Chapter 10
Comparison of Partial, Linear, and Weak Orders

10.1 Introduction

There are two reasons to compare different orders:

1. From (X, IB), a linear or weak order, called O_{poset}, can be derived, applying methods rendered in Chapter 9. By calculation of an index (Chapter 7), another linear or weak order can be obtained, which we call O_{Γ}. As there are the same set of indicators, IB, and the same set of objects, X, O_{poset}, and O_{Γ} should ideally be coincident. As can be suspected, this is not necessarily the case and we need measures to quantify the degree of coincidence between O_{poset} and O_{Γ}.
2. There are different indicator sets IB_1, IB_2 inducing partial orders (X, IB_1) and (X, IB_2). Will any comparability $x < y$ in one order be reproduced in another? If not, how can we compare these two partial orders?

In the following, we discuss five concepts which may be useful for comparisons.

10.2 Representation of O_{poset} and O_{Γ} by a Hasse Diagram

10.2.1 Method

Both orders O_{poset} and O_{Index} will be considered as indicators I_{poset} and I_{Γ}, with which we can order the objects. From I_{poset} and $\underline{I}_{\Gamma}$, a new partial order $(X, \{I_{\text{poset}}, I_{\Gamma}\})$ can be found.

A Hasse diagram of $(X, \{I_{\text{poset}}, I_{\Gamma}\})$ may show us the following:

1. Chains of $(X, \{I_{\text{poset}}, I_{\Gamma}\})$: Subsets of objects with coincident mutual orders.
2. Antichains of $(X, \{I_{\text{poset}}, I_{\Gamma}\})$: Pairs of objects for which object rank inversions are found.

R. Brüggemann, G.P. Patil, *Ranking and Prioritization for Multi-indicator Systems*, Environmental and Ecological Statistics 5, DOI 10.1007/978-1-4419-8477-7_10,

3. A linear order (see Chapter 2): The two orders O_{poset} and O_Γ coincide completely. The number of incomparabilities U will be 0.
4. A complete antichain: O_{poset} is the dual (Chapter 3) of O_Γ. $U = n^*(n-1)/2$, with n being the number of objects.

Therefore, the degree of coincidence d_{coin1} can be quantified by the following equation:

$$d_{\text{coin1}} := 1 - \frac{|\,U(X, \{I_{\text{poset}}, I_\Gamma\})\,|}{\binom{n}{2}} \tag{10.1a}$$

Instead of counting pairwise incomparability (Eq. (1.10)), we can measure the degree of coincidence by the extent of pairwise incomparability. We sum objectwise the absolute difference between ranks of each object due to I_{poset} and I_Γ and normalize this sum by its maximum attainable value of $1 + 2 + \ldots + n - 1 = n^*(n-1)/2 = \binom{n}{2}$. Thus we obtain another coincidence measure d_{coin2}:

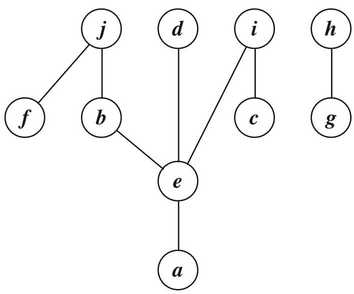
$$d_{\text{coin2}} := 1 - \frac{\sum_{x \in X} |\,\text{rank}(x, I_{\text{poset}}) - \text{rank}(x, I_\Gamma)\,|}{\binom{n}{2}} \tag{10.1b}$$

10.2.2 Illustrative Example

Figure 10.1 shows the Hasse diagram of 10 objects, three attributes, and the data matrix.

The order O_{poset} is obtained from the set of linear extensions and from them the averaged height, hav.

The order O_Γ is obtained by assuming the weight vector (0.5, 0.3, 0.2).



	q_1	q_2	q_3
a	1	2	3
b	3	2	5
c	2	5	2
d	3	3	4
e	1	2	4
f	4	2	2
g	6	1	1
h	6	1	3
i	2	6	5
j	4	2	6

Fig. 10.1 Hasse diagram of $(X, \text{IB}) = (X, \{q_1, q_2, q_3\})$ for demonstrating the calculation of d_{coin1}

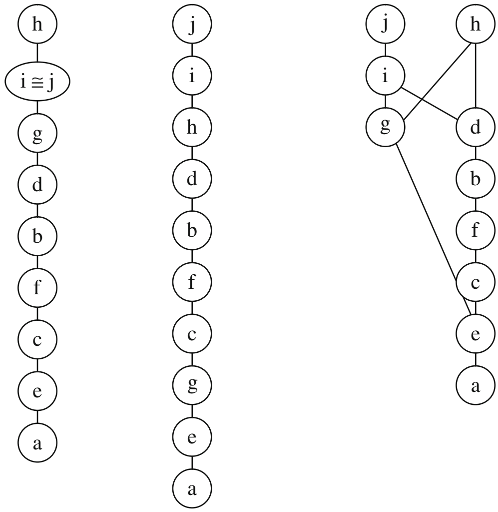


Fig. 10.2 (*LHS*) $(X, \{I_\Gamma\})$, (*middle*) $(X, \{I_{\text{poset}}\})$, (*RHS*) $(X, \{I_{\text{poset}}, I_\Gamma\})$

Both orders $(X, \{I_{\text{poset}}\})$ and $(X, \{I_\Gamma\})$ as well as $(X, \{I_{\text{poset}}, I_\Gamma\})$ are depicted in Fig. 10.2.

$U(X, \{I_{\text{poset}}, I_\Gamma\}) = 6$, hence $d_{\text{coin1}} = 1 - 0.1333 = 0.8666$, whereas $d_{\text{coin2}} = 1 - 0.27 = 0.73$. We see that both orders coincide pretty well. By looking at the Hasse diagram of $(X, \{I_{\text{poset}}, I_\Gamma\})$, we see that

- object g is incomparable to the objects d, b, f, c and
- object h is incomparable to the objects i, j.

What may be the reason for this? Looking back at Fig. 10.1, we see that the two objects g and h are forming own component in (X, IB). Therefore, g and h will contribute to be in $U(X, \{I_{\text{poset}}, I_\Gamma\})$ and be most prone to influences of the weight vector.

10.3 Spearman Correlation Analysis

10.3.1 Method

Spearman correlation applies if we have two linear or weak orders and we would like to know how far they are correlated. The correlation index tells us how much the two linear or weak orders coincide. So far as any single ranking position is considered, the Spearman correlation coefficient is pretty sensitive to rank inversion. We call ranks of the order 1 $R_1(i)$ and those of the order 2 $R_2(i)$, where $i = 1, \ldots, n$, with number of objects n.

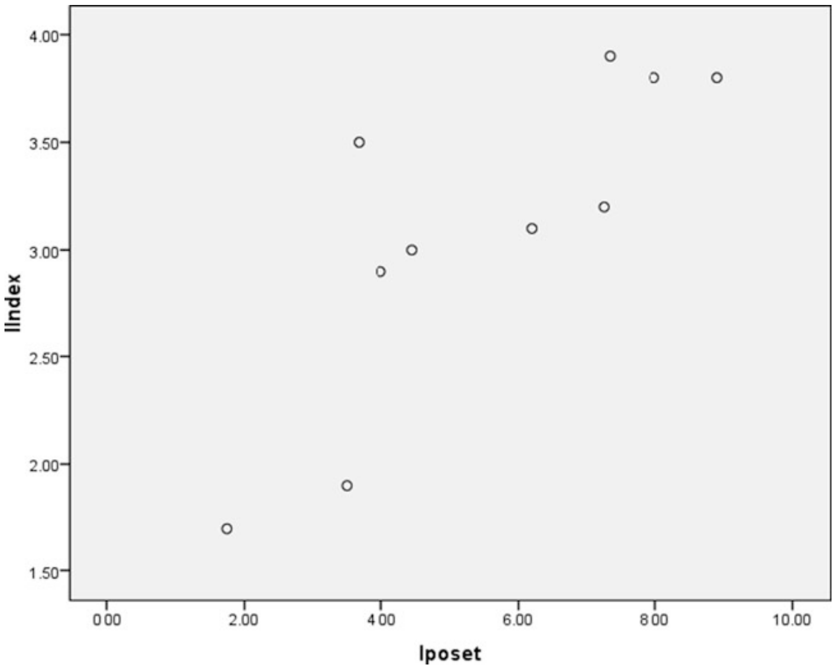


Fig. 10.3 Scatter plot of $I_\Gamma (= I_{\text{Index}})$ and I_{poset} of example in Section 10.2

Writing $d(i) := R_1(i) - R_2(i)$, the Spearman correlation is given by

$$r_s = 1 - \frac{6 \cdot \sum_{i=1}^{n} (d(i))^2}{n \cdot (n^2 - 1)} \tag{10.2}$$

$$r_s \in [-1, 1].$$

10.3.2 Illustrative Example

As order 1 we apply I_{poset}, and as order 2, I_Γ of the example in Section 10.2. The Spearman correlation coefficient is $r_s = 0.839$, indicating a good correlation, hence a good coincidence of both orders. In Fig. 10.3, the scatter plot of I_Γ (ordinate axis) vs I_{poset} (abscissa) is shown.

10.4 Concordance Analysis

10.4.1 Method

The concordance analysis is a robust and intuitive tool to compare one weak or linear order with another one. For example, using an index Γ, we may get five

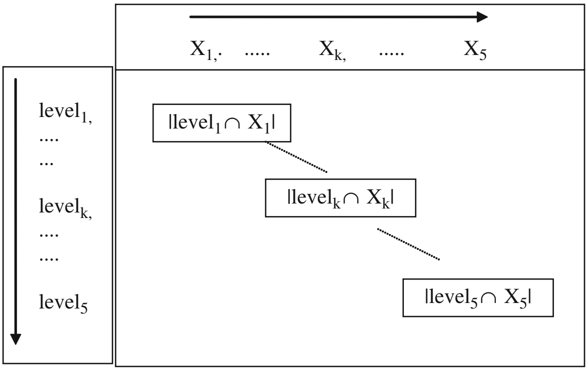


Fig. 10.4 Principle of concordance analysis

subsets $X_1, X_2, \ldots, X_5, X_i \subset X, X_i \cap X_j = \emptyset$, for $j \neq i$, which can be interpreted as "very bad," "bad," "medium," "good," "very good" and by the partial order, five levels (see Chapter 9) level_1 ("very bad"), level_2 ("bad"), ..., level_5 ("very good") are provided and we want to compare the level sets with the subsets. Patil (2005) introduces this intuitive technique to perform this comparison (Fig. 10.4). The arrows in Fig. 10.4 indicate the orientation (from very bad to very good) and we count the elements occurring in the main diagonal. In order to get a measure, we normalize this number by dividing by the number of objects.

The basic idea can be represented by a square $d \times d$ matrix cd (see Fig. 10.4) as follows: Let

$$P_{ki} = \{x \in X : \text{ height } i \text{ in } k\text{th order}\}, \text{ then } cd_{ij} = \left|P_{1\text{i}} \cap P_{2\text{j}}\right| \tag{10.3}$$

As motivated by Fig. 10.4, the two orders coincide more, as the entries in the diagonal of the concordance matrix are more. However, we accept some slight deviations as concordant: Therefore, the two diagonals neighboring the main diagonal are also taken into account, however with a lower weight κ (0.5 is recommended). Hence we arrive at

$$\text{con} = \left(\sum cd_{ii} + \kappa^* \sum cd_{|i-j|=1}\right) / n, n \text{ being } |\text{X}| \tag{10.4}$$

The expression $cd_{|i-j|=1}$ counts for the two diagonals neighboring the main diagonal and n is the number of objects. Patil (2005) finds by simulation studies a threshold value T.

$$T = 1/d, d \text{ being the number of subsets } X_i \text{ (levels)} \tag{10.5}$$

We speak of concordance between two rankings if

$$\text{con} > \text{T} \tag{10.6}$$

A fictitious example follows; real-life examples will be found in the case study chapters.

10.4.2 Illustrative Example

In Fig. 10.5, a Hasse diagram together with the data matrix is shown.

Assuming all weights equal, the index yields the weak order:

$$a < d < e < b \cong h < c < g \cong i < f$$

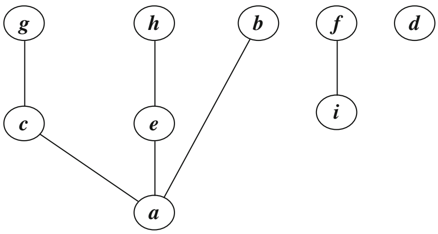
There are three levels (see Chapter 9):

$$\text{level}_1 = \{a\}, \text{level}_2 = \{c, e, i\}, \text{level}_3 = \{g, h, b, f, d\}$$

Correspondingly, we form three disjoint subsets:

$$X_1 = \{a, d, e\}, X_2 = \{b, h, c\}, X_3 = \{g, i, f\}$$

Corresponding to Fig. 10.4, we obtain Table 10.1.



	q_1	q_2	q_3
a	1	2	3
b	3	2	5
c	4	6	3
d	1	7	0
e	2	3	4
f	0	8	9
g	6	6	3
h	2	3	5
i	0	7	8

Fig. 10.5 Illustrative Hasse diagram and the data matrix

Table 10.1 Intersections (top) and concordance matrix *cd* (bottom)

	$\{a, d, e\}$	$\{b, c, h\}$	$\{f, g, i\}$
$\{a\}$	$\{a\}$	Ø	Ø
$\{c, e, i\}$	$\{e\}$	$\{c\}$	$\{i\}$
$\{b, d, f, g, h\}$	$\{d\}$	$\{b, h\}$	$\{f, g\}$

	$\{a, d, e\}$	$\{b, c, h\}$	$\{f, g, i\}$
$\{a\}$	1	0	0
$\{c, e, i\}$	1	1	1
$\{b, d, f, g, h\}$	1	2	2

10.5 Intersection of Partial Orders

10.5.1 Motivation

Often two posets arise from different sets of indicators and it is of interest to relate the two posets. One way to do this is to combine both indicator sets and to get an extended information base $= \mathrm{IB}_1 \cup \mathrm{IB}_2$, then see what comparabilities remain in $(X,\ \mathrm{IB}_1 \cup \mathrm{IB}_2)$.

10.5.2 Method

In $(X,\ \mathrm{IB}_1 \cup \mathrm{IB}_2)$ only those comparabilities occur, which are present at the same time in (X, IB_1) and (X, IB_2). In other words

$$(X, \mathrm{IB}_1 \cup \mathrm{IB}_2) = (X, \mathrm{IB}_1) \cap (X, \mathrm{IB}_2) \tag{10.6b}$$

The two partial orders coincide more as larger the intersection in Eq. (10.6b) is.

The evaluation of the intersection operation in Eq. (10.6b) is best done if the representation of posets by ordered sets is applied (Section 2.5): If $x < y$ in (X, IB), then $(x, y) \in \mathbf{X} = (X, \mathrm{IB})$.

As in Chapter 7, V is the number of comparabilities. We call V_1 the number of comparabilities of (X, IB_1), V_2 of (X, IB_2), and V_{12} of $(X, \mathrm{IB}_1 \cup \mathrm{IB}_2)$.

Then obviously

$$V_{12} \leq \min(V_1, V_2) \tag{10.7}$$

If $(X, \mathrm{IB}_1) \subset (X, \mathrm{IB}_2)$, then $V_{12} = V_1$. Useful results by intersection can be expected if only one of the posets is included in the other one.

10.5.3 Illustrative Example

Consider the three Hasse diagrams in Fig. 10.6.

As in Chapter 2, we disregard the pairs which are in the diagonal of X^2:

$$(X, \mathrm{IB}_1) = \{(a,b),(a,d),(c,b),(c,d),(b,d)\}$$
$$(X, \mathrm{IB}_2) = \{(a,d),(c,b),(c,d),(b,d)\}$$
$$(X, \mathrm{IB}_3) = \{(c,b),(c,a),(c,d),(b,d),(b,a)\}$$

Inclusions:

$$(X, \mathrm{IB}_2) \subset (X, \mathrm{IB}_1) \text{ and } V_{12} = V_2 = 4$$

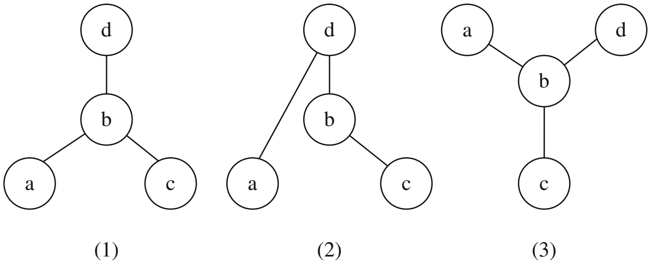


Fig. 10.6 Three visualizations: (*1*) (X, IB_1), (*2*) (X, IB_2), (*3*) (X, IB_3)

The ordered pair (b, a) is neither in (X, IB_1) nor in (X, IB_2). Furthermore, the ordered pair (a, d) which is in (X, IB_1) and (X, IB_2) is not in (X, IB_3). Therefore, no other $\subset$ relation can be found among the three posets.

Intersections:

$(X, IB_1) \cap (X, IB_2) = \{(a, d), (c, b), (c, d), (b, d)\} = (X, IB_2), V_{12} = 4$ as it must be.

$(X, IB_1) \cap (X, IB_3) = \{(c, b), (c, d), (b, d)\}, V_{13} = 3$, a chain $c < b < d$ results, object a is isolated.

$(X, IB_2) \cap (X, IB_3) = \{(c, b), (c, d), (b, d)\}, V_{23} = 3$, a chain $c < b < d$ results, object a is isolated.

10.6 Comparison of Two Partial Orders as a Multivariate Problem

10.6.1 Motivation

Here we want to count what is different between any two pairs (x, y) obtained from the one and the other partial order. We aim to visualize these counts in a histogram-like diagram.

10.6.2 Method

Let us take two elements x, $y \in X$, then the following constellations appear while comparing two empirical posets (Table 10.2).

From Table 10.2 we see that 4*4 different constellations are possible: (1a) with (2a), (1a) with (2b), etc.

Table 10.2 Possible constellations when two partial orders are compared, taken $x, y \in X$

Identifier	$x, y \in X$ in (X, IB_1)	Identifier	$x, y \in X$ in (X, IB_2)
1*a*	$x < y$	2*a*	$x < y$
1*b*	$x > y$	2*b*	$x > y$
1*c*	$x \parallel y$	2*c*	$x \parallel y$
1*d*	$x \cong y$	2*d*	$x \cong y$

We call any constellation out of the 16 situations a matching, m_i, $i = 1, \ldots 16$, and denote it by the corresponding symbols out of the list $\{<, >, \parallel, \cong\}$. (10.8)

For example, the constellation (1*a*), (2*c*) : $(<, \parallel)$.

As any pair $(x, y) \in X^2$ has exactly one of the 16 matchings, we can also write $m(x, y)$ in order to describe which concrete matching appears for pair x, y by comparing two partial orders. Only two matchings, namely (<, <) and (>, >), will contribute to the partial order $(X, \text{IB}_1 \cap \text{IB}_2)$. The information about 14 other matchings will be lost. Therefore, the theoretical idea is to count the frequencies of matchings as we browse through the object set:

$$F(\text{matching } i) = \text{count of } m_i, \text{ for all } (x, y) \in X^2 \tag{10.9}$$

The evaluation of Eq. (10.9) can best be explained as in Table 10.3.

Counts of some matchings like (<, <) and (>, >) as well as (>, <) and (<, >) separately are not meaningful if we have a comparison in mind. Therefore, instead of taking care of all 16 matchings, we group them in "behavior classes," $B_1, \ldots, B_5$, as follows (Fig. 10.7).

We describe now the degree of coincidence of two partial orders by performing the following steps (Eq. (10.9)):

1. Count the m_i by checking a matrix like that in Table 10.3.
2. Add up those m_i which belong to behavior class B_1.
3. Repeat step 2 for B_2, B_3, B_4, and B_5.
4. Denote the frequency of B_i by $F(B_i)$.
5. Normalize $F(B_i)$ by dividing by $n^*(n-1)$ and call the resulting number $f(B_i)$.

Table 10.3 Counting matchings

Objects	Objects				
	1	. . .	*k*	. . .	*n*
1	$m(1, 1)$	. . .	$m(1, k)$	. . .	$m(1, n)$
. . .	....	. . .	. . .	. . .	...
k	$m(k, 1)$		–		$m(k, n)$
. . .	...	. . .	. . .	. . .	...
n	$m(n, 1)$	. . .	$m(n, k)$	. . .	$m(n, n)$

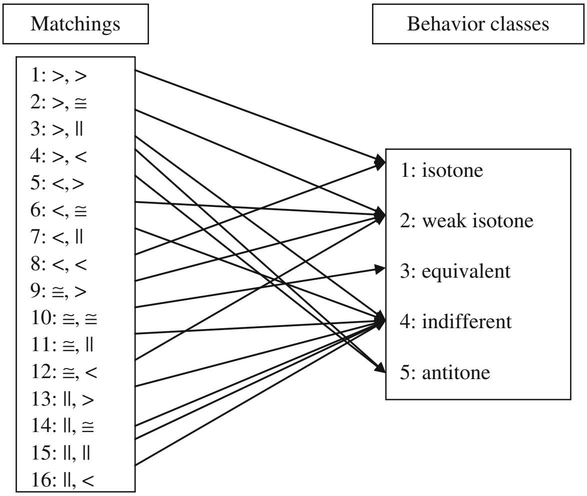


Fig. 10.7 Assignment of matchings m_i to behavior classes B_i

There are the following typical cases:

- $(X, \mathrm{IB}_1) \subset (X, \mathrm{IB}_2)$: $F(B_5) = 0, F(B_4) \neq 0, F(B_1) = 2^* V_1$ "good coincidence"
- $(X, \mathrm{IB}_2) \subset (X, \mathrm{IB}_1)$: $F(B_5) = 0, F(B_4) \neq 0, F(B_2) = 2^* V_2$ "good coincidence"
- $|(X, \mathrm{IB}_1) \cap (X, \mathrm{IB}_2)| = V_{12} << \min(V_1, V_2)$ (Eq. (10.7)): $F(B_5) \neq 0, F(B_4) \neq 0, F(B_1) = 2^* V_{12}$ "medium to poor coincidence"
- $(X, \mathrm{IB}_1) =$ dual of (X, IB_2) (see Chapter 4): $F(B_5) = 2^* V_1 = 2^* V_2, F(B_4) \neq 0, F(B_1) = 0$ "no coincidence, countercurrent behavior"

In order to describe the behavior of two partial orders in a compact way, we use the wording:

- isotone: matchings (<, <) and (>, >)
- antitone: matchings (>, <) and (>, <)
- weak isotone: the following matchings: (<, ≅), (>, ≅), (≅, <), (≅, >)
- indifferent: all matchings where || is part of the pair
- equivalent: matching (≅, ≅)

It is convenient to present the comparison of two partial orders by a bar diagram of $f(B_i)$. This multivariate consideration of the comparison of partial orders is called "proximity analysis."

10.6.3 Illustrative Example

In Fig. 10.8 two Hasse diagrams (a) and (b) are shown which visually present two partial orders (X, IB_1) and (X, IB2). What is the result of the proximity analysis?

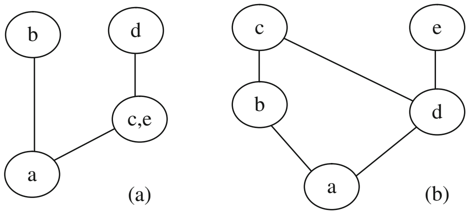


Fig. 10.8 Example for proximity analysis with two Hasse diagrams (a) and (b)

In Table 10.4 we fill out the cells by the matchings $m_1, \ldots, m_{16}$: In the cells first the result of the Hasse diagram (a) and then that of (b) are written. The matchings of the same type are counted and assigned to one of the five B_i classes, following the lines of Eq. (10.9).

The normalized frequencies $f(B_i)$ are presented as a bar diagram (Fig. 10.9).

Table 10.4 Comparison matrix among the five objects of the Hasse diagrams shown in Fig. 10.8 (only the upper triangle is shown)

	a	*b*	*c*	*d*	*e*
a	–	<, <	<, <	<, <	<, <
b		–	||, <	||, ||	||, ||
c			–	<, >	=, ||
d				–	>, <
e					–

B_1, isotone: 4; B_2, weak isotone: 0; B_3, equivalent: 0; B_4, indifferent: 4; B_5, antitone: 2

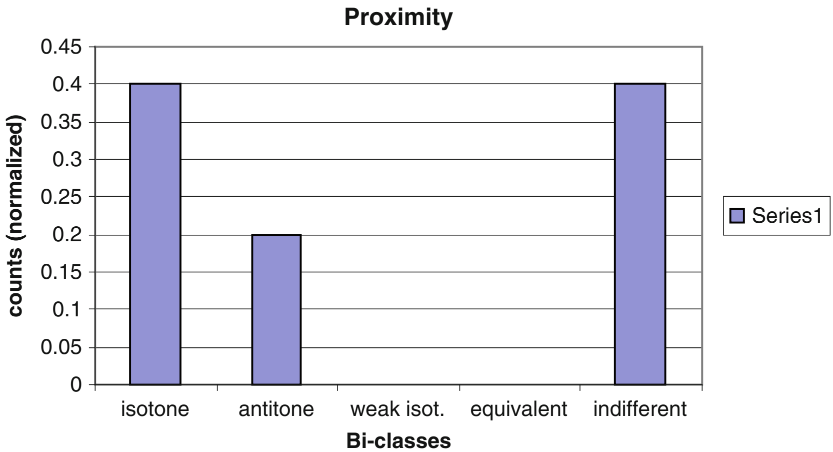


Fig. 10.9 Proximity analysis of the two Hasse diagrams of Fig. 10.8

How the normalized counts are combined to get a final scalar, expressing the proximity, depends on the application. For a real-life example, see Chapters 14 and 15.

10.7 Summary and Commentary

So, actually the tools at hand to compare orders (see Table 10.5) are the following:

1. Hasse diagram of two linear or weak orders, d_{coinc}
2. Spearman correlation analysis, r_{s}
3. Concordance analysis, con
4. Proximity analysis of two partial orders

There is another way of comparing linear or weak orders which we do not mention so far but for which here is a good place.

Imagine that the empirical composite index is formulated by a set of weights, which we abbreviate as g(emp). This set of weights bears all the experiences of stakeholders and decision makers. Another set of weights called g(poset) may be derived such that an index with weights g(poset) induces the same order as that derived by methods explained in Chapter 9, i.e., by LPOM, Bubley Dyer, CRF, and lattice theoretical methods. If such a set of weights g(poset) can be derived, we can compare it with g(emp).

Table 10.5 Advantages and disadvantages of comparison tools

	Advantages	Disadvantages
Hasse diagram of two linear or weak orders	Coincidences and non-coincidences can be visualized. Objects which are incomparable to many others can be easily identified	Sensitive to any single ranking inversion
Spearman correlation analysis	Easily applicable. Software is generally available Test statistic is provided	Sensitive to any single ranking inversion
Concordance analysis	Intuitive concept Test statistic A PyHasse module is available	Need of finding the same number of subsets for both rankings. The parameter κ may influence the result
Proximity analysis	Detailed information about the matchings arising from two partial orders. A PyHasse module is available	There is no single number telling us about the proximity of two partial orders

Now we are confronted with four cases:

(1) Such a set of *g*(poset) does not exist, which draws serious attention to the data matrix (indicator selection, measurement errors, and rounding procedures).
(2) By a suitable distance measure, both sets of weights are sufficiently close to each other. Then the partial order method justifies the set of weights *g*(emp).
(3) The distance is large. Then a re-examination of the data matrix and the weighting procedure to obtain the composite indicator is recommended.

As a software-supported approach is still not available, we let this kind of comparison be open for future work.

Reference

Patil, G.P. (2005). Cross-disciplinary class room notes. Center for Statistical Ecology and Environmental Statistics, Penn State University.