## POLISH ACADEMY OF SCIENCES

Systems Research Institute

# DEVELOPMENT OF METHODS AND TECHNOLOGIES OF INFORMATICS FOR PROCESS MODELING AND MANAGEMENT 

Editors:<br>Jan Studzinski<br>Olgierd Hryniewicz

DEVELOPMENT OF METHODS AND TECHNOLOGIES OF INFORMATICS FOR PROCESS MODELING AND MANAGEMENT

# Polish Academy of Sciences - Systems Research Institute Series: SYSTEMS RESEARCH Vol. 50 

Series Editor:

Prof. Jakub Gutenbaum

# DEVELOPMENT OF METHODS AND TECHNOLOGIES OF INFORMATICS FOR PROCESS MODELING AND MANAGEMENT 

Editors:

Jan Studzinski<br>Olgierd Hryniewicz

This publication was supported by Polish Ministry of Science and Higher Education

This book consists of papers describing applications of informatics in process modeling and management and in environmental engineering. Problems presented in the papers concern development of methods supporting process management, development of calculation methods for process modeling and development of technologies of informatics for solving some problems of environmental engineering. In several papers results of the research projects supported by the Polish Ministry of Science and Higher Education are presented.

Paper Reviewers:<br>Prof. Olgierd Hryniewicz<br>Dr. Jan Studzinski

Text Editor: Anna Gostynska

Copyright © Systems Research Institute of Polish Academy of Science, Warsaw 2006

Systems Research Institute of Polish Academy of Science Newelska 6, PL 01-447 Warsaw<br>Section of Scientific Information and Publications<br>e-mail: biblioteka@ibspan.waw.pl

## CHAPTER 3

## Tools of informatics in environmental engineering

# EVALUATION OF THE POLLUTION STATE IN BADEN-WUERTTEMBERG BY A LOCAL ANALYSIS 

Rainer BRÜGGEMANN ${ }^{1)}$, Kristina VOIGT ${ }^{2)}$<br>${ }^{1)}$ Leibniz-Institute of Freshwater Ecology and Inland Fisheries<br>[brg@igb-berlin.de](mailto:brg@igb-berlin.de)<br>${ }^{2)}$ GSF Research Center for Environment and Health Institute for Biomathematics and Biometry<br>[kvoigt@gsf.de](mailto:kvoigt@gsf.de)


#### Abstract

The visualization of partially ordered sets by Hasse Diagrams is a useful tool and can be considered as a first step in the evaluation of the pollution status of regions. Although much information can be drawn from Hasse diagrams, the decision makers usually may nevertheless be unhappy to get results which does not allow a unique decision. Here in continuing our series about local analysis in partial order, we discuss the estimation of the probability that one of two incomparable objects may nevertheless be considered as dominating. Starting from the model - poset of a double chain, an exact formula is derived. The result motivates to test an empirical equation, which uses the same input parameter as the double chain model but can be used to all types of empirical posets. At least with the example of the pollution status of Baden - Wuerttemberg, Germany we obtained satisfactory results.


Keywords: Partial Order, Hasse diagram, pollution, probability, geoinformatics.

## 1. Introduction

The visualisation of partially ordered sets by Hasse Diagrams is a useful tool (compare for example (Brüggemann,Voight, Steinberg, 1995; Brüggemann, Simon, Mey, 2005) and can be considered as a first step in the evaluation of the pollution status of regions (Brüggemann, Welzl, Voight, 2003; Brüggemann, Steinberg, 2000). In Baden Wuerttemberg, Germany, regions were selected and monitored with respect to $\mathrm{Pb}, \mathrm{Cd}, \mathrm{Zn}$ and S pollution in the herb layer. Taken these data as the basis of an evaluation study, partial order relations can be defined as follows:

Let $G$ be the set of regions, characterized by 4 integers $q_{1}, \ldots, q_{4}$, which quantify the pollution by $\mathrm{Pb}, \mathrm{Cd}, \mathrm{Zn}$ and S . Furthermore the range values of $\mathrm{Pb}, \mathrm{Cd}$ and Zn is $[0,1,2]$ and for $\mathrm{S}[0,1]$ according to recommendations of the Environmental Protection Agency of Baden Wuerttemberg. Then we call ( $G, \leq$ ) a partially ordered set, as follows:

$$
\begin{equation*}
x, y \in G . x \leq y: \Leftrightarrow q_{i}(x) \leq q_{i}(y) \text { for all } i=1, \ldots, 4 \tag{1}
\end{equation*}
$$

As the partial order is defined by the indicators and their realizations, we also write $(G, I B), I B=\{\mathrm{Pb}-, \mathrm{Cd}-, \mathrm{Zn}-, \mathrm{S}$-scores in the herb-layer $\}$. Sometimes the symbol $\perp$ is used for denoting a comparability between $x$ and $y$ without indicating an orientation. The notation $x \| y$ is used if $x$ is incomparable to $y$. For more details about Hasse diagrams, see Brüggemann et al., (2001).

The Hasse diagram is shown in Figure 1. It gives a series of useful information, namely:

1. pollution status [Brg Steinberg, Brg, Pudenz],
2. correlation behaviour,
3. geochemical aspects,
4. dimension,
5. ranking of the objects derived from the set of linear extensions (see below).


Figure 1. Partial order of regions in Baden-Wuerttemberg, Germany.
We call the aspects 1-5 global aspects of partial order as the diagram as a whole is to be examined. Although much useful information can be drawn from a Hasse diagram like that of Figure 1 and although there are many publications demonstrating these five aspects, the decision makers usually may nevertheless not be satisfied to get a result which does not give a unique decision. For example there are 5 out of 14 regions which are of high priority. Therefore the decision maker may select another evaluation method, like PROMETHEE (Brans, Vincke, 1985; Brans, Vincke, Mareschal, 1986) or ELECTRE (Roy, 1972; 1990) or METEOR (Voight, Brüggemann, 2005) to get a unique result. Usually these decision support systems need subjective inputs like weights or preference functions etc. In recent publications effort was also made to analyze the statistical behaviour of the so-called linear
extensions (Lerche, Sørensen, 2003). The aim of this statistical analysis is to get a linear order without the difficult process of weightings. In this paper we are interested in the latter aspect. Incomparabilities are thought of as awkward, because they hamper the unique decision.

Here we discuss the possibilities, just to concentrate oneself on one incomparable pair of objects $x \| y$ and to derive an estimation of probability for the preference of $x$ above $y$. As this aspect is focusing the behaviour of just two objects out of a whole set of objects, we call this a local analysis.

## 2. Towards a local analysis

Concept of a canonical representation of objects: Once a partial order is found from a data set one may find the set of linear extensions and from this the averaged rank (Rkav) which induces a linear order (Winkler, 1982). I.e.: by Rkav firstly an equivalence relation is defined (see e.g. Brüggemann, Simon, Mey, 2005) and secondly the elements of the corresponding quotient set are linearly ordered. This specific partially ordered set we denote as ( $G, \leq_{\text {Rkav }}$ ). We consider the averaged ranks as a canonical representation of the decision problem. Introducing subjective preferences, may change the ranking. Nevertheless a comparison with the canonical representation may help to identify the role of subjectivity and the role of the data values. The linear order derived from Rkav encompasses all objects of the object set $G$. Often, however one is only interested in the comparison of only two objects $x, y$, namely: What is the probability that object $x>y$ in $\left(G, \leq_{\text {Rkav }}\right)$ ? Once ( $G, \leq_{\text {Rkav }}$ ) is found, this question can be easily answered, here however, we are interested in an estimation of the probability, without first deriving the poset ( $G, \leq_{\text {Rkav }}$ ).

Definition of mutual probability: Any partial order ( $G, I B$ ) can be presented by a set of linear extensions, $L E$. LE is the set of all linear orders which can be found from $(G, I B)$ by order preserving maps. Let now $x, y$ be two objects of the ground set $G$.

There are two cases:
Case 1: If $x \perp y$ then $x \leq y$ or $y \leq x$ in all linear extensions.
Case 2: If however $x \| y$ then for some linear extensions $x \leq y$ and for the rest of linear extensions $y \leq x$. The number $|L E(y \leq x)|$ is the count how often $y \leq x$ is found within $L E:|L E(y \leq x)|$ can be compared with $L T$ the total number of linear extensions (i.e.: $L T=|L E|$ ) and we thus arrive at an estimation of the mutual probability, $p m(y \leq x)$ as follows:

$$
\begin{equation*}
p m(y \leq x)=|L E(y \leq x)| / L T \tag{2}
\end{equation*}
$$

In applying equation (2) we assume that all linear extensions are considered as equally probable. The quantity $p m(y \leq x)$ is useful, as in management problems
often not the whole ground set $G$ is of interest, but only some few incomparable elements.

So far, the local analysis to derive a preference relation without any prejudice can be successfully applied, just by setting up $L E$ and count those linear extensions for which e.g. $y \leq x$. However there are considerable computational problems: A crude upper estimation of the number of linear extensions is $N$ ! Assuming a standard equipment of computational facilities then an object set of more than 20 elements takes more than 12 hours (Lerche, Sørensen, Brüggemann, 2003)! Hence approximate analytical expressions for $\mathrm{pm}(y \leq x)$ would be extremely useful. There is still another idea behind the derivation of an analytical expression: We want to understand which graph-theoretical structures in the digraph of the poset, i.e. the Hasse diagram are important for pm . In Table 1 some notation is listed.

Table 1. Notation with respect to the mutual probability.

| Symbol | Explanation | Remark |
| :---: | :---: | :---: |
| pm | mutual probability without specification. Just the concept. | We use pm like an abbreviation |
| $p m(W)$ | Calculation of $p m$ by counting all linear extensions of interest and of $L T$ according to eq. 2. Exact values. | This task can be done applying the software WHASSE. |
| $p m(M,<\ldots>)$ | Calculation of $p m$ by assuming model posets. In <...> the model posets are specified; for example $p m(M, C C)$. | We derived many model posets and found analytical expressions for pm . |
| CC,ACAC, Q | Abbreviations for the considered model posets or empirical relations | see text |
| $p m(e s t)$ | Taken $(p m(W), p m(M,<\ldots>)$ as data points one may examine the linear regression. $p m(e s t)=a+b^{*} p(M,<\ldots>)$ <br> $r_{2} D F$, F-statistics, $a \rightarrow 0$, and $b \rightarrow 1$ are considered as quality parameters. | $p m(e s t)$ is found by usual least square minimization of $\Sigma(p m(H)-p m(M,<\ldots>)) 2$. |
| Lxy | $L x y:=\|L E(y \perp x)\|$ | Either $\geq$ or (exclusively) $\leq$ may be considered.). If one of these two possibilities is explicitly meant, then we write for example $L E(x \leq y)$ |
| LT | $L T=\|L E\|$ | italic notation: a set |
| $N$ | $N=\|G\|$ |  |

Approach to derive an analytical expression for pm: There are two crucial points within the concept of mutual probabilities: (1) to determine $L T$ and (2) to find out the number $L x y$ of a poset. Could we derive an analytical formula for pm ?

An empirical poset is hardly to be described without counting explicitly all linear extensions. Therefore for a derivation of an analytical expression we need a simple but sufficiently representative posets (Figure 2; the lines above and below x and y resp. symbolize additional elements). The four parameters $n, n p, m, m p$ count the elements above, below $x$ and above and below $y$, resp.. A simple appropriate model system would be a poset, consisting of two chains, because then both typical characteristics are included: Incomparabilities, and numbers of lower and upper neighbors in a Hasse diagram. There are four parameters on which pm will certainly depend: $n$, $n p, m$ and $m p$ (see Figure 2(a)). If we want to know, for example, how likely $x \geq y$, then an extension of the poset (Figure 2(b)) can be found, expressing the fact $x \geq y$. Now, taking from the starting poset (Figure 2(a) all its linear extensions, then we get $L T$, and deriving the number of linear extensions of the poset, shown in Figure 2(b), then we get $|L E(x>y)|$. Hence for posets, consisting of double chains, it is very easy to derive $\mathrm{pm}(y \leq x)$. According to Table 1, we call this quantity $p m(M, C C)$, because the analytical expression is derived from the assumption of a double chain (" $\mathrm{C}^{\prime \prime}$ ) poset.


Figure 2. Model poset to derive $L E(y \leq x)$.
For the sake of convenience of the reader, we describe briefly how to derive the analytical expression for pm. A more detailed explanation can be found in Brüggemann, Lerche, Sørensen (2003). In the Hasse diagram shown in Figure 2(b) element $x$ can take $m+1$ positions. Let $i$ be an index, running from 0 to $m$ then $i=0$ means, $x$ covers $y$, whereas $i=m$ means $x$ is located at the top of the chain of $m$ elements above $y$. Let us now keep $i \geq 0$ fixed. How many linear extensions are possible for the configuration at the right side of Figure 2? The $n$-chain can only be
inserted into the chain containing $m-i$ elements, whereas the $n p$ elements can be located into a chain of $m p+1+i$ elements. Hence for a fixed $i$ we find:

$$
\begin{equation*}
\mid L E(x \text { in ith position }>y) \left\lvert\,=\frac{(n p+m p+1+i)!\cdot(n+m-i)!}{n p!n!\cdot(m p+1+i)!(m-i)!}\right. \tag{3}
\end{equation*}
$$

As index $i$ can take $m+1$ positions, we have to sum up and arrive at:

$$
\begin{equation*}
\left|L E_{C C}(x>y)\right|=\sum_{i=0}^{i=m} \frac{(n p+m p+1+i)!(n+m-i)!}{n p!n!\cdot(m p+1+i)!(m-i)!} \tag{4a}
\end{equation*}
$$

Note that (4a) can also be written more compactly as:

$$
\left|L E_{C C}(x>y)\right|=\sum_{i=0}^{i=m}\binom{n+m-i}{m-i} \cdot\binom{n p+m p+1+i}{m p+1+i}
$$

$L T$ for a double chain $\left(L T_{C C}\right)$ is:

$$
\begin{equation*}
L T_{C C}=\frac{((n+1+n p)+(m+1+m p))!}{(n+1+n p)!\cdot(m+1+m p)!} \tag{4b}
\end{equation*}
$$

Hence:

$$
\begin{equation*}
p m(M, C C(x>y))=\left|L E_{C C}(x>y)\right| / L T_{C C} \tag{4c}
\end{equation*}
$$

The equations 4 can easily be simplified, if the mutual probabilities among two maximal or two minimal elements are to be determined. If for example the mutual probability of two minimal elements of the double chain is of interest, one finds applying the combinatorial identity (Stanley, 1986).

$$
\begin{align*}
& \sum_{j}^{w}\binom{x+j}{j}=\binom{x+w+1}{w}  \tag{5}\\
& p m(M, C C(x(\text { minimal })>y(\text { minimal })))=(m+1) /(n+m+2) \tag{6}
\end{align*}
$$

Furthermore it can be shown that $p m(x>y)$ in Hasse diagrams in a configuration as shown in Figure 3 can also be calculated by equations (4a)-(4c), when $n, n p, m$ and $m p$ are appropriately redefined, as Figure 3 motivates. Although the expressions for $\left|L E_{\mathrm{ACAC}}(x>y)\right|$ and $L T(A C A C)$ are modified, the probability $p m(M, x>y)$ is the same as found for the double chain system:

$$
\begin{equation*}
L T_{A C A C}=F \cdot \frac{(n+n p+m+m p+2)!}{(n+n p+1)!(m+m p+1)!} \tag{7a}
\end{equation*}
$$

$$
\begin{equation*}
|L E, A C A C(x>y)|=F \cdot \sum_{i=0}^{m} \frac{(m-i+n)!\cdot(m p+1+i+n p)!}{(m-i)!\cdot(m p+1+i)!n!n p!} \tag{7b}
\end{equation*}
$$

$$
\begin{equation*}
F=n!\cdot n p!m!m p! \tag{7c}
\end{equation*}
$$



Figure 3. "Double - sandwich-system" ( $A C A C$ ).
If the probability pm is calculated from the quotient as in equation (4c) then the factor $F$ is eliminated. We conclude that as long as $\mathrm{n}, \mathrm{np}, \mathrm{m}$ and mp are describing the upper and lower neighbors of $x$ and $y$ resp. and there are two not connected subgraphs, one around $x$ and one around $y$ then $\operatorname{pm}(M, x>y)$ can be calculated from $n$, $n p, m$ and $m p$, alone using equations (4a)-(4c).

Empirical approach: Although equations 4 and 7 are easily programmed and depend on structural information about the Hasse diagram which can easily be obtained, equations 4 and 7 are disadvantageous, because of their relative complex forms. The role, how $n, m, n p$ and $m p$ are influencing the mutual probability is not directly recognizable. Beyond this, the equations 4 and 7, resp. can only be an approximation for more general (i.e. empirical) posets, we derived an empirical equation just by a trial and error - procedure. The input quantities for the approximated mutual probability $p m(e s t)(x>y)$ are the same as for the $C C$ and $A C A C$ resp. model; they are, however more precisely defined in Table 2.

Table 2. Parameters of the empirical model (Q-model, see below).

| Symbol | Explanation |
| :--- | :--- |
| $n$ | number of elements above $x$ and not comparable with $y$ |
| $n p$ | number of elements below $x$ and not comparable with $y$ |
| $m$ | number of elements above $y$ and not comparable with $x$ |
| $m p$ | number of elements below $y$ and not comparable with $x$ |

We define:

$$
\begin{equation*}
Q(x):=(n+1) /(n p+1) \tag{8a}
\end{equation*}
$$

and similarly:

$$
\begin{equation*}
Q(y)=(m+1) /(m p+1) \tag{8b}
\end{equation*}
$$

If an expression for $p m(M)$ is to be found, one has to take into account that there are constraints, namely:
$p m(M, Q(x>y))+p m(M, Q(y>x))=1$ and $1 \geq p m(M, Q(x>y)) \geq 0$
and $1 \geq p m(M, Q(y>x)) \geq 0$.
Technically, $p m(M, \ldots)$ can be considered as the solution of a weak functional equation together with constraints (Reich, 2005). A simple, but by no means the only one realization (the equation 4 is an example for another realization), which fulfills these constraints is:

$$
\begin{equation*}
p m(M, Q(x>y))=\frac{Q(y)}{Q(y)+Q(x)} \tag{9a}
\end{equation*}
$$

For an interpretation equation (9a) can be rearranged as follows:

$$
\begin{align*}
& p m(M, Q(x>y))=\frac{1}{1+R(x, y)}  \tag{9b}\\
& R(x, y):=\frac{(n+1) \cdot(m p+1)}{(n p+1) \cdot(m+1)} \tag{9c}
\end{align*}
$$

If equation (9b) is applied on the minimal objects, the same expression, as in equation 4 is found. Hence the limiting cases of the double chain model are well described by equation (9a).

The probability that $x$ dominates y is the more pronounced, the less $R(x, y)$ is. One may compare this with the picture of a balance (Figure 4). In Figure 4, The elements above $x$ and $y$ exert a pressure down, whereas the elements below $x$ and $y$ exert a pressure above. En passant we state that if $Q(x)=Q(y)$ i.e. the relation of predecessors of $x$ to the successors of $x$ is the same as for the element $y$ (only the disjoint elements are counted) then equation 9 predicts no preference of $x$ about $y$.
$p m(M, Q(x>y)=p m(M, Q(y>x)=0.5$. This observation fits pretty well into the picture of a balance.


Figure 4. Analogy of equation (9b) with a balance.

## 3. Results

Test of the model equations: We apply the different approaches for pm on all pairs of incomparable objects of $G$. There are 45 such pairs, which get a quite diverse pattern in terms of $n, m, n p$ and $m p$ (see Table 3).

Table 3. Characteristics of the Baden-Wuerttemberg dataset.

|  | $n$ | $n p$ | $m$ | $m p$ |
| :--- | :--- | :--- | :--- | :--- |
| $\boldsymbol{\operatorname { m i n }}$ | 0 | 0 | 0 | 0 |
| $\boldsymbol{\operatorname { m a x }}$ | 4 | 7 | 6 | 5 |
| mean | 0.7 | 1.6 | 2.3 | 1.1 |

One may think that the four characteristic quantities are inter-correlated, if a specific poset is to be analyzed. Therefore additionally in Table 4 the correlation matrix (after Pearson) is shown:

Table 4. Pearson correlation of characteristics of the Hasse diagram, shown in Figure 1.

|  | $n$ | $m$ | $n p$ | $m p$ |
| :--- | :--- | :--- | :--- | :--- |
| $n$ | 1 | 0.025 | $-0.329^{* *}$ | -0.189 |
| $m$ |  | 1 | -0.225 | $-0.404^{* *}$ |
| $n p$ |  |  | 1 | -0.287 |
| $m p$ |  |  |  | 1 |

**: Significant correlation, two-sided, confidence level: 0.01
Table 4 shows that the pairs $n-n p$ and $m-m p$ are negatively correlated because e.g. $n$ can only increase at the cost of $n p$ and the other way round, if a specific poset is considered. whose $N$ is fixed. The statistical results in application of the two approaches for $p m$ are summarized in Table 5.

Table 5. Comparison of the models for the mutual probability with the exact $p m$-values $(p m(W(x>y))(45$ cases $)$.

|  | $r_{D F}^{2}$ | $F$ | $a$ | $b$ |
| :--- | :--- | :--- | :--- | :--- |
| $p m(M, C C(x>y))$ | 0.903 | 409 | 0.087 | 0.867 |
| $p m(M Q(x>y)$, | 0.896 | 380 | 0.054 | 0.932 |

Table 5 does not allow a clear preference between the empirical model, based on the $Q$-concept and the $C C$ and $A C A C$-model: The variance parameters $r_{D F}^{2}$ and F favour the $C C$-model, whereas the bias-related quantities, $a$ and $b$, favour slightly the $Q$-model. For the sake of simplicity we give the model $p m(M, Q)$ the preference and apply it to a typical question, related to the Hasse diagram of Figure 1.

Application of pm to typical question, related to the data of BadenWürttemberg: In Figure 1 Weinheim (region no 57) and Pfalzgrafenweiler (region no 35) are maximal objects and are incomparable. Which of these two should be remediated? If there are no socioeconomic arguments, but only the knowledge due to the pollution then a decision which region is to be selected may be based on the mutual probability, i.e. $p m(57>35)$. Applying equation 2 we find $p m(35<57)=0.77$. That means that although both regions are incomparable, in the linear orders derived from the partial order the region 35 is more often ranked below 57 than in the reverse order. Similarly other pairs $(x, y)$ can be selected for which $x \| y$. Hence, by the count of successors and predecessors a preference between $x$ and $y$ can be nevertheless obtained as far as $p m \neq 0.5$.

## 4. Summarizing, Discussion and Conclusion

The representation of partial orders by Hasse diagrams is a useful tool in environmental sciences. Many conclusions can directly be drawn from the graphical
scheme. An important mathematical tool is the construction of the set of linear extensions, $L E$ and to derive from that the mutual probabilities for preferences between any two incomparable elements out of the whole set of objects. Empirical formulas were developed, based on simple characteristics of any of the objects of $G$, if the poset ( $G, I B$ ) is known. This paper continues the concept of local approaches, especially here in focusing on two objects which are incomparable in ( $G, I B$ ), but from which one would like to know, how probable object $x$ can be considered as dominating object $y$, if a linear order would exist. As such the local approach can be put into a more general scheme, namely:

1. Analysis by Hasse diagrams.
2. Performing a global analysis which means:

- What are the priority elements,
- Are there striking graph theoretical structures which indicate interesting data structures?
- Does the Hasse diagram a high degree of rank correlation (in the sense of Spearman correlation)?
- What is the dimension of the poset? Consequently a method has to be developed to find latent variables which describe the evaluation problem in a lower dimensionality?

3. Performing a local analysis:

- Identification of peculiar elements,
- Derivation of averaged ranks,
- Derivation of mutual probabilities between any two objects, which are incomparable in ( $G, I B$ ).

These benefits, however, have their price: The multivariate problem due to a multitude of attributes like the chemical concentration of $\mathrm{Pb}, \mathrm{Cd}, \mathrm{Zn}$ and S in the herb layer, is reduced to a comparison of upper and lower neighbors within the directed graph. Whether or not a relation $x \leq y$ or $x \| y$ is found depends on the numerical values. Even if numerical differences among attribute values of different objects are statistically insignificant, they will contribute to the structure of the Hasse diagram and may influence the counting of objects as demonstrated above. Hence it is a major task in the future to find procedures for a robustification, for example by a good classification scheme or a sophisticated rounding technique (see e.g. Helm, 2003). If the data values are discrete, as supposed here then the final decision which object of $x \| y$ may dominate can considered as a question, how many other objects $x$ (or $y$ ) is representing in the $<$ - or the $>$-sense.

There is another question to be posed: Are there posets, for which the $n, n p$, $m, m p$-concept must necessarily fail? Indeed posets which can be characterized by
a scheme, as shown in Figure 5 can typically not described by this four-parameterscheme. Furthermore it remains unclear how far objects, which are at the same time greater or less than $x$ and $y$ should be counted (see Figure 6 for an explaining scheme).

By an empirical study we found that with

$$
\begin{aligned}
& \text { gup }:=\mid\{z: z \geq x \text { and } z \geq y\} \mid \\
& \text { gdown }:=\mid\{z: z \leq x \text { and } z \leq y\} \mid \\
& k:=k 1+k 2 \\
& k u p(x):=k \cdot \frac{n+g u p}{n+g u p+n p+g d o w n+1}
\end{aligned}
$$

and

$$
k d o w n(x):=k \cdot \frac{n p+g d o w n}{n+g u p+n p+g d o w n+1}
$$



Figure 5. Type of posets where beside $n, n p, m$ and $m p$ other parameters (like $k_{1}$ and $k_{2}$ ) are needed to find expressions for $p m$.

$$
k u p(y):=k \cdot \frac{m+g u p}{m+g u p+m p+g d o w n+1}
$$

and

$$
k d o w n(y):=k \cdot \frac{m p+g d o w n}{m+g u p+m p+g d o w n+1}
$$

together with

$$
Q(x, k):=\frac{n+g u p+k u p(x)+1}{n p+g \operatorname{down}+k \operatorname{down}(x)+1} \text { and } Q(y, k):=\frac{m+g u p+k u p(y)+1}{m p+g d o w n+k d o w n(y)+1}
$$

and finally

$$
p m(M, Q(x, k), Q(y, k))=\frac{Q(y, k)}{Q(y, k)+Q(x, k)}
$$

gave slightly better results. I.e. taking into account the additional structure parameter, gup, gdown, kup, kdown the variance and bias parameter were improved $\left(r_{D F}^{2}=0.92, F=491, a=-0.0005, b=0.97\right)$ testing the 45 incomparable pairs taken from $G$. Nevertheless we are not convinced that the afford of getting those additional parameters is justified. Hence, still further research is needed, outgoing from the classification, as suggested by the scheme in Figure 6.


Figure 6. Schematic example for a classification of objects. Note that not necessarily all $k$-type objects are comparable to $n, n p, m$, or $m p$-objects.

Checking the Hasse diagram shown in Figure 1, and selecting the incomparable objects 41 and 22 for an analysis, we would arrive at Table 6:

Table 6. Classification applied to a specific pair of objects in $G$ (refer to Figure 1).

| Contributing to.... | objects |
| :--- | :--- |
| gup | 14,57 |
| gdown | 17,30 |
| $k$ | $9,34,38,45,48$ |
| $n$ | 18,35 |
| $n p$ | - |
| $m$ | - |
| $m p$ | 6 |

We hope that this classification is a good starting point for further investigations.

## References

Brans J.P., Vincke P.H. (1985) A Preference Ranking Organisation Method (The PROMETHEE Method for Multiple Criteria Decision - Making). Management Science, 31: 647-656.

Brans J.P., Vincke P.H., Mareschal B. (1986) How to select and how to rank projects: The PROMETHEE method. European Journ.Oper. Research, 24: 228-238.

Brüggemann R., Halfon E., Weizl G., Voigt K., Steinberg C (2001) Applying the concept of partially ordered sets on the ranking of near-shore sediments by a battery of tests. J. Chem. Inf. Comp. Sc., 41: 918-925.

Brüggemann R., Lerche D., Sørensen P.B. (2003) First attempts to relate structures of Hasse diagrams with mutual probabilities. In: Sørensen. P.B. et al. (Eds.) Order Theory in Environmental Sciences - Integrative Approaches. Proceedings of the 5th workshap held at NERI. 2002. NERI. Ministry of the Environment. Denmark. Roskilde. 7-51.

Brüggemann R., Simon U., Mey S. (2005) Estimation of averaged ranks by extended local partial order models. Match - Commun. Math. Comput. Chem. 54: 489-518.

Brüggemann R., Steinberg C. (2000) Einsatz der Hassediagrammtechnik zur vergleichenden Bewertung von Analysendaten - am Beispiel der Umweltuntersuchungen in den Regionen Baden-Württembergs-. In: Günzler. H. (Ed.) Analytiker Taschenbuch 21. Springer-Verlag. Berlin. 3-33

Brüggemann R., Welzl G., Voigt K. (2003) Order Theoretical Tools for the Evaluation of Complex Regional Pollution Patterns. J. Chem. Inf. Comp. Sc. 43:1771-1779

Helm,D. (2003) Bewertung von Monitoringdaten der Umweltprobenbank des Bundes mit der Hasse-Diagramm-Technik. UWSF - Z. Umweltchem. Ökotox. 15: 85-94.
Klein J., Brüggemann R., Voigt K., Steinberg C. (1995) Advances in comparative analysis of adverse effects in aquatic ecosystems with emphasis on studies with humic substances and on progress in mathematical analysis technique. Wat. Res., 29: 2261-2268.

Lerche D., Sørensen P.B. (2003) Evaluation of the ranking probabilities for partial orders based on random linear extensions. Chemosphere, 53: 981-992.
Lerche D., Sørensen P.B., Brüggemann R. (2003) Improved estimation of the ranking probabilities in partial oOrders using random linear extensions by approximation of the mutual probability. J. Chem. Inf. Comp. Sc. 53: 1471-1480.
Reich L. (2005) Karl-Franzens-University Graz, Mathematical Institute, personal communication.
Roy B. (1972) Electre III: Un Algorithme de Classements fonde sur une representation floue des Preferences En Presence de Criteres Multiples. Cahiers du Centre d'Etudes de Recherche Operationelle, 20: 32-43.

Roy B. (1990) The outranking approach and the foundations of the ELECTRE methods. In: Bana e Costa (ed) Readings in Multiple Criteria Decision Aid. Springer-Verlag. Berlin. 155-183.

Stanley R.P. (1986) Enumerative Combinatorics. Volume I, Wadsworth \& Brooks/cole, Monterey. 1-306.

Voigt K., Brüggemann R. (2005) Water contamination with pharmaceuticals: data availability and evaluation approach with hasse diagram technique and METEOR. Match - Commun. Math. Comput. Chem. 54: 671-689.
Winkler P. (1982) Average height in a partially ordered set. Discrete Math. 39: 337-341.

Jan Studzinski, Olgierd Hryniewicz (Editors)

## DEVELOPMENT OF METHODS AND TECHNOLOGIES OF INFORMATICS FOR PROCESS MODELING AND MANAGEMENT

The purpose of this publication is to popularize application of informatics in process modeling and management and in environmental engineering. The papers published are thematically selected from the works presented during the conference 'Multi-accessible Computer Systems' organized by the Systems Research Institute and the University of Technology and Agriculture in Bydgoszcz for several years already in Ciechocinek. Problems presented in the papers concern: development of quality and quantity methods supporting the process management, development of quantity methods for process modeling and simulation, development of technologies of informatics for solving problems of environmental engineering. In several papers results of research projects supported by the Polish Ministry of Science and Higher Education are presented.

## ISBN 83-894-7507-3 9788389475077

