Multiple Criteria Ranking Enterprises Based on Inconsistent Estimations[1]

Alexey B.Petrovsky

Institute for System Analysis
Russian Academy of Sciences
Prospect 60 Let Octyabrya, 9
Moscow 117312, Russia
pab@isa.ru

Abstract: The paper considers a technique for ordering objects that are described with many quantitative and/or qualitative attributes and may exist in several copies. Such multi-attribute objects may be represented as multisets or sets with repeating elements. Object ordering is based on the theory of multiset metric spaces and applied to ranking enterprises, which are estimated differently by several decision makers/experts under multiple criteria.

1 Introduction

Traditionally in economic analysis, the business activity and enterprises' positions of in any market sector are evaluated with various economical, financial, organizational, operational, and other indicators. Some of indicators are estimated by experts, whereas the others have to be measured or calculated [GO01]. A list of indicators and their values depend on the aim of analysis. Often in real-life situations, one needs to compare and arrange enterprises from the best to the worst by their activity indicators. Ranking enterprises, which are estimated differently by several experts under multiple criteria, is one of the main decision aiding problems. But how to solve this problem when manifold indicators describe an enterprise activity and a convolution of numerical and/or verbal estimates is either impossible or mathematically incorrect.

Let $A=\{A_1,...,A_n\}$ be a collection of n objects (enterprises, companies, etc) which are evaluated under m quantitative and/or qualitative criteria $Q_1,Q_2,...,Q_m$. For example, in the case of ranking enterprises, an enterprise activity is evaluated as follows: Q_1.Total profit; Q_2.Gross sales; Q_3.Net sales; Q_4.Number of projects; Q_5.Project importance; Q_6.Number of employees; Q_7.Professional skills of staff, and so on. Different criteria

[1] This work is partially supported by the Russian Academy of Sciences, projects of Scientific Programs "Basic Problems of Informatics and Information Technologies", "Bases of Information Technologies and Systems", the Russian Foundation for Basic Research, projects 04-01-00290, 05-01-00666, 06-07-89352.

may have different relative importance (weights) in various problem situations. A scale of criterion estimates $Q_s=\{q_s^1, q_s^2,\ldots, q_s^{h_s}\}$, $s=1,\ldots,m$ may be discrete or continuous. Qualitative estimates are nominative or ordinal. Ordinal estimates are supposed to be ordered from the best to the worst as $q_s^1 \phi q_s^2 \phi \ldots \phi q_s^{h_s}$. Thus, the scale of the criterion Q_7."Professional skills of staff" may look as follows:

q_7^1 – a staff with a very high qualification level and experience;
q_7^2 – a staff with a high qualification level and experience;
q_7^3 – a staff with a medium qualification level and experience;
q_7^4 – a staff with a small qualification level and experience.

Sometimes, for a simplicity and convenience of object's evaluation and comparison, it is useful to transform a quantitative continuous scale into a qualitative discrete scale with a small number of grades. For instance, the criterion scale Q_s may be the following: q_s^1 – very large (high); q_s^2 – large (high); q_s^3 – medium; q_s^4 – small. On the other hand, verbal estimates are never converted in numerical ones.

When an enterprise A_i is evaluated several times (for instance, monthly) or by several experts, a description of enterprise activity consists of many different quantitative and/or qualitative estimations $q_s^{e_s} \in Q_s$, $e_s=1,\ldots,h_s$. And these estimations may be repeated within the description and be diverse, inconsistent and even contradictory. Various methods for ranking multi-attribute objects have been developed [KS72], [Mi80], [HY81], [ZZG84], [PS94], [Roy96], [LM97]. But the most of methods operates with numerical attributes and don't pay attention to a discordance of objects' descriptions. So, special techniques are to be elaborated for ranking multi-attribute objects with peculiarities of such kinds.

In the paper, a method for ordering multi-attribute objects that are described with inconsistent and contradictory verbal attributes is considered. Multi-attribute objects are represented as multisets or sets with repeating elements and arranged by their closeness with regard to any "ideal" object in a multiset metric space [Pe94], [Pe03a,b]. The suggested technique is applied to a practical problem of ranking enterprises [Who00].

2 Problems of ranking multi-attribute objects

Let us discuss some of ranking techniques. Direct sorting objects is very popular due to its simplicity for a human being. Every object, which is estimates under a numerical criterion, is assigned one of the given classes immediately. In the case of several persons, the final ordering objects is more complicate and to be constructed by the Kemeny median or weighted averages, if a concordance of estimates is acceptable [KS72].

Binary relations are widely used for an objects' arrangement. In the pairwise comparisons, the final ordering objects will be complete if all pairs of objects are comparable, and decision maker/expert's preferences are transitive. If objects are incomparable, then ordering will be partial. In the case of multiple criteria and/or several persons, a set of comparison matrixes will be very large. So, the final arrangement of objects by processing matrixes is rather difficult [HY81].

Ranking objects may be made according to their ranks $r_1 < r_2 < \ldots < r_k$, where the rank value is calculated or evaluated by decision maker/expert. In the case of one criterion and a small collection of objects, an objects' arrangement is not so difficult for a human being The more a number of objects, criteria and persons, the more complicate and hard a procedure of ranking due to persons' errors and inconsistencies [Mi80], [LM97].

In order to overcome complexities and contradictions of such kinds, the other techniques are elaborated. In ELECTRE methods [Roy96], objects are compared and arranged by the outranking relation based on the special indexes of concordance and discordance. In ZAPROS methods [LM97], the so-called joint ordered scales are constructed for an objects' arrangement. The special procedures are used to elicit decision maker's preferences, to check out assessments' transitivity, and to correct possible contradictions.

In the techniques mentioned above, the final ordering objects is being found by an aggregation or concordance of various estimates. But often in real applied problems, expert's judgments are inconsistent and may be contradictory, for instance, when several experts estimates objects independently under multiple criteria. So, one is in need of method for ranking multi-attribute objects that incorporates different individual estimates of all persons simultaneously without a compromise between them.

3 Multi-attribute object representation

Usually a multi-attribute object A_i is represented as a vector or cortege $q_i=(q_{i1}^{e_1},\ldots,q_{im}^{e_m})$ in the Cartesian space of attributes $Q=Q_1\times\ldots\times Q_m$, where $Q_s=\{q_s^{e_s}\}$ is a continuous or discreet scale of s-th attribute. As discussed above, if one and the same object A_i is evaluated k times or by k experts, then this object may be represented as a collection of k vectors $\{q_i^{(1)},\ldots,q_i^{(k)}\}$. Here $q_i^{(j)}=(q_{i1}^{e_1(j)},\ldots,q_{im}^{e_m(j)})$, $j=1,\ldots,k$ is a j-th copy of object A_i. The collection of vectors is to be considered as a whole in spite of a possible incomparability of separate vectors $q_i^{(j)}$. Such complex collection has an overcomplicated structure that is very difficult for an analysis. For instance, in group decisions [HY81], a collection of k vectors $\{q_i^{(1)},\ldots,q_i^{(k)}\}$ is replaced usually by one vector that aggregates all of vectors. And properties of all initial vectors $q_i^{(1)},\ldots, q_i^{(k)}$ will be lost.

Now, instead of the space $Q=Q_1\times\ldots\times Q_m$, let us consider a set $G=Q_1$Y...YQ_m, which consists of m attribute groups $Q_s=\{q_s^1, q_s^2,\ldots, q_s^{hs}\}$, $s=1,\ldots,m$. The set G characterizes the object properties. So, an object A_i may be represented as the following set of many repeating attributes or multiset over the domain G:

$$A_i = \{k_{Ai}(q_1{}^1)\circ q_1{}^1,\ldots,k_{Ai}(q_1{}^{h_1})\circ q_1{}^{h_1},\ldots,k_{Ai}(q_m{}^1)\circ q_m{}^1,\ldots, k_{Ai}(q_m{}^{h_m})\circ q_m{}^{h_m}\}. \qquad (1)$$

Here $k_A(q_s{}^{e_s})$ is a number of attribute $q_s{}^{e_s}$ occurrence within the description of object A_i that is denoted by the sign $\circ$.

New types of operations with multisets (see Appendix) allow us to combine an object description by various ways. For instance, the description of object A_i may be aggregated as an addition $\boldsymbol{A}_i=\sum_j \boldsymbol{A}_i^{(j)}$, union $\boldsymbol{A}_i=\mathrm{Y}_j \boldsymbol{A}_i^{(j)}$ or intersection $\boldsymbol{A}_i=\mathrm{I}_j \boldsymbol{A}_i^{(j)}$ of multisets $\boldsymbol{A}_i^{(j)}$, which represent the object copies. The multiset $\boldsymbol{A}_i$ may be also formed as a linear combination of corresponding multisets $\boldsymbol{A}_i=\sum_j b_j\bullet \boldsymbol{A}_i^{(j)}$, $\boldsymbol{A}_i=\mathrm{Y}_j b_j\bullet \boldsymbol{A}_i^{(j)}$ or $\boldsymbol{A}_i=\mathrm{I}_j b_j\bullet \boldsymbol{A}^{(j)}{}_i$, $b_j>0$. When the object description is formed as a multiset addition, all properties of the object copies (all values of attribute) are combined, and $k_A(q_s{}^{e_s})$ is equal to a number of times that the attribute $q_s{}^{e_s}$ has been evaluated by experts. In the case of union or intersection of multisets, the best properties (maximal values of attribute) or the worth properties (minimal values of attribute) of the individual object copies are intensified. The theoretical model of multisets is very appropriated to represent and analyze a collection of objects that are described with many inconsistent attributes.

4 Ordering multi-attribute objects

When an activity of company A_i is represented as a multiset (1) the problem of ordering multi-attribute objects is transformed into the problem of ordering multisets. The simplest way for ordering multisets is a multiset arrangement by a binary relation of inclusion. In this case, an i-th object A_i is better than a j-th object A_j ($A_i \phi A_j$), if the inclusion $\boldsymbol{A}_i \supset \boldsymbol{A}_j$ is fulfilled for corresponding multisets. The last statement is equal to the condition $k_{Ai}(q_s{}^{e_s})>k_{Aj}(q_s{}^{e_s})$ for all attribute values $q_s{}^{e_s}\in G$. But this option appears very seldom in practice, because usually the most of such multi-attribute objects are incomparable.

Let us consider now multi-attribute objects as points of multiset metric space $(\boldsymbol{A},d)$, for instance, with the Hamming-type distance $d_{11}(\boldsymbol{A},\boldsymbol{B})=m(\boldsymbol{A}\Delta\boldsymbol{B})$ (see Appendix). In our case, the distance $d_{11}(\boldsymbol{A},\boldsymbol{B})$ may be written in the following form

$$d_{11}(\boldsymbol{A},\boldsymbol{B}) = \sum_{s=1}^{m} w_s \sum_{e_s=1}^{h_s} | k_A(q_s{}^{e_s}) - k_B(q_s{}^{e_s})|,$$

where $w_s>0$ is a relative importance of the criterion Q_s. If all criteria are equally important, then all $w_s=1$.

Let us determine two objects (possibly, not existent really) that have the highest and the lowest estimates by all criteria Q_s: the best one $A_{\max}$ and the worst one $A_{\min}$. These objects may be represented as the following two multisets:

$$\boldsymbol{A}_{\max} = \{k\circ q_1{}^1,0,\ldots,0,\ k\circ q_2{}^1,0,\ldots,0,\ldots,\ k\circ q_m{}^1,0,\ldots,0\},$$
$$\boldsymbol{A}_{\min} = \{0,\ldots,0,k\circ q_1{}^{h_1},\ 0,\ldots,0,k\circ q_2{}^{h_2},\ldots,\ 0,\ldots,0,k\circ q_m{}^{h_m}\},$$

where k is the total number of times or experts, who estimate objects. The multisets $\boldsymbol{A}_{\max}$ and $\boldsymbol{A}_{\min}$ are named as the reference points in the multiset space $(\boldsymbol{A},d)$.

Objects will be compared and arranged with respect to their closeness to the best object $A_{\max}$. An object A_i is said to be better then an object A_j $(A_i \phi A_j)$, if a multiset $\boldsymbol{A}_i$ is more close to the multiset $\boldsymbol{A}_{\max}$, that is $d_{11}(\boldsymbol{A}_{\max},\boldsymbol{A}_i)<d_{11}(\boldsymbol{A}_{\max},\boldsymbol{A}_j)$. All objects are arranged by values of their distances from the best object $A_{\max}$. If $d_{11}(\boldsymbol{A}_{\max},\boldsymbol{A}_i)=d_{11}(\boldsymbol{A}_{\max},\boldsymbol{A}_j)$ for certain objects A_i and A_j, then these objects will be equivalent or incomparable. So the obtained descending arrangement will be nonstrict and partial.

While each object A_i is estimated by k experts under all m criteria, it's easy to show that the distance between multisets $\boldsymbol{A}_{\max}$ and $\boldsymbol{A}_i$ may be rewritten as:

$$d_{11}(\boldsymbol{A}_{\max}, \boldsymbol{A}_i) = 2 \sum_{s=1}^{m} w_s \; [k - k_{Ai}(q_s^{\,1})].$$

Now, the condition for objects' comparisons looks as follows: an object A_i is better then an object A_j $(A_i \phi A_j)$, if $S_{Ai}^{\,1} > S_{Aj}^{\,1}$. Here $S_{Ai}^{\,1}=\sum_s w_s k_{Ai}(q_s^{\,1})$ is the weighted sum of the first (best) objects' estimates by all criteria Q_s.

Equalities $d_{11}(\boldsymbol{A}_{\max},\boldsymbol{A}_{i1})=\ldots=d_{11}(\boldsymbol{A}_{\max},\boldsymbol{A}_{it})$ may exist for some objects A_{ir}, $r=1,\ldots t$. In this case, we have a partial ordering objects $A_{i1},\ldots,A_{it}$ within the group. In order to arrange these equivalent or incomparable objects from the best to the worst, one may find the weighted sums $S_{Air}^{\,2}=\sum_s w_s k_{Air}(q_s^{\,2})$ of the second objects' estimates by all criteria Q_s. So, an object A_{iu} is better then an object A_{iv}, if $S_{Aiu}^{\,2}>S_{Aiv}^{\,2}$. If sums $S_{Air}^{\,2}$ will be equal for some objects A_{irp} within the subgroup, then these objects are to be ordered from the best to the worst by the values of weighted sum $S_{Airp}^{\,3}=\sum_s w_s k_{Airp}(q_s^{\,3})$ of the third objects' estimates by all criteria Q_s. This procedure is being repeated until all objects of collection $A=\{A_1,\ldots,A_n\}$ will be ordered.

An ascending arrangement of multi-attribute objects with respect to their closeness to the worst object $A_{\min}$ is constructed analogously. An object A_i is said to be better then an object A_j $(A_i \phi A_j)$, if a multiset $\boldsymbol{A}_i$ is more far from the multiset $\boldsymbol{A}_{\min}$, that is $d_{11}(\boldsymbol{A}_{\min},\boldsymbol{A}_i)>d_{11}(\boldsymbol{A}_{\min},\boldsymbol{A}_j)$. As above, the objects A_i and A_j will be equivalent or incomparable, if $d_{11}(\boldsymbol{A}_{\min},\boldsymbol{A}_i)=d_{11}(\boldsymbol{A}_{\min},\boldsymbol{A}_j)$. Note that ordering objects with respect to the best and to the worst ones may be diverse. The final ordering will be constructed as a combination of the descending and ascending arrangements.

5 Conclusion

In this paper, we have suggested the simple and transparent tool for group multicriteria decision aiding. This technique for ordering multi-attribute objects, when several copies of objects may exist, is based on the theory of multiset metric spaces. The multiset approach allows us to discover, present and utilize the available information that is contained in the verbal descriptions of objects, to analyze the obtained results and their peculiarities, especially for inconsistencies of objects' properties and contradictions of decision maker's preferences. Underline once more that qualitative (verbal) attributes are not transformed in or replaced by any quantitative (numerical) ones as, for instance, in MAUT and TOPSIS methods [HY81], and in fuzzy set theory [ZZG84].

The method described here was applied to rating Russian companies that work in the area of information and telecommunication technologies [Who00]. 50 experts estimated about 50 companies under dozen multiple criteria. 30 companies were selected as the leading high-tech companies, 10 companies as the mostly dynamic developing companies, and 10 companies as the leading developers of software.

Appendix. Basis notions of multiset theory

A multiset (also called a bag) is the known notion in combinatorial mathematics. Let us review briefly a theory of multisets and multiset metric spaces [Kn69], [Ya86], [Pe03a,b]. A multiset $\boldsymbol{A}$ drawn from a crisp set $G=\{x_1,x_2,...,x_j,...\}$ with different elements, which is called a domain, is defined as a collection of elements' groups

$$\boldsymbol{A}=\{k_A(x_1)\circ x_1, k_A(x_2)\circ x_2,\ldots\}=\{(k_A(x)\circ x)|x\in G, k_A\in \mathbf{Z}_+\}.$$

Here $k_A{:}G\rightarrow\mathbf{Z}_+=\{0,1,2,\ldots\}$ is called a counting function of multiset, which defines the number of times that the element $x_i\in G$ occurs in the multiset $\boldsymbol{A}$, and this is indicated with the symbol $\circ$. A multiset $\boldsymbol{A}$ becomes an ordinary set when $k_A(x)=\chi_A(x)$, where $\chi_A(x)=1$, if $x\in\boldsymbol{A}$, and $\chi_A(x)=0$, if $x\notin\boldsymbol{A}$. The cardinality $|\boldsymbol{A}|=\Sigma_x k_A(x)$ and dimensionality $/\boldsymbol{A}/=\Sigma_x\chi_A(x)$ of the finite multiset $\boldsymbol{A}$ are defined as a total number of all element copies, and as a total number of different elements. The maximal value of the counting function $\text{alt}\boldsymbol{A}=\max_{x\in G}k_A(x)$ is called a height of the multiset $\boldsymbol{A}$, and a crisp set $\text{Supp}\boldsymbol{A}=\{x|x\in G, \chi_{\text{Supp}A}(x)=\chi_A(x)\}$ is named as a support set of the multiset $\boldsymbol{A}$. The multiset is called the empty multiset $\varnothing$, if $k_\varnothing(x)=0$, and the maximal multiset $\boldsymbol{Z}$, if $k_Z(x)=\max_{A\in A} k_A(x)$, $\forall x\in G$. We shall call the multiset $\boldsymbol{A}_{[h]}$ constant if $k_{A[h]}(x)=h=\text{const}$, $\forall x\in\boldsymbol{A}_{[h]}$. So the empty multiset $\varnothing$ is the constant multiset $\boldsymbol{A}_{[0]}$ of the height 0, and any crisp set A, including the domain G, and support sets, is the constant multiset $\boldsymbol{A}_{[1]}$ of the height 1.

Consider the rules for comparing multisets. Multisets $\boldsymbol{A}$ and $\boldsymbol{B}$ are said to be equal ($\boldsymbol{A}=\boldsymbol{B}$) if $k_A(x)=k_B(x)$, $\forall x\in G$. For equal multisets we have $|\boldsymbol{A}|=|\boldsymbol{B}|$, $/\boldsymbol{A}/=/\boldsymbol{B}/$, $\text{alt}\boldsymbol{A}=\text{alt}\boldsymbol{B}$, $\text{Supp}\boldsymbol{A}=\text{Supp}\boldsymbol{B}$. Multisets $\boldsymbol{A}$ and $\boldsymbol{B}$ are called equicardinal if $|\boldsymbol{A}|=|\boldsymbol{B}|$; equidimensional if $/\boldsymbol{A}/=/\boldsymbol{B}/$; and equivalued if multisets are equicardinal and equidimensional. The equal multisets are equivalued. The converse is invalid, in general. We say that multiset $\boldsymbol{B}$ is contained or included in a multiset $\boldsymbol{A}$ ($\boldsymbol{B}\subseteq\boldsymbol{A}$) if $k_B(x)\leq k_A(x)$, $\forall x\in G$. Then the multiset $\boldsymbol{B}$ is called a submultiset of the multiset $\boldsymbol{A}$, and the multiset $\boldsymbol{A}$ is called an overmultiset of the $\boldsymbol{B}$. Obviously, $|\boldsymbol{B}|\leq|\boldsymbol{A}|$, $/\boldsymbol{B}/\leq/\boldsymbol{A}/$, $\text{alt}\boldsymbol{B}\leq\text{alt}\boldsymbol{A}$, $\text{Supp}\boldsymbol{B}\subseteq\text{Supp}\boldsymbol{A}$.

Multisets $\boldsymbol{A}$ and $\boldsymbol{B}$ are said to be homogeneously equivalent or S-equivalent ($\boldsymbol{A}\simeq\boldsymbol{B}$), if their support sets coincide $\text{Supp}\boldsymbol{A}=\text{Supp}\boldsymbol{B}$, and there exists a one-to-one correspondence

f between the multiset components of the same name: $k_B(x)=f(k_A(x))$, $\forall x\in G$. Multisets $\boldsymbol{A}$ and $\boldsymbol{B}$ are said to be heterogeneously equivalent or D-equivalent ($\boldsymbol{A}\approx\boldsymbol{B}$), if their support sets are equivalent Supp$A$~Supp$B$ or equal, and there exists a one-to-one correspondence f between the multiset components of different names: $k_B(x_i)=f(k_A(x_j))$, $x_i,x_j\in G$. Here f is an integer-valued function with the range $\mathbf{Z}_+$. The S- and D-equivalent multisets are equidimensional $/\boldsymbol{B}/=/\boldsymbol{A}/$, their heights are connected as alt$\boldsymbol{B}=f($alt$\boldsymbol{A})$. One of two S-equivalent multisets is always a submultiset of the other, but this statement is invalid for D-equivalent multisets. D-equivalent multisets can be transformed into S-equivalent multisets by renaming the elements $x_i\rightarrow x_j$ in one of the multisets. Special cases of S-equivalency are equal multisets; shifted multisets, for which $k_B(x)=k_A(x)+s$; and stretched or proportional multisets, for which $k_B(x)=qk_A(x)$, $s\geq 0$, $q\geq 1$ are integers. Any constant multiset $\boldsymbol{A}_{[h]}$ is a multiset that is shifted by $h-1$ units, or stretched by h times with respect to its support set Supp$\boldsymbol{A}_{[h]}$. A special case of D-equivalency is equicomposed multisets with equal heteronymous components $k_A(x_i)=k_B(x_j)$, $x_i,x_j\in G$. Equal multisets are equicomposed, whereas the converse is invalid.

The following operations with multisets are defined:

union	$\boldsymbol{A}\mathrm{Y}\boldsymbol{B}=\{k_{A\mathrm{Y}B}(x)\circ x \mid k_{A\mathrm{Y}B}(x)=\max(k_A(x),k_B(x))\}$;
intersection	$\boldsymbol{A}\mathrm{I}\boldsymbol{B}=\{k_{A\mathrm{I}B}(x)\circ x \mid k_{A\mathrm{I}B}(x)=\min(k_A(x),k_B(x))\}$;
arithmetic addition	$\boldsymbol{A}+\boldsymbol{B}=\{k_{A+B}(x)\circ x \mid k_{A+B}(x)=k_A(x)+k_B(x)\}$;
arithmetic subtraction	$\boldsymbol{A}-\boldsymbol{B}=\{k_{A-B}(x)\circ x \mid k_{A-B}(x)=k_A(x)-k_{A\cap B}(x)\}$;
symmetric difference	$\boldsymbol{A}\Delta\boldsymbol{B}=\{k_{A\Delta B}(x)\circ x \mid k_{A\Delta B}(x)=\lvert k_A(x)-k_B(x)\rvert\}$;
complement	$\overline{\boldsymbol{A}}=\boldsymbol{Z}-\boldsymbol{A}=\{k_{\overline{A}}(x)\circ x \mid k_{\overline{A}}(x)=k_Z(x)-k_A(x)\}$;
multiplication by a scalar	$b\bullet\boldsymbol{A}=\{k_{b\bullet A}(x)\circ x \mid k_{b\bullet A}(x)=b\cdot k_A(x),\ t\in\mathbf{N}\}$;
arithmetic multiplication	$\boldsymbol{A}\bullet\boldsymbol{B}=\{k_{A\bullet B}(x)\circ x \mid k_{A\bullet B}(x)=k_A(x)\cdot k_B(x)\}$;
raising to an arithmetic power	$\boldsymbol{A}^n=\{k_{A^n}(x)\circ x \mid k_{A^n}(x)=(k_A(x))^n\}$;
direct product	$\boldsymbol{A}\times\boldsymbol{B}=\{k_{A\times B}\circ(x_i,x_j) \mid k_{A\times B}=k_A(x_i)\cdot k_B(x_j),\ x_i\in\boldsymbol{A},\ x_j\in\boldsymbol{B}\}$;
raising to a direct power	$(\times\boldsymbol{A})^n=\{k_{(\times A^n)}\circ(x_1,\ldots,x_n) \mid k_{(\times A^n)}=\prod_{i=1}^{n}k_A(x_i),\ x_i\in\boldsymbol{A}\}$.

Linear combinations of union, intersection, arithmetic addition, multiplication, and direct product with a multiplication by a scalar are determined as follows:

$$\mathrm{Y}_{i\in I}(b_i\bullet\boldsymbol{A}_i),\quad \mathrm{I}_{i\in I}(b_i\bullet\boldsymbol{A}_i),\quad \Sigma_{i\in I}(b_i\bullet\boldsymbol{A}_i),\quad \Pi_{i\in I}(b_i\bullet\boldsymbol{A}_i),\quad (b_1\bullet\boldsymbol{A}_1)\times\ldots\times(b_n\bullet\boldsymbol{A}_n).$$

There exists a duality of operations with multisets that likes de Morgan laws for sets:

$$\overline{\boldsymbol{A}\,\mathrm{Y}\,\boldsymbol{B}}=\overline{\boldsymbol{A}}\,\mathrm{I}\,\overline{\boldsymbol{B}},\quad \overline{\boldsymbol{A}\,\mathrm{I}\,\boldsymbol{B}}=\overline{\boldsymbol{A}}\,\mathrm{Y}\,\overline{\boldsymbol{B}};$$
$$\overline{\boldsymbol{A}+\boldsymbol{B}}=\overline{\boldsymbol{A}}-\boldsymbol{B}=\overline{\boldsymbol{B}}-\boldsymbol{A},\quad \overline{\boldsymbol{A}-\boldsymbol{B}}=\overline{\boldsymbol{A}}+\boldsymbol{B},\quad \overline{\boldsymbol{A}}-\overline{\boldsymbol{B}}=\boldsymbol{B}-\boldsymbol{A}.$$

Some properties of operations, which exist for sets, are absent for multisets. And new properties of operations that have no analogues for sets arise simultaneously, for instance,

$$\boldsymbol{A}+\overline{\boldsymbol{A}}=\boldsymbol{Z},\ \boldsymbol{Z}-\overline{\boldsymbol{A}}=\boldsymbol{A},\ \boldsymbol{A}-\overline{\boldsymbol{A}}\neq\varnothing,\ \boldsymbol{A}\cup\overline{\boldsymbol{A}}\neq\boldsymbol{Z},\ \boldsymbol{A}\cap\overline{\boldsymbol{A}}\neq\varnothing.$$

In general, the operations of arithmetic addition, multiplication by a scalar, arithmetic multiplication, and raising to arithmetic powers are not defined in the theory of sets. Analogues of these operations may be operations with vectors $\boldsymbol{a}+\boldsymbol{b}=(a_1+b_1,\ldots,a_n+b_n)$, $h\cdot\boldsymbol{a}=(ha_1,\ldots,ha_n)$, and matrixes $\mathrm{A}+\mathrm{B}=\|a_{ij}+b_{ij}\|_{m\times n}$, $h\cdot\mathrm{A}=\|h\cdot a_{ij}\|_{m\times n}$, $\mathrm{A}\cdot\mathrm{B}=\|a_{ij}\cdot b_{ij}\|_{m\times n}$. The last operation is different from a traditional matrix multiplication. The operation of multiset selection suggested in [Ya86] is a special case of multiset arithmetic multiplication, where one of the factors is an ordinary set. When multisets are reduced to sets, the operations of arithmetic multiplication and raising to arithmetic powers degenerate into a set intersection, but the operations of set arithmetic addition and set multiplication by a scalar will be impracticable.

Different classes of metric spaces $(\boldsymbol{A},d)$ on a multiset family $\boldsymbol{A}$ have the following distances:

$$d_{1p}(\boldsymbol{A},\boldsymbol{B})=[m(\boldsymbol{A}\Delta\boldsymbol{B})]^{1/p};\quad d_{2p}(\boldsymbol{A},\boldsymbol{B})=[m(\boldsymbol{A}\Delta\boldsymbol{B})/m(\boldsymbol{Z})]^{1/p};\ d_{3p}(\boldsymbol{A},\boldsymbol{B})=[m(\boldsymbol{A}\Delta\boldsymbol{B})/m(\boldsymbol{A}\cup\boldsymbol{B})]^{1/p},$$

where $p>0$ is integer. A measure m of multiset $\boldsymbol{A}$ is a real-valued non-negative function that is defined on the algebra $L(\boldsymbol{Z})$ of multisets. A measure of multiset has a properties of strong additivity $m(\sum_i\boldsymbol{A}_i)=\sum_i m(\boldsymbol{A}_i)$; weak additivity $m(\cup_i\boldsymbol{A}_i)=\sum_i m(\boldsymbol{A}_i)$ for $\boldsymbol{A}_i\cap\boldsymbol{A}_j=\varnothing$; weak monotonicity $m(\boldsymbol{A})\leq m(\boldsymbol{B})\Leftrightarrow\boldsymbol{A}\subseteq\boldsymbol{B}$; continuity $\lim_{i\to\infty} m(\boldsymbol{A}_i)=m(\lim_{i\to\infty}\boldsymbol{A}_i)$; symmetry $m(\boldsymbol{A})+m(\overline{\boldsymbol{A}})=m(\boldsymbol{Z})$; elasticity $m(b\bullet\boldsymbol{A})=bm(\boldsymbol{A})$; and $m(\varnothing)=0$. Note that due to the continuity of the multiset measure, the distance $d_{3p}(\boldsymbol{A},\boldsymbol{B})$ is not defined for $\boldsymbol{A}=\boldsymbol{B}=\varnothing$. So, $d_{3p}(\varnothing,\varnothing)=0$ by the definition.

The measure of multiset $\boldsymbol{A}$ may be determined in the various ways, for instance, as a linear combination of counting functions: $m(\boldsymbol{A})=\sum_i w_i k_A(x_i)$, $w_i>0$. In this case, the distances are as follows:

$$d_{1p}(\boldsymbol{A},\boldsymbol{B})=\left(\sum_{x_i\in G} w_i\,|k_A(x_i)-k_B(x_i)|\right)^{1/p};$$

$$d_{2p}(\boldsymbol{A},\boldsymbol{B})=\left(\sum_{x_i\in G} w'_i\,|k_A(x_i)-k_B(x_i)|\right)^{1/p},\quad w'_i=1/\sum_{j=1}^{h} w_j;$$

$$d_{3p}(\boldsymbol{A},\boldsymbol{B})=\left(\frac{\sum_{x_i\in G} w_i\,|k_A(x_i)-k_B(x_i)|}{\sum_{x_i\in G} w_i\max[k_A(x_i),k_B(x_i)]}\right)^{1/p}.$$

The distance $d_{1p}(\boldsymbol{A},\boldsymbol{B})$ is a Hamming-type distance between objects, which is traditional for many applications. The distance $d_{2p}(\boldsymbol{A},\boldsymbol{B})$ characterizes a difference between two objects related to common properties of all objects as a whole. And the distance $d_{3p}(\boldsymbol{A},\boldsymbol{B})$ reflects a difference related to properties of only both objects. For any fixed p, the metrics d_{1p} and d_{2p} are the continuous and uniformly continuous functions, the metric d_{3p} is the piecewise continuous function almost everywhere on the corresponding metric space. For p=1 in the case of sets, $d_{11}(A,B)=m(A\Delta B)$ is called the Fréchet distance, and $d_{31}(A,B)=m(A\Delta B)/m(AYB)$ is called the Steinhaus or biotopic distance.

Bibliography

[GO01] Godlevsky M.D., Orekhov S.V. An Enterprise Financial State Estimation Based on Data Mining. In (Godlevsky M., Mayr H. Eds.): Information Systems Technology and its Applications. Gesellschaft für Informatik, Bonn, 2001, pp. 229-235.

[HY81] Hwang C.L, Yoon K. Multiple Attribute Decision Making – Methods and Applications: A State of the Art Survey. Springer-Verlag, New York, 1981.

[Kn69] Knuth D.E. The Art of Computer Programming. Vol.2. Seminumerical Algorithms. Addison-Wesley, Reading, 1969.

[KS72] Kemeny J.G., Snell J.L. Mathematical Models in the Social Sciences. The MIT Press, Cambridge, 1972.

[LM97] Larichev O.I., Moshkovich H.M. Verbal Decision Analysis for Unstructured Problems. Kluwer Academic Publishers, Boston, 1997.

[Mi80] Mirkin B.G. Analysis of Qualitative Attributes and Structures. Statistika, Moscow, 1980 (in Russian).

[Pe94] Petrovsky A.B. An Axiomatic Approach to Metrization of Multiset Space. In (Tzeng G.H., Wang H.F., Wen U.P., Yu P.L. Eds.): Multiple Criteria Decision Making. Springer-Verlag, New York, 1994, pp. 129-140.

[Pe03a] Petrovsky A.B. Spaces of Sets and Multisets. Editorial URSS, Moscow, 2003 (in Russian).

[Pe03b] Petrovsky A.B. Cluster Analysis in Multiset Spaces. In (Godlevsky M., Mayr H. Eds.): Information Systems Technology and its Applications. Gesellschaft für Informatik, Bonn, 2003, pp. 109-119.

[PS94] Pawlak Z., Slowinski R. Rough Set Approach to Multi-Attribute Decision Analysis. European Journal of Operational Research, 1994, 72, pp. 443-459.

[Roy96] Roy B. Multicriteria Methodology for Decision Aiding. Kluwer Academic Publishers, Dordrecht, 1996.

[Who00] Who is the Most Intellectual in Russia? Rating Leading Russian Developers of Hi-Tech. Company, 2000, N 47(143), pp.38-39 (in Russian).

[Ya86] Yager R.R. On the Theory of Bags. International Journal of General Systems. 1986, Vol.13, pp. 23-37.

[ZZG84] Zimmerman H.J., Zadeh L.A., Gaines B.R. Fuzzy Sets and Decision Analysis. North-Holland, Amsterdam, 1984.