Mathematical aspects of ranking theory

A. AMELJAŃCZYK

andrzej.ameljanczyk@wat.edu.pl

Military University of Technology, Faculty of Cybernetics
Institute of Computer and Information Systems
Kaliskiego Str. 2, 00-908 Warsaw, Poland

The paper covers the theoretical grounds for defining of rankings, basing on the terms taken from the *relation space* theory. One presented an array of new definitions which allow establishing rankings without the necessity of using typical ranking functions. Moreover, one introduced the term *precedence ranking relation* (not necessarily *order relation*), and demonstrated general algorithms to establish rankings on the basis of definitions of extreme elements.

1. Introduction

The role of rankings has increased worldwide in recent years. They become common method of making and justifying decisions on distribution of resources over various types of social and economic activities which utility is assessed with respect to many aspects (multi-criteria assessment). Moreover, rakings as the objective (scientifically proven) decision support method are the most frequently applied tool of "simplified optimisation". As a result of the application of rankings financial resources, often in huge amounts, are distributed, e.g. rankings applied in processes of the best offer selection in public procurement procedures concerning services or supplies, rankings of high schools, tertiary education schools, or various products, ranking-based selection in R & D call for proposals, etc. What is more, a new raking-related problem occurred leading to the so-called object categorisation, i.e. the problem of distribution of a set of objects among classes (clusters, categories). The division of a set into classes on the basis of arbitrarily-determined ranges of ranking values for given objects proved to be unsatisfactory. The thing is that clusters should constitute subsets of objects which are equivalent in some way. Such expectations require broadening of the term "ranking" beyond just a mere score ranking. Rankings owe their popularity mostly due to: the common opinion about their objectivity, most of all, resulting, in particular, from the application of allegedly scientifically-proven procedures; and the simplicity of designing and implementing of ranking procedures.

Even the simplest ranking (score ranking) calms down both: a ranking's organiser and its participants – it becomes a specific legal alibi of its settlements. One could not be more wrong. An ill-designed or deliberately manipulated ranking can bring nothing but losses and damages people often do not realize. Therefore, rankings as decision support tools can be extremely dangerous because they can result in uncritical decisions stemming from the belief in their unique, even magical, objective indication of the best order of elements subject to assessment. It is extremely rare that one carries out a reliable qualitative analysis of partial criteria applied in rankings (their correspondence to the aim of an endeavour) and mechanisms of replacing of many, often non-comparable, partial scores with a complex ranking score. "The Scoring approach", or "the weighted scoring approach" in the best case, has ruled in rankings for ages. Regardless of their definition, ranking procedures constitute, however, specific multi-criteria optimisation tasks.

The establishment of rankings of the set of objects subject to multi-criteria assessment differs from optimisation in the classic sense in the way that the former determines "the quality" of all objects – from "the best" one till "the worst" one – whereas the latter defines only "the best" element [1, 2, 8, 9, 10]. Subsequent chapters presented the formal grounds of the ranking theory and the general ranking algorithms applying the concept "extreme elements" when constructed.

2. The formal definition of ranking

The term *ranking* is widespread and broad with multiple of definitions from very intuitive [5, 7] to formal ones [3, 4, 6]. Below one presents general, formal definitions of notions related to rankings, and further part of the paper describes the most commonly applied ranking procedures based on the theory of space with relations where extreme elements are used. Let Y is a given finite, non-empty set of elements (objects). Its cardinality is equal to $M = |Y|$. The symbol $\mathcal{K}$ stands for the subset of natural numbers $\{1, 2, ..., k, ...K\}$. By analogy, the symbol $\mathcal{M}$ stands for the set $\{1, 2, ..., m, ...M\}$. Suppose $K \leq M$.

Definition 2.1
A partition of the set Y is any family of non-empty, disjoint subsets of the set Y which sum gives the set Y (is equal to the set Y).
The partition of the set Y is denoted by $P(Y)$.

Definition 2.2
A ranking of the set Y is any sequence $r(Y)$ of subsets $Y(k)$ of the set Y constituting *the partition* of the set Y (is any *ordered partition* of a set Y)

$$r(Y) = (Y(1), ..., Y(k), ..., Y(K)) \qquad (1)$$

So, it is the sequence of non-empty subsets $Y(k)$ for $k \in \mathcal{K}$ of the set Y that

$$Y(k) \cap Y(m) = \varnothing \; dla \; k \neq m$$

$$\bigcup_{k \in \mathcal{K}} Y(k) = Y$$

Definition 2.3
A set $Y(k)$, $k \in \mathcal{K}$ is called a kth element of the ranking $r(Y)$ (a kth *cluster* or a kth *category* of the ranking $r(Y)$).

Definition 2.4
A linear ranking of the set Y is a ranking r(Y) where $|Y(k)| = 1$ for every $k \in \mathcal{K}$.
Therefore, the following equation $K = M = |Y|$ applies for the *linear ranking*.

Definition 2.5
A trivial ranking (pseudo ranking) of the set Y is a ranking $r(Y)$ where $K = 1$ (all elements of the set Y are "of the same importance").

Definition 2.6
A non-linear ranking is a ranking $r(Y)$ which is not linear, i.e. $K < M$.
The term ranking is connected with the notion "equivalence relation" (it results from its definition using the term partition).

Definition 2.7
The *equivalence relation R* in the set Y is the relation $R \subset Y \times Y$ which is reflexive, transitive, and symmetric.
Let **R**(Y) be the set of all equivalence relations in the set Y. Every *equivalence relation R* belonging to the set generates a given *partition* $P_R(Y)$ of the set Y, and the other way around: every *partition* of the set Y fixes a certain *equivalence relation*. The number of all *partitions* of the set of its cardinality equal to M is determined with the *Bell's number*, obtained by the recurrence formula:

$$B_{m+1} = \sum_{k=0}^{m} \binom{m}{k} B_k, \;\; m \in \mathcal{M}, \;\; B_0 = 1 \quad (2)$$

If $P(Y)$ is *the partition* of the set Y, the *equivalence relation* determined by *the partition* is the relation:

$$R_{P(Y)} = \{(x, y) \in Y \times Y \,|\, \text{exists } A \in P(Y),$$
$$\text{that } (x \in A \text{ and } y \in A)\} \quad (3)$$

Every *partitition* $P(Y)$ generates a set $r(P(Y))$ of rankings of the set Y (the set of the *permutations* of set $P(Y)$).

Definition 2.8
A ranking procedure is the transformation of the set Y in the sequence $r(Y)$ of the subsets constituting *the partition* of a set Y.

Example 2.1
Suppose Y = {△,○,□}. How many and what kind of rankings can be obtained from the elements of the set Y?
According to the Definition 2.2, one gets 5 partitions for the set, and then, 13 rankings of the set Y.
1) {{△,○,□}} → one ranking (*trivial*)
    ({△,○,□})

2) $\{\{\triangle\},\{\circ,\square\}\}$ → two rankings $(\{\triangle\},\{\circ,\square\}), (\{\circ,\square\},\{\triangle\})$

3) $\{\{\circ\},\{\triangle,\square\}\}$ → two rankings $(\{\circ\},\{\triangle,\square\}), (\{\triangle,\square\},\{\circ\})$

4) $\{\{\square\},\{\triangle,\circ\}\}$ → two rankings $(\{\square\},\{\triangle,\circ\}), (\{\triangle,\circ\},\{\square\})$

5) $\{\{\triangle\},\{\circ\},\{\square\}\}$ → six linear rankings.

Ad 1) *Trivial ranking* (pseudo ranking).
All elements of the ranking belong to the same *category* (*cluster*) – all of them "*are of the same importance*" (*equally good*, etc.)".

Ad 2) In the ranking $r_1(Y) = (\{\triangle\},\{\circ,\square\})$ – only the element "△" belongs to "the first category" (to the no. 1 *cluster*). Both elements ○ and □ belong (ex equo) to the second category.
The opposite in the ranking $r_2(Y) = (\{\circ,\square\},\{\triangle\})$.

Ad 3), 4) – by analogy.

Ad 5) The case (such *partition* of the set Y) gives typical linear rankings. For example, $r_3(Y) = (\{\circ\},\{\square\},\{\triangle\})$ means that the element "○" is the *first category* – it is "the first element" in the ranking; the element "□" forms the *second category* – the "second element" in the ranking; the element "△" gives the *third category* – the "third (and the last) elemen" in the ranking. For example, the second *partition* gives the following equivalence relation:

$R_{P_2(Y)} = \{(\circ,\square), (\square,\circ),(\triangle,\triangle),(\circ,\circ),(\square,\square)\}$.

For *linear ranking* the brackets standing for "a set" were omitted for the purpose of simplification. And $r_2(Y) = (\{\circ\},\{\square\},\{\triangle\})$ was simplified to $r_2(Y) = (\circ, \square, \triangle)$.
The formal definitions of rankings were presented above. In practice rankings are determined to achieve global (aggregated) information on quality (meaning, usefulness) of elements of the set Y.
For this purpose one defines the so-called *ranking precedence (or preference) model* in the form of ranking preference relations (*precedence ranking relation*) or ranking functions (e.g. *scoring functions*) [5, 7].

3. Precedence ranking relations in defining of ranking procedures

The previous chapter presented general definitions of terms related to rankings. For the fixed, finite set Y one determined the set of all possible rankings. However, how to choose the ranking one is interested in from the set? How to create a procedure leading to such ranking? The key is ranking establishment is to determine a way how the set is divided into equivalence classes (for example, quality (*importance*) of objects) and how *the order* of the classes is determined. In practice, the most frequent ground for ranking establishment is the information on the way of comparison of elements of the investigated set with respect to their quality. Depending on *the complexity* of elements of the set Y, the task to compare elements can pose more or less difficulties. The most general approach is to define the proper *preference ranking relation*. The relations in certain cases can be easily defined by *ranking* and *scoring functions* widely applied in practice [3]. They are often *order relations* or at least *partial order relations*. Generally, one should assume that a *precedence ranking relation* can be any relation that $R \subset Y \times Y$.

Example 3.1
Figure 1 demonstrates a general relation R on the set $Y = \{a,b,c,d,e\}$ in the form of a digraph. Formula $(y,z) \in R$ means that "*y* precedes *z* (in the ranking)", so "*y* is before *z* (in the ranking)" or "*y* is better than *z*" [1, 2, 3].

$$R = \{(b,c),(a,c),(d,a),(d,b),(d,c),$$
$$(e,a),(e,b),(e,c),(e,d)\} \subset Y \times Y$$

The graphic presentation is the following:

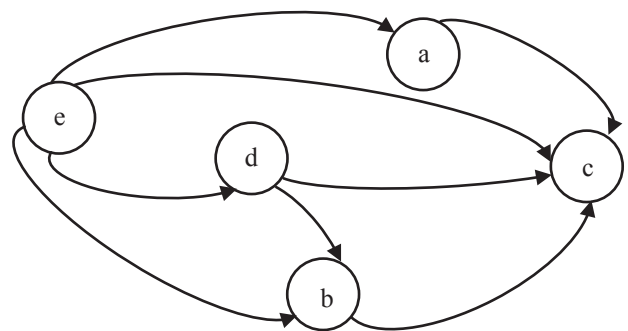Fig. 1. *Precedence ranking relation* R

For example, an object "e" precedes objects: "a, b, c, d"; an object "b" precedes only an object "c"; an object "c" does not precede any object – however, it is preceded by objects "b, d, e", etc. The ranking establishment task can be treated as a recurrence sequence of tasks of the extreme element selection from the ranking set Y. In the first step, one determines a set of

these elements from the set Y which precede all the other elements, i.e. the set $Y_{\inf}^R(1) = Y_D^R = \{e\}$ (see definitions of the set of smallest (dominating) elements or minimal (non-dominated) elements [2, 3]). Then, out of the remaining elements, i.e. the elements from the set $Y - Y_{\inf}^R(1) = \{a,b,c,d\}$ one determines the elements which precede all the other elements. The set is as follows:

$$Y_{\inf}^R(2) = \{a,b,c,d\}_{\inf}^R = \{d\}$$

Next, from the set of remaining elements, i.e. from the set:

$$(Y - Y_{\inf}^R(1)) - Y_{\inf}^R(2) = \{a,b,c\}$$

one determines the elements which precede all the other elements. It will be an empty set: $Y_{\inf}^R(3) = \{a,b,c\}_{\inf}^R = \varnothing$. The subset $\{a,b,c\}$ could not be *ordered* in the aforementioned way. However, one can determine a certain *partition* of the set Y (see (4)), as follows:

$$P_{\inf}(Y) = \{\{e\},\{d\},\{a,b,c\}\}$$

So, on the basis of the *partition* one can establish a ranking of the set Y by determining the order of its elements according to recurrence (the sequence of the operations performed above), as follows

$$r_{\inf}(Y,R) = \left(\{e\},\{d\},\{a,b,c\}\right)$$

The formula for other set establishing (ranking elements) basing on the notion *the smallest element* (*the first element*) is the following, (assuming that $k = 0$, $Y_{\inf}^R(0) = \varnothing$) [3]:

1. determine $\quad Y_{\inf}^R(1) = Y_{\inf}^R \qquad (4)$

2. determine $\quad Y_{\inf}^R(2) = \left(Y - Y_{\inf}^R(1)\right)_{\inf}^R$, etc.

3. expressed as:

$$Y_{\inf}^R(k) = \left(Y - \bigcup_{m=0}^{m=k-1} Y_{\inf}^R(m)\right)_{\inf}^R, k = 1,2,... \qquad (5)$$

The algorithm ends in the step $k = K$ that $Y_{\inf}^R(K) = \varnothing$

If:

$$Y - \bigcup_{m=0}^{m=K-1} Y_{\inf}^R(m) \neq \varnothing, \qquad (6)$$

one substitutes:

$$Y_{\inf}^R(K) = Y - \bigcup_{m=0}^{m=K-1} Y_{\inf}^R(m). \qquad (7)$$

The similar procedure can be applied by using the term *minimal elements* of the set Y. In this case ranking is determined in the following way. Firstly, one determines these elements of the set

Y which are not preceded by any other element of the set Y. The set is as follows:

$$Y_{\min}^R(1) = Y_N^R = \{e\}$$

Next, out of the set of the remaining elements $Y - Y_{\min}^R(1) = \{a,b,c,d\}$ one determine these elements which are not preceded by any other element of the set. The result is the following:

$$Y_{\min}^R(2) = \{a,b,c,d\}_{\min}^R = \{d\}$$

In the next step, one determines the set $Y_{\min}^R(3)$, etc. One gets:

$$Y_{\min}^R(3) = \{a,b\}, \ Y_{\min}^R(4) = \{c\}, \quad Y_{\min}^R(5) = \varnothing$$

The aforementioned procedure (*minimal elements*) resulted in the following *partition* of the set Y: $P_{\min}(Y) = \{\{e\},\{d\},\{a,b\},\{c\}\}$ and the ranking (in the recurrence order):

$$r_{\min}(Y,R) = \left(\{e\},\{d\},\{a,b\},\{c\}\right)$$

The element "e" is the best element of the ranking, "c" is the worst, whereas the elements "a" and "b" are *ex aequo* on the third place. The above procedure (by analogy to (1) – (4)) is formally expressed as (assuming that $Y_{\min}^R(0) = \varnothing$):

$$Y_{\min}^R(k) = \left(Y - \bigcup_{m=0}^{m=k-1} Y_{\min}^R(m)\right)_{\min}^R \qquad (8)$$

Generally, one should demonstrate that the sets (5) or (8) create *the partition* of the set Y (see Definition 2.2).

In practice ($Y \subset \mathcal{R}^N$ for most cases), there is no problem if definitions of the smallest elements $Y_{\inf}^R$ or the *minimal elements* $Y_{\min}^R$ should be used in establishing ranking of the set Y. If a ranking relation is a quasi-order, there should be *minimal elements* – so, one applies the definition of the *smallest elements*; if a ranking relation is an order or a linear order, one uses minimum elements. In these cases sets generally give *the partition*. It is common that "objects" of the definite set Y subject to ranking are elements $y \in \mathcal{R}^N$ (multi-criteria assessment [1, 2, 3, 4, 5, 7, 8]. Then, the following theorems are true:

Theorem 3.1

Suppose $Y \subset \mathcal{R}^N$ – a finite ranking set, $R \subset Y \times Y$ is the *Pareto relation* [1]. The sequence of sets:

$$Y_{\min}^R(k) = \left(Y - \bigcup_{m=0}^{m=k-1} Y_{\min}^R(m)\right)_{\min}^R,$$

$k = 1, 2, ..., K$ gives *the partition* of the set Y, and also the proper ranking

$$r_{\min}(Y, R) = \left(Y_{\min}^R(k), k = 1, 2, ..., K\right)$$

Proof

For any set Y, meeting the theorem assumptions, there is a *lexicographic relation* $\mathcal{L}$ of a *linear order* that $R \subset \mathcal{L}$ [1, 2]. Then, $Y_N^{\mathcal{L}} \subset Y_N^R$, and $Y_N^{\mathcal{L}} = Y_D^{\mathcal{L}} \neq \varnothing$. It means that for every Y meeting the theorem assumptions there is $Y_N^R \neq \varnothing$. The sets $Y_{\min}^R(k)$ from the definition are pair-wise disjoint. Then $Y_{\min}^R(k)$ give *the partition* of the set Y and the proper ranking (Definition 2.2). ∎

Correspondent theorems can be proven for the *Hurwicz relation*, in particular of the *Optimist* and *Pessimist relations* [1, 2].

Theorem 3.2

Suppose $Y \subset \mathcal{R}^N$ – a finite ranking set, $R_O \subset Y \times Y$ is the *Optimist relation* [1].

The sequence of sets $Y_{\inf}^{R_O}(k)$ of the *smallest elements* forms *the partition* of the set Y and the proper ranking

$$r_{\inf}(Y, R_O) = \left(Y_{\inf}^{R_O}(k), k = 1, 2, ..., K\right)$$

Theorem 3.3

Suppose $Y \subset \mathcal{R}^N$ – a finite ranking set, $R_P \subset Y \times Y$ is the *Pessimist relation* [1].

The sequence of sets $Y_{\inf}^{R_P}(k)$ f the *smallest elements* forms *the partition* of the set Y and the proper ranking:

$$r_{\inf}(Y, R_P) = \left(Y_{\inf}^{R_P}(k), k = 1, 2, ..., K\right)$$

Ranking relations other than *an order* are very rarely encountered in practice and require "individual" treatment. Generally, rankings can be establish "with respect to the smallest (*dominating*) or minimal (*non dominated*) elements – generally, with respect to extreme elements of the set Y for ranking preference relation R (the so-called *externalizing* of the set Y)".

Definition 3.1

A *ranking cluster* (*category*) *of k-degree* is the set: $Y_{ext}^R(k)$, $k = 1, 2, ...$, where an index *ext* is understood as *inf*, or *min*, or *sup*, or *max*.

Definition 3.2

A ranking of the set Y with the *precedence relation* R is the sequence of sets

$$r_{ext}(Y, R) = \left(Y_{ext}^R(1), ... Y_{ext}^R(k), ..., Y_{ext}^R(K)\right) \quad (9)$$

constituting *the partition* of the set Y.

$Y_{ext}^R(k)$ is a kth subset (*category, cluster*) of extreme elements determined by means of the ranking preference relation R. The first cluster in the ranking is the cluster no. 1, then no. 2, etc....

Definition 3.3

A *ranking procedure* applying extreme elements is a set determination procedure $Y_{ext}^R(k)$, $k = 1, 2, ..., K$. *A stop criterion* is such a step K that $Y_{ext}^R(K) = \varnothing$. When $Y - \bigcup_{m=0}^{m=K-1} Y_{ext}^R(m) \neq \varnothing$ one substitutes:

$$Y_{ext}^R(K) = Y - \bigcup_{m=0}^{m=K-1} Y_{ext}^R(m).$$

If there are no extreme elements in a given subset of the ranking set Y, the adopted stop criterion allows stopping the procedure at the moment when "a partial ranking (approximated ranking)" is established (see Example 3.1) by including the subset without extreme element into the last cluster. Very interesting results for ranking building with any precedence relation, can be achieved, using a hybrid ranking algorithm (applying all four classes extreme elements: *inf, min, sup, max* in the same algorithm).

4. Conclusion

The paper presents new definitions of ranking procedure-related terms which allow establishing rankings for any sets of objects without the necessity of preparing classic scores expressed in numbers (ranking functions) (Definition 2.2). Chapter three touches upon precedence (preference) ranking relations. One does not have to assume that they are order or quasi-order relations. One proposed a very general definition for a ranking procedure (9) which allows establishing rankings on the basis of any precedence relation by means of recurrence term "extreme elements" (the smallest or the greatest or minimal (maximal), with respect to the adopted precedence ranking relation). Generally, a ranking is a sequence of subsets (clusters) of

the ranking set Y, giving a *partition*. Elements of given clusters are equivalent with respect to equivalence relation resulting from the recurrence partition of the set Y obtained according to the procedure basing on the definition of extreme elements. The cardinality and properties of clusters are a function of properties of the elements of the set Y, and most of all a function of properties of the adopted precedence ranking relation. Moreover, the paper covers examples of theorems concerning given (widely applied) ranking relations, such as the Pareto relation and the generalized Hurwicz relation [1, 6]. Application of any classic ranking functions to define ranking relations allows using the introduced terms, in particular general ranking procedure-related terms (9).

5. Bibliography

[1] Ameljańczyk A., *Multiple optimization*, WAT, Warszawa, 1986.

[2] Ameljańczyk A., *Optymalizacja wielokryterialna w problemach sterowania i zarządzania*, Ossolineum, Wrocław, 1984.

[3] Ameljańczyk A., "Metoda podziału zbioru obiektów na wielokryterialne klastry jakościowe", *Biuletyn Instytutu Systemów Informatycznych*, Nr 12, 1–8 (2013).

[4] Bouyssou D., Marchant T., "An axiomatic approach to noncompensatory sorting methods in MCDM, I: The case of two categories", *EJOR* ,178(1), 217–245 (2007).

[5] Brans J.P., Vincke Ph., "A preference ranking organization method: The PROMETHEE method for Multiple Criteria Decision-Making", *Management Science*, Vol. 31, No. 6, 647–656 (1985).

[6] Rasiowa H., *Wstęp do matematyki współczesnej*, PWN, Warszawa, 2005.

[7] Saaty T.L., "Rank from comparisons and from ratings in the analytic hierarchy/ network processes", *EJOR*, 168(2), 557–570 (2006).

[8] Seo F., Sakawa M., *Multiple Criteria Decision Analysis in Regional Planning*, D. Reidel-Kluwer, Dordrecht-Boston--Lancaster-Tokyo, 1988.

[9] Yu P.L., Leitmann G., "Compromise solutions, domination structures and Salukwadze's solution", *JOTA*, Vol. 13, 14–21(1974).

[10] Yu P.L., Leitmann G., "Nondominated decision and cone convexity in dynamic multicriteria decision problems", *JOTA*, Vol. 14, 195–203 (1974).

Matematyczne aspekty teorii rankingów

A. AMELJAŃCZYK

W pracy przedstawiono podstawy teoretyczne definiowania rankingów, bazujące na pojęciach teorii zbiorów i relacji. Zaprezentowano szereg nowych definicji pozwalających budować rankingi bez konieczności korzystania z typowych funkcji rankingowych. Wprowadzono pojęcie relacji rankingowego poprzedzania (niekoniecznie porządku) oraz przedstawiono ogólne algorytmy pozwalające budować rankingi w oparciu o definicje elementów ekstremalnych.

Słowa kluczowe: relacja rankingowego poprzedzania, pseudo ranking, ranking liniowy, elementy ekstremalne, klasteryzacja danych.