communications in mathematical
match
and in computer chemistry

no. 42, October 2000

ISSN 0340-6253

MATCDY (42) 119-143 (2000)

AN APPROACH TO COMBINE CLUSTER ANALYSIS WITH ORDER THEORETICAL TOOLS IN PROBLEMS OF ENVIRONMENTAL POLLUTION

B.Luther, R.Brüggemann*, S.Pudenz
Institute of Freshwater Ecology and Inland Fisheries
Department: Ecohydrology
E-Mail (Brüggemann): Brg@IGB-Berlin.de

*) Corresponding author

Abstract
To uniformly deal with data of environmental pollution, we represent objects (in this case regions) of an ecological examination as elements of a partially ordered set. The results are visualized by Hasse diagrams. To reduce the complexity of our posets, methods of cluster analysis are used. There is also a question of statistical significance of the ordinal relations. In particular we present an algorithm that first calculates an unsharp partition (using a method of fuzzy clustering) of an object set. Next the unsharp partition is transformed to a sharp one with the aim to define a new order. By splitting the quality function it can be shown that objects which cannot be assigned to clusters forming ensembles of more than one element contribute to the method's error. A strategy minimizing that error is proposed.

Keywords: environmental pollution, chemicals, evaluation, posets, Hasse diagrams, fuzzy, cluster analysis

1 Introduction

Often one is faced with the evaluation of objects: What is the price of an article, what are the costs of some engineering constructions? In environmental sciences typically the evaluation has to be performed regarding several criteria [1]. For example, objects of such an environmental evaluation may be chemicals (as shown in [1]), or geographical areas. In any case, in order to evaluate objects, a tuple of data is needed. Those data are considered as helpful to describe the objects with respect to the criteria by which they are to be evaluated. The evaluation requires a comparison of objects, therefore often a ranking index is introduced. However the use of a ranking index implies that the combination of criteria can be quantitatively described. Typically in environmental sciences there is no consensus on how to do this [2-4]. Therefore the concept of a partial order appears to be extremely helpful, to perform at least a comparative evaluation. That means that the objects are partially ordered corresponding to the componentwise order of their tuples of data. Some details are discussed in Brüggemann and Bartel [5]. The properties useful for evaluation are called "(evaluative) attributes" and it is convenient to introduce a set, namely the information basis of evaluation, IB. The objects form the object set O. The partially ordered set (poset) $(O,\leq)$ is then to be examined.

Several applications can be found in the literature. Here only some recent publications will be mentioned [6-17]. The main idea of such a data driven evaluation (i.e. without any subjectivism in finding a correct combination for the criteria of interest) is the generation of a set of order relations. This set may be represented as an acyclic directed graph (from now on: whenever we say "graph" we mean a *directed* graph). Orientation and a transitive reduction leads to the well known visualization of the poset by a Hasse diagram. See for details the prolegomenon of D.J. Klein, this issue.

If the number of objects to be evaluated is large, then often the Hasse diagrams are messy systems of lines, from where no information can be drawn directly. An evaluation project performed for an environmental protection agency (LfU Baden-Württemberg) [18] had to deal with such a complicated system of lines. The task was to evaluate 59 regions according to the content of lead, cadmium, zinc and sulfur in different matrices. The matrices were: herb layer, the leaf layer (leaves of trees), moss and earth worms. Each of these matrices were considered to be indicators for different kinds of pollution patterns. Here the leaf layer will serve as an example.

2 Defining the Problem

As outlined in the introduction, one drawback of applying Hasse diagrams in environmental evaluation (called the Hasse diagram technique, abbr.: HDT) is that large sets of objects lead to complex diagrams. The problem is not that only the appearance of the graphic, but also (and more important) that slight differences in data (we refer to them as "original" data) are ordinally interpreted. For example chains or antichains may lead to conclusions with respect to objects, important for decision makers, which are not based on significant numerical differences. An interesting approach based on the concept of probabilities and confidence levels was recently published by Sørensen et al. [13, 14]. As an alternative to the approach of Sørensen et al. one may classify each attribute and instead of investigating the original data, indices of classes are used [5]. This procedure can be generalized to methods of multivariate statistics, especially to different clustering methods [19,20]. Here the method of fuzzy clustering [20,21] will be further examined, with the aim to combine this method with the HDT. The strategy is to find a partitioning of the object set, and to order the clusters (or better: a representative of a cluster) instead of the objects themselves. This leads to a question concerning the relevance of such a "distortion" of the data concerning the evaluative aim. That means, the **assessment of error due to the transition to suitable representatives of a cluster** or **feasibility of a given partition** is to be examined more closely. Thus in this paper two different mathematical tools are of interest:

1. The **order theory** as a basis for every comparative evaluation.
2. The theory of **cluster analysis** to perform a robustification with respect to slight differences in data.

We will discuss a dataset (the loadings of the pollutants Pb, Cd, Zn, and S on leaves of trees (mg/kg dry weight)) which was evaluated within the scope of the project mentioned above [9]. In short we arrive at the following agreements:

1. O={1,...,59} is the set of objects to be classified (we simply label the regions).
2. IB={ Pb, Cd, Zn, S } denotes the attribute set or information basis.
3. The concentration data are transformed according to $z^{(k)} = \frac{x^{(k)} - \mu}{\sigma}$ where μ is the (estimated) mean and σ is the (estimated) standard derivation. This so-called z-transformation is order preserving and gears the values to a standard normal (0/1)-distribution on the interval $(-\infty,+\infty) = \mathbb{R}$.
4. We will denote the family of sets with potential possible values A={ IR, IR, IR, IR } (not quite correctly but unmistakably) sometimes as $\mathbb{R}^4$, too.

As clustering method, a fuzzy approach is used. By this, not only a quantity describing the partitioning (in some cluster methods, the ultrametric [19]) is available as a steering quantity, but also the membership function. In fact, the actual partitioning into clusters will be described by:

FCL: the number of **f**uzzy **cl**usters which are offered,

TMF: a **t**hreshold for the **m**embership **f**unction.

Applying these two parameters, instead of one partially ordered set $(O,\leq)$ the family $\mathfrak{F} = \{\ (O/\cong_{(FCL,TMF)}, \leq) \mid FCL \in \mathbb{N} \text{ and } TMF \in \left(0.5\ \ ,\ \ 1\right]\}$ of possible "clusterings" is to be investigated[1]. Thus a method is to be introduced to select an optimal $(O/\cong_{(FCL,TMF)},\leq)$ out of the elements of $\mathfrak{F}$.

3 Comparing Assessment – Partially Ordered Sets as Base and Cluster Methods as "Simplification Tool"

3.1 Partially Ordered Sets

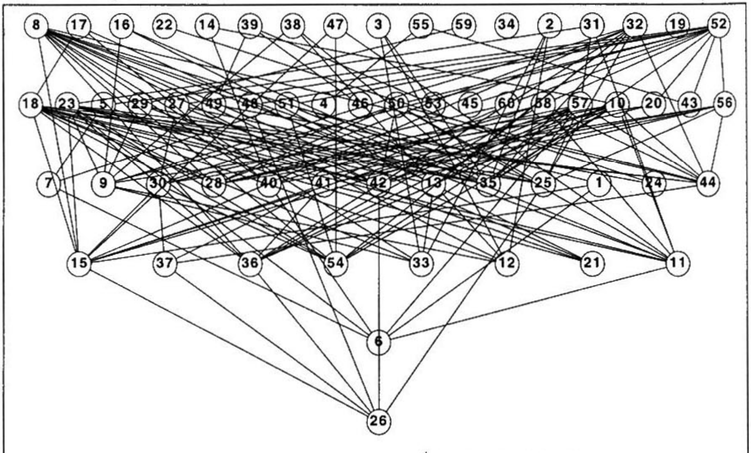


Figure 1: Diagram HD(O, IB, $\mathbb{R}^4$) using original data

We introduce a mapping $D: O \rightarrow \mathbb{R}^n$, which assigns to each object $o \in O$ of an object set some n-tuple $x^{(o)}:=D(o)$ of real numbers (in our case a quadruple $x:= (x_1,x_2,x_3,x_4)$ of transformed concentrations).

We define: $a\ ,b \in O \quad a \leq b :\Leftrightarrow x^{(a)} \leq x^{(b)}$

[1] The equivalence relation $\cong_{(FCL,TMF)}$ may be called "belonging to the same cluster" and results from specific values of the steering-parameters FCL and TMF.

Thus, a dataset $X:=D(O) \subseteq \mathbb{R}^n$ is introduced (written as: $X=\{x^{(1)},...,x^{(N)}\} \subset \mathbb{R}^n$ with $O = \{o^{(1)},...,o^{(N)}\}$). The dataset X is partially ordered by the product order $\leq^n := \underbrace{\leq \times \cdots \times \leq}_{\text{n times}}$ where "≤" denotes the usual order relation of the real numbers. By that convention we can consider $(O,\leq)$ with data from the leaf layer (that means to take a *specific* mapping $D_{leaf}: O \rightarrow \mathbb{R}^n$) and the resulting partial order can be visualized by a Hasse diagram, which we denote as $HD(O, IB, \mathbb{R}^4)$ according to the object set O, the attribute set IB and the data representation $\mathbb{R}^4$ (figure 1).

The $HD(O, IB, \mathbb{R}^4)$ shown in figure 1 clearly illustrates that the HDT reaches the limits of graphical representation (and interpretation), although the technique is quite suggestive. As already mentioned, the problem is that on the one hand the number of *objects* not longer allows a clear graphical representation and on the other hand slight numerical differences imply insignificant comparabilities and incomparabilities. Probably the number of attributes does not cause the bad "representability". That means: We do not want to change the underlying information basis (attribute set) IB={ Pb, Cd, Zn, S }; instead, a new object set O' is to be established, for which $|O'| << |O|$ holds. This will be done by gathering similar objects, combining them to "clusters", which can be treated as new objects – in simple words: using cluster methods as "simplification and robustification tool".

3.2 Cluster Methods

In order to start in a formal correct way, we want to introduce some basic notions:

3.2.1 (Sharp) Partitions.

(Sharp) Partition. Let O be a finite set (of objects), and $A_1,...,A_m \subseteq O$ with

- $(\forall i,j \in \{1,...,m\})\quad A_i \cap A_j = \emptyset$ and
- $\bigcup_{i=1}^{m} A_i = O$.

Then we call $(A_1,...,A_m)$ a *(sharp) partition* of the object set O (of FCL clusters).

Singleton. Let $(A_1,...,A_m)$ be a *(sharp) partition* of the object set O and $A_i=\{o\}$ a cluster with only one object. Then we call o a singleton object and A_i a singleton cluster.

3.2.2 Unsharp Partitions.

Unsharp Partition. Let $O=\{o^{(1)},...,o^{(N)}\}$ be a finite set (of objects), and $U\in[0,1]^{FCL\times N}$ a (membership-) matrix, i.e.

- $(\forall\ k\in\{1,...,N\})\quad \sum_{i=1}^{FCL} u_{ik}=1$.

Then we call U an *unsharp partition* of the object set O. (Each object is assigned to all clusters, "with different memberships u_{ik}".)

Remark: In these terms a *(sharp) partition* $(A_1,...,A_{FCL})$ can be considered as unique *unsharp partition* $U\in[0,1]^{FCL\times N}$ with the additional property $U\in \mathbf{Z}^{FCL\times N}$ (**Z** is the set of whole numbers – the entries u_{ik} of U are then 0/1-valued and they are given by $u_{ik}=\begin{cases}1 & \text{if } o^{(k)}\in A_i\\ 0 & \text{if } o^{(k)}\notin A_i\end{cases}$). In this sense the concept of an *unsharp partition* is a generalization of the concept of a *(sharp) partition* – we can talk about the *matrix* A even if A is a sharp partition.

Purity. Let $U\in[0,1]^{FCL\times N}$ be an unsharp partition (of a finite object set $O=\{o^{(1)},...,o^{(N)}\}$), $o^{(k)}\in O$. Then we call $\text{pur}(o^{(k)}) := \max\{u_{ik}\mid i\in\{1,...,m\}\}$ the *purity* of the assignment of $o^{(k)}$.

Hybrids. Let $U\in[0,1]^{FCL\times N}$ be an unsharp partition (of a finite object set $O=\{o^{(1)},...,o^{(N)}\}$). Then we call $o^{(k)}$ a *relative hybrid* iff $\text{pur}(o^{(k)})\in(0.5\ ,\ 1)$ – we call it an *absolute hybrid* iff $\text{pur}(o^{(k)})\in(0\ ,\ 0.5]$. In general we call $o^{(k)}$ a TMF-*hybrid* or a *hybrid* with corresponding treshhold TMF (TMF $\in(0\ ,\ 1]$) iff $\text{pur}(o^{(k)})\le$ TMF. (We do not forbid the term 1-*hybrid* for sharp assigned objects.) On the other hand: Iff $o^{(k)}$ is no TMF-*hybrid* we consequently call $o^{(k)}$ TMF-*pure* . Iff $\text{pur}(o^{(k)})\in(0.5\ ,\ 1]$ $o^{(k)}$ is *relative pure*. After all $\text{pur}(o^{(k)})=1$ implies that $o^{(k)}$ is *sharp assigned* or *pure* at all. To sum all this up we make a small decision tree (figure 2).

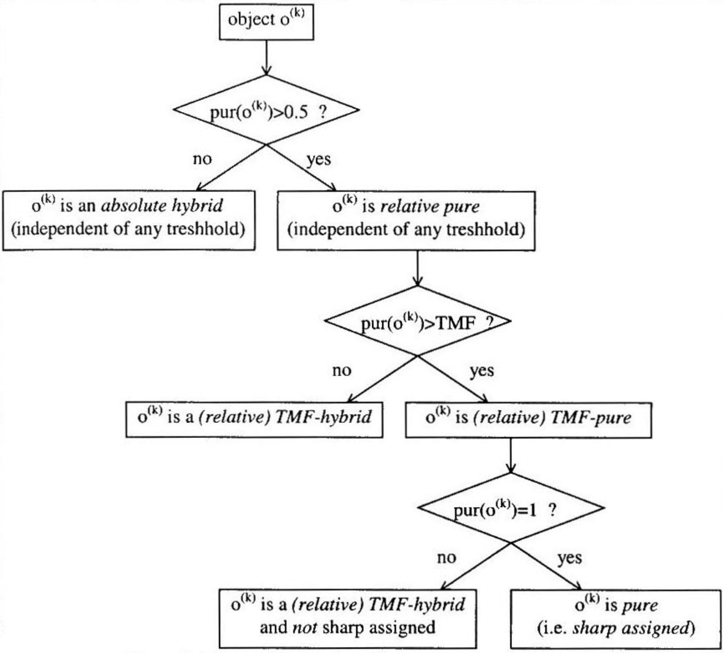

object $o^{(k)}$
$pur(o^{(k)})>0.5$?
no
yes
$o^{(k)}$ is an *absolute hybrid* (independent of any treshhold)
$o^{(k)}$ is *relative pure* (independent of any treshhold)
$pur(o^{(k)})>TMF$?
no
yes
$o^{(k)}$ is a *(relative) TMF-hybrid*
$o^{(k)}$ is *(relative) TMF-pure*
$pur(o^{(k)})=1$?
no
yes
$o^{(k)}$ is a *(relative) TMF-hybrid* and *not* sharp assigned
$o^{(k)}$ is *pure* (i.e. *sharp assigned*)

Figure 2: How to characterize objects in terms of their purity

It is easy to illustrate these terms geometrically. Assume that we want to get an unsharp partition for a set O with at most three clusters. The feasible set of possible "membership vectors" is then given by the 2-dimensional simplex (embedded in the $\mathbb{R}^3$) whose corners represent the three possible sharp assignments (figure 3) – the usual geometric representation of points on the 2-dimensional simplex is done by ternary diagrams (figure 4). We use this form of visualization because it shows our terms very well: The small triangle inside of the simplex forms the set of absolute hybrids – any object that is assigned to a point *out of this region* is only a *relative hybrid* because it can be considered to be a member of a *unique* cluster (corner) "according to its maximal membership $pur(o^{(k)})$".

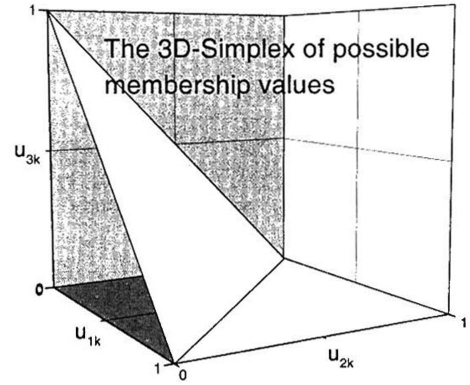


Figure 3: A link to polyhedron theory

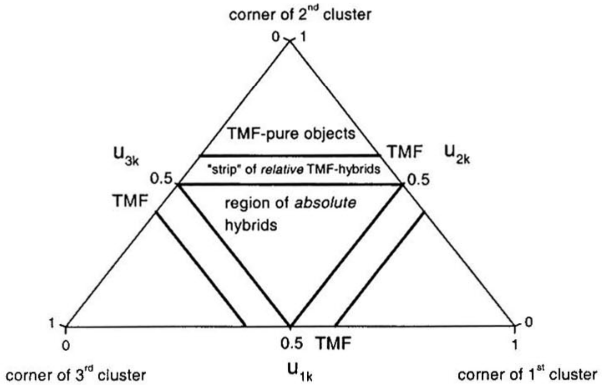


Figure 4: Good and bad regions on the "simplex of memberships"

3.3 Algorithm "SHARP"

In mathematical terms we have the following task:

Let O={$o^{(1)}$, ..., $o^{(N)}$} again be a finite object set which is described by some measurements to the finite dataset X={$x^{(1)}$,...,$x^{(N)}$} ⊂ $\mathbb{R}^n$. Furthermore let FCL be the number of clusters, in which the object set O is dispersed and TMF a threshold between 0.5 and 1, which describes when an unsharp classified object can nearly be assigned sharply to one unique cluster[2]. Then consider the procedure

```
SHARP(O,D,FCL,TMF)                                  ( $FCL \in \mathbb{N}$ and $TMF \in (0.5 \;,\; 1]$ )
1. Generate an optimal[3] unsharp partition U:= UNSHARP(O,FCL) of the object set O by
   means of $D(O)=X=\{x^{(1)},...,x^{(N)}\}$.
2. For k=1...N do:   If     ($\exists\, i \in \{1,...,FCL\}$) $u_{ik} > TMF$        ($\Leftrightarrow pur(o^{(k)}) > TMF$ )
                     Then   Assign $o^{(k)}$ to cluster $A_i$ .
                     Else   Assign $o^{(k)}$ to the set of TMF-hybrids (called Hyb(TMF))
3. Return $A:=\{A_i \mid A_i \neq \varnothing\}$ and Hyb(TMF).
```

This is our heuristic approach aiming to combine cluster analysis with order theory – as could be read in the last footnote, the meaning of the procedure **UNSHARP** which calculates the matrix elements u_{ik} of an unsharp partition U will briefly be explained later (section 4), up to now we only have to know that **UNSHARP** delivers an unsharp partition.

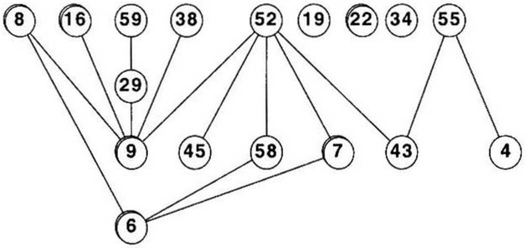
At this point the output of **SHARP(O,D,FCL,TMF)** may be seen as a partition of the *whole* set O: We may formally assign each $h \in Hyb(TMF)$ to a cluster which contains only one element considering the partition $B:=A \cup \{\, \{h\} \mid h \in Hyb(TMF)\,\}$

Even the case $|A_i|=1$ is possible, however there is a great difference between the clusters A_i with the property $|A_i|=1$ and the clusters $\{h\}$ with $h \in Hyb(TMF)$, even if we can call all of them *singleton clusters*. An $a \in O$ with $A_i=\{a\}$ is nearly *optimal* for the unsharp partition – the assignment to a singleton cluster is not arbitrary. But our dealing with TMF-hybrids *is* arbitrary – it is just an emergency measure because we do not get enough information about a suitable assignment from the algorithm.

The sharp partition A drops out all TMF-hybrids – they are collected in Hyb(TMF). However, the evaluation by the HDT must include *all* singletons even those of set Hyb(TMF), because their x-tuples may reveal a strange pattern of pollution which may has an influence on suitable remediation procedures. Consequently partition B is the starting point to find a new order. As an example we plot the diagram of figure 5 that is a planar[4] graph. (The conclusion

[2] See Bock [21] for details of unsharp clustering.
[3] We will describe below, what "optimal" means.
[4] A planar graph can be drawn in the plane without intersections (of edges).

that therefore the dimension of the poset is 2, and that therefore an embedding into $\mathbb{R}^2$ must be possible, will be suppressed here.)



HD(SHARP(O,D,6,0.6)**, IB,** $\mathbb{R}^4$ **)**

Eleven 0.6-hybrids:	{4, 19, 29, 34, 38, 43, 45, 52, 55, 58, 59}
nontrivial equivalent classes:	{**6**,15,28,40,42,37,26,36,54,25,11,44};
(6 clusters, 49 assigned objects)	{**8**,17,18,23,14,57,10,31,32};
	{**7**,30,5,41,27,47,48,51,50,60,1,20,56};
	{**9**,13,35,33,12,21}; {**16**,3,2}; {**22**,39,49,46,53,24}

Figure 5: Hasse diagram of clusters, instead of objects. The labeling corresponds to a representative of the equivalence relation ≅(6, 0.6). (Nontrivial equivalence classes are indicated in the diagram.)

If we consider this HD as "suitable" for a description of our regions/objects, it offers much more information about the relationships between the objects, than the first one:

- **Isolated objects:** An isolated object (vertex) of a graph is an object which is not connected with any other. In the HD on figure 5 we have three isolated objects (=clusters): {19}, {34} and {**22**,39,49,46,53,24}. If we can consider our clusters as "well seperated", we have to regard this clusters as indicators for specific patterns of loading.
- **Articulation points:** An articulation point (vertex) of a graph is an object which removing would split the graph. For example {**9**,13,35,33,12,21}, {43}, {52} and {55} are the articulation points on figure 5. Such points give information about the data structures, too.
- **Chains:** A chain of a graph is a *directed* sequence of connected objects (vertices), i.e. in figure 5 "9→29→59" is a chain but "9→52→45" is *not*. In our case a chain describes a si-

multaneous increasing (decreasing) of different loadings, so a chain gives rise to deterministic reasoning (see Brüggemann [24]).

- **Antichains:** An antichain of a graph is a subset of objects (vertices) which are mutually *not* connected by an edge. An antichain of a poset is a subset of mutually incomparable objects. In our example antichains correspond to the diversity of pollution patterns. For each pair of incomparable objects at least one attribute value increases at the cost of another one ("antagonisms" of attributes). There are two special antichains of a partial order that are always of special interest: The set of maximal and the set of minimal elements.

Of course we can search for chains or isolated objects in the first HD, too. But in the first diagram an isolated object may be the result of slightly antagonistic fluctuations which do not allow conclusions for the data structure and if we look at (anti-)chains or articulation points the situation is similar – one of the few things the first HD shows "pure" are the extreme elements (maximal and minimal elements).

However, we still need a measure for the quality of our clusterings, i.e. the main objective can now be defined: It is **to assess the clustering, depending on our methodological (steering) parameters FCL and TMF.**

In order to reach this goal, a quality-function is needed that assess our clustering.

4 Quality-Functions, Splitting of the Unsharp Partition ("good/bad"), the Sharp, the Unsharp and the Mixed Partition

As mentioned above, the basis for evaluation by partial orders and its visualization by Hasse diagrams is the partitioning B. After assigning objects to clusters corresponding to their u_{ik}-values and their relation to TMF the fuzzy clustering gets a "hard" one, i.e. a sharp partition B is generated. Therefore conventional quality functions (like the sum of the quadratic distances of the vectors $x^{(k)}$ to their cluster centers $\overline{z}^{(k)}$) may be appropriate:

$$g_{sharp}(FCL, TMF) := g_{sharp}(B) := \sum_{k=1}^{N} \left\| x^{(k)} - \overline{z}^{(k)} \right\|^2$$

An example of performing the algorithm **SHARP** with multiple FCL and TMF and $g_{sharp}(\bullet)$ as quality function is shown in section 5. The function $g_{sharp}(\bullet)$ seems not to be appropriate, because the effects of the TMF-hybrids are not visible – their error is "ignored" (it is zero because every TMF-hybrid is its own center), but these objects are *important*. However, quality functions, which explicitly take the membership values u_{ik} into account, may be more appropriate. Reference [21] gives us a measure for the quality of the unsharp partitions:

$$g_{\text{unsharp}}(FCL) := g_{\text{unsharp}}(U) := \sum_{i=1}^{FCL} \sum_{k=1}^{N} u_{ik}^{r} \cdot \left\| x^{(k)} - z^{(i)} \right\|^{2},$$

with $r > 1$ as a robustness parameter (we sometimes write briefly "g" instead of " g_{unsharp}"). Note that $g_{\text{unsharp}}(U)$ only depends on FCL and $g_{\text{unsharp}}(U) = g_{\text{sharp}}(U)$ holds for all sharp partitions U[5] (we then have: $u_{ik}=1 \Rightarrow z^{(i)} = \overline{z}^{(k)}$).

Remark 1: The hypothetical procedure **UNSHARP** solves the optimization problem: "min g(U), subject to: $U \in [0,1]^{FCL \times N}$ is an unsharp partition". That is the meaning of the formulation "*optimal* unsharp partition" in our description of algorithm **SHARP**. Although this optimization problem cannot be solved in an explicit form there is a way to approximate a local minimum by an iterative algorithm (Bock, [21]). The draft of Bock inspired our theorem 1 about mixed partitions in dependence of the steering parameter TMF (see below).

Remark 2: The "robustness parameter" r could be chosen to 1, but for r=1 an optimal unsharp partition refer to g_{unsharp} would be a sharp partition (as proved in [21]), so this choice makes no sense. The exponent r shall reduce the influence of small u_{ik} in contrast to the big ones.

Varying both steering parameters, different A- (and B-) partitions arise. In the case of the B-partitions not only FCL but FCL+|Hyb(TMF)| formal clusters are to be examined. The quantity |Hyb(TMF)| depends on FCL and TMF.

To understand the behaviour of our partitions as functions of FCL and TMF, we offer two propositions:

1. We are splitting the measure $g_{\text{unsharp}}(U)$ into two parts:

$$\text{good}_{\text{unsharp}}(U, TMF) := \sum_{i=1}^{FCL} \sum_{k \notin \Theta} u_{ik}^{r} \cdot \left\| x^{(k)} - z^{(i)} \right\|^{2}, \qquad \text{referring to partition A}$$

(with $\Theta := \left\{ k \in \{1,\dots,n\} \mid o^{(k)} \in \text{Hyb(TMF)} \right\}$); and

[5] Note the two different terms for the centers: $\overline{z}^{(k)}$ stands for the center of the cluster to which $x^{(k)}$ is assigned (makes only sense for sharp partitions) and belongs only on those objects, which belong to the cluster. $z^{(i)}$ generally describes the center of cluster i ($i \in \{1,\dots,FCL\}$) and depends on the contribution of all objects – this has to be done for unsharp partitions. Later we will use the notation $\overline{z}^{(k)}$ for the unique center of cluster i with $u_{ik} > 0.5$ too – such a cluster exists for pure objects.

$$\mathrm{bad}_{\mathrm{unsharp}}(\mathrm{U},\mathrm{TMF}) := \sum_{i=1}^{FCL}\sum_{k\in\Theta} \mathrm{u}_{\mathrm{ik}}^{\mathrm{r}} \cdot \left\| \mathrm{x}^{(\mathrm{k})} - \mathrm{z}^{(\mathrm{i})} \right\|^{2},$$ referring[6] to Hyb(TMF).

Therefore the contribution of the TMF-hybrids to the variance in $g_{\mathrm{unsharp}}(U)$ can be quantified: $g_{\mathrm{unsharp}}(\bullet) = \mathrm{good}_{\mathrm{unsharp}}(\bullet, \mathrm{TMF}) + \mathrm{bad}_{\mathrm{unsharp}}(\bullet, \mathrm{TMF})$.

As far as approximately $g_{\mathrm{unsharp}}(\bullet) \approx \mathrm{good}_{\mathrm{unsharp}}(\bullet, \mathrm{TMF})$ is valid, the use of the quality function g_{sharp} of hard clustering may be justified.

2. We are introducing a "mixed partition" $U_{\mathrm{mix}} := U_{\mathrm{mix}}(U,\mathrm{TMF})$, with $\mathrm{TMF} \in (0.5\ ,\ 1]$, i.e. we define:

   $U_{\mathrm{mix}} := (p_{ik})$ for fixed $i \in \{1,\ldots,FCL\}$ and $k \in \{1,\ldots,N\}$ formally as

$$p_{ik} := \begin{cases} 1, & \text{if} \qquad\qquad\qquad\qquad\qquad\quad u_{ik} > \mathrm{TMF}; \\ 0, & \text{if} \quad (\exists j \in \{1,\ldots,FCL\} \setminus \{i\}) \quad u_{jk} > \mathrm{TMF}; \\ u_{ik}, & \text{if} \qquad \forall j \in \{1,\ldots,FCL\} \qquad u_{jk} \le \mathrm{TMF}; \end{cases}$$

   What happens here is a "rounding" of the membership matrix, but only in some columns (each object corresponds to one column). If the value of the membership-function u_{ik} of an object $o^{(k)}$ exceeds the treshhold TMF, we increase it to 1 (sharp assignment) and have consequently to drop down all other u_{jk} ($j \neq i$) to 0. (See figure 4: all TMF-hybrids are "moved" to their (unique) corner of the simplex. See also: figure 6.) Therefore we get something like a sharp partition, but in contrast to the sharp partition we still include the TMF-hybrids as unsharp assigned objects (formally the mixed partition is considered as an unsharp partition).

Clearly the quality function $g_{\mathrm{unsharp}}(\bullet)$ of U_{mix} differs significantly (in some cases) from the assessment of A, because U_{mix} is a rounded optimal solution, where A would not even have the same format as U[7].

[6] Note that $\|x^{(k)}-z^{(i)}\|^2 \neq 0$ (i.e. $x^{(k)} \neq z^{(i)}$) holds for all hybrids $x^{(k)} \in \mathrm{Hyb(TMF)}$ and all centers $z^{(i)}$ ($i \in \{1,\ldots,FCL\}$).

[7] More formal by: $U \in [0,1]^{FCL \times N}$ is a $FCL \times N$-matrix (like U_{mix}). If the sharp partition A is formulated as Matrix, we get $A \in \{0,1\}^{(\,FCL+|\mathrm{Hyb(TMF)}|\,)\times N}$, so A is a (FCL+|HYB(TMF)|)×N-matrix, it has another format than U.

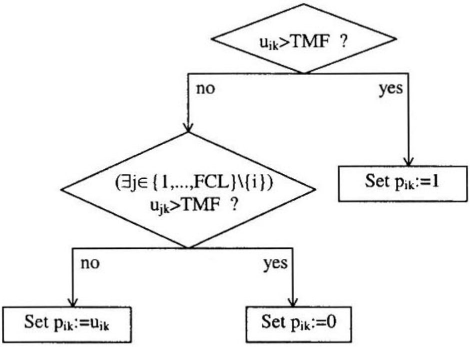


Figure 6: How to calculate the mixed partition

For the mixed partition U_{mix} the following theorem holds:

Theorem: Let $U=(u_{ik})$ be an optimal unsharp partition for $O = \{o^{(1)},..., o^{(N)}\}$ with $X = \{x^{(1)},..., x^{(N)}\} \subset \mathbb{R}^n$. Furthermore let $0.5 < \mathrm{tmf} \leq \mathrm{TMF} < 1$, let p be the mixed partition corresponding to tmf and let P be the mixed partition corresponding to TMF (and U). Then $g(P) \leq g(p)$.

Proof: We prove the equivalent relation $g(p)-g(P) \geq 0$. In order to do this, define a subset I of the set {1,...,N} of all indices as follows:

$I:=\{k\in\{1,...,N\} \mid (\exists\, j\in\{1,...,FCL\})\; u_{jk}\in \left(\mathrm{tmf} \quad , \quad \mathrm{TMF}\right]\}$

For each *relative pure* object $o^{(k)}$ we call its center $\overline{z}^{(k)}$ (the unique $z^{(i)}$ with $u_{ik} > 0.5$) (we will need no special notation for the *absolute hybrids*). We further introduce:

$$\left(\forall k \in \{1,...,N\}\right) \quad d^{(k)}:= \frac{1}{\sum_{j=1}^{FCL} ||x^{(k)} - z^{(j)}||^{-2(r-1)}}.$$

We then have:

$$g(p) - g(P) = \sum_{i=1}^{FCL} \sum_{k=1}^{N} (p_{ik}^{r} - P_{ik}^{r}) \cdot \left\|x^{(k)} - z^{(i)}\right\|^2 = \sum_{k=1}^{N} \sum_{i=1}^{FCL} (p_{ik}^{r} - P_{ik}^{r}) \cdot \left\|x^{(k)} - z^{(i)}\right\|^2$$

$= \sum_{k\in I} \sum_{i=1}^{FCL} (p_{ik}^{r} - P_{ik}^{r}) \cdot \left\|x^{(k)} - z^{(i)}\right\|^2$ because for all $k\notin I$ one gets $p_{ik}=P_{ik}=u_{ik}$ especially according to **theorem 2** of reference [21]. Further:

$$\sum_{k\in I}\sum_{i=1}^{FCL}(p_{ik}^{r} - P_{ik}^{r})\cdot\left\|x^{(k)} - z^{(i)}\right\|^2 = \sum_{k\in I}\left(\left\|x^{(k)} - \overline{z}^{(k)}\right\|^2 - \sum_{i=1}^{FCL} u_{ik}^{r}\cdot\left\|x^{(k)} - z^{(i)}\right\|^2\right)$$

$$= \sum_{k\in I}\left(\left\|x^{(k)} - \overline{z}^{(k)}\right\|^2 - \sum_{i=1}^{FCL}\left(\left(d^{(k)}\right)^{r}\cdot\left\|x^{(k)} - z^{(i)}\right\|^{-2r/(r-1)}\right)\cdot\left\|x^{(k)} - z^{(i)}\right\|^2\right)$$

$$= \sum_{k\in I}\left(\left\|x^{(k)} - \overline{z}^{(k)}\right\|^2 - \left(d^{(k)}\right)^{r}\cdot\sum_{i=1}^{FCL}\left\|x^{(k)} - z^{(i)}\right\|^{-2/(r-1)}\right)$$

$$= \sum_{k\in I}\left(\left\|x^{(k)} - \overline{z}^{(k)}\right\|^2 - \left(d^{(k)}\right)^{r-1}\right)$$

By definition of I:

$d^{(k)} \in \left(\text{tmf}\cdot\left\|x^{(k)} - \overline{z}^{(k)}\right\|^{2/(r-1)}\ ,\ \text{TMF}\cdot\left\|x^{(k)} - \overline{z}^{(k)}\right\|^{2/(r-1)}\right]$, i.e. every part of the sum above is in the interval $\left[\ (1-\ \text{TMF}^{r-1})\cdot\left\|x^{(k)} - \overline{z}^{(k)}\right\|^2\ ,\ (1-\ \text{tmf}^{r-1})\cdot\left\|x^{(k)} - \overline{z}^{(k)}\right\|^2\right)$ and therefore not negative. Q. e. d.

So we can restrict our consideration to a strip out of the region of absolute hybrids (see figure 7).

Remark: This theorem shows that choosing a higher TMF leads to a better mixed partition. On the other hand, if we increase FCL, we know what happens to the (optimal) *unsharp* partition – it can at most be improved. (Note, that this is not necessary true for *sharp* partitions.) This follows from the fact that every unsharp partition $U\in[0,1]^{FCL\times N}$ can be written as an unsharp partition $U\in[0,1]^{(FCL+\Delta FCL)\times N}$ with ΔFCL *more* clusters – call it a relaxation[8].

Illustration:

$$\begin{pmatrix} u_{1,1} & \cdots & u_{1,N} \\ \vdots & \ddots & \vdots \\ u_{FCL,1} & \cdots & u_{FCL,N} \end{pmatrix} \xrightarrow{\text{relaxation}} \begin{pmatrix} u_{1,1} & \cdots & u_{1,N} \\ \vdots & \ddots & \vdots \\ u_{FCL,1} & \cdots & u_{FCL,N} \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} \Delta FCL - times \end{matrix}$$

[8] A *relaxation* describes the transition to a greater set of feasible solutions in a problem of optimization, i.e.:

For A⊆B we have $\begin{matrix}\min/\max\ g(x)\\ \text{s.t. } x\in A\end{matrix} \xrightarrow{\text{relaxation}} \begin{matrix}\min/\max\ g(x)\\ \text{s.t. } x\in B\end{matrix}$.

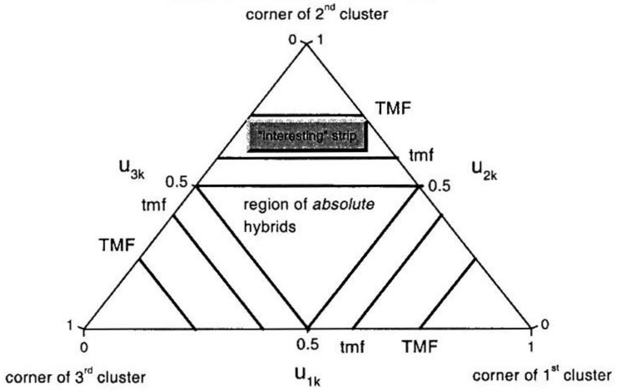


Figure 7: Illustration of the position of relevant objects (for three clusters)

Concluding:

With respect to the *mixed* partition we can justify the heuristic recommendation to choose both parameters as high as possible, if one wants a "good" mixed partition. As we have two perspectives to interpret the sharp partition, we further recommend two strategies:

1. If the singletons are not considered as important, choose an FCL where $g_{sharp}(\bullet)$ has a (maybe local) minimum. Select TMF as high as possible.
2. The more important the singletons, the more one should select values "nearer" to a minimum of $bad_{unsharp}(\bullet)$. The importance of singletons may be derived from their position in a Hasse diagram. If some singletons are (for example) extremal elements they are regarded as more important than others.

Figures 8 and 9 visualize the results of both strategies and show that a good choice with respect to the quality function does not imply a clear and readable diagram. Figure 9 is better readable than figure 8, whereas the HD shown in figure 8 which is similar to that of the original one (figure 1). The strong difference between these two HDs results from the fact, that the first strategy tries to minimize g_{sharp} (different from but *similar* to $good_{unsharp}$) while the second strategy tries to minimize $\text{bad}_{\text{unsharp}}$ (*antagonistic* to $\text{good}_{\text{unsharp}}$). It is clear that a higher TMF leads to more objects in the HD (more hybrids) and a higher value for $\text{bad}_{\text{unsharp}}$. Therefore so in the second strategy we did not choose TMF "as *high* as possible" but "as *low* as possible" – this makes the second diagram more "readable" than the first but there is no "guarantee for readability". This is not surprising, because our quality function expresses the **feasibility** of our clustering. On the other hand TMF is a demand for the **separation** of the clusters (we will give no formal discussion about a quantitative description of **separation** or the **entropy/fuzzy measure** for they are explained for example in [21]– an intuitive idea of separation is still sufficient), so we have still said nothing about the resulting HDs themselves.

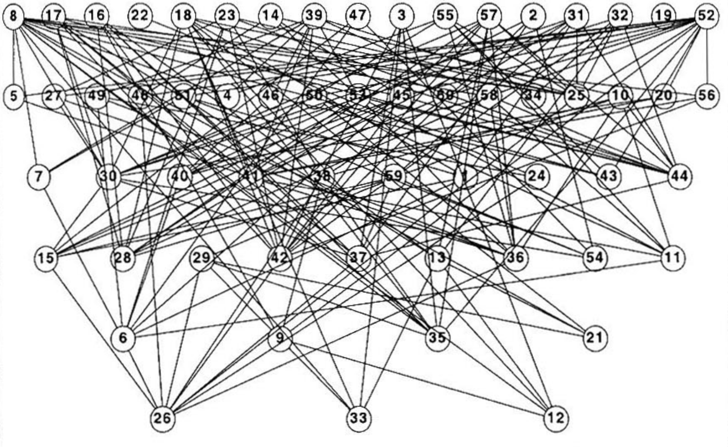


Figure 8: 1st Strategy → **HD(SHARP**(O,D,8,0.9, **IB**, $\mathbb{R}^4$)

For TMF=0.9 this diagram has the minimal value of g_{sharp} (and of $\text{good}_{\text{unsharp}}$). There are 37 0.9-hybrids. The nontrivial equivalent classes are:

{**17**,23,57}; {**9**,13,35,33,12,21}; {**16**,3}; {**22**,49,24}; {**5**,60,56}; {**27**,50,20}; {**36**,11}, {**42**}

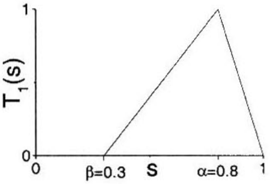
(8 clusters, 23 assigned objects)[9]

[9] There occurs a singleton object (42) which *is no hybrid.* The reason: In the unsharp partition some membership values of the other objects to this cluster were different from 0 so the cluster-center of 42 is not equal to the dataset of 42. After the rounding process all other objects were assigned to other clusters (or became singletons), so we got a one-element cluster.

What then is our intuitive idea about a HD's "readability"? Graph theory gives some attempts to answer this question. We have mentioned that the "readable" graph on figure 5 is **planar** and indeed the number of crossings may be seen as first measure for complexity. However, up to now there are only heuristic algorithms known, to calculate the crossings of a directed graph (see Sugiyama [24] and Deffland [25]), and even then there is no advice when a readable diagram would be obtained.

Therefore we introduced an own heuristic measure for readability of Hasse diagrams:

$$T_1(s) := \begin{cases} 0 & 0 \le s \le \beta \\ \dfrac{s-\beta}{\alpha-\beta} & \beta \le s \le \alpha \\ \dfrac{1-s}{1-\alpha} & \alpha \le s \end{cases}$$



$$s := \frac{\text{number of incomparabilities}}{\text{number of objects} \cdot (\text{number of objects} - 1)}$$

The quantity s is the ratio between the number of incomparabilities of the partial order and the number of edges of the corresponding complete directed graph (there is no "N" in the denominator of s, because in the partial order every *cluster* is considered as one object, i.e. "number of objects" ≤ N) – α and β are again steering parameters. The function T_1 is motivated by the observation that chains ($T_1(s)=s=0$) and antichains ($s=1$, $T_1(s)=0$) are extremely "readable", therefore complex HDs will probably appear for $s \in [\beta,1)$. Empirically we have set $\alpha=0.8$ and $\beta=0.3$ and then defined the product ("number of objects" * $T_1(x)$) to be the complexity of our partial order.

In figure 8 our "unreadability measure" delivers 45*0.87=39.09 which is much higher than (23*0.98 = 22.56), the value for figure 9. Both values are distinct lower than the value for the "original" HD on figure 1 (60*0.97=57.97) but higher than the value for the planar graph on figure 5 (17*0,63=10.63) as was to be expected.

Obviously both strategies lead to a considerable reduction of the HD's complexity compared to the complexity of the original HD. Our two optimization strategies however did not lead to the HD of figure 5. An improvement is to be expected if the optimization is performed under the constraints of a complexity measure (which may not necessarily identical with our heuristical one).

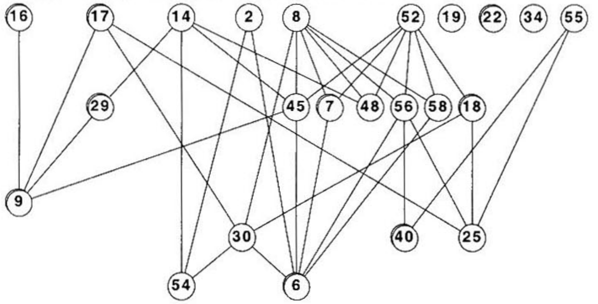


Figure 9: 2nd Strategy → **HD(SHARP(**O,D,9,0.6, **IB**, $\mathbb{R}^4$ **)**

For FCL=9 and TMF=0.6 this diagram has the minimal value of bad$_{unsharp}$. There are 14 0.6-hybrids. The nontrivial equivalent classes are:

{**6**,15,28,42,37,26,36,11,44}, {**7**,5,41,51,60,1}, {**17**,23,57,31,32}, {**9**,13,35,33,12,21}, {**16**,3}, {**22**,39,49,53,24}, {**18**,27,47,50,20}, {**40**,4,46,43}, {**29**,38,59,10}

(9 clusters, 46 assigned objects)

Finally we want to state clearly that it is not enough to compare two HDs if the (dis-) similarity of the underlying posets is the focus of interest. Two posets may be very similar but have very dissimilar HDs. A detailed discussion of this point would require the introduction of graph theoretical terms and measures which would go beyond the scope of this draft. We give just an example for "hidden similarities" of HDs: In figure 8 we see the comparability {6} < {47} which *seems* to disappear in figure 9. But in figure 9 the relation {**6**,15,28,42,37,26,36,11,44} < {30} < {18,27,**47**,50,20} holds (we consider the relation {6} < {47} as "preserved").

5 Exemplary Evaluation of the Quality Functions

Results of the proposal 1

For the dataset of the leaf layer we get – in dependence on TMF and FCL – the following form of the graph of $g_{sharp}(\bullet)$ (see figure 10). Contrary to the expectation that g_{sharp} should depend monotonically on FCL (more clusters, smaller variances), at higher FCL-values there is an increase in the value of the quality function. However, this graph is not helpful, because there is no differentiation between the partition A and Hyb(TMF). The unsharp partition, which depends by definition only on FCL, shows on the contrary a monotonous decreasing quality function $g_{unsharp}(\bullet)$ for increasing FCL (figure 11).

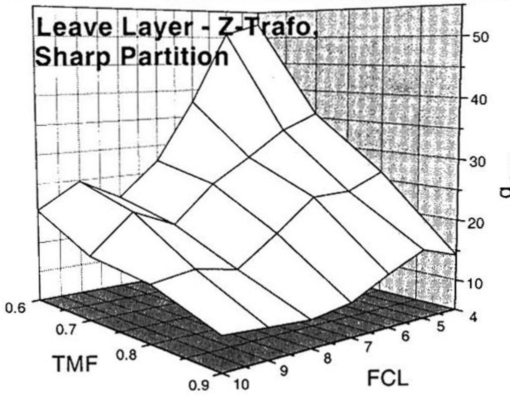


Figure 10: g_{sharp} as a function of TMF and FCL

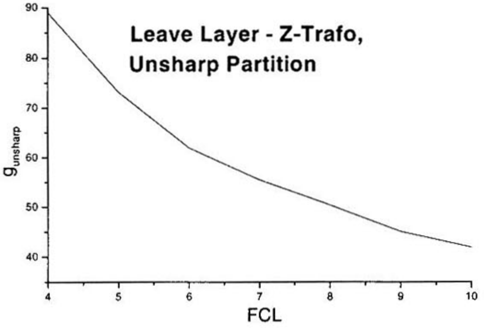


Figure 11: $g_{unsharp}$ as function of FCL

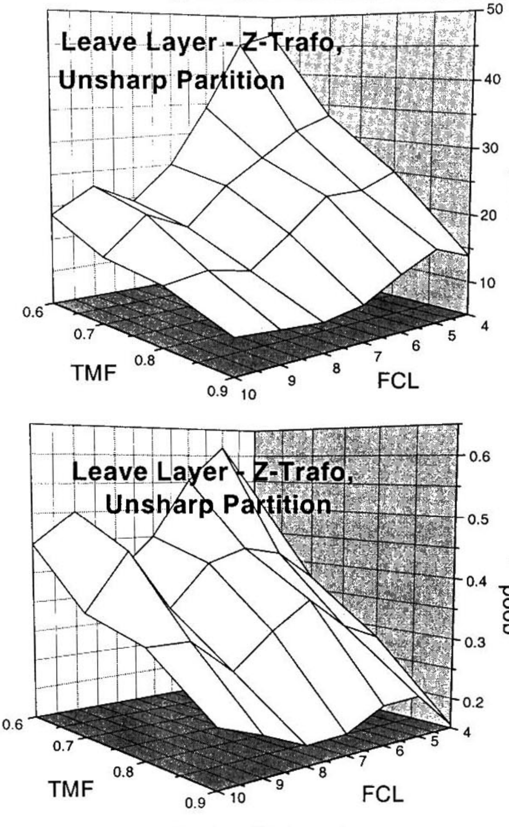


Figure 12: $good_{unsharp}$ as function of both steering parameters FCL and TMF
(In the second figure we consider the proportion $good_{unsharp}/g_{unsharp}$.)

The comparison of $good_{unsharp}(\bullet,\bullet)$ (figure 12) and $bad_{unsharp}(\bullet,\bullet)$ (figure 13) shows clearly that even for small values of TMF the "bad" part of the error is relatively big: the "bad" part for high TMF increases up to 75% (FCL=4, TMF=0.9), the contribution of the "good" part to only 62% of the whole variance(at FCL=5, TMF=0.6).

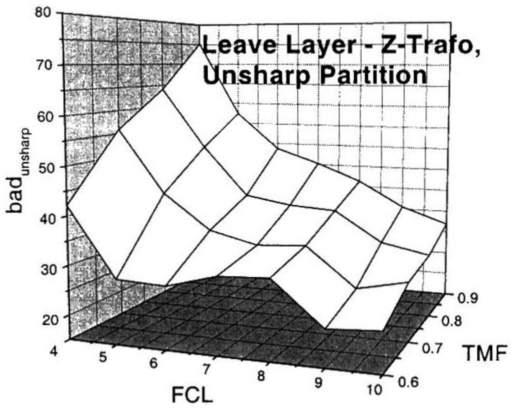


Figure 13: $\mathrm{bad}_{\mathrm{unsharp}}$ as function of the two parameters FCL and TMF. Note the scale used here in comparison to that of Fig. 12.

The main information we get from fig. 12 and 13 is that the form of the graph of $\mathrm{good}_{\mathrm{unsharp}}(\bullet,\bullet)$ is very similar to that from $g_{\mathrm{sharp}}(\bullet,\bullet)$. Therefore the non monotonous behaviour of g_{sharp} may be explained by the appearance of TMF-hybrids, i.e. by the effects of $\mathrm{bad}_{\mathrm{unsharp}}$.

Results of the proposal 2

The mixed partition is much "tamer" than the sharp one (in the sense that the 2-dimensional curve of the assessment of the optimal partition U can be recognized). Its form is plausible – the higher TMF, the "more" is the optimal (unsharp) solution rounded. (This means: The substitution of an unsharp assignment of an object to *all* clusters with a sharp assignment corresponds to a "rounding" of one column of the matrix of membership U to a 0/1-valued column vector). It is not obvious that generally the quality function has to be a monotonous decreasing function of TMF, but we proved it in the theorem 1.

It is interesting that the dependence of TMF on figure 14 is very small in relation to the dependence of FCL. That seems to affirm that our optimal unsharp solution is only slightly modified – it should be so because otherwise the rounding of our unsharp partitions would be a big distortion (i.e. our method is "empirically justified"). For TMF=1 we would get the quadratic distance of the unsharp partition (figure 11).

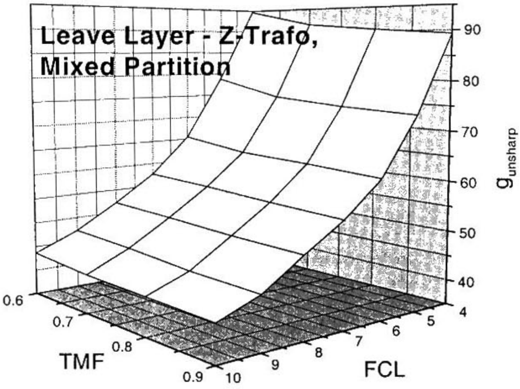


Figure 14: Mixed partition

6 Conclusions and Outlook

We started with the realization of the fact that large datasets need a reduction – in this or in other ways – if they shall be evaluated by using the HDT. Now we are at the starting point again. We ask: **How shall data be reduced in order to be still useful?**

The aggregation of different attributes in order to form a ranking index, which in turn induces a total order does not *solve* that problem – but we can think about some more (not so primitive) possibilities (like methods of clustering), and try to use their advantages. One of this advantages is that we **can quantify the total error**, and explain some details as shown in this paper. The quantifying of the error may be well done by the proposed method, however another important point was omitted: **The order preserving**. We do not expect this property from a statistical method – but instead of calling this a weakness of the method, we can say that this is a way to avoid over-interpretations (only insignificant relations are destroyed). Of course, in general order preserving methods are better than methods that are not order preserving. The problem in this point seems to be that there is still no uniform theory of cluster methods coupled with applications as shown here, even if we see efforts in this field, as for example the approach of Janowitz [22]. What we need is a formal combination of **order theory** *and* **cluster analysis**, especially in environmental sciences, where large datasets are to be evaluated. However a methodological way which is only defined by mathematical properties of the input data will be independent from its use. Such a method will be applicable to non-environmental data too, namely in all those cases, where several properties are to be considered at once. In that sense it is hoped to contribute to the puzzle "partial orderings in chemistry".

7 References

1. Brüggemann, R., Bücherl, C., Pudenz, S., and Steinberg, C. Application of the concept of Partial Order on Comparative Evaluation of Environmental Chemicals. *Acta hydrochim.hydrobiol.* 27(3):170-178, 1999.
2. Klöpffer, W. and Volkwein, S. Bilanzbewertung im Rahmen der Ökobilanz. In: *Management der Kreislaufwirtschaft*, edited by Thom'e-Kozmiensky and Joachim, K.EF-Verlag für Energie- und Umwelttechnik GmbH, 1995,p. 336-340.
3. Klöpffer, W. Subjectivity is not Arbitrary. *Int.J.LCA* 3(2):61-62, 1998.
4. Umweltbundesamt ***Methodik der produktbezogenen Ökobilanzen*** - *Wirkungsbilanz und Bewertung-*, Berlin:Umweltbundesamt, 1995.
5. Brüggemann, R. and Bartel, H.-G. A Theoretical Concept to Rank Environmentally Significant Chemicals. *J.Chem.Inf.Comp.Sc.* 39(2):211-217, 1999.
6. Behrendt, H., Altschuh, J., Sixt, S., Gasteiger, J., Höllering, R., and Kostka, T. A Unified Approach to Exposure Assessment by Computer Models for Degradation Reactions and Soil Accumulation: The Triazine Herbicide Example. *Chemosphere* 38(8):1811-1823, 1999.
7. Galassi, S., Provini, A., and Halfon, E. Risk Assessment for Pesticides and Their Metabolites in Water. *Intern.J.Environ.Anal.Chem.* 65:331-344, 1996.
8. Pflugmacher, S. and Sandermann, H.J. The use of Hasse-diagram technique in biochemistry and ecotoxicology. In: ***Proceedings of the Workshop on Order Theoretical Tools in Environmental Sciences***
   Berichte des IGB
   Heft 6, Sonderheft I, 1998, edited by Group Pragmatic Theoretical Ecology,Berlin:IGB, 1998,p. 49-52.
9. Brüggemann, R. and Halfon, E. Comparative Analysis of Nearshore Contaminated Sites in Lake Ontario: Ranking for Environmental Hazard. *J.Environ.Sci.Health* A32(1):277-292, 1997.
10. Pudenz, S., Brüggemann, R., Komossa, D., and Kreimes, K. An Algebraic / Graphical Tool to Compare Ecosystems with Respect to Their Pollution by Pb,Cd III: Comparative Regional Analysis by Applying a Similarity Index. *Chemosphere* 36(3):441-450, 1998.
11. Brüggemann, R., Pudenz, S., Voigt, K., Kaune, A., and Kreimes, K. A Generalized Order Concept - a Helpful Tool for Decision Support in Environmental Sciences. In: ***ECO-INFORMA'97*** *Information and Communication in Environmental and Health Issues,* edited by Alef, K., Brandt, J., Fiedler, H.,

Hauthal, W., Hutzinger, O., Mackay, D., Matthies, M., Morgan, K., Newland, L., Robitaille, H., Schlummer, M., Schüürmann, G., and Voigt, K.Bayreuth:ECO-INFORMA Press, 1997,p. 517-522.

12.Schrenk, C., Pflugmacher, S., Brüggemann, R., Sandermann, H.J., Steinberg, C., and Kettrup, A. Glutathione S-Transferase Activity in Aquatic Macrophytes with Emphasis on Habitat Dependence. *Ecotox.Environ.Saf.* 40:226-233, 1998.

13.Sörensen, P.B., Mogensen, B.B., Carlsen, L., and Thomsen, M. The role of uncertainty in Hasse diagram ranking. In: ***Proceedings of the Workshop on Order Theoretical Tools in Environmental Sciences***
Berichte des IGB
Heft 6, Sonderheft I, 1998, edited by Group Pragmatic Theoretical Ecology,Berlin:IGB, 1998,p. 71-84.

14.Sörensen, P.B., Mogensen, B.B., Gyldenkaerne, S., and Rasmussen, A.G. Pesticide Leaching Assessment Method for Ranking both Single substances and Scenarios of Multiple Substance Use. *Chemosphere* 36(10):2251-2276, 1998.

15.Trenkle, R. Vorschlag fuer den Gebrauch eines neuen Bewertungsschemas in Umweltverträglichkeitsstudien. *UVP report* 12(1):7-8, 1998.

16.Voigt, K. *Erstellung von Metadatenbanken zu Umweltchemikalien und vergleichende Bewertung von Online Datenbanken und CD-ROMs*, Aachen:Shaker Verlag, 1997.pp. 1-209.

17.Voigt, K. Hasse Diagram Technique as an Effective Tool in Environmental Information Management. In: ***Proceedings of the Workshop on Order Theoretical Tools in Environmental Sciences***
Berichte des IGB
Heft 6, Sonderheft I, 1998, edited by Group Pragmatic Theoretical Ecology,Berlin:IGB, 1998,p. 55-66.

18.Brüggemann, R., Voigt, K., Kaune, A., Pudenz, S., Komossa, D., and Friedrich, J. *Vergleichende ökologische Bewertung von Regionen in Baden-Württemberg GSF-Bericht 20/98*, Neuherberg:GSF, 1998.pp. 1-148.

19.Bock, H.H. *Automatische Klassifikation*, Göttingen:Vandenhoeck&Ruprecht, 1974.pp. 6-480.

20.Bock, H.H. ***Klassifikation und Erkenntnis III*** *Numerische Klassifikation Vorträge und Diskussionen zur Numerischen Klassifikation bei der 3. Fachtagung der Gesellschaft für Klassifikation e.V. , Königsstein /Ts., 4.-6.4.1979*, Frankfurt:Gesellschaft für Klassifikation, 1979.pp. 1-174.

21.Bock, H.H. Clusteranalyse mit unscharfen Partitionen. In: *Studien zur Klassifikation, Bd 6 Klassifikation und Erkenntnis III - Numerische Klassifikation*, edited by Bock, H.H.Frankfurt:Gesellschaft für Klassifikation, 1979,p. 137-163.

22. Janowitz, M.F. An Order Theoretic Model for Cluster Analysis. *SIAM J.Appl.Math.* 34(1):55-72, 1978.

23.Brüggemann, R., Kaune, A., Komossa, D., Kreimes, K., Pudenz, S., and Voigt, K. Anwendungen der Theorie partiell geordneter Mengen in Bewertungsfragen. *DGM* 41(5):205-209, 1997.

24.Sugiyama, K., Tagawa, S., and Toda, M. Methods for visual Understanding of Hierarchical System Structures. *IEEE Transactions on Systems,Man, and Cybernetics* 11(2):109-125, 1981.

25.Deffland, F. *Layoutalgorithmen für Graphen Diplomarbeit, Technische Universität Berlin Fachbereich Mathematik*, Berlin:Eigenverlag, 1996.pp. 1-83.