is defined by a multiset $\boldsymbol{E}_1 = \{16\circ z_1^0, 0\circ z_1^1, 0\circ z_1^2\}$. We show that, over a year, the pupil O_1 received sixteen high marks in all studied disciplines.

According to the fourth aggregation scheme (Fig. 1, *d*), the first and second stages are the same as in the first scheme. In the next stage, the indicators L_1, Physical-Mathematical disciplines, L_2, Chemical-Biological disciplines, L_3, Socio-Historical disciplines, and L_4, Philological disciplines form a final integral index N_2, Academic score: $N_2 = (L_1, L_2, L_3, L_4)$, which have a scale $Z_2 = \{z_2^0, z_2^1, z_2^2\}$ with the following verbal gradations: z_2^0 is 0/high mark, including estimates y_1^0, y_2^0, y_3^0, y_4^0; z_2^1 is 1/middle mark, including estimates y_1^1, y_2^1, y_3^1, y_4^1; z_2^2 is 2/low mark, including the estimates y_1^2, y_2^2, y_3^2, y_4^2.

Each object O_p is represented by a multiset

$$\boldsymbol{F}_p = \{k_{\boldsymbol{F}p}(z_2^0)\circ z_2^0,\ k_{\boldsymbol{F}p}(z_2^1)\circ z_2^1,\ k_{\boldsymbol{F}p}(z_2^2)\circ z_2^2\} \quad (22)$$

over the set Z_2 of gradations of the indicator scale N_2. The multiplicities of elements of the multiset $\boldsymbol{F}_p$ (22) form the rows of the matrix H_4 'Object–Attributes' (Table 4) and are determined by rule (16) for forming the scale of composite indicator N_2 from the scales of indicators L_1, L_2, L_3, L_4. In particular, the object O_1 is defined by a multiset $\boldsymbol{F}_1 = \{16\circ z_2^0, 0\circ z_2^1, 0\circ z_2^2\}$. We show that, over a year, the pupil O_1 received sixteen high marks in all studied disciplines.

Table 4
Matrices 'Objects–Attributes'

H_2 (the second aggregation scheme)							H_3 (the third aggregation scheme)				H_4 (the fourth aggregation scheme)			
$O\backslash U$	u_1^0	u_1^1	u_1^2	u_2^0	u_2^1	u_2^2	$O\backslash Z$	z_1^0	z_1^1	z_1^2	$O\backslash Z$	z_2^0	z_2^1	z_2^2
$\boldsymbol{D}_1$	8	0	0	8	0	0	$\boldsymbol{E}_1$	16	0	0	$\boldsymbol{F}_1$	16	0	0
$\boldsymbol{D}_2$	1	1	6	1	3	4	$\boldsymbol{E}_2$	2	4	10	$\boldsymbol{F}_2$	2	4	10
$\boldsymbol{D}_3$	0	2	6	3	1	4	$\boldsymbol{E}_3$	3	3	10	$\boldsymbol{F}_3$	3	3	10
$\boldsymbol{D}_4$	5	2	1	7	1	0	$\boldsymbol{E}_4$	12	3	1	$\boldsymbol{F}_4$	12	3	1
$\boldsymbol{D}_5$	7	1	0	8	0	0	$\boldsymbol{E}_5$	15	1	0	$\boldsymbol{F}_5$	15	1	0
$\boldsymbol{D}_6$	8	0	0	8	0	0	$\boldsymbol{E}_6$	16	0	0	$\boldsymbol{F}_6$	16	0	0
$\boldsymbol{D}_7$	2	2	4	1	3	4	$\boldsymbol{E}_7$	3	5	8	$\boldsymbol{F}_7$	3	5	8
$\boldsymbol{D}_8$	6	1	1	6	2	2	$\boldsymbol{E}_8$	12	3	1	$\boldsymbol{F}_8$	12	3	1
$\boldsymbol{D}_9$	1	3	4	0	6	2	$\boldsymbol{E}_9$	1	9	6	$\boldsymbol{F}_9$	1	9	6
$\boldsymbol{D}_{10}$	6	2	0	5	1	2	$\boldsymbol{E}_{10}$	11	3	2	$\boldsymbol{F}_{10}$	11	3	2

The indicators can also be aggregated in another way. For example, the attributes K_1, Mathematics, K_2, Physics, K_3, Chemistry, K_4, Biology form a composite indicator M_3. Natural disciplines: $M_3 = (K_1, K_2, K_3, K_4)$. The attributes K_5, Geography, K_6, History, K_7, Literature, K_8, Foreign language form a composite indicator M_4. Humanitarian disciplines: $M_4 = (K_5, K_6, K_7, K_8)$. The composite indicators M_3 and M_4 can either be considered as final indicators, or combined further into an integral index N_3, Academic score: $N_3 = (M_3, M_4)$. Other options for aggregating indicators are also possible. When forming aggregation schemes, it is advisable to combine the initial attributes into a composite indicator so that it makes sense, and the gradations of its scale consisted of a small number of combinations of the initial gradations.

So, in the transition from the initial data to aggregated indicators, the dimension of the transformed spaces decreases sequentially from 40 to 24, 12, 6, 3, but the total number of estimates in all studied subjects, that is expressed by the cardinality of the multisets $\boldsymbol{A}_p$ (17), $\boldsymbol{B}_p$ (18), $\boldsymbol{C}_p$ (19), $\boldsymbol{D}_p$ (20), $\boldsymbol{E}_p$ (21), $\boldsymbol{F}_p$ (22), does not change.

We can assume that the five constructed schemes for aggregating characteristics are the judgments of five independent experts. In this case, any multi-criteria choice problem becomes a collective choice problem, which is solved in various reduced spaces of attributes, using several different methods in each space. This ensures greater validity of the final results.

Let us illustrate the suggested technique, considering rankings of objects $O_1,...,O_{10}$ obtained with the multi-method technology PAKS-M (Progressive Aggregation of the Classified Situations by many Methods) for multi-criteria choice in the attribute space of large dimension [10]. Firstly, for each aggregation scheme, we built collective rankings of objects using three methods of group choice: ARAMIS, weighted sum of estimates, lexicographic ordering [7, 8].

The ARAMIS (Aggregation and Ranking of Alternatives close to Multi-attribute Ideal Situations) method allows to rank multi-attribute objects, evaluated by several experts upon many quantitative and/or qualitative criteria $K_1,\ldots,K_n$, without building individual rankings of objects. The objects are ordered in the Petrovsky metric space of multisets according to the value of the indicator $l(O_p)$ of the relative proximity of the object O_p to the best (possibly hypothetical) object O_+, to which all experts gave the highest grades by all criteria.

The method of weighted sums of estimates allows us to rank multi-attribute objects by the values of their value functions. The value of the object O_p is given by the sum $s(O_p)$ of the products of numbers of the scale gradations by the weight of gradation. In the example considered, the high gradation was assigned the weight 3, the middle gradation – the weight 2, the low gradation – the weight 1.

The method of lexicographic ordering allows us to rank multi-attribute objects by the total number of corresponding estimate gradations. The place $p(O_p)$ of the object O_p in the ranking is determined firstly by the number of the high marks; then by the number of the middle marks if several objects have the same number of high marks; further by the number of the low marks if several objects have the same number of middle marks, etc.

For all five schemes for aggregating characteristics, the results of data processing by each of the above methods turned out to be the same. They are shown in Table 3. In other words, the judgments of all five independent experts based on any method of choice coincided. This was a consequence of the additivity of rules (10) and (16) for the transformation of attribute scales. The collective rankings of objects, obtained according to any of the five schemes using the methods of ARAMIS $R_{\mathrm{A}}{}^{gr}$, weighted sum of estimates $R_{\Sigma}{}^{gr}$, lexicographic ordering $R_{\Lambda}{}^{gr}$, look like this:

$$R_{\mathrm{A}}{}^{gr} \Leftrightarrow [O_1, O_6 \succ O_5] \succ [O_4, O_8 \succ O_{10}] \succ [(O_9 \succ O_7) \succ O_3 \succ O_2],$$

$$R_{\Sigma}{}^{gr} \Leftrightarrow [O_1, O_6 \succ O_5] \succ [O_4, O_8 \succ O_{10}] \succ [O_7, O_9 \succ (O_3 \succ O_2)],$$

$$R_{\Lambda}{}^{gr} \Leftrightarrow [O_1, O_6 \succ O_5] \succ [O_4, O_8 \succ O_{10}] \succ [O_7 \succ O_3 \succ O_2 \succ O_9].$$

The rankings of objects according to the methods of ARAMIS, weighted sum of estimates, lexicographic ordering can also be considered as the judgments of three other experts. We will combine the opinions of these new experts using the Borda voting procedure [7], according to which the order of objects in the final group ordering is given by the sum $b(O_p)$ of Borda points (Table 3) The generalized group ranking of objects, that combines the rankings $R_{\mathrm{A}}{}^{gr}$, $R_{\Sigma}{}^{gr}$, $R_{\Lambda}{}^{gr}$, has the form:

$$R_{\mathrm{B}}{}^{gr} \Leftrightarrow [O_1, O_6 \succ O_5] \succ [O_4, O_8 \succ O_{10}] \succ [O_7 \succ (O_9 \succ O_3) \succ O_2].$$

Near objects are enclosed in the round brackets, groups of distant objects are enclosed in the square brackets.

Thus, the final orderings of objects obtained in different ways, or, equivalently, the collective preferences of many different groups of experts (several versions of objects, schemes for aggregating attributes, methods for choosing objects) almost completely coincide, with the exception of small differences in the placements of objects included in the last group. In all the ratings, there are identical groups of “good objects” O_1, O_6, O_5 with the high marks, “medium objects” O_4, O_8, O_{10} with the middle marks, almost identical groups of “bad objects” O_7, O_9, O_3 with the low marks.

According to the aggregated estimates of all experts upon all attributes, the best objects are O_1, O_6, which take the first places in all rankings. The worst object is O_2, which takes the last places in three rankings and the penultimate place in one ranking. The gaps between the groups "good objects", "medium objects" and "bad objects" are clearly expressed. Therefore, group orderings of objects can also be considered as group ordinal classifications, where the classes of objects and the places of objects into the classes are given by the corresponding rankings. Exactly the same rankings of objects $O_1,...,O_{10}$ were obtained by other author's method for reducing the dimension of attribute space [10].

In practical problems of multi-criteria choice, a DM/expert may encounter inconsistency and contradiction of the results. Such situations can be caused by various reasons, in particular, the formal combination of attributes, the unsuccessful formation of gradations on the scales of composite criteria and integral index, or poor semantic relationships between the attributes and indicators.

6. Conclusions

Solving the problems of multicriteria choice in the reduced attribute spaces significantly diminishes the labor costs of a decision maker/expert and substantively explains the choice made. The new method SOCRATES for reducing the dimension of the attributes space has certain universality, as it allows to operate simultaneously with both verbal (qualitative) and numerical (quantitative) data. An attractive feature of the method is the possibility to use it in combination with various decision-making methods and information processing technologies. And most attractively, the initially available information is not distorted or lost.

The proposed method SOCRATES is easily integrated into the original multi-stage technology PAKS (Progressive Aggregation of the Classified Situations) and multi-method technology PAKS-M (Progressive Aggregation of the Classified Situations by many Methods) [10] for solving problems of multi-criteria choice in large-dimensional spaces. These technologies provide greater validity for choosing the most preferable object and have the following important features. They form several schemes with different options for aggregating attributes, in which the gradations of the scale of a composite indicator are presented as combinations of grades of the original attributes. The problem considered is solved by several methods of multi-criteria choice. An understandable explanation of the obtained results helps a DM/expert to find the most suitable scheme for aggregating attributes, or to apply several schemes together.

Technologies for solving multi-criteria choice problems in large-dimensional spaces were used to evaluate the results of scientific research, rate organizations by the effectiveness of activities, select a prospective personal computing complex [10]. The use of the new SOCRATES method will vastly reduce the complexity and time of solving similar practical problems.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (projects 17-29-07021, 18-07-00132, 18-07-00280, 19-29-01047)

References

[1] S.A. Ayvazyan, V.M. Bukhshtaber, I.S. Enyukov, L.D. Meshalkin, Applied Statistics. Classification and Reduction of Dimension, Finansy i statistika, Moscow, 1989, in Russian.
[2] V.A. Glotov, V.V. Pavel'yev, Vector Stratification, Nauka, Moscow, 1984, in Russian.
[3] J.A. Hartigan, Clustering Algorithms. Wiley, New York, NY, 1975.
[4] D. Kahneman, P. Slovik, A. Tversky, Decision-Making in Uncertainty: Heuristics and Biases, Cambridge University Press, Cambridge, 1982.
[5] O.I. Larichev, Verbal Decision Analysis, Nauka, Moscow, 2006, in Russian.
[6] V.D. Noghin, Reduction of the Pareto Set. An Axiomatic Approach, volume 126 of Studies in Systems, Decision and Control, Springer International Publishing, 2018. doi:10.1007/978-3-319-67873-3.

[7] A.B. Petrovsky, Decision Making Theory, Publishing Center “Academiya”, Moscow, 2009, in Russian.
[8] A.B. Petrovsky, Group verbal decision analysis, in: F. Adam, P. Humphreys (Eds.), Encyclopedia of Decision Making and Decision Support Technologies, IGI Global, Hershey, NY, 2008, pp. 418-425.
[9] A.B. Petrovsky, Indicators of similarity and differences of multi-attribute objects in the metric spaces of sets and multisets, Scientific and Technical Information Processing, 45(5) (2018) 331-345. doi:10.3103/S0147688218050052.
[10] A.B. Petrovsky, Group Verbal Decision Analysis, Nauka, Moscow, 2019, in Russian.
[11] H. Samet, Foundation of Multidimensional and Metric Data Structures, Elsevier, Boston, 2006.