Chapter 11 Illustrative Case Studies

11.1 Overview

Here we demonstrate some of the tools outlined in Chapters 2, 3, 4, 5, 6, 7, 8, 9, and 10. The illustrative case studies are based on "real-life" data matrices. The sections are organized as follows (Table 11.1).

Tool	Sections	Case study	Related to
Hasse diagram	2.5	Pollution in	Environmental
Discretization	6.3	Baden-Wuerttemberg	chemistry
p-algorithm	6.5		
POSAC	3.5	Internet sources about drinking water quality	Environmental health
Attribute-related sensitivity	4.2	Fish communities in wetlands	Biology
Ambiguity and antagonism	4.3, 4.4, 5.5.2	Ranking of high-production chemicals	Environmental chemistry
Attribute value-related sensitivity	6.6	Ecological value of communes	Ecology
Dominance among subsets	5.6	(a) Chemicals in a river(b) Human environment index	(a) Environmental chemistry(b) Environmental sciences
Separability	5.4, 5.6	Management in a river basin	Hydrology
Stepwise aggregation	7.4	Management in a river basin	Hydrology
Navigation	3.6, 5.3	Human environment index	Environmental Sciences
Fuzzy partial order	6.4	Biomanipulation in a lake	Limnology

Table 11.1 Organization of Chapter 11

11.2 Illustrative Case Study: Pollution in Baden-Wuerttemberg (Environmental Chemistry)

11.2.1 Introduction

For monitoring the pollution status of the German state Baden-Wuerttemberg, the Environmental Protection Agency divided Baden-Wuerttemberg into 60 regions, which are the objects of our analysis (EPA Baden-Wuerttemberg, 1994; Bruggemann et al., 1997).

11.2.1.1 Hasse Diagram of (X, IB)

- X, the object set, consists of 59 regions, where data are available.
- IB, the information base, consists of four indicators, *c*_{Pb}, *c*_{Cd}, *c*_{Zn}, and *c*_S, the concentrations of Pb (lead), Cd (cadmium), Zn (zinc), and S (sulfur) in the herb layer in mg/kg dry mass. As shorthand notation, we write simply: IB = {Pb, Cd, Zn, S}.
- Orientation: The larger the concentration, the larger the loading of the region.

For convenience, the Hasse diagram based on the original data matrix, already shown in Chapter 3, is displayed once again in Fig. 11.1.

11.2.2 Further Tools of Partial Order Analysis

11.2.2.1 Minimum Rank Graph

In Fig. 11.2, the minimum rank graph is shown. We select two regions, namely regions 48 and 52, both of which are proper maximal elements (see Chapter 2).



Fig. 11.1 Baden-Wuerttemberg. 59 regions. Herb layer, indicators Pb, Cd, Zn, and S (mg/kg dry weight)



Fig. 11.2 Minimum rank graph of the two regions 48 and 52. Canonical sequence: $Pb\cong S>Cd>Zn$

Both regions are maximal elements; however the minimum rank graph shows remarkable differences: Region 48 has more successors compared to region 52, because the concentration of Pb in region 48 is larger than that in region 52. Adding indicator S reduces |O(48)| drastically, because the concentration of sulfur is pretty low. Adding further indicators has only slight effects on |O(48)| and therefore on its minimum rank.

Region 52 has less successors compared to region 48 because of its lower concentration of Pb. Adding the indicator S and Cd reduces |O(52)| slightly, whereas adding indicator Zn to the data matrix has the strongest effect on region 52. The reason is that the concentration of Zn in region 52 is pretty low. Therefore, many regions are eliminated as elements of the down set of 52 after the last step.

11.2.2.2 Dominance Degrees

There is information available on the following characteristics:

- 1. Contribution forests
- 2. Degree of agricultural activity
- 3. Industrial density
- 4. Traffic density
- 5. Settlement density

The experts scored these five characteristics from 1 to 3, whereby value 3 indicates a bad status, i.e., a high pressure (in the sense of the DPSIR concept of the OECD; see Kristensen, 2004) on the ecosystem. We select regions having in

Table 11.2 Pressures

Pressure	Identifier	Regions
In three characteristics	H: high pressure	18, 52
In two characteristics	M: medium pressure	19, 22, 59, 1, 30, 47, 49, 56
In one characteristic	L: Low pressure	54,4,7,16,41,45,50,55,60,3,43,44,31,39,42

Fig. 11.3 Directed graph. Evaluation of the edges by the dominance degrees $\neq 0$



one, two, or three characteristics a score 3. Table 11.2 informs about the loadings of regions according to the pressure indicators.

The calculation of the dominance degrees of contextual defined subsets $X_i \subset X$ can be conveniently displayed by a directed graph and is shown in Fig. 11.3.

Figure 11.3 (number of regions belonging to different degrees of pressure is given in bold numbers; see Table 11.2) shows that Dom(H, M) = 0.44. Therefore, we may hypothesize that the degree of pressure governs the pollution level. However, if we compare regions of medium pressure to those of low pressure, the dominance analysis does not support this hypothesis. The reason may be that the pressure indicators provide a description on a regional scale, whereas the pollution may also be affected by long-range airborne transport processes.

11.2.2.3 Transformations of the Data Matrix

Discretization Scheme

The K values (see Chapter 6) were given by the scientists of the Environmental Protection Agency (Dr Kreimes: personal communication) as follows:



Fig. 11.4 Partial ordering of the regions of Baden-Wuerttemberg (see text)

K(Pb) = K(Cd) = K(Zn) = 3 with the scores 0, 1, 2, K(S) = 2.

The minimum and maximum values are taken from the columns of the normalized data matrix (Table A.3).

11.2.2.4 Hasse Diagram

Figure 11.4 shows the resulting partial order. Instead of circles and identifiers as graphical elements, we draw small bar diagrams, which indicate the pollution profile.

11.2.2.5 Interpretation

In Fig. 11.4 the numbers to the right of the pollution profile are identifiers of the regions and the bars show their pollution status with respect to lead (Pb), cadmium (Cd), zinc (Zn), and sulfur (S) in the herb layer (mg/kg dry mass). For the equivalence classes, see Table A.4). The Hasse diagram makes evident that the following:

• Pollution increases in different manner when starting from the bottom (region 30) and moving upward along the lines.

- Any bifurcation tells us that the pollution pattern (like moving upward from region 17 to 41 on the one hand and 22 on the other hand) is qualitatively different: 41 is additionally loaded by cadmium, whereas 22 is loaded by lead.
- Any union of lines tells us that the pollution load is so large that qualitative differences in the lower part of the diagram are of no concern for all objects upward.
- In comparison to Fig. 11.1, we identify the separated subset {regions 34 and 09}. These regions cannot be compared with the most other regions, because they have high scores of the metal Zn, together with a medium score of the rather toxic but broadly technically used metal Cd. This pollution pattern will not be found for other regions. The graphical display by a Hasse diagram shows this peculiar property. The causal background for this high degree of separability may be the mining activities in former centuries in those regions (Kreimes, 1996).
- There is a triangular shape (see Chapter 5): Going upward on the Hasse diagram, one sees how the different profiles evolve, and how more and more incomparabilities appear and we can see why an object is and where it is located because we can see the data profile.
- There is a need to remedy the pollution status of many regions, but the strategy of remediation will be different, as each maximal element has another pollution profile.
- Finally, the component {12, 9} found in the Hasse diagram, Fig. 11.1 is not reproduced. Indeed the numerical differences leading to the two nontrivial components in Fig. 11.1 are very small.

11.2.2.6 p-Algorithm

We selected the cut values q_{i0} such that with respect to each attribute, 90% of the objects are sent into the swamp (see Chapter 6). In Fig. 11.5, the resulting Hasse diagram (equipped with the profiles of the maximal elements) is shown.



Fig. 11.5 The same Hasse diagram after applying the p-algorithm. Equivalence classes: {8, 45, 20}, {38, 57, 24}, {19, 43}, and the swamp SW (see text)

Following the lines of Chapter 5, we give some examples of navigation: Starting from the maximal elements of the left part, we see there is no Zn (in terms of the p-algorithm). Applying Eq. (5.9), we identify the objects in the left part of the diagram as being not polluted by Zn. Similarly the right part is that there is no Pb pollution.

11.3 Illustrative Case Study: Internet Sources About Drinking Water Quality (Environmental Health, IT)

11.3.1 Introduction

In a study performed by Voigt and Welzl (2002), the question was about the extent of informativeness of Internet sources about drinking water quality. Drinking water is extremely important. Its contamination by active pharmaceutical substances becomes a serious matter (see Freier et al., 2007). Hence it is of interest to know as to how we get information about the quality of drinking water. Voigt and Welzl (2002) investigated Internet sources of the 16 states of Germany. Five indicators were worked out:

- 1. Number of chemicals reported, NU
- 2. Kinds of chemicals, KC
- 3. Other properties like microbial studies, hardness of water, OP
- 4. Number of monitoring sites in cities, MS
- 5. Degree of explanations and the contexts, EX

Hence the information base is $IB = \{NU, KC, OP, MS, EX\}$.

11.3.2 Data Matrix and Partial Order

The indicator values are 0, 1, 2 and the orientation is that high value indicates a good information status. The object set $X = \{BAW, BAY, BER, BRE, BRA, HAM, HES, MEC, NIE, NOR, RHE, SAA, SAN, SHO, THU\}$. In Table 11.3, the data are summarized.

In short, Fig. 11.6 shows that a unique ranking of subsets of the 16 states is possible, although more than one indicator is needed to characterize them. For example, BRE < HES < BRA < SAN < BAW < BAY. We see furthermore that SHO and BAW have good positions as they are near the top of the Hasse diagram. However, they differ in their attribute values: SHO has value 2 in four of five indicators but 0 in indicator MS. BAW, however, has only good values in NU and KC, and medium values for the residual indicators, but no worst value. The Hasse diagram makes us aware of such data profiles.

Id	State in Germany	NU	KC	OP	MS	EX
BAW	Baden-Württemberg	2	2	1	1	1
BAY	Bayern	2	2	2	1	2
BER	Berlin	1	2	1	2	1
BRA	Brandenburg	0	1	1	0	0
BRE	Bremen	0	0	0	0	0
HAM	Hamburg	0	0	1	0	1
HES	Hessen	0	0	1	0	0
MEC	Mecklenburg	0	0	0	0	0
NIE	Niedersachsen	0	0	1	0	0
NOR	Nordrhein-Westfalen	2	2	1	1	1
RHE	Rheinland-Pfalz	2	2	1	1	1
SAA	Saarland	0	0	0	0	0
SAC	Sachsen	1	1	1	1	1
SAN	Sachsen-Anhalt	1	2	1	0	1
SHO	Schleswig-Holstein	2	2	2	0	2
THU	Thueringen	0	0	1	1	0

Table 11.3 Sixteen states of Germany and the information status about drinking water



. / >			
ES	Equivalence	Equivalence	Equivalence
	class	class	class
RE	{BAW,NOR,RHE}	{BRE,MEC,SAA}	{HES,NIE}
-		•	•

Fig. 11.6 Hasse diagram of (X, IB), together with a table of its equivalence classes

Several other chains of this length can be found. If we consider the representative elements of equivalence classes (see Chapter 2), then we find two linear extensions (Chapter 3) as follows:

le(1): (BRE, HES, THU, HAM, BRA, SAC, SAN, BER, BAW, SHO, BAY) and le(2): (BRE , HES, BRA, HAM, SAN, SHO, THU, SAC, BAW, BAY, BER)

Do these two linear extensions reproduce $(X/\cong, IB)$ implying dim $(X/\cong, IB) = 2$? With PyHasse software (Chapter 17) this can easily be checked and answered with "yes."



Fig. 11.7 A POSAC-like diagram, drawn after Voigt and Welzl (2002)

The two linear extensions allow a complete embedding of the objects in a twodimensional coordinate system. The ranks of the two linear extensions can be taken as coordinate values. For example, THU would get (3, 7) and SAN (7, 5).

Now POSAC tries to find two coordinates preserving the comparabilities as much as possible (see Section 3.5). The procedure is based on an optimization algorithm, hence it may obtain the correct coordinate-wise representation only approximately. Indeed, the POSAC routine, as provided in the statistical software SYSTAT, finds a solution with 98% accuracy. The POSAC result is shown in Fig. 11.7.

Figure 11.7 shows the representatives. One can see that BRA is comparable to HAM, according to the POSAC diagram. However, in reality, it is incomparable.

This example shows that dimension theory (see Chapter 3, Section 3.5) has some useful applications here. Voigt and Welzl show that LOV(1) is mainly determined by KC, whereas for LOV(2), the indicators MS, OP, and KC contribute nearly to the same extent.

11.4 Illustrative Case Study: Fish Communities in Wetlands (Biology)

11.4.1 Introduction

Wetlands of Gosen are located southeast of Berlin, capital of Germany. There are 12 creeks whose maintenance is cost intensive. There are many considerations for being "important," for example, rare macrophytes and the riparian zone as ecotone.

Here, however, the fish communities are of main interest, because some of the fish species belong to the "red list species." That means there is a risk that closing a creek may cause the extinction of rare fish species. Scientists of the Leibniz-Institute of Freshwater Ecology and Inland Fisheries investigate these 12 creeks with respect to their fish communities (Bruggemann et al., 2002).

Hence:

- Object set $X = \{gs, gv, gl, gm, ga, gz, A, F, G, K, M, T\}$.
- IB consists of nine indicators measuring the abundance of nine fish species.
- Orientation of the attributes: A creek with more individuals of any of the nine fish species is considered as more important than a creek with less.

Thus a 12×9 matrix of abundance measurements was obtained and is to be evaluated.

11.4.2 Attribute-Related Sensitivity Study

Figure 11.8 shows the Hasse diagram of this data matrix. There are seven maximal elements. These seven creeks should be maintained because there are no other creeks that are better.¹

There clearly arises one question: Which fish species influences that result most? Perhaps the Hasse diagram collapses to only some few maximal elements if the indicator of one fish species could be deleted from the data matrix. The attribute-related sensitivity study by means of the matrix *W* showed that the most important indicator is the number of individuals of a small fish, called "crucian carp."

Why is this little fish so important for the structure of the Hasse diagram? The reason is that the crucian carp is, on the one hand, very weak in competition for nutrients compared to the other eight fish species. On the other hand, the crucian



Fig. 11.8 Creeks of the wetlands of Gosen, ordered by the abundances of nine fish species

¹Whether or not the remaining five creeks can be closed depends on the topology of the networks of creeks and is here of minor interest.

carp tolerates bad conditions, i.e., very shallow and warm, i.e. oxygen-deficient creeks. Other fish species avoid these creeks. Hence there is a trade-off: If the crucian carp is present, then the other fish species are not or almost not present, or when other fish species prefer a certain creek, then the crucian carp tends to stay in other ones. Therefore, this small fish induces many incomparabilities. Deleting the column "abundance of crucian carp" from the data matrix will replace many of these incomparabilities. Hence, the abundance of the crucian carp is the most important indicator for the Hasse diagram.

11.5 Illustrative Case Study: Ranking of High Production Volume Chemicals (HPVCs) (Environmental Chemistry)

11.5.1 Introduction

The evaluation of marketed chemicals is confronted with around 1,00,000 already used chemicals. Yearly around 1,000 chemicals are appearing newly on the market. The current number of existing substances marketed in volumes above 1 ton is estimated to be 30,000. With REACH (registration, evaluation and authorization of chemicals) a new impact was given to rank chemicals (see, for instance, Ahlers et al., 2008; Führ and Bizer, 2007): Before sophisticated but rather expensive and time-consuming risk assessment studies (for instance, by application of the simulation model EUSES; see, e.g., Attias et al., 2005) are performed, one wants to rank with easily available indicators in order to find out the most important chemicals to save time and costs.

11.5.2 Partial Order

As an example, 12 high production volume chemicals (HPVC) were selected (Lerche et al., 2002). Let us repeat the main questions and steps in a ranking study:

What is the aim of ranking? Environmental hazard

How can we describe the environmental hazard?

By attributes like the production volume (prod.vol), toxicity (tox.), accumulation tendency (acc), and probable lifetime of a chemical (degrad.).

What orientation should be selected?

- Production volume: We let it as it is; large values indicate an hazard
- *Toxicity*: We have to revert the data (for example, by multiplication with -1)
- Accumulation: We let it as it is; large values indicate an hazard
- *Degradation*: In the literature, one finds biodegradation in percentage per day. That means the larger the number, the higher the degree of degradation, and the lower the lifetime of a chemical in our environment, the lesser their adverse impact on the environment. Therefore, we have to revert the data.



Fig. 11.9 Hasse diagram of high production volume chemicals (HPVCs) under accumulation, toxicity, and lifetime

What are the objects?

Twelve chemicals with a rather high production volume, for which abbreviations like MAL, CHL, GLY, and ISO are used to identify them in the Hasse diagram.

The following example should illustrate the role and the use of CAM, the cumulative ambiguity maximum. Therefore, we simulate the effect of adding an indicator. We omit for the moment the information production volume and see how the Hasse diagram looks like and which value we get for CAM (Fig. 11.9).

11.5.3 Cumulative Ambiguity Maximum (CAM)

Software WHASSE renders the value 0.621 for CAMobject set.

Following the lines of Chapter 4, an impact on the Hasse diagram is to be expected if an indicator is either deleted from or added to the data matrix. Therefore, it might be good to look for further indicators to characterize the objects. Indeed we have one indicator left in our background, the production volume.

Figure 11.10 shows the Hasse diagram, which is obtained by inserting the corresponding data column "production volume" in the data matrix.

Now, $CAM_{object set} = 0.76$.

CAM increased from 0.621 to 0.76. Hence the impact on a Hasse diagram by adding new indicators is decreased. Inspection of Fig. 11.10 shows three components (compare Chapter 2). Therefore, the high number of incomparabilities is plausible. More attributes would further reduce the number of comparabilities and introduce more ambiguity into the ranking study. Thus – if there are no urgent contextual requirements for the introduction of further indicators – we do not need to find more of them.



Fig. 11.10 Hasse diagram of HPVCs under production volume, accumulation, toxicity, and lifetime

11.5.4 Antagonism

Inspecting the Hasse diagram in Fig. 11.10, we may wonder as to how the separation $X_1 = \{\text{DIA}, \text{LIN}, 4\text{NP}\}$ from $X_{\text{res}} = X - X_1$ can be explained. Using methods explained in Chapter 5

$$Sep(X_1, X_{res}, \{PV, \log KOW\}) = 0.815$$

Therefore, we can display the order relations by an approximate scatter plot (Fig. 11.11).

In Fig. 11.11, the objects of X_{res} which are not minimal, maximal, or isolated elements are located within the grey field. As the separation is not complete, some comparabilities among elements of X_1 and X_{res} are still present. Some of these relations are indicated by \bullet lines.



Fig. 11.11 The objects of X_1 are denoted as *diamonds*, whereas the *circles* are the elements of MIN, MAX, and ISO of (X_{res} , IB) (see text)

The indicator BD (lifetimes of chemicals) completes the separation. The block arrows indicate how adding BD breaks the comparabilities such that

AIB = {PV, log KOW, BD} and Sep $(X_1, X_{res}, AIB) = 1$.

11.6 Illustrative Case Study: Ecological Value of Communes (Ecology)

11.6.1 Introduction

For a study, 15 out of 108 communes in Italy (Val Baganza near Parma) were arbitrarily selected and denoted as a, b, \ldots, o .

In Val Baganza, the most frequent CORINE Biotope habitats are

- · lowland hay meadows
- Medio-European rich soil thickets
- subalpine thermophile siliceous grasslands
- northern Apennine Mesobromion grasslands.

The task is to estimate the ecological value of the 15 communes. However, there is no measure to quantify the ecological value. Hence we need proxies by which the unknown ecological value may be quantified (Fig. 11.12, hierarchy of proxies).

Rossi et al. (2008) defined the following indicators as proxies for ecological value (see Rossi, 2001 for background material):

- size_ha: All the other things being equal, the biodiversity tends to increase with the habitat size and is therefore indicating an ecological value.
- vert_rich: All the other things being equal, the value of a habitat seems to be positively correlated with the number of vertebrates, whose ranges cover the habitat.
- h_rarity: Value of a Corine Biotope habitat within the size of the area studied.



- v_rarity: Involvement of the Corine Biotope habitats in those areas which host rare vertebrates.
- soil_rough: Value of a habitat with reference to its soil roughness. The irregular topography of each Corine Biotope habitat has been computed as a coefficient of variation (CV):

$$CV = [std. dev. (altitude)/(mean(altitude))]^*100$$
 (11.1)

11.6.2 Partial Order

The data matrix consists of 15 rows and 5 indicators, |X| = 15, |IB| = 5. In Fig. 11.13, the Hasse diagram of (X, IB) is shown.



Fig. 11.13 Hasse diagram of 15 Italian communes, IB = {VRI, HRA, VRA, Sro, Sz}



Fig. 11.14 Hasse diagram of 15 arbitrarily selected communes (out of 108) and five indicators {VRI, HRA, VRA, Sro, SZ}, each one with the scores 0,1,2

Following the principle of "ordinal modeling," we also perform a discretization with the discretization scheme K = 3, with minimum and maximum values taken from the data matrix (see Chapter 6) and Table A.6. The resulting Hasse diagram is shown in Fig. 11.14.

We see that the Hasse diagram (Fig. 11.14) has many comparabilities, hence the selection of the five indicators seems to describe a common concept, which we interpret as "ecological value" of these communes.

We see furthermore that there is one commune, c, which is isolated, hence this commune may be more intensively analyzed, applying, for example, the tool of antagonism (Chapter 5). As we already applied antagonism in several application chapters, and as the isolation of commune c may just arise from the arbitrariness of the selection of the communes, we exclude this analysis.

11.6.3 Attempts to Improve the Position in the Partial Order (Attribute Value Sensitivity)

Now assume that the communes would like to try to improve their ecological status. Let us select those communes which are presently located at the fifth level (one level below the top one).

The representative communes are *h*, *k*, *i*.

Indicators which already have the score 2 are not changed, because they are in the best state. In Table 11.4, applying $\Delta = 1$, the actual and the simulated tuples of indicators are shown together with the effects on successors and predecessors.

Figure 11.15 visualizes the results for the commune h following the data in Table 11.4.

Figure 11.15 shows the change for commune *h* in terms of its successors and its predecessors if the indicator value is increased by $\Delta = 1$.

Object	Standard VRI, HRA, VRA, Sro, Sz	Simulated tuple	Change of successors ¹	Change of predecessors
h	21210	2 2 210	5	-1
		212 2 0	2	-1
		21211	2	0
k	22110	22 2 10	5	-1
		221 2 0	0	-1
		22111	1	-1
i	21011	2 2 011	2	-1
		21111	0	0
		210 2 1	0	0
		2102 2	0	0

Table 11.4 Simulated data and their effects on characteristics of the partial order

The simulated value is in bold literals

¹Equivalent elements are counted as successors



Fig. 11.15 Attribute value sensitivity (avs) (Chapter 4) of the commune h

From Table 11.4 one can deduce that for commune h, the most important effect is to change for the attribute HRA from score 1 to 2. In this case it gets five more successors. Hence, in any index ranking, these five additional successors must be below commune h. Somewhat worse but still efficient is the change of the other two attributes, Sro and Sz. Commune k: Most efficient is the change in attribute VRA, where k wins five additional successors. One additional successor is obtained if Sz is enhanced from 0 to 1.

Commune *i*: Here, only attribute HRA improves the situation. Commune *i* gets two more successors and becomes a maximal element.

11.7 Illustrative Case Study: Chemicals in a River (Environmental Chemistry)

11.7.1 Introduction

The background may be that an accident happens which releases a mixture of substances. Here we select chemicals whose systematic name is "alkanes." In Fig. 11.16, some simplified chemical formulas are shown.





Fig. 11.17 (*LHS*) Hasse diagram of the quotient set of alkanes. (*RHS*) The resulting dominance graph (see text)

Restrepo et al. (2007) analyzed the chemical class of alkanes. In Fig. 11.17 (LHS), the Hasse diagram of (X, IB) is shown. (The numbers within the circles indicate different molecules. The rectangles indicate the compounds having the same sum formula C_nH_m .)

In the partial order (X, IB), X is the object set consisting of alkanes with carbon number between 5 and 8, and IB is the set of indicators describing the sedimentation, downstream transport, and volatilization of the chemicals. The indicator values are calculated by means of EXWAT, assuming some environmental data of a river (Matthies et al., 1989; Bruggemann and Drescher-Kaden, 2003; Bruggemann et al., 2006). Orientation: The larger the values of the indicators, the higher the hazard the chemicals exert.

11.7.2 Dominance

We pose the question: Is the fate of chemicals in the river governed by molecular weight, or branching of a chemical structure, or the maximal length of the carbon chain? To answer this question, classify the chemicals into subsets of compounds according to their sum formula. Thus molecules with different chemical structure belong to the same class, because they have the same number of carbon and hydrogen atoms. Thus the chemical structural information allows us to assign an additional structure to the Hasse diagram, namely that of chemical classes (rectangles in the diagram). Now among these chemical classes, the dominance values for any pair can be calculated applying the software.

PyHasse: The result is shown in Fig. 11.17 (RHS). The directed edges are evaluated by $\max(\text{Dom}(X_1, X_2), \text{Dom}(X_2, X_1))$. Broken lines correspond to transitivity relations. We find the dominance sequence $C_8H_{18} > C_7H_{16} > C_6H_{14} > C_5H_{12}$.

In detail:

- The directed edges indicate the prevailing fraction of order relations. For members of the chemical class C₆H₁₄, *x*, and members of C₅H₁₂, *y* in 93.3% of all cases *x* ≥ *y*.
- We interpret the diagram as follows: Independent of the branching or other structural peculiarities of the molecules, the environmental loading is mainly determined by the molecular weight.
- The larger the weight, the more hazardous the chemical with respect to its fate descriptors in rivers. The dominance analysis established a canonical sequence of classes of molecules corresponding to their molecular weight.

11.8 Illustrative Case Study: Management in a River Basin

11.8.1 Background Information

Let us consider an example from water management in the river Elbe basin which is located in middle/east Europe (Behrendt et al., 2002).

The water management in the watershed of river Elbe is faced with

- regional consequences (as one of the sustainability principles)
- different loadings (pressure indicators, see OECD: http://de.wikipedia.org/ wiki/DPSIR, see also Kristensen (2004): http://enviro.lclark.edu:8002/servlet/ SBRead?ResourceServlet?rid=1145949501662_742777852_522)
- attempts to apply different measures, aiming toward a reduction of pressures on the river Elbe and its tributaries.

Forecasting the impacts of different scenarios of water management strategies on the Elbe river basin cannot realistically be simulated by field experiments. Therefore, a mathematical simulation model is necessary. The model MONERIS (Behrendt et al., 2002) is a regional water quality model, which takes into account different inputs within a watershed and estimates retardation factors and degrees of elimination by reactions or volatilization. Eight indicators, concerning different speciations and pathways of nitrogen into surface waters, were calculated in order to characterize the scenarios. Following the DPSIR concept of the OECD, the eight indicators are classified as pressure indicators. Their orientation is as follows: High pressures are expressed by high values of the indicators and indicate a bad status of the surface waters in the river Elbe basin.

11.8.2 The Hasse Diagram

Following Behrendt (personal communication), 14 scenarios of water management strategies are selected (see Table 11.5).

For any scenario, the eight indicators are estimated by means of MONERIS. Additionally the measured eight indicators for the years 1985, 1995, and 1999 were included into the final data matrix as reference years. The structure of the 17×8 data matrix as basis for the partial order analysis is shown in Fig. 11.18.

Scenario	Туре	Explanation (Behrendt, Opitz)
1	Rural	10% reduction of tile drained area
2	Rural	20% reduction of tile drained area
3	Rural	25% of arable land will be cultivated without plough
4	Rural	50% of arable land will be cultivated without plough
5	Rural	75% of arable land will be cultivated without plough
6	Marketing	Detergents with P will be replaced by P-free detergents in Czech Republic
7	Technical/urban	All particulate sewage from population not connected to sewers is transported to wastewater treatment plants (wwtp's)
8	Technical/urban	99% of population with sewer connection are connected to wwtp's
9	Technical/urban	50% storage for combined sewers (50% storage corresponds to 11.6 m ³ storage volume per hectare paved urban area)
10	Technical/urban	100% storage for combined sewers (100% storage corresponds to 23.2 m ³ storage volume per hectare paved urban area)
11	Wwtp	All wwtp's are in agreement with the EU wastewater guideline
12	Wwtp	All wwtp's with more than 1,00,000 inhabitants implement an additional microfiltration. P-effluent concentration lower than 0.050 mg/l P
13	Retention	In addition to the surface waters, all wetland areas (according to Corine 346 km ²) are used for retention
14	Retention	0.5% of agricultural area (according to Corine 450 km ²) is transferred to retention areas

 Table 11.5
 Fourteen water management scenarios



	I ₁ , I ₂ , I ₃ ,I ₈
scenario 1 scenario 14	Estimation by MONERIS
year 1985 year 1995 year 1999	Measured values





Thus the analysis of (X, IB) is based on (i) object set X of 17 scenarios; (ii) information base IB of eight indicators; and (iii) orientation: The larger the value, the higher the pressure. Figure 11.19 shows the Hasse diagram.

Figure 11.19 shows the following:

- There is a simultaneous decrease of pressure values for the chain 85 > 99 > 3 > 4 > 5.
- There are isolated elements, namely the year 1995 and the scenario 14.
- The scenarios, intended to improve the water quality in the future, are not necessarily better than reference years.
- Many scenarios are incomparable to the recent year, 1999, of measurements.
- Rural scenarios {1, 2, 3, 4, 5} and technical scenarios {7, 8, 9, 10} are separated subsets (Chapter 5).

11.8.2.1 Use of Superindicators

By an antagonism study, the following result was obtained: $X_1 = \{1, 2, 3, 4, 5\}$, the rural scenarios, $X_2 = \{7, 8, 9, 10\}$, the technical scenarios, $AIB(X_1, X_2) = \{I_3, I_4, I_7\}$. If the set X_1 is reduced to X_1' to have only the rural scenarios 3, 4, 5, with 25, 50, or 75% of arable land cultivated without plough, then

Sep(X_1' , X_2 , { I_4 , I_7 }) = 1. The two indicators I_4 and I_7 describe the nitrogen release into surface waters by erosion (I_4) and from wastewater treatment plants (I_7). As these two indicators are responsible for the separation of X_1' and X_2 , a combination of both to a superindicator (see Chapter 7) may allow us to compare scenarios of X_1' with those of X_2 . The normalized indicators of nitrogen input due to erosion, I_4 , and due to point sources, I_7 , were combined as follows:

case 1: $0.2^*I_4 + 0.8^*I_7$, case 2: $0.4^*I_4 + 0.6^*I_7$, case 3: $0.6^*I_4 + 0.4^*I_7$, case 4: $0.8^*I_4 + 0.2^*I_7$

Therefore, four new data matrices were obtained, where the remaining six indicators { I_1 , I_2 , I_3 , I_5 , I_6 , I_8 } are not changed but the four superindicators included from four weight combinations. From Fig. 11.20, we see that the relative configuration of scenarios 8, 9, and 10 remains unchanged, while the weights used for the superindicator are varied. Scenario 9 is in all cases worse than the rural scenarios of X_1' . Scenario 7 is comparable with all three rural scenarios and gets its best position relative to the scenarios 3, 4, 5 in case 1, where the indicator I_7 (nitrogen release due to wastewater treatment) has a large weight. Obviously then the good values of I_7 can compensate bad values of I_4 . We suspect that the technical needs to improve wastewater treatment plants are easier to fulfill than the needs for improving the



Fig. 11.20 Hasse diagrams of the scenarios {3, 4, 5, 7, 8, 9, 10} corresponding to the four cases (see above)

input due to erosion, which is a process depending on factors, some of which cannot be influenced by technology.

We see that a stepwise aggregation to superindicators (the very idea of METEOR, Chapter 7) can be helpful. The partial order is enriched without averaging over all indicators.

11.9 Illustrative Case Study: The Human Environment Interface Index (HEI) (Environmental Sciences)

11.9.1 Overview and Data Matrix

The human environment interface (HEI) index (Singh, 2008; Patil and Taillie, 2004) is a composite index based on three leading indicators, namely the following:

- Land indicator (L): The land indicator L considers the change in the percentage of the forested land to the total land area with respect to the reference (base) year's characteristic.
- *Air indicator* (*A*): The amount of carbon dioxide emissions per capita of a country for a year denotes the level of the air indicator for that year.
- *Water indicator* (*W*): This indicator is the arithmetic mean of the percentage of population with sustainable access to an improved water source and improved sanitation.

The indicators capture the human progress toward the environmental management. Being expressed in different units, these indicators are first transformed into dimensionless indices. Then an aggregation with equal weights yields the HEI. Lower value reveals less of a country progress toward environmental protection goal. Hence, HEI reflects the managerial efforts a country makes to maintain and improve the greenness of land, blueness of sky, and cleanness of drinking water (Patil, 2000). The ranking-based information is expected to stimulate the countries for the improvement of the existing environment.

All in all the number of countries is 151 for which the HEI can be computed, hence the object set X consists of 151 countries and the information base (IB) encompasses three indicators. The case study aims at comparing HEI with the outcomes of partial order.

The complete data matrix, the countries, and the abbreviations used, as well as the HEI, are shown in the appendix (Table A.5).

11.9.2 Ordinal Approaches

The analysis is based on the raw data matrix. The Hasse diagram of (X, IB) is shown in Fig. 11.21:



Fig. 11.21 Hasse diagram for 151 countries, based on the indicators L, A, W

- The object set *X* consists of 151 countries.
- The information base IB is $\{L, A, W\}$ and indicators are normalized.
- The orientation: The larger the indicator value, the better the status of the country.

11.9.2.1 Shape

First of all, its shape is striking: The Hasse diagram seems to have a triangular shape (see Chapter 5). With greater values in the indicators, the number of incomparabilities among the countries seems to increase. As we are aware that the shape also depends on the convention of how to draw a Hasse diagram, we apply the analysis tool U(L(i)) and the result is shown in Fig. 11.22.

This diagram shows that there is no trend in the number of incomparabilities if we compare the different levels. Indeed there are more minimal than maximal elements.

11.9.2.2 Navigation

The tools are up sets and down sets. For example, when we want to know the position of Germany with respect to other countries, we can look for up sets and down sets (Fig. 11.23).



Fig. 11.22 Average incomparabilities per level



There are 19 other countries in which all three indicators are better (F(GE)), and there are 10 countries in which all three indicators are worse (O(GE)) than those in Germany. There are no connections to 121 other countries (U(GE) = 121), therefore the profile of Germany based on *L*, *A*, and *W* must be rather peculiar.

11.9.3 Dominance Degree of Contextual Subsets

Consider, for example, the European nations (Table 11.6). We pose the following question: Is the highly industrialized middle Europe dominated by the lower industrialized and hence less polluted south Europe? We address this question by applying the dominance-separability approach as explained in Chapter 5.

We calculate the dominance degree and separabilities. In Table 11.7, the dominance degrees (Dom(i,j) and Dom(j,i)) and the separabilities Sep(i,j) can be found (i,j = SE, ME, ..., NE).

Country	Subset	Country	Subset	Country	Subset
Albania	SE	Iceland	WE	Germany	ME
Austria	ME	Ireland	WE	Greece	SE
Belgium	WE	Italy	SE	Hungary	OE
Bulgaria	OE	Netherlands	WE	Switzerland	ME
Denmark	NE	Norway	NE	United Kingdom	WE
Finland	NE	Portugal	SE	Slovakia	SE
France	WE	Rep. of Moldova	OE	Spain	SE
Sweden	NE	Ukraine	OE	1.	

Table 11.6 Countries of Europe, classified to belong to north Europe (NE), west Europe (WE), middle Europe (ME), south Europe (SE), and East Europe (OE)

Table 11.7 First figure Dom(i, j) (row *i* dominates column *j*), Dom(j, i) (column *j* dominates row *i*), and the symmetric separability

	SE	ME	OE	WE	NE
SE ME	-	0.5, 0.05, 0.44	0.125, 0.08, 0.79 0.08, 0.17, 0.75	0.61, 0.02, 0.36 0.55, 0.28, 0.17	0.67, 0, 0.33
OE WE	-	-	_ _	0.21, 0, 0.79	0.25, 0, 0.75 0.33, 0.25, 0.42

If $Dom(x, y) \ge \varepsilon$, with filter ε being 0.5, we draw a directed edge from x to y and obtain the following directed graph (i.e., an extremely small network), see Fig. 11.24.

As can be seen from Fig. 11.24, there is an isolated element, OE. For the other subsets, two maximal chains (dominance sequences, compare Chapter 5) can be found: SE > ME > NE and SE > ME > WE.

In any case, south Europe (SE) has a top position among the countries of Europe.



Fig. 11.24 Based on the indices *L*, *A*, and *W*, a dominance diagram is constructed (filter $\varepsilon = 0.5$). For details, see Tables A.19 and A.20



Fig. 11.25 Hasse diagram, based on the data of Table 11.8 (software PyHasse)

11.10 Illustrative Case Study: Analysis of Lake Restoration and Biomanipulation: Example Phytoplankton Community Structure (Biology)

11.10.1 Introduction

The status of a lake (Feldberger Haussee, Germany) needs to be improved (Krienitz et al., 1996). There is a high phytoplankton population and one wants to reduce it by inserting predator fish into the lake. The motivation for this biomanipulation lies in food web theory, but does it work in practice? To answer this question, we applied partial order on the data matrix, where the years 1987 (start of biomanipulation) until 1997 are the objects. The concentrations of five phytoplankton species were selected as indicators. The response of phytoplankton to the changed conditions reflects the success of the biomanipulation measures.

Year	Cl	Ba	Cr	Di	Су	Sum
87	5.5	2.3	5.3	0.8	4.0	17.9
88	3.4	0.6	1.3	1.8	7.0	14.1
89	3.6	1.6	1.3	0.2	6.0	13.7
90	3.2	0.2	0.3	0.1	15.0	18.8
91	3.5	1.8	1.7	0.1	13.0	20.1
92	2.7	0.4	3.9	0.1	21.0	28.1
93	6.7	1.4	1.3	0	8.0	17.4
94	2.6	0.3	2.8	0	2.0	7.7
95	0.8	1.4	3.5	1.7	7.5	14.9
96	1.7	1.5	1.2	0	2.0	6.4
97	1.8	0.2	0.6	0.2	0.1	2.9

 Table 11.8
 Annual mean value of phytoplankton group concentrations (mg/l)

Clearly, it is of interest as to how the phytoplanktonic biomass is reduced. First of all, we can see by means of an indicative function (sum of the biomass over phytoplankton species, see Table 11.8) that the phytoplanktonic mass did not monotonically decrease. Second, the reduction may not impact all species in the same way: Our focus, however is not on the phytoplankton species distribution and their intricate dynamics but on how we can rank the years due to the responding phytoplankton biomass. By means of Hasse diagrams, we can do both: We can rank the years but also make evident that the phytoplanktons respond differently to the grazing stress due to the enhanced zooplanktonic population. Incomparable years are just the order theoretical expression for that different phytoplankton behavior.

11.10.2 Methods and Materials

In this study, we have compiled the annual mean values of the main phytoplankton groups: Chlorophyceae (Cl), Bacillariophyceae (Ba), Cryptophyceae (Cr), Dinophyceae (Di) and Cyanophyceae (Cy); the abbreviations of these phytoplankton groups are indicated in parenthesis. For details of phytoplankton successions in Feldberger Haussee, see Krienitz et al. (1996).

11.10.2.1 Data in the Poset Approach

The years of biomanipulation are considered as our objects and will be comparatively evaluated. Hence the object set is $X = \{87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97\}$.

The concentrations of five main phytoplankton groups (mg/l) are selected as the indicators, by which the years of biomanipulation are characterized: Therefore, the information base is $IB = \{CI, Ba, Cr, Di, Cy\}$. The phytoplankton data set as well as the indicative sum of their biomass is shown in Table 11.8.

Related on the sum (Table 11.8 last column) the ranking of the years would be

$$92 > 91 > 90 > 87 > 93 > 95 > 88 > 89 > 94 > 96 > 97$$

The order of years indicates that the expectation of the longer the biomanipulation, the smaller the phytoplanktonic biomass is not correct. Was biomanipulation successful?

11.10.2.2 Partial Order

Based on the data set of Table 11.8, the Hasse diagram encompassing all years is shown in Fig. 11.25.

We see two levels and three isolated objects (90, 93, 95). The year 1987, when the biomanipulation started, is in the same level as six subsequent years. Only three years, 94, 96, and 97, show the desired dependency. What happens? Once again: Was the biomanipulation successful? However, we must take care of the ordinal character of Hasse diagrams. So we follow the idea of ordinal modeling (see Chapter 6): We could perform a discretization of data or a fuzzy analysis. Here we will apply the concept of fuzzy partial order.

11.10.3 Results of Fuzzy Partial Order

In fuzzy partial order (Chapter 6), the degree of what should be considered as "noise" is mapped onto the tolerance level α . Correspondingly, we can find a series of Hasse diagrams, beginning with a very high degree of tolerance and ending with the lowest degree of tolerance, where any measurement detail is ordinal interpreted.

There are 31 α cuts with the lowest at 0.406 and the highest at 1:

- With the α less than 0.406, we consider all data differences as irrelevant. Hence, all years are put into one equivalence class. We do not see any differentiation and hence we cannot decide whether the biomanipulation was successful.
- If we select $\alpha = 0.45$, then all differentiations are still neglected, except the greatest one: The years 87, 88, ..., 96 form one equivalence class. Year 97 is a singleton. The Hasse diagram based on this tolerance level (see Fig. 11.26) tells us that within this crude ordinal modeling, indeed there was an improvement from the starting year 1987 compared with the final year 1997!
- If we selected $\alpha = 0.55$, then a chain of three elements is obtained 97 < 94 < 87 (Fig. 11.26).
- Further increasing α values leads to a nonlinear structure of the partial order which indicates that nature does not react linearly, but due to different phytoplankton types differently on the grazing pressure. Still a chain of years corresponding to the expectation 97 < 96 < 94 < 88 < 87 can be identified.

If α is selected higher or equal to 0.9, then we see a breakdown of the order. We can no more conclude that the biomanipulation works as expected over the years.



Fig. 11.26 Evolution of (not labeled) Hasse diagrams (PyHasse, see Chapter 17)

There are many possible reactions which even lead to incomparability between the last two years, i.e., between 96 and 97. Even worse, isolated objects occur. Hence we may say that now the final two Hasse diagrams are a result of the biological complexity, which masks the general trend of the years of biomanipulation.

In Fig. 11.26, the Hasse diagram corresponding to $\alpha = 1$ is what we already have seen: It is just the Hasse diagram of the original data, i.e., if any measurement detail is brought into evidence by the partial order concept. Taking the results shown in Fig. 11.26, together with the tendency to render 97 as a minimal element, there is a clear evidence that the biomanipulation should be considered successful.

References

- Ahlers, J., Stock, F. and Werschkun, B. (2008). Integrated testing and intelligent assessment New challenges under REACH. ESPR-Environ. Sci. Pollut. Res., 15, 565–572.
- Alkanes. http://www.docbrown.info/page06/DFalkanes.htm.
- Attias, L., Boccardi, P., Boeije, G., Brooke, D., de Bruijn, J. and Comber, M., et al. (2005). European Union system for the evaluation of substances: The second version. Chemosphere, 59, 473–485.
- Behrendt, H., Kornmilch, M., Opitz, D., Schmoll, O. and Scholz, G. (2002). Estimation of the nutrient inputs into river systems – Experiences from German rivers. Reg. Environ. Change, 3, 107–117.
- Bruggemann, R. and Drescher-Kaden, U. (2003). Einfuehrung in die modellgestuetzte Bewertung von Umweltchemikalien Datenabschaetzung, ausbreitung, verhalten, wirkung und Bewertung. Berlin: Springer.
- Bruggemann, R., Fredrich, F., Wolter, C., Pudenz, S. and Steinberg, C. (2002). Partielle ordnungen: Ein hilfsmittel zur beschreibung von artengemeinschaften. In A. Gnauck (Ed.), Theorie und modellierung von oekosystemen, Workshop Koelpinsee 2000 (pp. 173–200). Aachen: Shaker-Verlag.
- Bruggemann, R., Pudenz, S. and Kreimes, K. (1997). Oekologische Bewertung der Regionen Baden-Wuerttembergs mit neueren mathematischen Techniken. In Landesanstalt fuer Umweltschutz Baden-Wuerttemberg (Ed.), 5. Statuskolloquium Projekt "angewandte Oekologie" 18. und 19. Maerz 1997 Schloss Ettlingen (pp. 335–343). Karlsruhe: Landesanstalt fuer Umweltschutz Baden-Wuerttemberg.

Bruggemann, R., Restrepo, G. and Voigt, K. (2006). Structure – fate relationships of organic chemicals derived from the software packages E4CHEM and WHASSE. J. Chem. Inf. Model., 46, 894–902.

DPSIR. http://de.wikipedia.org/wiki/DPSIR.

- Freier, K.P., Bruggemann, R. and Nützmann, G. (2007). Pharmazeutisch aktive Substanzen in kleinen Flieβgewaessern – Modellierung unter Beruecksichtigung des Einflusses von Grundwasser. UWSF – Z.Umweltchem.TMkotox., 19(3), 189–196.
- Führ, M. and Bizer, K. (2007). REACh as a paradigm shift in chemical policy responsive regulation and behavioural models. J. Cleaner Prod., 15, 327–334.
- Kreimes, K. (1996). Oekologisches Wirkungskataster Baden-Wuerttemberg Bewertung und zusammenfassende Darstellung von Untersuchungsergebnissen. In U. Arndt, A. Fomin, and S. Lorenz (Eds.), Bio-indikation; neuere Entwicklungen – Nomenklatur – Synoekologische Aspekte; Beitraege und Diskussion 1. Hohenheimer Workshop zur Bioindikation am Kraftwerk Altbach-Deizisau, 1995 (pp. 160–169). Ostfildern: Guenter Heimbach.
- Krienitz, L., Kasprzak, P. and Koschel, R. (1996). Long term study on the influence of eutrophication, restoration and biomanipulation on the structure and development of phytoplankton communities in Feldberger Haussee (Baltic Lake District, Germany). Hydrobiologia, 330, 89–110.
- Kristensen, P. (2004). The DPSIR framework. http://enviro.lclark.edu:8002/servlet/ SBReadResourceServlet?rid=1145949501662_742777852_522.
- Landesanstalt fuer Umweltschutz Baden-Wuerttemberg (Ed.). (1994). Signale aus der Natur, 10 Jahre Oekologisches Wirkungskataster Baden-Wuerttemberg. Ettlingen: Kraft Druck GmbH.
- Lerche, D., Bruggemann, R., Sorensen, P.B., Carlsen, L. and Nielsen, O.J. (2002). A comparison of partial order technique with three methods of multicriteria analysis for ranking of chemical substances. J. Chem. Inf. Comput. Sci., 42, 1086–1098.
- Matthies, M., Bruggemann, R., Muenzer, B., Schernewski, G. and Trapp, S. (1989). Exposure and ecotoxicity estimation for environmental chemicals (E4CHEM): Application of fate models for surface waters and soils. Ecol. Model., 47, 115–130.
- Patil, G.P. (2000). UNEP workshop on human environment interface index. Washington, DC.
- Patil, G.P. and Taillie, C. (2004). Multiple indicators, partially ordered sets, and linear extensions: Multi-criterion ranking and prioritization. Environ. Ecol. Stat., 11, 199–228.
- Restrepo, G., Bruggemann, R. and Voigt, K. (2007). Partially ordered sets in the analysis of alkanes fate in rivers. Croatica Chem. Acta, 80(2), 261–270.
- Rossi, O. (Ed.) (2001). Cartografia multiscalare della natura testo delle relazioni. IX Congresso Nazionale Della Societa Italiana Di Ecologia, Lecce Settembre 1999, University of Parma, Parma, Italy.
- Rossi, F., Pecci, A., Amadio, V., Rossi, O. and Soliani, L. (2008). Coupling indicators of ecological value and ecological sensitivity with indicators of demographic pressure in the demarcation of new areas to be protected: The case of the Oltrepo Pavese and the Ligurian–Emilian Apennine area (Italy). Landsc. Urban Plann., 86, 12–26.
- Singh, A. (2008). UNEP workshop on human environment interface index. Nairobi, Kenya.
- Voigt, K. and Welzl, G. (2002). Drinking water analysis systems in German cities: An evaluation approach combining Hasse diagram technique with multivariate statistical methods. In K. Voigt and G. Welzl (Eds.), Order theoretical tools in environmental sciences – Order theory (Hasse diagram technique) meets multivariate statistics (pp. 113–127). Aachen: Shaker-Verlag.