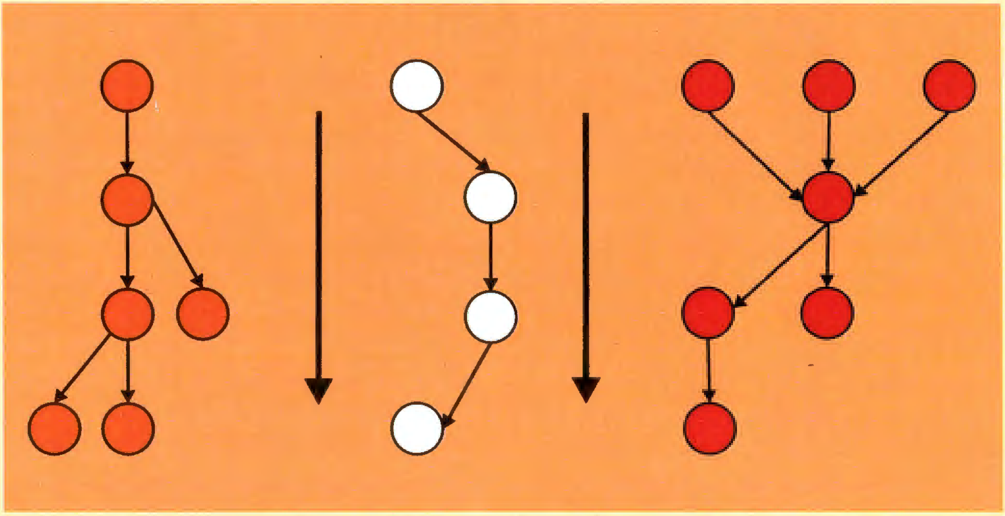
SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES

MULTICRITERIA ORDERING AND RANKING:
PARTIAL ORDERS, AMBIGUITIES AND APPLIED ISSUES



Jan W. Owsiński and Rainer Brüggemann
Editors

Warsaw 2008

SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES

MULTICRITERIA ORDERING AND RANKING: PARTIAL ORDERS, AMBIGUITIES AND APPLIED ISSUES

Jan W.Owsiński and Rainer Brüggemann
Editors

Warsaw 2008

This book is the outcome of the international workshop held in Warsaw in October 2008 within the premises of the Systems Research Institute. All papers were refereed and underwent appropriate modification in order to appear in the volume. The views contained in the papers are, however, not necessarily those officially held by the respective institutions involved, especially the Systems Research Institute of the Polish Academy of Sciences.

ISBN 83-894-7521-9
EAN 9788389475213

Technical editing and typesetting:
Jan W.Owsiński, Anna Gostyńska, Aneta M.Pielak

Theoretical Developments

Multicriteria Ordering and Ranking: Partial Orders, Ambiguities and Applied Issues
Jan W. Owsiński and Rainer Brüggemann, Editors

Properties of Mutual Rank Probabilities in Partially Ordered Sets

Karel De Loof[1], Bernard De Baets[2] and Hans De Meyer[1]

[1]Department of Applied Mathematics and Computer Science,
Ghent University, Krijgslaan 281 (S9), B-9000 Gent, Belgium,
(karel.deloof{hans.demeyer}@ugent.be)
[2]Department of Applied Mathematics, Biometrics and Process Control,
Ghent University, Coupure links 653, B-9000 Gent, Belgium,
(bernard.debaets@ugent.be)

An overview of certain properties of mutual rank probabilities in partially ordered sets (posets) together with some new theoretical and experimental results are given. In a first part, transitivity properties are studied. It is shown that the type of transitivity shown by mutual rank probabilities nicely fits into the cycle-transitivity framework tailor-made for expressing transitivity of reciprocal relations. In a second part, so-called linear extension majority cycles (LEM cycles) which can occur in posets with $n \geq 9$ elements are studied. Minimum cutting levels to avoid such LEM cycles are derived. In a last part approximation formulae for the mutual rank probabilities are established and their accuracy is compared for posets on up to 11 elements.

Keywords: posets, mutual rank probabilities, cycle-transitivity, linear extension majority cycles, cycle-free cuts, approximations

1. Overview

In many situations, one attempts to rank objects according to some well-defined properties. In an environmental context one could for example aim to rank chemical substances according to their environmental impact or to rank regions according to their soil pollution. However, frequently, in such a ranking attempt only a partial ranking of the objects can be established, since one e.g. lacks objective information to solve incomparabilities between objects. A possible approach is to insist on deriving a linear order by trying to remove those incomparabilities, such as using average ranks of the objects as to ensure comparability of each pair of objects. However, in some cases knowing the probability that one object is ranked higher than another can be sufficient. In this context, mutual rank probabilities are used as an objective and valuable quantification of such probabilities.

It is known that in the probability graph in which the vertices are the elements of a partially ordered set and the directed edges the ordered pairs of elements for which their mutual rank probability is at least $1/2$, cycles can occur. Such cycles are called linear extension majority cycles or LEM cycles. This observation has raised the interest in the type of transitivity exhibited by the mutual rank probabilities. In a first part of this contribution, we focus on this transitivity which is not yet fully characterized. We show, however, that a generalization and therefore weaker type of transitivity than mutual rank transitivity nicely fits into the cycle-transitivity framework. This cycle-transitivity framework has been tailor-made for expressing transitivity of reciprocal relations and has already shown in the past to be a powerful tool to concisely express transitivity of various reciprocal relations.

In a second part we then take a closer look at the LEM cycles themselves and present some experimental results. We specifically search for posets with n elements (for $n = 9, 10, 11, 12$) such that the minimum probability corresponding to an edge in a k-cycle (for $k = 3, 4, 5$) is maximum for all posets of size n.

Computing the mutual rank distribution is a computationally hard task and for posets of considerable size out of reach with current technology. Therefore, for practical purposes, one is interested in obtaining sufficiently good approximations for the mutual rank probabilities. In a third and last part we focus on such approximations and present an experiment in which accuracy is verified for all posets with up to 11 elements.

2. Preliminaries

2.1. Posets

A binary relation $\leq_P$ on a set P is called a *(partial) order relation* if it is a reflexive ($x \leq_P x$), antisymmetric ($x \leq_P y$ and $y \leq_P x$ imply $x =_P y$) and transitive

($x \leq_P y$ and $y \leq_P z$ imply $x \leq_P z$) relation. A *linear order relation* $\leq_P$ is an order relation in which every two elements are comparable ($x \leq_P y$ or $y \leq_P x$). If $x \leq_P y$ and $x \neq y$, we write $x <_P y$. If neither $x \leq_P y$ nor $x \geq_P y$, we say that x and y are *incomparable* and write $x \,||_P\, y$. A couple $(P, \leq_P)$, where P is a set of objects and $\leq_P$ is an order relation on P, is called a partially ordered set or *poset* for short. If no distinction between order relations has to be made, the index P in $\leq_P$ can be omitted. Moreover, if the order relation is clear from the context, we can simply denote the poset as P. A *chain* of a poset P is a subset of P in which every two elements are comparable. Dually, an *antichain* of a poset P is a subset of P in which every two elements are incomparable. A poset $(Q, \leq_Q)$ is called the *dual poset* of $(P, \leq_P)$ if $P = Q$ and $x \leq_Q y$ iff $x \geq_P y$ for all $x, y \in P$.

From here on, we consider only finite posets $(P, \leq_P)$. For elements $x, y \in P$ we say that y *covers* x, denoted as $x \prec_P y$, if $x <_P y$ and there exists no $z \in P$ such that $x <_P z <_P y$. In other words, x is smaller than y, and no third element is situated in between x and y.

A poset $(P, \leq_P)$ can be conveniently represented by a *covering graph* or so-called *Hasse diagram*, displaying the covering relation $\prec_P$. Note that $x <_P y$ if and only if there is a sequence of connected lines upwards from x to y. We call the elements in the Hasse diagram *vertices* and the lines representing the covering relation *edges*.

A permutation of the elements $x_1, x_2, \ldots, x_n$ of P which is consistent with the partial order, that is, such that $x_i <_P x_j$ implies $i < j$, is called a *linear extension* of P. More generally, a poset $(Q, \leq_Q)$ is called an *extension* of $(P, \leq_P)$ if $Q = P$ and if $x \leq_P y$ implies that $x \leq_Q y$. A linear extension is an extension in which every two elements are comparable. Each linear extension of P corresponds to a possible ranking of the elements of P which obeys the order relation $\leq_P$. Both concepts are essentially the same.

We denote by $p(x_i < x_j)$ the fraction of linear extensions of P in which x_i precedes x_j. If the space of all linear extensions of P is equipped with the uniform measure, the position of $x \in P$ in a linear extension can be regarded as a discrete random variable X with values in $\{1, \ldots, n\}$. Since $p(x_i < x_j) = \text{Prob}(X_i < X_j)$, it is called a mutual rank probability.

2.2. Reciprocal relations

A fuzzy relation R on A is an $A^2 \to [0, 1]$ mapping that expresses the degree of relationship between elements of A: $R(a,b)=0$ means a and b are not related at all, $R(a,b)=1$ expresses full relationship, while $R(a,b)\in]0,1[$ indicates a partial degree of relationship only. For such relations, the concept of T-transitivity is very natural.

Let T be a t-norm. A fuzzy relation R on A is called T-transitive if for any $(a,b,c)\in A^3$ it holds that $T(R(a,b),R(b,c))\leq R(a,c)$. The three basic t-norms are $T_{\mathbf{M}}$, the minimum operator, $T_{\mathbf{P}}$, the ordinary product, and $T_{\mathbf{L}}$, the Łukasiewicz t-norm.

Another class of $A^2\to[0,1]$ mappings are the reciprocal relations Q satisfying $Q(a,b)+Q(b,a)=1$, for any $a,b\in A$. They arise in the context of pairwise comparison. Though the semantics of reciprocal relations and fuzzy relations are different, the concept of T-transitivity is sometimes formally applied to reciprocal relations as well. However, more often the transitivity properties of reciprocal relations can be characterized as of one of various kinds of stochastic transitivity. The following general formulation of stochastic transitivity has been proposed in De Baets and De Meyer (2005).

Let g be a commutative increasing $[1/2,1]^2\to[1/2,1]$ mapping. A reciprocal relation Q on A is called stochastic transitive w.r.t. g if for any $(a,b,c)\in A^3$ it holds that $(Q(a,b)\geq 1/2 \wedge Q(b,c)\geq 1/2) \Rightarrow Q(a,c)\geq g(Q(a,b),Q(b,c))$.

This definition includes strong stochastic transitivity when $g=\max$, moderate stochastic transitivity when $g=\min$, weak stochastic transitivity when $g=1/2$, and λ-transitivity, with $\lambda\in[0,1]$, when $g=\lambda\max+(1-\lambda)\min$. Clearly, strong stochastic transitivity implies λ-transitivity, which implies moderate stochastic transitivity, which, in turn, implies weak stochastic transitivity.

3. Transitivity properties

3.1. Transitivity of mutual ranking probabilities

For any poset $P=\{x_1,\ldots,x_n\}$,

$$Q_P(x_i,x_j)=\mathrm{Prob}(X_i<X_j)=p(x_i<x_j)\,. \tag{1}$$

defines the reciprocal relation Q_P, which is the relational expression of the mutual rank probabilities of P.

The problem of characterizing the transitivity of Q_P was already raised by Fishburn (1973). For any $u,v\in[0,1]$, define $\delta(u,v)$ as

$$\delta(u,v)=\inf\{p(x_i<x_k)\mid p(x_i<x_j)\geq u, p(x_j<x_k)\geq v\}\,,$$

where the infimum is taken over all choices of P and distinct x_i, x_j, x_k. Fishburn proved that

$$\begin{aligned}
&\delta(u,v) = 0 \text{ if } u+v<1\,,\\
&u+v-1 \leq \delta(u,v) \leq \min(u,v)\,,\\
&\delta(u,1-u) \leq 1/e\,,\\
&\delta(u,v) \leq 1-(1-u)(1-v)(1-\ln[(1-u)(1-v)])\,. \qquad (2)
\end{aligned}$$

A non-trivial lower bound on δ was proved in Kahn and Yu (1998) via geometric arguments. Define

$$\gamma(u,v) = \inf\{\text{Prob}(Y_i < Y_k) \mid \text{Prob}(Y_i < Y_j) \geq u, \text{Prob}(Y_j < Y_k) \geq v\}\,,$$

where the infimum is taken over $Y = (Y_1, \ldots, Y_n)$ chosen uniformly from some n-dimensional compact convex subset of $\mathbb{R}^n$. Since $\delta(u,v) \geq \gamma(u,v)$, the function γ provides a lower bound for δ. Kahn and Yu (1998) proved that

$$\gamma(u,v) = \begin{cases} 0 & \text{, if } u+v<1\,,\\ \min(u,v) & \text{, if } u+v-1 \geq \min(u^2,v^2)\,,\\ \dfrac{(1-u)(1-v)}{u+v-2\sqrt{u+v-1}} & \text{, otherwise}\,. \end{cases} \qquad (3)$$

Since the mutual rank probability relation Q_P is a reciprocal relation, we want to investigate how these transitivity results translate into the transitivity framework called cycle-transitivity.

3.2. Cycle-transitivity framework

In the cycle-transitivity framework (De Baets et al., 2006), for a reciprocal relation Q on A, the quantities

$$\alpha_{abc} = \min(Q(a,b), Q(b,c), Q(c,a))\,,\ \beta_{abc} = \text{med}(Q(a,b), Q(b,c), Q(c,a))\,,$$

$$\gamma_{abc} = \max(Q(a,b), Q(b,c), Q(c,a))\,,$$

are defined for all $(a,b,c) \in A^3$. Obviously, $\alpha_{abc} \leq \beta_{abc} \leq \gamma_{abc}$. Also, the notation $\Delta = \{(x,y,z) \in [0,1]^3 \,|\, x \leq y \leq z\}$ will be used.

A function $U : \Delta \rightarrow \mathbb{R}$ is called an upper bound function if it satisfies:

(i) $U(0,0,1) \geq 0$ and $U(0,1,1) \geq 1$;

(ii) for any $(\alpha,\beta,\gamma) \in \Delta$:

$$U(\alpha,\beta,\gamma) + U(1-\gamma,1-\beta,1-\alpha) \geq 1\,. \tag{4}$$

The function $L : \Delta \to \mathbb{R}$ defined by

$$L(\alpha,\beta,\gamma) = 1 - U(1-\gamma,1-\beta,1-\alpha) \tag{5}$$

is called the dual lower bound function of a given upper bound function U. Inequality (4) simply expresses that $L \leq U$.

A reciprocal relation Q on A is called cycle-transitive w.r.t. an upper bound function U if for any $(a,b,c) \in A^3$ it holds that

$$L(\alpha_{abc},\beta_{abc},\gamma_{abc}) \leq \alpha_{abc} + \beta_{abc} + \gamma_{abc} - 1 \leq U(\alpha_{abc},\beta_{abc},\gamma_{abc})\,, \tag{6}$$

where L is the dual lower bound function of U.
Due to the built-in duality, it holds that if (6) is true for some (a,b,c), then this is also the case for any permutation of (a,b,c). In practice, it is therefore sufficient to check (6) for a single permutation of any $(a,b,c) \in A^3$. Alternatively, due to the same duality, it is also sufficient to verify the right-hand inequality (or equivalently, the left-hand inequality) for two permutations of any $(a,b,c) \in A^3$ (not being cyclic permutations of one another), e.g. (a,b,c) and (c,b,a). Hence, (6) can be replaced by

$$\alpha_{abc} + \beta_{abc} + \gamma_{abc} - 1 \leq U(\alpha_{abc},\beta_{abc},\gamma_{abc})\,. \tag{7}$$

Note that a value of $U(\alpha,\beta,\gamma)$ equal to 2 will often be used to express that for the given values there is no restriction at all (indeed, $\alpha+\beta+\gamma-1$ is always bounded by 2). In Fig.1 the relationships between some types of (cycle-)transitivity are shown.

3.3. Cycle-transitivity of mutual rank probabilities

If we translate the bounds (2) and (3) into the cycle-transitivity framework, we obtain that (3) provides an upper bound $U(\alpha,\beta,\gamma)$ on $\alpha+\beta+\gamma-1$, whereas (2) provides a lower bound, which is, however, less stringent than the lower bound function $L(\alpha,\beta,\gamma)$ associated to $U(\alpha,\beta,\gamma)$ by (5). Surprisingly, the upper bound function $U(\alpha,\beta,\gamma)$ which is the equivalent of (3), is very simple.

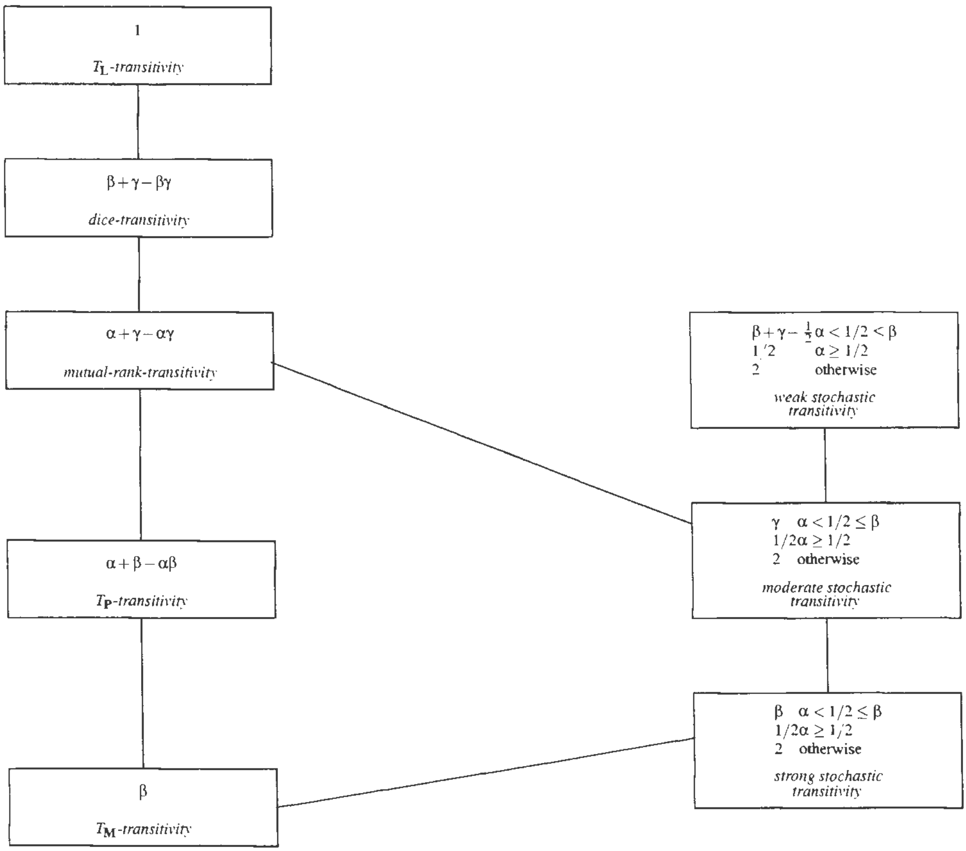


Fig.1: Hasse-diagram with different types of cycle-transitivity characterized by their upper bound function $U(\alpha,\beta,\gamma)$. Bottom up (partial) ordering of transitivity types is in agreement with the 'stronger than' relation.

Proposition 1. The reciprocal relation Q_P generated by the mutual rank probabilities in a poset P is cycle-transitive w.r.t. to the upper bound function U given by $U(\alpha,\beta,\gamma)=\alpha+\gamma-\alpha\gamma$.

In view of the above result, cycle-transitivity w.r.t. the upper bound function U given by $U(\alpha,\beta,\gamma) = \alpha+\gamma-\alpha\gamma$ is called mutual-rank-transitivity.

Inspection of Fig.1 reveals that mutual-rank-transitivity is stronger than dice-transitivity but weaker than $T_{\mathbf{P}}$-transitivity and also weaker than moderate stochastic transitivity. Also note that mutual rank transitivity does not imply weakly stochastic transitivity.

4. Cycle-free cuts

4.1. Linear extension majority cycles

The linear extension majority (LEM) relation of a poset P is the binary relation $\prec_{LEM}$ such that $x \prec_{LEM} y$ if $\mathrm{p}(x<y) > \mathrm{p}(y<x)$.

Aigner (1988), Fishburn (1973, 1974, 1976), Ganter et al. (1987), and Gehrlein and Fishburn (1990a) have given example posets P where the relation $\prec_{LEM}$ contains cycles, *i.e.* where a subset $\{x_1,x_2,...,x_m\}$ of elements of P exists such that $x_1 \prec_{LEM} x_2 \prec_{LEM} \cdots \prec_{LEM} x_m \prec_{LEM} x_1$. These cycles are referred to as LEM cycles on m elements, or m-cycles for short.

By using exhaustive enumeration (Gehrlein and Fishburn, 1990a), it was found that no poset with less than 9 elements contains a LEM cycle. In Gehrlein and Fishburn (1990b), the likelihood of LEM cycles up to $n = 12$ is estimated by generating random partial orders. However, to our best knowledge no exact counts for $n \geq 9$ are known. Counting the number of posets with $n \geq 9$ elements having a LEM cycle is no trivial task. The number of posets with n elements quickly explodes for increasing n, and even worse, the number of linear extensions of each poset can be exponential in n. Brinkmann and McKay (2002) describe a very efficient method to construct pairwise non-isomorphic posets, which allows them to enumerate posets on up to 16 points. As an illustration of the size of the problem, the number of (unlabeled) posets of size 9 to 16 are shown in Table 1.

In order to be able to count the number of posets with LEM cycles for n on up to 12 we used the enumeration algorithm of Brinkmann and McKay (2002) to enumerate all posets of size n. Subsequently, the algorithm developed by the present authors (De Loof et al., 2006) based on the so-called lattice of ideals representation of a poset has been used to compute all mutual rank probabilities. This algorithm avoids enumerating all linear extensions for the computation of the mutual rank probabilities.

n	number of posets
9	183231
10	2567284
11	46749427
12	1104891746
13	33823827452
14	1338193159771
15	68275077901156
16	4483130665195087

Table 1: Number of unlabeled posets for $n = 9, 10, \ldots 16$.

n	3-cycles	4-cycles	5-cycles
9	5	0	0
10	138	6	0
11	5439	89	0
12	204935	2677	5

Table 2: The number of n-element posets that contain LEM cycles.

Its running time is linear in the number of ideals, which can still be exponential in n. However, in most cases the number of ideals is much smaller than the number of linear extensions. The results of this counting procedure are summarized in Table 2.

4.2. Minimum cutting levels to remove LEM-cycles

We now want to determine a minimum cutting level δ such that the graph of the crisp relation, obtained from the mutual rank probability relation by setting its elements smaller than or equal to δ equal to 0 and its other elements equal to 1, is free from cycles. In other words, we want to obtain the minimum δ such that at least one mutual rank probability in any LEM cycle is smaller than or equal to δ.

The strict cut at value $c \in [1/2, 1[$ of a reciprocal relation Q defined on a set A, is the crisp relation Q^c defined by

$$Q^c(x,y) = \begin{cases} 1\,, & \text{if } Q(x,y) > c\,, \\ 0\,, & \text{otherwise.} \end{cases}$$

The m-cutting-level is the smallest number c_m such that for any finite poset the strict cut Q^{c_m} of the mutual rank probability relation Q is free of LEM cycles of length $\leq m$.

From the previous section it is known that the mutual rank probability relation of any poset is mutual-rank-transitive. This property allows us to derive an upper bound for the m-cutting-level. Denote by $\mathcal{MR}$ the set of all mutual rank probability relations, and by $\mathcal{CR}$ the set of all finite reciprocal relations that are mutual-rank-transitive. Clearly $\mathcal{MR} \subseteq \mathcal{CR}$.

Let us first consider the case $m = 3$. The 3-cutting-level is the largest value that can be obtained for the minimum of three relational elements forming a 3-cycle in a mutual rank probability relation, i.e.:

$$c_3 = \sup_{Q \in \mathcal{MR}} \max_{i,j,k} \min(q_{ij}, q_{jk}, q_{ki})\,.$$

An upper bound $\bar{c}_3$ for c_3 is

$$\bar{c}_3 = \sup_{Q \in \mathcal{CR}} \max_{i,j,k} \min(q_{ij}, q_{jk}, q_{ki})\,.$$

To obtain $\bar{c}_3$, it is therefore sufficient to consider the mutual-rank-transitive reciprocal relations with 3 elements and to find a set of values of q_{12}, q_{23}, q_{31} such that $\min(q_{12}, q_{23}, q_{31})$ is maximal. From symmetry considerations it follows that the investigation may be restricted to relations for which $q_{12} = q_{23} = q_{31} = q$. Expressing that such a relation be mutual-rank-transitive, yields the condition $q^2 \leq 3q - 1 \leq 2q - q^2$, or, equivalently $(3 - \sqrt{5})/2 \leq q \leq (\sqrt{5} - 1)/2$. As a consequence we have that

$$c_3 \leq \bar{c}_3 = (\sqrt{5} - 1)/2 = 0.618034\,.$$

In general, for $m \geq 3$, let us introduce the following notations. We number the nodes of the complete graph from 1 to m and consider the cycle of length m in which the nodes appear in the natural order. The edges of this cycle are attributed equal weight $a_1^{(m)}$. We need to find the maximal value of $a_1^{(m)}$ such that the reciprocal relation which underlies the graph is mutual-rank-transitive. By symmetry, we attribute to

m	$a_1^{(m)} = \bar{c}_m$	$a_2^{(m)}$	$a_3^{(m)}$	polynomial equation
3	0.61803			$x^2+x-1=0$
4	0.66667	0.50000		$3x^2-2x=0$
5	0.70711	0.58579		$2x^2-1=0$
6	0.72361	0.61803	0.50000	$5x^2-5x+1=0$
7	0.74227	0.65270	0.53209	$3x^3-3x+1=0$

Table 3: Weights of graphs with m nodes that represent max-optimal mutual-rank-transitive reciprocal relations whose strict cuts at $a_1^{(m)}$ are m-cycle free.

the edges starting at node i and ending at node $(i+j) \mod m$, irrespective of i, the same weight $a_j^{(m)}$, where $j \in \{1,2,\ldots,m-1\}$. Clearly, since the reciprocal relation underlying this graph should be reciprocal, it holds that $a_j^{(m)} = 1 - a_{m-j}^{(m)}$ for all $j \in \{1,2,\ldots,m-1\}$. We call these weighted graphs max-optimal.

In the first column of Table 3 the maximum values of $a_1^{(m)}$, which also yield the values $\bar{c}_m$, are listed for $m \in \{3,4,5,6,7\}$. In the other colums the values are shown of $a_j^{(m)}$ for $j = 2,\ldots,\lfloor m/2 \rfloor$ (the remaining ones can be found by complementation). In the final column, we mention the polynomial equation whose largest real root provides the value of $a_1^{(m)}$. The reader can easily verify that with these weights one obtains graphs whose underlying reciprocal relation is mutual-rank-transitive.

At present, we have not yet been able to compute the limiting value of $a_1^{(m)}$ as $m \to \infty$. On the other hand, it is known from the work of Yu (1998) that the strict cut of any mutual rank probability relation at the value

$$\rho = \frac{1+(\sqrt{2}-1)\sqrt{2\sqrt{2}-1}}{2} = 0.78005$$

yields a crisp relation that is transitive, and thus obviously m-cycle free for any $m > 0$. Therefore, it must hold that

$$\lim_{m\to\infty} c_m \leq \rho\,.$$

Note, however, that this inequality does not imply $\lim_{m\to\infty} \bar{c}_m \leq \rho$, since $\bar{c}_m$ is only an upper bound for c_m.

These theoretical considerations on strict cutting levels concern the classes of either all mutual rank probability relations or all mutual-rank-transitive relations. Of course, the given bounds are not sharp when we restrict to posets of size less than given constant.

4.3. Experimental computation of minimal cutting levels

Again, by using the technique sketched in Section 3.1, we perform some tests on all posets of size n (with $n = 9, \ldots, 12$) to obtain the minimum cutting level needed to obtain cycle-free relations. In Table 4, the minimum cutting levels to avoid m-cycles in n-element posets ($m = 3, 4, 5$ and $n = 9, \ldots, 12$) is shown. Note that the minimum cutting levels for avoiding m-cycles are monotonically increasing for increasing n, since one can easily establish a poset of $n+1$ elements from an n-element poset with equal minimum cutting level by adding e.g. an element which is smaller than all n elements. Note that for $n = 11$ no higher minimum cutting level for avoiding 4-cycles is found than for $n = 10$. In Fig. 2-7 all posets giving rise to the non-trivial minimum cutting levels listed in Table 4 are depicted.

n	3-cycles	4-cycles	5-cycles
9	0.5031447	0.5	0.5
10	0.5039683	0.5028490	0.5
11	0.5061947	0.5028490	0.5
12	0.5073505	0.5086657	0.5003979

Table 4: Minimum cutting level to avoid cycles in posets of size $n = 9, \ldots, 12$.

12-element poset with the highest cutting level to avoid 4-cycles: it holds that $\text{Prob}(7 < 5) = \text{Prob}(5 < 8) = \text{Prob}(8 < 6) = \text{Prob}(6 < 7) = \frac{7396}{14540}$.

5. Approximating mutual ranking probabilities

Since computing the mutual rank probabilities quickly becomes infeasible for larger posets (De Loof et al., 2006), the question arises whether good approximations can be made. By using the so-called Markov Chain Monte Carlo method, one could sample almost uniformly from the set of linear extensions of a given poset (Bubley and

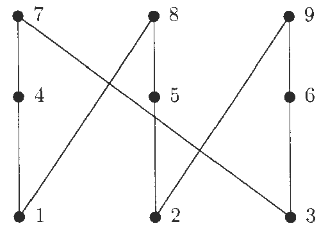


Fig.2: 9-element poset with the highest cutting level to avoid 3-cycles; it holds that $\mathrm{Prob}(2 < 1) = \mathrm{Prob}(1 < 3) = \mathrm{Prob}(3 < 2) = \frac{720}{1431}$, and due to symmetry also that $\mathrm{Prob}(8 < 7) = \mathrm{Prob}(7 < 9) = \mathrm{Prob}(9 < 8) = \frac{720}{1431}$.

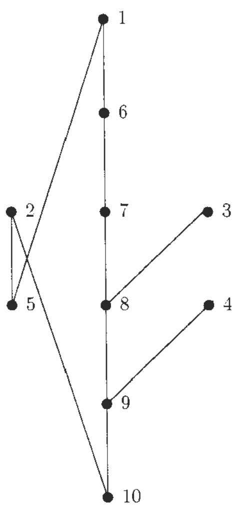


Fig. 3: 10-element poset with the highest cutting level to avoid 3-cycles (dual poset not depicted); it holds that $\mathrm{Prob}(3 < 6) = \frac{512}{1008}$ and $\mathrm{Prob}(6 < 2) = \mathrm{Prob}(2 < 3) = \frac{508}{1008}$.

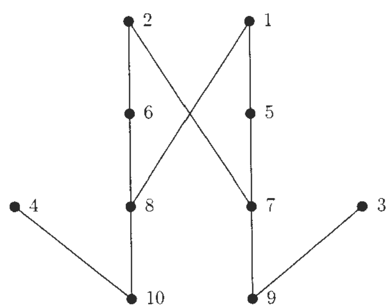


Fig. 4: 10-element poset with the highest cutting level to avoid 4-cycles (dual poset not depicted); it holds that $\mathrm{Prob}(3 < 5) = \mathrm{Prob}(5 < 4) = \mathrm{Prob}(4 < 6) = \mathrm{Prob}(6 < 3) = \frac{1765}{3510}$.

Dyer, 1999). This (almost) uniform sample could then be used to estimate the rank probabilities. With this approach, one is able to obtain good approximations, albeit

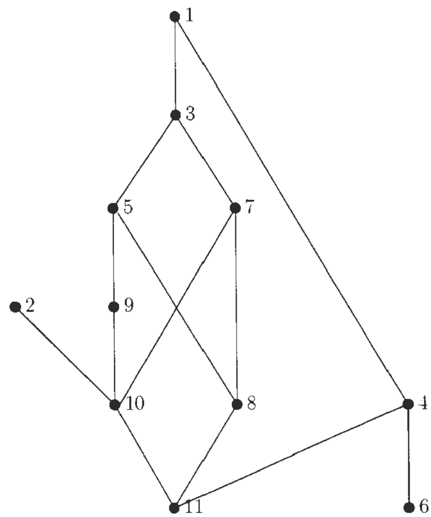


Fig. 5: 11-element poset with the highest cutting level to avoid 3-cycles; it holds that $\mathrm{Prob}(2 < 5) = \frac{1144}{2260}$, $\mathrm{Prob}(5 < 4) = \frac{1146}{2260}$ and $\mathrm{Prob}(4 < 2) = \frac{1145}{2260}$.

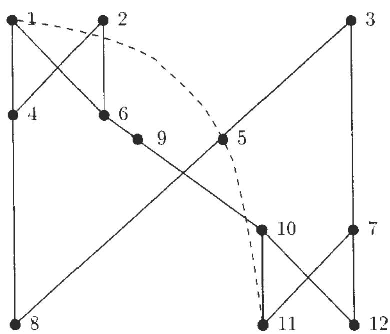


Fig. 6: 12-element poset with the highest cutting level to avoid 3-cycles; it holds that $\mathrm{Prob}(7 < 9) = \mathrm{Prob}(9 < 4) = \frac{6214}{12244}$ and $\mathrm{Prob}(4 < 7) = \frac{6212}{12244}$.

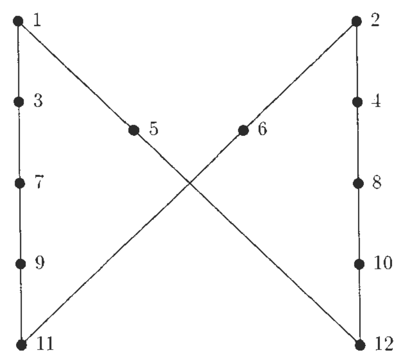


Fig. 7: 12-element poset with the highest cutting level to avoid 4-cycles; it holds that $\mathrm{Prob}(7 < 5) = \mathrm{Prob}(5 < 8) = \mathrm{Prob}(8 < 6) = \mathrm{Prob}(6 < 7) = \frac{7396}{14540}$.

at the expense of a good deal of computation time. It is worthwhile to investigate whether rough but quick approximations can be suggested. Ideally, one would like to have an approximation formula for $\mathrm{p}(x < y)$ combining variables that can be computed efficiently. Brüggemann et al. (2004) have suggested such a formula. First, they define $Q(x,y) = (N_u(x,y)+1)/(N_d(x,y)+1)$, where $N_u(x,y)$ is the number of objects above x which are not at the same time above y, and $N_d(x,y)$ is the number of objects under x which are not at the same time under y. The probability that x is ranked higher than y is then approximated as $\hat{\mathrm{p}}(x > y) = Q(y,x)/(1+Q(x,y))$. It is easily checked that

$$\hat{\mathrm{p}}_1(x > y) = \frac{[N_u(y,x)+1]\cdot[N_d(x,y)+1]}{[N_u(x,y)+1]\cdot[N_d(y,x)+1]+[N_u(y,x)+1]\cdot[N_d(x,y)+1]}. \quad (8)$$

In what follows we will suggest two variants of (8) and compare their accuracy for n up to 11. Finally, we show how they can be applied iteratively.

As a first variant of (8) we chose to include elements at the same time above x and y and elements at the same time under x and y. If we define $N_d(x)$ as the number of objects above x, and $N_u(x)$ as the number of objects under x, we can write

$$\hat{\mathrm{p}}_2(x > y) = \frac{[N_u(y)+1]\cdot[N_d(x)+1]}{[N_u(x)+1]\cdot[N_d(y)+1]+[N_u(y)+1]\cdot[N_d(x)+1]}. \quad (9)$$

The denominator in (8) is an approximation of the total number of linear extensions, while the numerator is an approximation of the number of linear extensions in which x is ranked higher than y (up to a common factor α in both numerator and denominator). Let us denote U as the set of elements above y, and D as the set of elements under x. If one sees $N_u(y)$ as an approximation of the number of linear extensions of U, and $N_d(x)$ as an approximation of the number of linear extensions of D (up to a common factor α), $[N_u(y,x)+1]\cdot[N_d(x,y)+1]$ approximates the number of linear extensions of the decomposable poset where each element of U is made greater than each element of D by adding an element, say z, such that $z < u$ for all $u \in U$ and $d < z$ for all $d \in D$. However, if one decides to approximate, again up to a common factor α, the number of linear extensions by the number of elements, it would be more natural to rewrite $[N_u(y,x)+1]\cdot[N_d(x,y)+1]$ as $N_u(y,x)+N_d(x,y)+1$. If we adapt the denominator correspondingly, we obtain a second variant

$$\hat{\mathrm{p}}_3(x > y) = \frac{N_u(y,x)+N_d(x,y)+1}{N_u(x,y)+N_d(y,x)+N_u(y,x)+N_d(x,y)+2}. \quad (10)$$

A third variant is now obtained by combining both ideas:

$$\hat{\mathrm{p}}_4(x > y) = \frac{N_u(y) + N_d(x) + 1}{N_u(x) + N_d(y) + N_u(y) + N_d(x) + 2}. \tag{11}$$

Given a poset P, let us define a mutual rank matrix $M(P)$ containing the exact mutual ranks $\mathrm{p}(x < y)$ for each $(x, y) \in P^2$, and four approximate mutual rank matrices $\hat{M}_i(P)$, for $i = 1, \ldots, 4$, containing the approximated mutual ranks $\hat{\mathrm{p}}_i(x < y)$ for each pair $(x, y) \in P^2$ of incomparable elements, *i.e.* each pair $(x, y) \in P^2$ for which $x||y$, and 0 or 1 for each pair $(x, y) \in P^2$ of comparable elements depending upon whether $x < y$ or $y < x$. In order to have an idea of the approximation error,

(i) the maximal absolute componentwise difference between $M(P)$ and $\hat{M}_i(P)$ over all posets P of size n (Table 5),

(ii) the maximal absolute componentwise difference between $M(P)$ and $\hat{M}_i(P)$ averaged over all posets P of size n (Table 6) and

(iii) the 1-norm of the absolute difference $|M(P) - M\hat{(P)}_i|$ (Table 7),

are computed for each n ($n = 1, \ldots, 4$) and the approximations (8)–(11).

n	$\hat{\mathrm{p}}_1(x > y)$	$\hat{\mathrm{p}}_2(x > y)$	$\hat{\mathrm{p}}_3(x > y)$	$\hat{\mathrm{p}}_4(x > y)$
4	0.067	0.067	0.083	0.167
5	0.104	0.095	0.104	0.208
6	0.135	0.141	0.150	0.233
7	0.173	0.167	0.178	0.264
8	0.207	0.193	0.222	0.285
9	0.233	0.215	0.257	0.307
10	0.270	0.247	0.289	0.326
11	0.303	0.273	0.320	0.342

Table 5: Maximal absolute componentwise difference between $M(P)$ and $\hat{M}_i(P)$ over all posets P of size n for each approximation formula

As can be seen from Tables 5-7, approximation formula (9) seems to perform best for $n \geq 7$. Also note that the approximation errors grow quite rapidly with larger n.

n	$\hat{p}_1(x>y)$	$\hat{p}_2(x>y)$	$\hat{p}_3(x>y)$	$\hat{p}_4(x>y)$
4	0.011	0.015	0.014	0.047
5	0.034	0.038	0.038	0.092
6	0.062	0.062	0.066	0.134
7	0.090	0.082	0.093	0.167
8	0.114	0.098	0.115	0.191
9	0.133	0.111	0.133	0.210
10	0.150	0.122	0.148	0.223
11	0.165	0.131	0.161	0.234

Table 6: Maximal absolute componentwise difference between $M(P)$ and $\hat{M}_i(P)$ averaged over all posets P of size n for each approximation formula

Each of the formulae (8)–(11) can be used in an iterative manner by "fuzzifying" the notions "above" and "under" in the calculation of N_d and N_u. The iterative scheme corresponding to $\hat{p}_1(x>y)$ is given as an example. Schemes corresponding to the other variants are analogous and left to the reader.
We define

$$C_0(x,y) = \begin{cases} 1 & \text{if } x \leq y, \\ 0 & \text{if } x > y \text{ or } x \| y. \end{cases} \tag{12}$$

For $i \geq 1$ and $(x,y) \in P^2$ define

$$N_{d,i}(x,y) = \sum_{z \in P} C_{i-1}(z,x) \cdot [1 - C_{i-1}(z,y)] \tag{13}$$

and

$$N_{u,i}(x,y) = \sum_{z \in P} C_{i-1}(x,z) \cdot [1 - C_{i-1}(y,z)]. \tag{14}$$

We then define

$$C_i(x,y) = \frac{[N_{u,i}(y,x)+1] \cdot [N_{d,i}(x,y)+1]}{[N_{u,i}(x,y)+1] \cdot [N_{d,i}(y,x)+1] + [N_{u,i}(y,x)+1] \cdot [N_{d,i}(x,y)+1]}. \tag{15}$$

n	$\hat{p}_1(x>y)$	$\hat{p}_2(x>y)$	$\hat{p}_3(x>y)$	$\hat{p}_4(x>y)$
4	0.004	0.004	0.005	0.016
5	0.009	0.008	0.011	0.023
6	0.014	0.013	0.017	0.030
7	0.018	0.016	0.022	0.036
8	0.022	0.019	0.026	0.041
9	0.025	0.022	0.029	0.046
10	0.028	0.023	0.031	0.050
11	0.030	0.025	0.033	0.053

Table 7: The 1-norm of the absolute difference $|M(P) - \hat{M}_i(P)|$ for each approximation formula

n	(1)		(2)		(3)	
	single	10 iter.	single	10 iter.	single	10 iter.
4	0.083	0.083	0.014	0.021	0.005	0.007
5	0.104	0.094	0.038	0.038	0.011	0.010
6	0.150	0.102	0.066	0.054	0.017	0.012
7	0.178	0.113	0.093	0.067	0.022	0.014
8	0.222	0.134	0.115	0.075	0.026	0.016
9	0.257	0.157	0.133	0.081	0.029	0.018
10	0.289	0.181	0.148	0.086	0.031	0.019
11	0.320	0.204	0.161	0.090	0.032	0.021

Table 8: Comparison between applying $\hat{p}_3(x>y)$ one single time or in 10 iterations; **(1)** maximal absolute componentwise difference between $M(P)$ and $\hat{M}_i(P)$ over all posets P of size n, **(2)** maximal absolute componentwise difference between $M(P)$ and $\hat{M}_i(P)$ averaged over all posets P of size n and **(3)** The 1-norm of the absolute difference $|M(P) - \hat{M}_i(P)|$

Remark that e.g. the expression in the summation of (13) is a fuzzification of the logical expression

$$z \leq x \wedge \neg(z \leq y). \tag{16}$$

Using the iterative versions of $\hat{\mathrm{p}}_1(x,y)$ and $\hat{\mathrm{p}}_4(x,y)$ one obtains less precise results than the corresponding non-iterative formulae, and for $\hat{\mathrm{p}}_2(x,y)$ only slightly better results. However, as Table 8 shows, the iterative application of $\hat{\mathrm{p}}_3(x,y)$ gains more than 50% in accuracy after iterating 10 times. Iterating more than 10 times only slightly improves approximation accuracy, thus no convergence towards the exact value can be observed.

Presently, we are investigating whether still better approximation formulae can be suggested.

References

Aigner M. (1988), *Combinatorial Search*, Wiley-Teubner, Chichester.

Brinkmann G. and McKay B.D. (2002) Posets on up to 16 points, *Order***19**147–179.

Brüggemann R. , Lerche D.B. and Sørensen P.B. (2004) First attempts to relate structures of Hasse diagrams with mutual probabilities, In: *Order Theory in Environmental Sciences. Integrative Approaches. The 5th workshop held at the National Environmental Research Institute (NERI), Roskilde, Denmark, November 2002*, 7–51.

Bubley R. and Dyer M. (1999) Faster random generation of linear extensions, *Discrete Mathematics* **201**, 81–88.

Chiclana F. , Herrera F. , Herrera-Viedma E. and Martínez L. (2003) A note on the reciprocity in the aggregation of fuzzy preference relations using OWA operators, *Fuzzy Sets and Systems* **137**, 71–83.

David H.A. (1963)*The Method of Paired Comparisons. Griffin's Statistical Monographs & Courses***12**, Charles Griffin & Co. Ltd., London.

De Baets B. and De Meyer H. (2005) Transitivity frameworks for reciprocal relations: cycle-transitivity versus FG-transitivity, *Fuzzy Sets and Systems* **152**, 249–270.

De Baets B. and De Meyer H. (2008) On the cycle-transitive comparison of artificially coupled random variables,*International Journal of Approximate Reasoning* **47**, 306–322.

B. De Baets, H. De Meyer, B. De Schuymer and S. Jenei, *Cyclic evaluation of transitivity of reciprocal relations*, Social Choice and Welfare **26** (2006), 217–238.

De Loof K. , De Meyer H. and De Baets B. (2006) Exploiting the lattic of ideals representation of a poset, *Fundamenta Informaticae* **71**, 309–321.

K. De Loof, B. De Baets, H. De Meyer and R. Brüggemann (2008) A hitchhiker's guide to poset ranking, *Combinatorial Chemistry and High Throughput Screening* in press.

B. De Schuymer, H. De Meyer and B. De Baets (2007) Extreme copulas and the comparison of ordered lists, *Theory and Decision* **62**, 195–217.

De Meyer H. , De Baets B. and De Schuymer B. (2007) On the transitivity of comonotonic and countermonotonic comparison of random variables, *Journal of Multivariate Analysis* **98**, 177–193.

De Schuymer B. , De Meyer H. , De Baets B. and Jenei S. (2003) On the cycle-transitivity of the dice model, *Theory and Decision* **54**, 261–285.

De Schuymer B. , De Meyer H. and De Baets B. (2005) Cycle-transitive comparison of independent random variables, *Journal of Multivariate Analysis* **96** , 352–373.

Dutta B. and Laslier J.-F. (1999) Comparison functions and choice correspondences, *Social Choice and Welfare* **16**, 513–532.

Fishburn P. (1973) Proportional transitivity in linear extensions of ordered sets, *Journal of Combinatorial Theory* **10**, 327–352.

Fishburn P. (1974) On the family of linear extensions of a partial order, *Journal of Combinatorial Theory* **17**, 240–243.

Fishburn P. (1976) On linear extension majority graphs of partial orders, *Journal of Combinatorial Theory* **21**, 65–70.

Fishburn P. (1985)*Interval Orders and Interval Graphs*, Wiley, New York.

Fishburn P. (1986) *Binary choice probabilities: On the varieties of stochastic transitivity, Journal of Mathematical Psychology* **B41**, 48–60.

Ganter B. , Häfner G. and Poguntke W. (1987) On the linear extensions of ordered sets with a symmetry, *Discrete Mathematics* **63**, 153–156.

Gehrlein W.V. and Fishburn P.C. (1990a) Linear extension majority cycles of small ($n \neq 9$) partial orders, *Computers and Mathematics with Applications* **20**, 41–44.

Gehrlein W.V. and Fishburn P.C. (1990b) Linear extension majority cycles on partial orders, *Annals of Operations Research* **23**, 311–322.

Herrera-Viedma E. , Herrera F. , Chiclana F. and Luque M. (2004) Some issues on consistency of fuzzy preference relations, *European Journal of Operational Research* **154** 98–109.

Kahn J. and Yu Y. (1998) Log-concave functions and poset probabilities, *Combinatorica* **18**, 85–99.

Laslier J.-F. (1997) *Tournament Solutions and Majority Voting. Studies in Economic Theory* **7**, Springer-Verlag, Berlin.

Nelsen R. (2005)*An Introduction to Copulas. Lecture Notes in Statistics* **139**, second edition, Springer-Verlag, New York.

Yu Y. (1998) On Proportional Transitivity of Ordered Sets, *Order* **15**, 87–95.

This book is a collection of papers, prepared in connection with the 8th International Workshop on partial orders, their theoretical and applied developments, which took place in Warsaw, at the Systems Research Institute, in October 2008. The papers deal with software developments (PYHASSE and other existing software), theoretical problems of ranking and ordering under various assumed analytic and decision-making-oriented conditions, as well as experimental studies and down-to-earth pragmatic questions.

ISBN 83-894-7521-9
EAN 9788389475213