

Chapter 5

Structures of Partial Orders

5.1 Motivation

While visualizing a poset (X, IB) with a Hasse diagram, it is initially interesting to observe the following:

1. Whether or not there is a messy system of lines.
2. Whether the Hasse diagram resembles a
 - a. triangular shape or
 - b. rectangular shape.
3. Whether there are different components or approximate components.

As (X, IB) is based on a data matrix, we want to relate these three aspects in the Hasse diagram with the properties of the data matrix. That is, we want to discover properties of the data matrix through the structure of Hasse diagrams. Therefore, [Chapter 5](#) is organized as follows: (1) we revisit the concept of levels, (2) we show how down sets and up sets are related to attribute properties, (3) the concept of separation of object subsets is introduced, (4) we explain why and how far structures of a poset and properties of a data matrix are related. Finally, (5) the concept of dominance of object subsets is discussed.

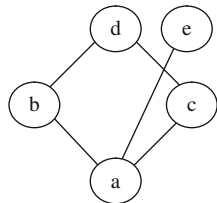
5.2 Levels and Shapes of a Hasse Diagram

5.2.1 Width and Height

In [Chapter 2](#), height and width of a poset have been introduced. Both numbers describe the shape of a Hasse diagram by inscribing it into a rectangle with height and width. The determination of width may be difficult (at least in complex Hasse diagrams) as the following example ([Fig. 5.1](#)) shows.

In [Fig. 5.1](#), the height is 3, and following the drawing protocol of Hasse diagrams (by different software packages like WHASSE and PyHasse, see [Chapter 17](#)) we

Fig. 5.1 Width and antichains



identify the antichains $\{b, c\}$, $\{d, e\}$ both having two elements. However, the width of the poset in Fig. 5.1 is 3, because the maximum antichain is $\{b, c, e\}$. In messy Hasse diagrams, it is difficult to identify the maximum antichain by visual inspection. Therefore, we use the concept of levels as a visual proxy for the discussion of shapes of the Hasse diagrams.

5.2.2 Level

The concept of levels is very useful:

- Due to the level concept, a weak order can be found among the objects.
- Levels are descriptive tools as they allow a partitioning of the objects even in messy Hasse diagrams.
- Levels are the starting point for a visualization technique, developed by Myers and Patil (2008), suitable for a huge number of objects.

5.2.2.1 Construction

Let $MAX \subseteq X$ be the set of maximal elements of a poset (see Chapter 2). (5.1)

$lg =$ number of cover relations in the maximum of all maximal chains. (5.2)

An element $x \in MAX$ gets the level number $lev(x) = lg + 1 = \text{height}$. (5.3)

Now perform a partitioning of X as follows.

Eliminate MAX from X and determine the new MAX . This new set gets a level number reduced by 1.

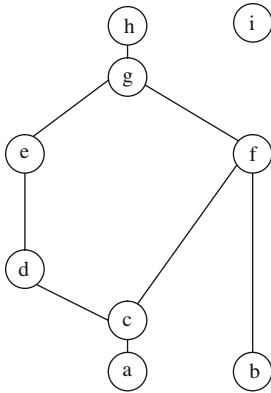
Continue the elimination process until X is exhausted.

The level sets are the equivalence classes under the equivalence relation:

$$lev = lev(x) = lev(y), \text{ with } lev \in \{1, 2, \dots, \text{height}\} \quad (5.4)$$

$$\text{Notation: } level_{lev} \text{ is the level set with the level number } lev. \quad (5.5)$$

In Fig. 5.2, an example follows.



Start
 $lg = 5$
 because the maximum of the maximal chains is $a \leq c, c \leq d, d \leq e, e \leq g, g \leq h$.

$MAX = \{h, i\}$ and gets as level set:
 $Level_6 = \{h, i\}$

Elimination process:
 After elimination of $Level_6$ from X the new MAX is: $\{g\}$, and gets level number 5.
 Hence $level_5 = \{g\}$
 The next elimination renders $\{e, f\}$ as the set with level number 4.

Levels (as sets):
 $level_6 = \{h, i\}$
 $level_5 = \{g\}$
 $level_4 = \{e, f\}$

Level numbers:
 $lev(c) = 2$
 $lev(e) = lev(f) = 4$
 $lev(a) = 1$
 $lev(b) = 3$

Fig. 5.2 Illustrative example of the concept “level”

By the number lev , the levels are enumerated from the bottom to the top of a Hasse diagram. By introducing the equivalence relation among objects “having equal lev ” (RL), the quotient set $X/_{RL}$ consists of the levels and the levels are strictly ordered due to increasing lev . In terms of the object set X , we obtain a weak order.

5.2.3 Shapes of Hasse Diagrams

5.2.3.1 Motivation

Roughly, we can identify the following types of shapes of Hasse diagrams:

1. With increasing level number, $|level_{lev}|$ is constant: rectangular shape. The number of incomparabilities is approximately constant.
2. With increasing level number, $|level_{lev}|$ is increasing: The vertex of the triangle is at the bottom of the Hasse diagram. The incomparabilities are increasing with lev .
3. With increasing level number, $|level_{lev}|$ is decreasing: The vertex of the triangle is at the top of the Hasse diagram. The incomparabilities are decreasing with lev .

For most practical purposes, these three basic shapes are sufficient.

In order to motivate the role of shapes, let us think of a class of students, being evaluated with respect to different disciplines:

- *Rectangular shape*: Independent of the level of skill, the disparity in the performance in single disciplines remains the same.
- *Triangle with the vertex at the bottom*: With increasing skill, the students show more and more disparity in the performance in single disciplines.
- *Triangle with the vertex at the top*: The better, in general, the student, the lesser the disparity in the performance in different disciplines.

5.2.3.2 Sharpening the Concept

Incomparabilities are not only among the elements of the levels but also between those of different levels such as the objects *b* and *e* in Fig. 5.2:

$$U(\text{level}_i) := \sum_{x \in \text{level}_i} |U(x)| \tag{5.6}$$

From Eq. (5.6), we can calculate the average number of incomparabilities of the *i*th level:

$$\bar{U}(\text{level}_i) = \frac{U(\text{level}_i)}{|\text{level}_i|} \tag{5.7}$$

Applying mainHD16.py of the PyHasse (see Chapter 17) delivers three histograms (Fig. 5.3).

The outcomes shown in Fig. 5.3 confirm that its Hasse diagram can be considered as having a triangular shape.

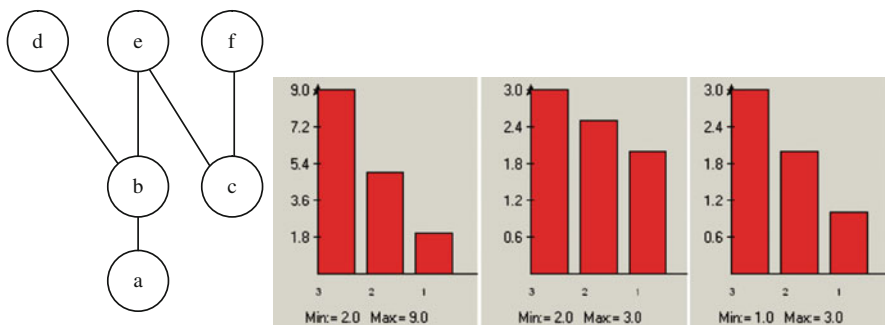


Fig. 5.3 Top: Hasse diagram; bottom: (LHS) $U(\text{level}_i)$, (middle) $\bar{U}(\text{level}_i)$, and (RHS) $|\text{level}_i|$ as function of the lev

5.2.4 Shapes and Weak or Linear Orders

The shape of a Hasse diagram allows to establish a relation

- through $\overline{U}(\text{level}_i)$ between lev and $\Delta h_{\Gamma}^{\max}(x)$, $x \in \text{level}_{\text{lev}}$, i.e., between lev and the ranking intervals and
- between $\text{lev}(x)$ and $\min(h_{\Gamma}(x))$ (Chapter 3) because in general $|O(x)|$ increases with $\text{lev}(x)$

as follows:

Rectangular shape: Increasing $\text{lev}(x)$ has no strong influence on $\Delta h_{\Gamma}^{\max}(x)$, $x \in \text{level}_{\text{lev}}$, because $\overline{U}(\text{level}_i)$ does not change much with lev .

Triangular shape (vertex at the bottom): Increasing $\text{lev}(x)$ implies increasing $\Delta h_{\Gamma}^{\max}(x)$, $x \in \text{level}_{\text{lev}}$ and $\min(h_{\Gamma}(x))$, because $\overline{U}(\text{level}_i) < \overline{U}(\text{level}_{i+1})$ and in general $|O(x)|$ becomes larger with lev .

Triangular shape (vertex at the top): Increasing $\text{lev}(x)$ implies decreasing $\Delta h_{\Gamma}^{\max}(x)$, $x \in \text{level}_{\text{lev}}$ but increasing $\min(h_{\Gamma}(x))$, because $\overline{U}(\text{level}_i) > \overline{U}(\text{level}_{i+1})$ and $|O(x)|$ increases with lev .

5.3 Down Sets and Up Sets Related to Properties of the Data Matrix

5.3.1 Idea

So far down sets and up sets are introduced to (i) simplify the Hasse diagram and (ii) relate $\Delta h_{\Gamma}^{\max}(x)$ with $|U(x)|$.

In this section, we render a relation between principal down sets and principal up sets and properties of the data matrix.

5.3.2 Realization

Through Eqs. (2.3) and (2.6)

$$y \in O(x) \Rightarrow q_i(y) \leq q_i(x) \text{ and } y \in F(x) \Rightarrow q_i(y) \geq q_i(x) \quad (5.8)$$

and

$$y \in \cap O(x_i) \Rightarrow q_j(y) \leq \min(q_j(x_i)) \text{ and } y \in \cap F(x_i) \Rightarrow q_j(y) \geq \max(q_j(x_i)) \quad (5.9)$$

Equations (5.8) and (5.9) couple properties of a poset (down sets and upsets) with some properties of the data matrix. Hence, navigation through a Hasse diagram, keeping Eqs. (5.8) and (5.9) in mind, renders some insights into the data matrix:

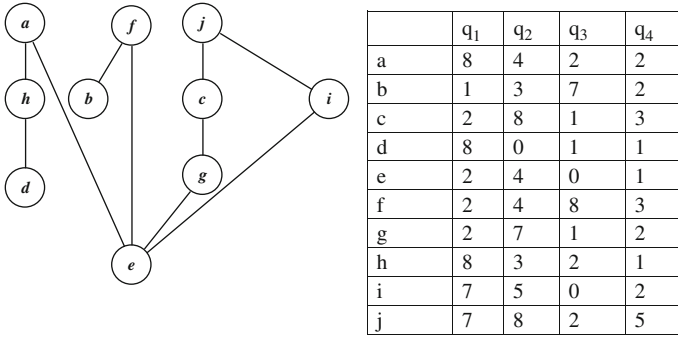


Fig. 5.4 Hasse diagram of (X, IB) , $X = \{a, b, c, d, e, f, g, h, i, j\}$, $IB = \{q_i: i = 1, \dots, 4\}$

Application of Eqs. (5.8) and (5.9) is best done for minimal or maximal objects (Chapter 2) in order to obtain useful, nonempty down sets and up sets.

5.3.3 Illustrative Example

In Fig. 5.4, a Hasse diagram together with its data matrix is shown.

- *Object j*: $q_3(j) = 2$. The range of $q_3(x)$ is $0, \dots, 8$. Equation (5.8) tells us that every object in $O(j)$ must have values in q_3 which are less than or equal to 2.
- *Object f*: $q_1(f) = 2$. The range of $q_1(x)$ is $1, \dots, 8$. Equation (5.8) tells us that for objects b and e (being elements of $O(f)$), $q_1 \leq 2$.
- *Object a*: $q_4(a) = 2$. The range of $q_4(x)$ is $1, \dots, 5$. Equation (5.8) tells us that for objects h, e , and d (being elements of $O(a)$), $q_4 \leq 2$.
- *Object e*: $e \in (O(f) \cap O(j) \cap O(a))$. Equation (5.9) tells us that e must have low values simultaneously in q_1, q_3 , and q_4 . Especially $e \in O(i)$, hence $q_3(e) = 0$ as $q_3(i) = 0$.
- *Object e*: $q_2(e) = 4$. Equation (5.8) tells us that $q_2(x) \geq 4$ for all elements $x \in F(e) = \{e, a, f, g, c, j, i\}$.

5.4 Separation of Object Subsets

5.4.1 Motivation

By a Hasse diagram, an ordinal representation of an $n \times m$ matrix in a plane is possible, even if m (the number of attributes (columns of the matrix)) is larger than 2. As such, it is a convenient visualization taking care of the order relations among the objects due to Eq. (2.3) or (2.6). However, the data profiles (Section 3.5) cannot directly be seen. Even if POSAC allows an approximate two-dimensional scatter

plot based on latent order variables, the relation to the original attributes is difficult to establish. So, why not try a projection of object subsets to a two-dimensional plane based on the original attributes preserving order theoretical information as much as possible? Hence the questions are the following:

- (1) Which projection? and (2) Which object subsets?

We begin with the second question, give then an answer to the first one. Finally we discuss an intimate relation between the Hasse diagram and the data matrix.

5.4.2 Separated Subsets, an Illustrative Example

Figure 5.5 shows the Hasse diagram of 11 objects and three attributes.

Naturally the following questions arise:

- (1) Why are all the objects of the subset $X_1 = \{m_{34}, m_{33}, m_{32}, m_{31}, m_3\}$ not comparable with all those of the subset $X_2 = \{m_1, m_2, m_{12,1}, m_{12,2}, m_{12,3}\}$? What are the common properties of X_1 and X_2 responsible for their separation in the Hasse diagram?

Beyond this, the second question in Section 5.4.1 is still open: How do we find such separated subsets in a messy Hasse diagram?

5.4.3 Articulation Points and Separated Object Subsets

Here we are going to answer the second question of Section 5.4.2.

Let n_{H1} be the number of components in poset (X, IB) and x_a be an element of X/\cong such that $(X/\cong - \{x_a\}, IB)$ has n_{H2} components with $n_{H2} > n_{H1}$, then x_a is called an articulation point. (5.10)

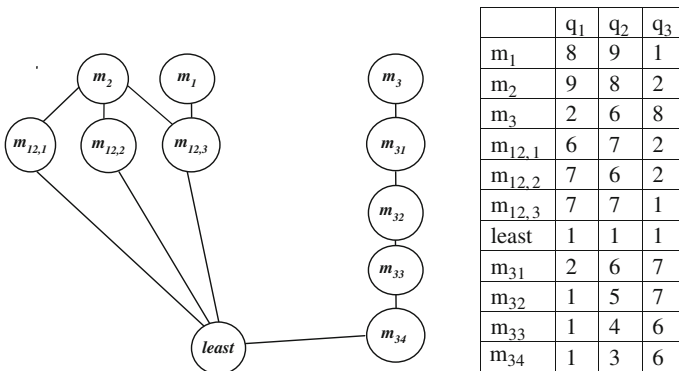
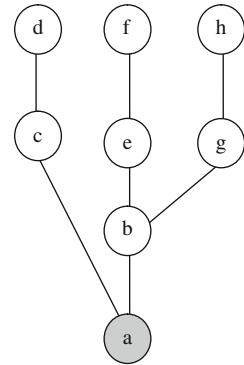


Fig. 5.5 (LHS) Hasse diagram of the data matrix (RHS)

Fig. 5.6 Hasse diagram exemplifying the concept “articulation point”



In Fig. 5.5 the object “least,” in Fig. 5.6 object a is an articulation point. By deletion of the row of object a (Fig. 5.6) in the data matrix, we obtain two disjoint subsets $X_1 = \{d, c\}$ and $X_2 = \{b, e, f, g, h\}$ which are components in the partial order. However, object b is not an articulation point, because $\{g, h\}$ and $\{e, f\}$ are still connected with object a through the transitivity of order relations (Chapter 2).

Let us identify two disjoint subsets X_1 and X_2 such that for all $x \in X_1$ and all $y \in X_2$, $x \parallel y$. We call such disjoint object subsets separated object sets and the identification of articulation points is a tool to find separated object subsets, because their presence is the reason for what we called “approximate components.”

In Fig. 5.6, the subsets $\{c, d\}$ and $\{b, e, f, g, h\}$ are approximate components. Deletion of the articulation point (object a) generates two components.

5.4.4 Separability

5.4.4.1 Motivation

The concept of separability goes the other way round: Instead of trying to find separated subsets, it is supposed that two candidate subsets are found, and we want to assess their degree of separation.

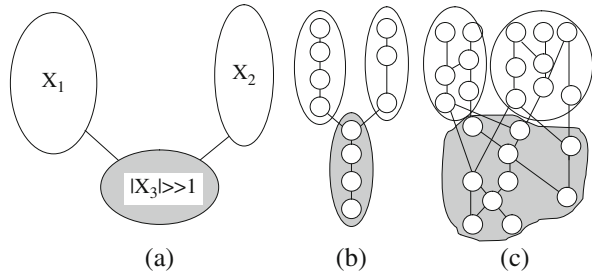
5.4.4.2 Concept

Let us identify two disjoint subsets of X/\cong : X_1 and X_2 . The possible number of relations (i.e., of $<$ or \parallel relations) $N(X_1, X_2)$ between X_1 and X_2 is

$$N(X_1, X_2) = |X_1| * |X_2| \quad (5.11)$$

Let $x \in X_1$ and $y \in X_2$, then $x \parallel y$ or $x < y$ or $y < x$. We count the \parallel relations as follows:

Fig. 5.7 (a) Separated subsets in a schematic presentation of Hasse diagrams. (b) and (c) Examples for which the scheme (a) may stand



$$U(X_1, X_2, IB') = \{(x, y) : x \parallel_{IB'} y, x \in X_1, y \in X_2, X_1 \cap X_2 = \emptyset\}, IB' \subseteq IB \tag{5.12}$$

We define the separability, $Sep(X_1, X_2, IB')$, as follows:

$$Sep(X_1, X_2, IB') := |U(X_1, X_2, IB')|/N(X_1, X_2) \tag{5.13}$$

We note that $Sep(X_1, X_2, IB') = Sep(X_2, X_1, IB')$.

The separability allows us to characterize any disjoint pair of subsets $X_i, X_j \subset X$ and to find separated subsets without checking the Hasse diagram for articulation points (Fig. 5.7).

5.4.4.3 Illustrative Example

Figure 5.8 shows a Hasse diagram together with three subsets X_1, X_2 , and $X_3, X_i \subset X$. We demonstrate the calculation of $Sep(X_i, X_j)$ ¹:

$$\begin{aligned} |X_1| * |X_2| &= 4, |U(X_1, X_2)| = 4, Sep(X_1, X_2) = 1 \\ |X_1| * |X_3| &= 6, |U(X_1, X_3)| = 6, Sep(X_1, X_3) = 1 \\ |X_2| * |X_3| &= 6, |U(X_2, X_3)| = 4, Sep(X_2, X_3) = 4/6 = 0.666 \end{aligned}$$

In the first two cases, the subsets X_1 and X_2 are separated, whereas subsets X_2 and X_3 are not separated.

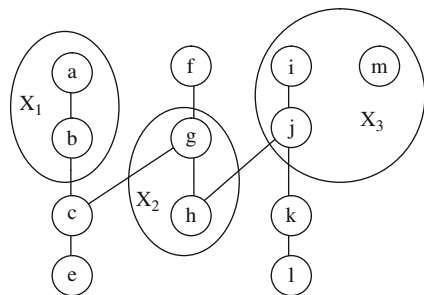


Fig. 5.8 Hasse diagram to demonstrate the calculation of separability

¹We sometimes simplify the notation by omitting the reference to IB' .

5.5 Data Matrix and Separation of Object Subsets in Partial Order

5.5.1 Motivation

So far we have discussed how to find separated subsets. These subsets are found by applying partially ordered object set and are not necessarily an expression of external classification. For example, by inspection of a Hasse diagram, two separated subsets may be identified which consist of both countries of Asia and Europe. Thus the interest is in properties of the data matrix that are responsible for this separation.

In this section, our focus is to find a best projection (question 1 in Section 5.4.1) and how we can find approximate solutions.

5.5.2 Antagonism

5.5.2.1 Concept

It should be possible to relate structural properties of the Hasse diagram, like the appearance of separated object subsets to properties related to the data matrix.

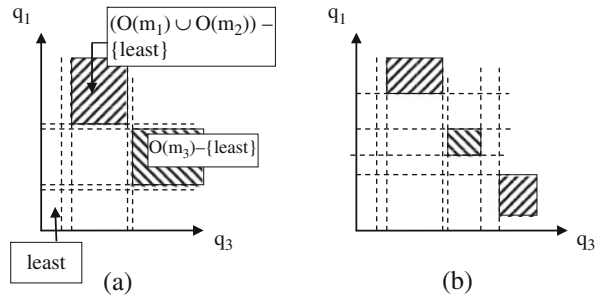
Let us consider $x, y \in X$ and $x || y$. The singletons $\{x\}$ and $\{y\}$ are the simplest example of separated object subsets. In case of $x || y$, there are two attributes q_i and q_j , $i \neq j$ such that $q_i(x) < q_i(y)$ and $q_j(x) > q_j(y)$. We say, the separation of x and y is due to q_i and q_j . Let us now consider two separated object subsets X_1 and X_2 with $|X_1|$ or $|X_2| > 1$, then it may be possible that not just one pair of attributes breaks all comparabilities simultaneously among the (unordered) pairs of $X_1 \times X_2$. Hence, we have to search for the smallest subset of attributes which simultaneously breaks all comparabilities of $(x, y) \in X_1 \times X_2$.

If IB' exists such that $x ||_{IB'} y$ for all $x \in X_1$ and all $y \in X_2$ with $X_1, X_2 \subset X$ and $\text{Sep}(X_1, X_2) = 1$ and $IB' \neq \emptyset$, $IB' \subseteq IB$, then we call IB' the set of antagonistic attributes/indicators and abbreviate it by $AIB(X_1, X_2)$ (antagonistic information base) and we often write AIB if there is no confusion possible (Simon, 2003; Simon et al., 2004a, b). AIB contains those attributes which are causing the separation of subsets X_1 and X_2 : While some attributes of AIB may have large values for objects of X_1 and small values for those of X_2 , some other attributes have low values for objects of X_1 and large ones for X_2 . The attributes of AIB separate X_1 and X_2 because they are “antagonistic.”

The smallest possible AIB is a pair $\{q_i, q_j\}$ such that for all $x \in X_1$ and all $y \in X_2$, we obtain $x || y$. (5.14)

1. This is the most desirable result of antagonism study because then a reasonable graphical display by a two-dimensional scatter plot may be possible. We also write that the attributes of AIB “explain” the separation of X_1 and X_2 . The search for AIB is a computational task and is a tool in the software WHASSE (Bruggemann et al., 1999), as well as in PyHasse (Bruggemann and Voigt, 2009).

Fig. 5.9 $|AIB| = 2$. (a) A scatter plot of X_1 and X_2 (Fig. 5.5); (b) a more complex pattern of two separated subsets



Example 1: Two attributes are sufficient to explain the separation of two subsets. We return to the Hasse diagram of Fig. 5.5 and select the subsets $X_1 = \{m_{34}, m_{33}, m_{32}, m_{31}, m_3\}$ and $X_2 = \{m_1, m_2, m_{12,1}, m_{12,2}, m_{12,3}\}$. We note that $Sep(X_1, X_2) = 1$. Indeed AIB contains only two attributes q_1 and q_3 , so we are able to construct a scatter plot (Fig. 5.9).

Figure 5.9 demonstrates the usefulness of the concept of antagonistic attributes: We see that q_1 has large values for X_2 and low values for X_1 , whereas q_3 has low values for X_2 but large values for X_1 , thus explaining the separation of the two subsets.

It may however be possible that we need more than two attributes to explain the separation of object subsets (Example 2), and it is possible that even with $|AIB| = 2$, the pattern of the separated subsets X_1 and X_2 is more complex (Fig. 5.9b).

Example 2 (real-life example):

Scientists of the Canadian Center of Inland Waters (CCIW) have developed a test battery (see Dutka et al., 1986). The responses of this test battery (our indicators) indicate the status of water samples or sediment samples with respect to their adverse impact on humans and on the environment.

The test battery includes as attributes (i) one fecal test (fecal coliforms, FC), (ii) two hygienic tests (test for coprostanol and coliforms *Escherichia coli*), CP and CH, (iii) one test for acute toxicity, MT (Microtox[®] test), and (iv) a genotoxicity test. Fifty sediment sites, labeled by numbers, were analyzed by applying this test battery. Figure 5.10 shows the Hasse diagram.

By inspection, we select two subsets X_1 and X_2 with $Sep(X_1, X_2) = 1$:

$$X_1 := \{5, 25, 27, 31, 95\} \text{ and } X_2 := \{7, 9, 18, 23, 32\}$$

How many and which attributes out of the five responses of the test battery explain that separation? Figure 5.11 shows how $Sep(X_1, X_2, IB_i)$ increases, depending on the number of attributes.

Figure 5.11 demonstrates that four out of five attributes are necessary to explain the separation of X_1 and X_2 . Therefore in Section 5.5.3, we pose the question: “if $|AIB| > 2$, then what?”

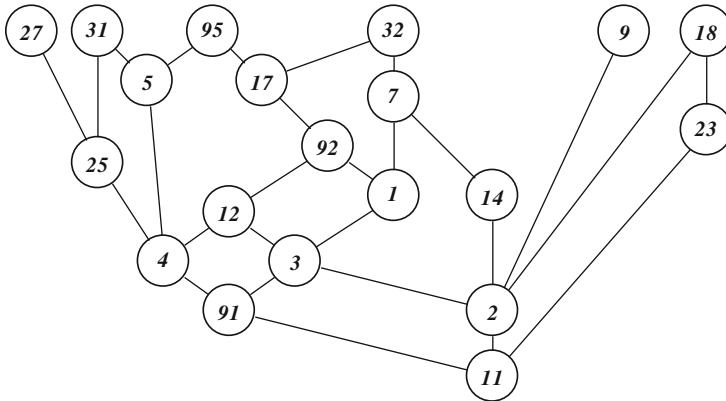
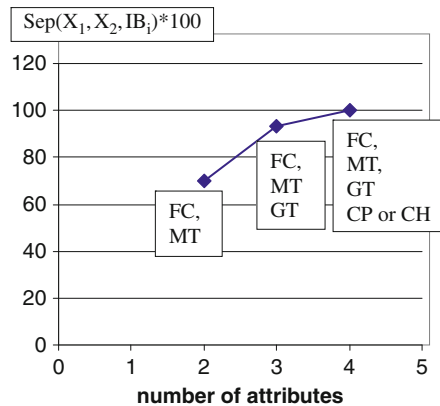


Fig. 5.10 Hasse diagram of sediment samples of Lake Ontario, based on a test battery

Fig. 5.11 Antagonism of attributes, Lake Ontario. Ordinate explained separation in % when IB_i is increasing, according to natt



5.5.3 If $|AIB| > 2$, Then What?

5.5.3.1 Motivation

In the case of $|AIB| = 2$, there is often a nice pictorial representation possible, like that shown in Fig. 5.9a. We recover diagrams of this kind several times in the application part of this monograph. However, if $|AIB| > 2$, then a 2D scatter plot allows only an insufficient view on the properties of the data matrix. Nevertheless, some general insights are possible even if $|AIB| > 2$: Let us assume that $|AIB| = 3$ and $1 > \text{Sep}(X_1, X_2, \{q_1, q_2\}) > 0.5$.

We see that $\{q_1, q_2\}$ does not completely explain the separation of X_1 and X_2 . However, the separability degree is large enough to assume that a scatter plot based on q_1 and q_2 is a good starting point. Obviously, some few object pairs (one object taken from X_1 and the other one from X_2) are only incomparable if a third attribute is introduced. So, one may find graphical techniques to indicate the role of the third attribute to break the remaining comparabilities. We will present several examples in

the application part (for example, we will construct a 3D scatter plot in the watershed case study, [Chapter 14](#)).

In the following, we will not display possible visualization techniques but demonstrate by an example that the appearance of separated subsets in the partial order implies some constraints on the attributes of the data matrix.

5.5.3.2 Structures in the Hasse Diagram Imply Constraints on the Data Matrix

Let us think of a scatter plot where the separation is not complete, like in Fig. 5.12.

There are the subset $X_2 = X_{20} \cup X_{21} \cup X_{22}$ and the subset X_1 . The subsets X_1 and X_{20} are large in comparison to X_{21} and X_{22} . X_1 and X_{20} alone would be completely separated by the attributes q_1 and q_2 . However, X_{21} and X_{22} contain objects which are comparable with some of X_1 , thus causing an incomplete separation of X_1 and X_2 . The third attribute q_3 has to break these comparabilities. A scatter plot (Fig. 5.12) will serve as an example.

The following observations are based on the assumption that a geometrical configuration as in Fig. 5.12 holds. In our experience, this kind of scatter plot is quite common. We define

$$\begin{aligned}
 X_i < X_j &: \Leftrightarrow, \text{ for all } x \in X_i, \text{ for all } y \in X_j : x < y \\
 X_i \parallel X_j &: \Leftrightarrow, \text{ for all } x \in X_i, \text{ for all } y \in X_j : x \parallel y
 \end{aligned}$$

To accomplish a complete separation by one and only one attribute q_3 , the attribute must necessarily lead to the following order relations:

$$X_{21} >_{q_3} X_1 \quad \text{and} \quad X_{22} <_{q_3} X_1 \tag{5.15}$$

Before we show how the structure of the Hasse diagram (existence of separated subsets) implies constraints on the data matrix, we need a compact notation:

$$\begin{aligned}
 q_3(X) &:= \{q_3(x), x \in X\} \text{ and } q_3(X_1) > q_3(X_2) : \Leftrightarrow, \text{ for all } x \in X_1 \text{ and} \\
 &\text{all } y \in X_2 : q_3(x) > q_3(y)
 \end{aligned}$$

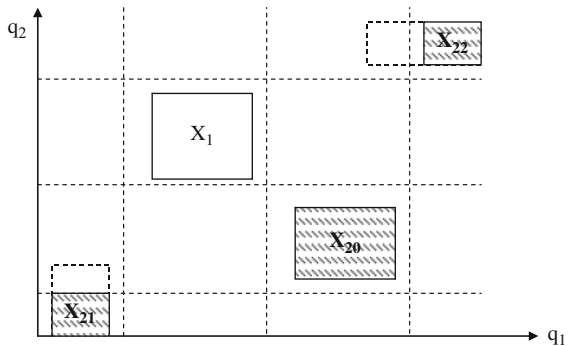
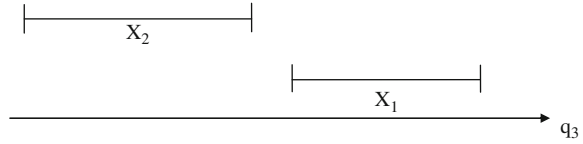


Fig. 5.12
 $X_2 = X_{20} \cup X_{21} \cup X_{22}$. X_1 is separated from X_{20} completely. There may be some overlap of X_{20} with X_{21} and X_{22} (indicated by broken lines)

Fig. 5.13 Presentation of the order relation with respect to q_3 on the line of real numbers: $q_3(X_2) < q_3(X_1)$



We can represent $q_3(X_1) > q_3(X_2)$ as closed intervals on the line of real numbers (Fig. 5.13).

We show that the assumption (a) $q_3(X_1) > q_3(X_{20})$ or (exclusively (b) $q_3(X_1) < q_3(X_{20})$) together with Eq. (5.15) leads to a contradiction of the assumption $|AIB| = 3$.

In the case of (a), we find $X_{21} >_{q_3} X_1, X_{21} <_{q_1} X_1, X_{22} <_{q_3} X_1, X_{22} >_{q_1} X_1, X_{20} <_{q_3} X_1$ and $X_{20} >_{q_1} X_1$. Hence $X_1 \parallel_{\{q_1, q_3\}} (X_{21} \cup X_{22} \cup X_{20})$.

Similarly in the case of (b), we find $X_1 \parallel_{\{q_2, q_3\}} (X_{21} \cup X_{22} \cup X_{20})$.

Assumptions (a) and (b) imply that only two attributes would explain the separation to 100% which contradicts $|AIB| = 3$.

Therefore

$$|AIB| = 3 \Rightarrow q_3(X_1) \cap q_3(X_{20}) \neq \emptyset \text{ or } q_3(X_{20}) \cap q_3(X_1) \neq \emptyset \quad (5.16)$$

Together with Eq. (5.15), we arrive at Fig. 5.14, which summarizes the result.

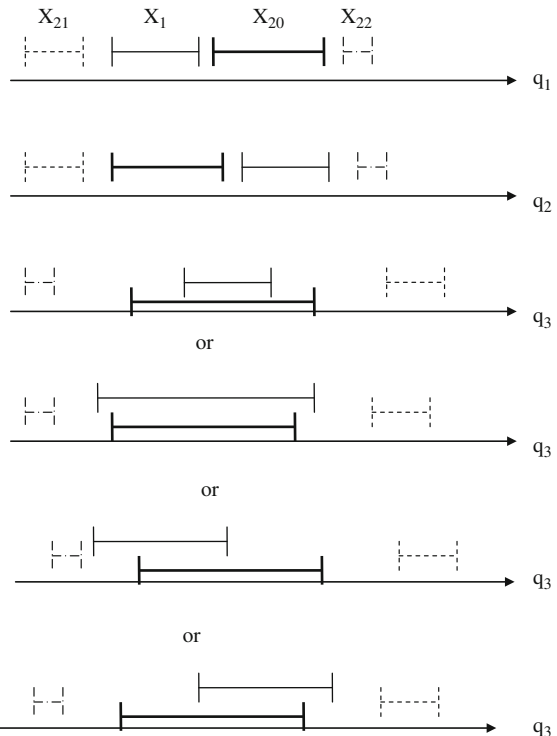


Fig. 5.14 Schematic representation of the intervals of the object sets if $|AIB| = 3$

Assuming the geometrical configuration such as in Fig. 5.12, we see that separated subsets with $|AIB| = 3$ imply that $q_3(X_{20})$ or $q_3(X_{21})$ must be within an interval, with an upper limit by the minimum value of $q_3(X_{21})$ and a lower limit by a maximum value of $q_3(X_{22})$, and that the intervals $q_3(X_{21})$ and $q_3(X_{20})$ must have a common intersection.

A 3D model within a real case study can be seen in Chapter 14 (Fig. 14.3).

5.6 Dominance and Separability

5.6.1 Motivation

Let us think of a poset (X, IB) with many objects. Often there is an additional information available by which a partitioning of X is possible. For example, students in a class may be evaluated by their knowledge in different disciplines. The set of students can be partitioned by the regions from where they come. Is it possible to rank the regions on the basis of the order relations among the students? This question will normally be answered by an appropriate aggregation by which attribute values of students of a certain region are transformed for the corresponding region (by forming means, or medians or adding up, etc.). Here we outline that a procedure is available which does not need to define an aggregation function to perform the transition from the microscale (the students and their evaluations in different disciplines) to a macroscale (regions and their evaluation with respect to different disciplines).

5.6.2 Concept

For any two subsets $X_1, X_2 \subset X/\cong$, $X_1 \cap X_2 = \emptyset$, $\text{Sep}(X_1, X_2, IB)$ may be between 0 and 1. We define

$$\begin{aligned} \text{Dom}(X_1, X_2) &:= |\{(x, y) \in X_1 * X_2, x \geq y\}| / (|X_1| * |X_2|) \\ \text{Dom}(X_2, X_1) &:= |\{(x, y) \in X_1 * X_2, x \leq y\}| / (|X_1| * |X_2|) \end{aligned} \quad (5.17)$$

and

$$\text{Sep}(X_1, X_2) := |U(X_1, X_2)| / (|X_1| * |X_2|) = |\{(x, y) \in X_1 * X_2, x \parallel y\}| / (|X_1| * |X_2|)$$

then

$$\begin{aligned} \text{Dom}(X_1, X_2) + \text{Dom}(X_2, X_1) + \text{Sep}(X_1, X_2) &= 1 \\ \text{Dom}(X_1, X_2) \neq \text{Dom}(X_2, X_1), \text{Dom}(X_i, X_j) &\in [0, 1] \end{aligned} \quad (5.18)$$

We speak of X_i dominates X_j to the degree $\text{Dom}(X_i, X_j)$. The dominance relation can be represented as a directed graph (digraph) as follows: (i) each subset X_i is

drawn as a vertex labeled with i , (ii) vertices are connected by a directed edge (i, j) if $\text{Dom}(X_i, X_j) > 0$, and (iii) the directed edges are weighted by $\text{Dom}(X_i, X_j)$ and pointing from vertex i to vertex j . The directed and weighted graph – a network – can be transferred into a simple digraph as follows:

If $\text{Dom}(X_i, X_j) > \varepsilon$, then (i, j) are connected by an edge starting from i and pointing to j . If $\text{Dom}(X_i, X_j) \leq \varepsilon$, then there is no connection from vertex i to vertex j . For an example, see Section 5.6.4.

5.6.3 Is the Dominance Relation a Partial Order?

Is the $\text{Dom}(X_1, X_2) > 0$ and $\text{Dom}(X_2, X_3) > 0$ sufficient to call dominance relations among subsets a partial order among subsets? The following example (Fig. 5.15) shows that a dominance relation is not necessarily transitive.

The number of elements in each of the three subsets is $|X_1| = 3$, $|X_2| = 6$, and $|X_3| = 3$.

$\text{Dom}(X_1, X_2) = 9/18$, $\text{Dom}(X_2, X_3) = 9/18$; however, $\text{Dom}(X_1, X_3) = 0$.

Restrepo and Bruggemann (2008) show that, for $\varepsilon \geq 0.5$, the digraph is a partial order.

5.6.4 Illustrative Example

In Fig. 5.16, a Hasse diagram is shown. Furthermore, three sets are defined by encircling objects: $X_1 = \{h, f, b\}$, $X_2 = \{e, a\}$, and $X_3 = \{d, g, c\}$.

By counting one finds

$$\begin{aligned} \text{Dom}(X_1, X_2) &= 5/6, \text{Dom}(X_2, X_1) = 0, \text{Sep}(X_1, X_2) = 1/6 \\ \text{Dom}(X_2, X_3) &= 0, \text{Dom}(X_3, X_2) = 3/6, \text{Sep}(X_2, X_3) = 3/6 \\ \text{Dom}(X_1, X_3) &= 1/9, \text{Dom}(X_3, X_1) = 0, \text{Sep}(X_1, X_3) = 8/9 \end{aligned}$$

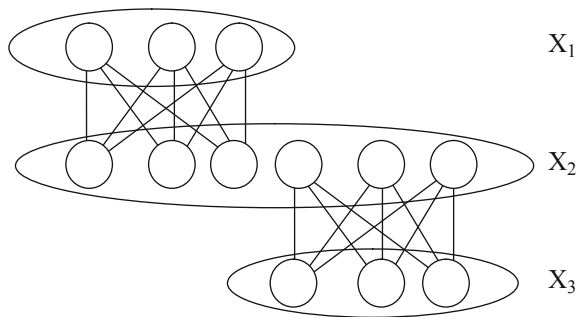


Fig. 5.15 Three subsets X_1 , X_2 , X_3 mutually disjoint

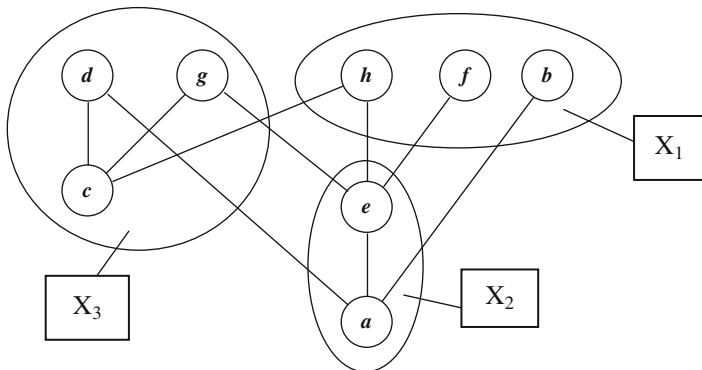


Fig. 5.16 Dominance of subsets of X due to the order relations of their elements

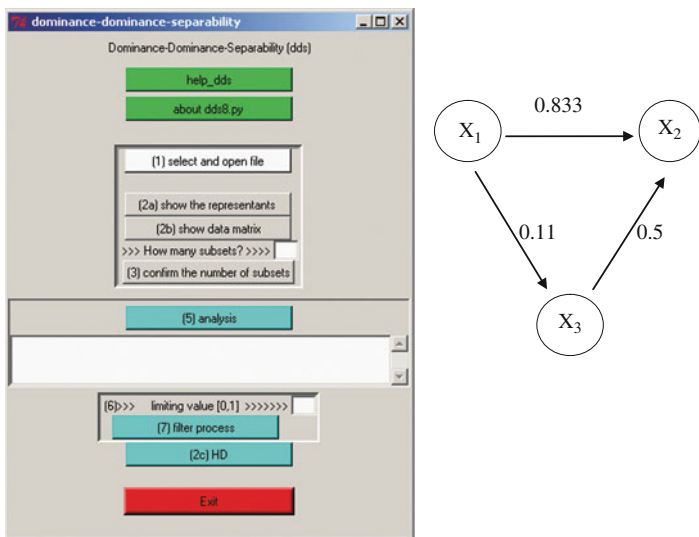


Fig. 5.17 (LHS) Graphical user interface of PyHase module dds8.py and the weighted directed graph (RHS) corresponding to Fig. 5.16

We apply the PyHase program “dds8.py” (see Chapter 17). Figure 5.17 shows its graphical user interface (LHS). After the user input of ϵ , the program dds8.py provides the corresponding directed graph (RHS).

From the directed graph (also called a dominance diagram), we can derive the dominance sequence: $X_1 > X_3 > X_2$.

The articulation point search or depth-first search for graph theoretical components of the Hasse diagram may be helpful. Whenever promising subsets are found, their separation can be assessed by examining the separability. Once separated subsets are found, we can identify the corresponding smallest “antagonistic indicator base,” AIB, capable of explaining their separation. We provide an example in which

$|AIB| > 2$ has implications on the data matrix. If a partition of the object set X is available by external knowledge, we can calculate not only the separability but also dominance. Instead of searching for order relations referring to objects, we can scale up and search for relations among the subsets of the partition. Introduction of a threshold ε leads to a digraph in which subsets relate to each other. This digraph is a partial order if $\text{Dom}(X_i, X_j) > \varepsilon \geq 0.5$.

5.7 Summary and Commentary

Partial orders can be very complex; therefore we need different tools to perform an adequate analysis. The shape of the Hasse diagrams and the analysis of the incompatibilities per level allow an overview about the data matrix with respect to the order relation it is inducing. It turns out that the concept of level is very useful. Another tool is provided by down sets or up sets because they allow some insights into the data profiles. An “ideal object” may just be found by an appropriate application of Eq. (5.9).

By the visualization of partial order by Hasse diagrams, the concept of a structure of a partial order was motivated. The vague concept of structure of posets can be sharpened by the concept of separated subsets. In general it is not easy to find separated subsets.

References

- Bruggemann, R., Pudenz, S., Voigt, K., Kaune, A. and Kreimes, K. (1999). An algebraic/graphical tool to compare ecosystems with respect to their pollution. IV: Comparative regional analysis by Boolean arithmetics. *Chemosphere*, 38(10), 2263–2279.
- Bruggemann, R. and Voigt, K. (2009). Analysis of partial orders in environmental systems applying the new software PyHasse. In W. Jochen and F. Michael (Eds.), *Simulation in Umwelt- und Geowissenschaften, Workshop Potsdam 2009* (pp. 43–55). Aachen: Shaker Verlag.
- Dutka, B.J., Walsh, K., Kwan, K.K., El Shaarawi, A., Liu, D.L. and Thompson, K. (1986). Priority site selection from degraded areas based on microbial and toxicant screening tests. *Water Pollut. Res. J. Can.*, 21(2), 267–282.
- Myers, W.L. and Patil, G.P. (2008). Semi-subordination sequences in multi-measure prioritization problems. Chapter 7. In M. Pavan and R. Todeschini (Eds.), *Ranking methods: Theory and applications, Data Handling in Science and Technology* (pp. 159–168, Vol. 27). Amsterdam, The Netherlands: Elsevier.
- Restrepo, G. and Bruggemann, R. (2008). Dominance and separability in posets, their application to isoelectronic species with equal charge. *J. Math. Chem.*, 44, 577–602.
- Simon, U. (2003). *Multikriterielle bewertung von wasserwirtschaftlichen Maßnahmen aus gewässeroekologischer Sicht – Beispiel Berlin*. Berlin: Tenea.
- Simon, U., Bruggemann, R. and Pudenz, S. (2004a). Aspects of decision support in water management – Example Berlin and Potsdam (Germany) I – Spatially differentiated evaluation. *Water Res.*, 38, 1809–1816.
- Simon, U., Bruggemann, R. and Pudenz, S. (2004b). Aspects of decision support in water management – Example Berlin and Potsdam (Germany) II – Improvement of management strategies. *Water Res.*, 38, 4085–4092.