

# *Multiple indicators, partially ordered sets, and linear extensions: Multi-criterion ranking and prioritization*

G.P. PATIL and C. TAILLIE

*Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Penn State University, University Park, PA 16802*

Received January 2002; Revised February 2003

---

This paper is concerned with the question of ranking a finite collection of objects when a suite of indicator values is available for each member of the collection. The objects can be represented as a cloud of points in indicator space, but the different indicators (coordinate axes) typically convey different comparative messages and there is no unique way to rank the objects while taking all indicators into account. A conventional solution is to assign a composite numerical score to each object by combining the indicator information in some fashion. Consciously or otherwise, every such composite involves judgments (often arbitrary or controversial) about tradeoffs or substitutability among indicators.

Rather than trying to combine indicators, we take the view that the relative positions in indicator space determine only a partial ordering and that a given pair of objects may not be inherently comparable. Working with Hasse diagrams of the partial order, we study the collection of all rankings that are compatible with the partial order (linear extensions). In this way, an interval of possible ranks is assigned to each object. The intervals can be very wide, however. Noting that ranks near the ends of each interval are usually infrequent under linear extensions, a probability distribution is obtained over the interval of possible ranks. This distribution, called the rank-frequency distribution, turns out to be unimodal (in fact, log-concave) and represents the degree of ambiguity involved in attempting to assign a rank to the corresponding object.

Stochastic ordering of probability distributions imposes a partial order on the collection of rank-frequency distributions. This collection of distributions is in one-to-one correspondence with the original collection of objects and the induced ordering on these objects is called the cumulative rank-frequency (CRF) ordering; it extends the original partial order. Although the CRF ordering need not be linear, it can be iterated to yield a fixed point of the CRF operator. We hypothesize that the fixed points of the CRF operator are exactly the linear orderings. The CRF operator treats each linear extension as an equal “voter” in determining the CRF ranking. It is possible to generalize to a weighted CRF operator by giving linear extensions differential weights either on mathematical grounds (e.g., number of jumps) or empirical grounds (e.g., indicator concordance). Explicit enumeration of all possible linear extensions is computationally impractical unless the number of objects is quite small. In such cases, the rank-frequencies can be estimated using discrete Markov chain Monte Carlo (MCMC) methods.

**Keywords:** CRF operator, Hasse diagram, MCMC, partial order, pre-order, poset, rank-interval, rank-frequency distribution, weighted distributions, zeta matrix

## 1. Introduction

This paper is concerned with the question of ranking a finite collection of objects when a suite of indicator values is available for each member of the collection. The objects can be represented as a cloud of points in indicator space, but the different indicators (coordinate axes) typically convey different comparative messages and there is no unique way to rank the objects while taking all indicators into account. A conventional solution is to assign a composite numerical score to each object by combining the indicator information in some fashion. Consciously or otherwise, every such composite involves judgments (often arbitrary or controversial) about tradeoffs or substitutability among indicators.

Rather than trying to combine indicators, we take the view that the relative positions in indicator space determine only a partial ordering and that a given pair of objects may not be inherently comparable. Working with Hasse diagrams of the partial order, we study the collection of all rankings that are compatible with the partial order (linear extensions). In this way, an interval of possible ranks is assigned to each object. The intervals can be very wide, however. Noting that ranks near the ends of each interval are usually infrequent under linear extensions, a probability distribution is obtained over the interval of possible ranks. This distribution, called the rank-frequency distribution, turns out to be unimodal (in fact, log-concave) and represents the degree of ambiguity involved in attempting to assign a rank to the corresponding object.

Stochastic ordering of probability distributions imposes a partial order on the collection of rank-frequency distributions. This collection of distributions is in one-to-one correspondence with the original collection of objects and the induced ordering on these objects is called the cumulative rank-frequency (CRF) ordering; it extends the original partial order. Although the CRF ordering need not be linear, it can be iterated to yield a fixed point of the CRF operator. We hypothesize that the fixed points of the CRF operator are exactly the linear orderings. The CRF operator treats each linear extension as an equal “voter” in determining the CRF ranking. It is possible to generalize to a weighted CRF operator by giving linear extensions differential weights either on mathematical grounds (e.g., number of jumps) or empirical grounds (e.g., indicator concordance). Explicit enumeration of all possible linear extensions is computationally impractical unless the number of objects is quite small. In such cases, the rank-frequencies can be estimated using discrete Markov chain Monte Carlo (MCMC) methods.

## 2. Combining multiple indicators

Most scientific concepts are multi-faceted and can be quantified in a variety of ways. Even such an everyday notion as a person’s “size” can be assessed by “height,” by “weight,” by “girth,” by “arm-length,” etc. The different ways of quantifying a single underlying concept will be referred to as views or indicators. While there is generally a positive association among different views, the association is not perfect and different indicators can provide different comparative assessments. Although in many ways these views are neither comparable nor combinable, it remains a strong and almost irresistible human urge to combine them into a single view and a corresponding linear ordering of the objects under consideration.

In this section, we examine the issues, challenges, and difficulties encountered in trying to combine multiple indicators into a single index. We consider a collection  $S$  of objects where each object has an associated suite,  $(I_1, I_2, \dots, I_p)$ , of real-valued indicators. We suppose that all indicators are consistently oriented so that small values indicate “poor” conditions and large values indicate “good” conditions.

The elements in  $S$  will be denoted by  $a, b, c, \dots$ . We would like to make comparative statements about two given objects  $a$  and  $a'$  based on their indicator values  $(I_1, I_2, \dots, I_p)$  and  $(I'_1, I'_2, \dots, I'_p)$ , respectively. If it happens that  $I'_j \geq I_j$  for all  $j$ , then we say that  $a'$  is intrinsically “better” or “bigger” than  $a$  (in the loose sense) and we write

$$a' \geq a \quad \text{or} \quad a \leq a'.$$

When, on the other hand, the indicators are not unanimous in comparing  $a$  and  $a'$ , we have an ambiguous situation in which different investigators might rank  $a$  and  $a'$  differently. Here there is no consensus ranking. The possibilities are indicated in Fig. 1 in the case of  $p = 2$  indicators. Object  $a$  divides indicator space into four quadrants. Objects  $a'$  falling in the first quadrant (including its boundary) are intrinsically better than  $a$  and those falling in the third quadrant are intrinsically worse than  $a$ . The second and fourth quadrants (excluding their boundaries) are regions of ambiguity; objects falling here are not intrinsically comparable with  $a$ . When  $p > 2$ , indicator space is divided into  $2^p$  “hyper-quadrants” of which  $2^p - 2$  are regions of ambiguity.

Resolution of ambiguity can be accomplished (mathematically) by combining the indicators into an index:

$$\text{index} = H(I_1, I_2, \dots, I_p) = H(a).$$

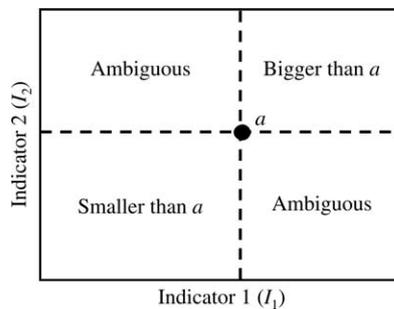
We will use letters like  $H$  and  $G$  to denote such combinations. The simplest combination is linear,

$$H = w_1 I_1 + w_2 I_2 + \dots + w_p I_p.$$

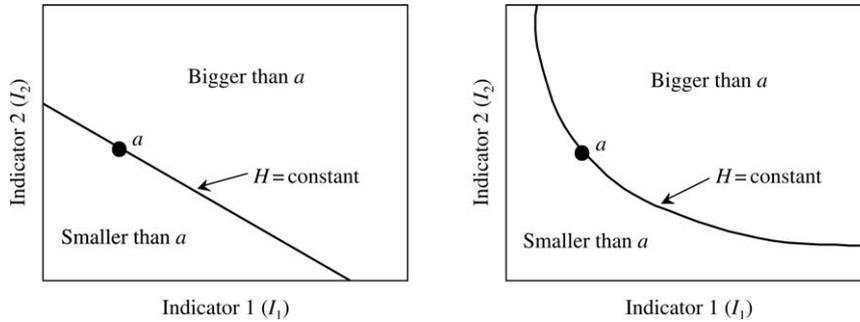
Each index  $H$  defines a linear ordering on the set of objects by the rule:

$$a \leq_H a', \quad \text{if and only if } H(a) \leq H(a').$$

The induced linear ordering can be displayed pictorially in terms of the contour of  $H$  that



**Figure 1.** With two indicators, each object  $a$  divides indicator space into four quadrants. Objects in the second and fourth quadrants are ambiguous in making comparisons with  $a$ .



**Figure 2.** Contour of index  $H$  passing through object  $a$ . A linear index is shown on the left and a non-linear index on the right.

passes through  $a$ . The contour divides indicator space into two regions; objects in the upper right-hand region are intrinsically bigger than  $a$  while those in the lower left-hand region are intrinsically smaller than  $a$ . See Fig. 2.

However, for an index  $H$  to be considered valid, its induced ordering should be consistent with the intrinsic ordering, i.e., we need to require that

$$a \leq a' \Rightarrow H(a) \leq H(a').$$

Pictorially, this means that the contour of  $H$  that passes through object  $a$  must lie entirely within the ambiguous regions for  $a$ . Fig. 3 shows some valid contours and also some invalid contours.

The mathematical conditions for an index to be valid are very simple:

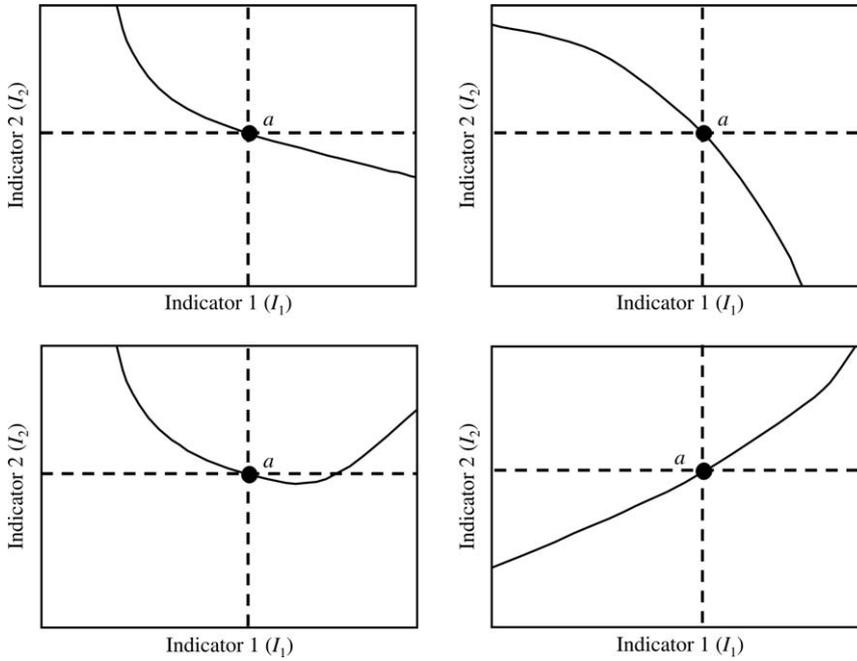
- An index  $H(I_1, \dots, I_p)$  is valid if and only if  $H$  is monotone increasing in each variable separately.
- A differentiable index  $H(I_1, \dots, I_p)$  is valid if and only if  $\partial H / \partial I_j \geq 0$  for all  $j$ .
- A linear index  $H = w_1 I_1 + w_2 I_2 + \dots + w_p I_p$  is valid if and only if  $w_j \geq 0$  for all  $j$ .

Fig. 3 suggests that validity is related to monotonicity of the contours when  $p = 2$ . Regarding  $I_2$  as a function of  $I_1$ , we have

$$\frac{dI_2}{dI_1} = -\frac{\partial H / \partial I_1}{\partial H / \partial I_2},$$

which implies that valid indexes have contours that are monotone decreasing. The converse is not quite true since both  $H$  and  $-H$  have the same contours; one has to choose the sign so that  $H$  is increasing (instead of decreasing) toward the upper right-hand corner of indicator space.

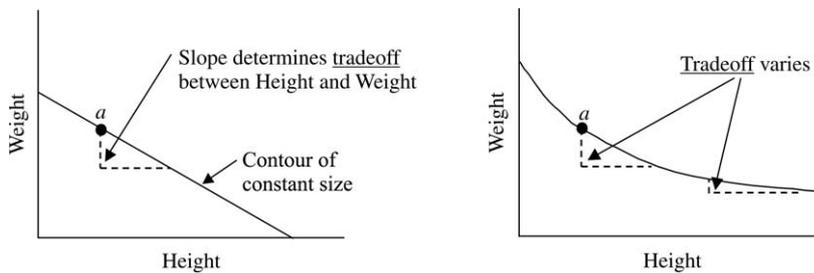
Validity is thus a mild restriction and still leaves a lot of freedom in choosing an index. Any proposed choice has to be considered in light of the “tradeoffs” or “substitutions” that are implied by the index’s contours. As an example, suppose we wish to compare “size” of people, using height ( $I_1$ ) and weight ( $I_2$ ) as indicators. If we adopt a linear index,  $H = w_1 I_1 + w_2 I_2$ , then we need to choose numerical values for the coefficients  $w_1$  and  $w_2$ . In part, these numerical values reflect our units of measurement; if we measure



**Figure 3.** The top two diagrams depict valid contours while the bottom two diagrams depict invalid contours.

weight in kilograms instead of pounds, the value of  $w_2$  must change accordingly. Given the units of measurement, the ratio  $w_2/w_1$  determines the tradeoff between height and weight in assessing size (Fig. 4). The situation becomes even more complex with a nonlinear index or when there are more than two indicators.

If one can argue persuasively for specific tradeoff value(s), then it makes a lot of sense to use the corresponding index. Typically, though, an index is adopted on grounds of mathematical convenience or simplicity (e.g., an average) with little effort to justify or even discuss the implied tradeoffs. For this reason, we spend the rest of this paper discussing the intrinsic ordering and its properties without combining indicators into an index.



**Figure 4.** The tradeoff or substitutability between height and weight in assessing the size of a person. The tradeoff is constant with a linear index (left) but varies across indicator space with a nonlinear index (right).

### 3. Partially ordered sets (posets)

A pre-order on a set  $S$  is a binary relation (written as  $\leq$ ) on  $S$  that satisfies the following two conditions:

- *Reflexivity*:  $a \leq a$  for all  $a \in S$ .
- *Transitivity*:  $a \leq b$  and  $b \leq c$  implies that  $a \leq c$ .

A pre-order is a partial order if the following holds:

- *Antisymmetry*:  $a \leq b$  and  $b \leq a$  implies that  $a = b$ .

A set  $S$  equipped with a partial order is called a partially ordered set or a poset. In the case of a poset, we write  $a < b$  if  $a \leq b$  but  $a \neq b$ . We do not define the relation  $a < b$  for a pre-order since there is ambiguity as to whether it should mean ( $a \leq b$  but  $a \neq b$ ) or ( $a \leq b$  is true but  $b \leq a$  is false). For additional information on partially ordered sets, see Fishburn (1985), Neggers and Kim (1998), and Trotter (1992).

Consider the intrinsic ordering on  $S$  that was defined in Section 2 ( $a \leq b$  means  $I_j(a) \leq I_j(b)$  for  $j = 1, 2, \dots, p$ ). Since this relation is clearly reflexive and transitive, the intrinsic order is a pre-order. Unfortunately, it does not have to be a partial order; in fact, antisymmetry fails whenever there exists a distinct pair of objects  $a, b \in S$  for which  $I_j(a) = I_j(b)$  for all  $j$ . Such a pair of objects plot to the same point in indicator space and can therefore be described as tied. The purpose of the antisymmetry condition is to rule out such ties. Since antisymmetry plays a fundamental role, we assume until further notice that antisymmetry holds true and that the intrinsic ordering is a partial order. We also assume that the set  $S$  is finite—this is not an issue in most scientific applications.

We need one further relation in a poset. Object  $b$  is said to cover object  $a$  provided (i)  $a < b$  and (ii) there is no object  $x$  for which  $a < x < b$ . Note that all the inequalities in this definition are strict. We write  $a < b$  when  $b$  covers  $a$ .

#### 3.1 Representations of posets

In this section, we describe three ways of portraying partially ordered sets:

- Hasse diagrams;
- zeta matrices; and
- cover matrices.

The Hasse diagram is a planar graph whose vertices are in one-to-one correspondence with the objects in  $S$  and whose pattern of edges determines the order relation. Hasse diagrams are excellent for visualization purposes—provided  $S$  is not unduly large. On the other hand, the zeta matrix is better for analytic purposes—in fact many of the operations on posets can be expressed by matrix multiplication. The cover matrix is a variant of the zeta matrix. One of our goals in this section is to explain how the Hasse diagram can be obtained formally from the zeta matrix.

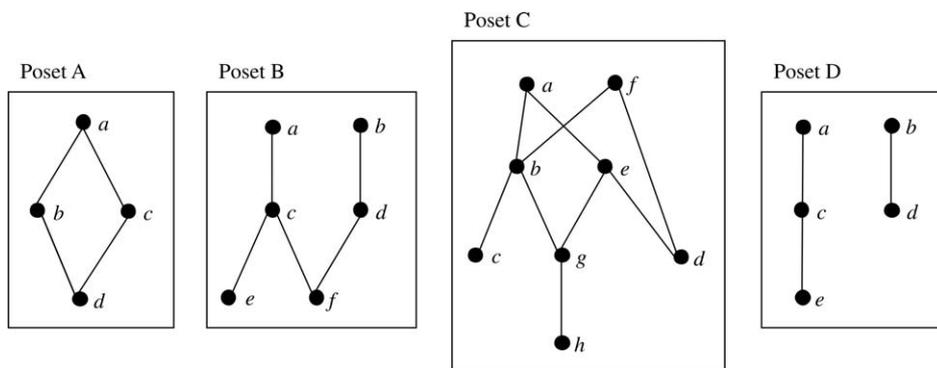
### 3.1.1 Hasse diagram

We describe how a Hasse diagram might be drawn on a sheet of paper. First, a point (or vertex) is plotted for each object  $a \in S$  so that  $b \in S$  is located higher on the paper whenever  $a < b$ . Second,  $a$  and  $b$  are connected by a straight line segment—an edge—whenever  $a$  is covered by  $b$  (notationally, when  $a \prec b$ ). There is considerable freedom in locating the vertices so that the Hasse diagram is far from unique. Considerable research has gone into finding “pleasing” ways of drawing graphs, including Hasse diagrams. See Di Battista *et al.* (1999) and Sugiyama (2002). Below, we give a systematic way of removing some of the non-uniqueness from the Hasse diagram. Fig. 5 gives Hasse diagrams for four different posets. It is important to note that edges are drawn to represent cover relationships only. In Poset A, for example, no edge joins  $d$  and  $a$  even though  $d < a$ ; this latter relationship is implied by the two covers  $d < c$  and  $c < a$ , each of which has an edge in the Hasse diagram. Also, observe that edges are permitted to cross one another as in the diagram for Poset C. Crossings are usually necessary if the Hasse diagram is to be drawn in a plane. See Neggers and Kim (1998) for detailed discussion of Hasse diagrams.

Poset D is of particular interest since its Hasse diagram consists of two disjoint pieces called connected components. In general, the connected components partition the poset into disjoint subsets such that if  $x$  is an arbitrary member of one component and  $y$  is an arbitrary member of a different component, then  $x$  and  $y$  are not comparable. Each of the other three posets in Fig. 5 has only one connected component and is accordingly said to be connected. We will later see how to determine the connected components of a poset from its zeta matrix.

The Hasse diagrams of Fig. 5 have been drawn in levels from the top down. The Hasse diagram of Poset B, for example, has three levels:

- Level 1 (top level):  $\{a, b\}$
- Level 2:  $\{c, d\}$
- Level 3:  $\{e, f\}$



**Figure 5.** Hasse diagrams for four different posets. Poset D has a disconnected Hasse diagram with two connected components  $\{a, c, e\}$  and  $\{b, d\}$ .

Similarly, Poset C has four levels:

- *Level 1 (top level):*  $\{a, f\}$
- *Level 2:*  $\{b, e\}$
- *Level 3:*  $\{c, g, d\}$
- *Level 4:*  $\{h\}$

We can determine these levels very easily. A member  $x$  of a poset is maximal if the poset contains no (strictly) larger elements, i.e., if there is no member  $y$  such that  $x < y$ . The top level consists of the maximal elements. Level 2 is obtained by removing the elements in the top level and determining the maximal elements of the resulting poset. Level 3 consists of the maximal elements after Levels 1 and 2 have been removed from the poset. This process is continued until all the levels have been determined. The Hasse diagram can then be drawn by distributing the members of each level across an invisible horizontal line (in some arbitrary order). Finally, an edge is drawn for each cover relationship in the poset.

The various levels can be characterized in the following way: if  $x$  is in Level  $k + 1$ , then there is an edge from  $x$  up to at least one vertex in Level  $k$ . Edges can skip levels (e.g., the  $df$  edge in Poset C) but there is always at least one edge up to the next level (e.g., edge  $de$  in Poset C).

The level in which a given element falls can be characterized directly. Element  $x$  is said to have depth  $k$  if  $k$  is the largest integer for which there is a sequence,

$$y_1 > y_2 > \cdots > y_k \equiv x.$$

Notice that, when  $k$  is maximal, each of the relationships  $y_i > y_{i+1}$  is a cover. It is now easy to see that  $x$  is in Level  $k$  if and only if  $x$  has depth equal to  $k$ . This is a useful fact, but not very practical for actually determining the levels. (Sometimes zero-indexing is used so that members of level 1 have depth 0.)

We always use the above top-down construction of the levels, but there is also a bottom-up construction using the notion of minimal elements from a poset. For clarity, we use the term co-level for each of the bottom-up levels. For example, Poset C has four co-levels:

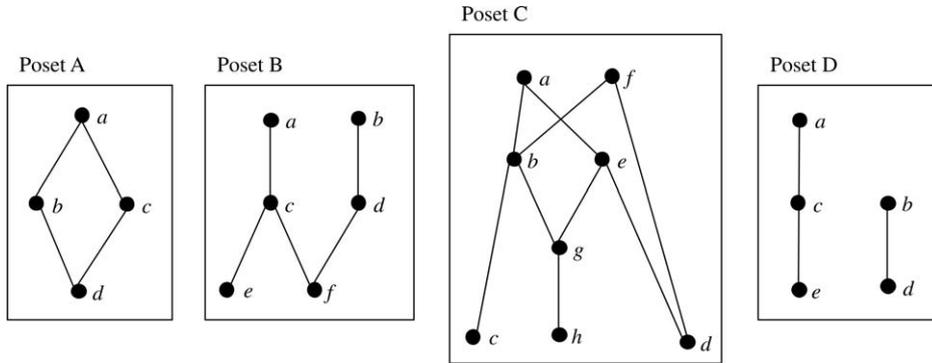
- *Co-level 1 (bottom level):*  $\{c, h, d\}$
- *Co-level 2:*  $\{g\}$
- *Co-level 3:*  $\{b, e\}$
- *Co-level 4:*  $\{a, f\}$

Bottom-up versions of the Hasse diagrams for the posets of Fig. 5 are shown in Fig. 6. The top-down and bottom-up versions are completely equivalent even though they can sometimes give different visual impressions. In the top-down version, vertices “bubble up” to the highest possible level, whereas they “sink” to the lowest possible level in the bottom-up version.

A member of Co-level  $k$  is said to have height  $k$ . The height of  $x$  can be characterized, as the largest integer  $k$  for which there is a sequence,

$$y_1 < y_2 < \cdots < y_k \equiv x.$$

The height of a poset is the number of its levels (or co-levels). The height of a poset also



**Figure 6.** Bottom-up Hasse diagrams for the posets of Fig. 5. Hasse diagrams for Posets A and B are unchanged.

equals the largest height of any of its members as well as the largest depth of any of its members. If  $x$  is any member of poset  $S$  then

$$\text{height}(x) + \text{depth}(x) \leq \text{height}(S) + 1.$$

(This relationship is more elegant when zero-indexing is used.) In Poset C, we have

$$\text{height}(d) = 1, \quad \text{depth}(d) = 3, \quad \text{but } \text{height}(S) = 4,$$

so this is an example in which the above inequality is strict. Equality holds in the above inequality for all  $x$  if and only if the top-down and bottom-up diagrams have exactly the same level sets.

### 3.1.2 Zeta matrix

The zeta matrix of a poset is a square matrix whose rows and whose columns are labeled by the members of the poset. The entries of the zeta matrix are either 0 or 1 as follows:

$$\zeta_{a,b} = 1, \quad \text{if } a \leq b; \quad \text{otherwise } \zeta_{a,b} = 0.$$

When we read across row  $a$  of the zeta matrix, the presence of a 1 means that the column label is greater than or equal to  $a$ . Similarly, reading down column  $b$ , each occurrence of a 1 means that the row label is less than or equal to  $b$ . We have borrowed the term ‘‘zeta’’ matrix from Budach *et al.* (1988); other authors use different names for the same matrix. As illustrations, the zeta matrix of Poset A in Fig. 5 is

$$\zeta_A = \begin{array}{c|cccc} & a & b & c & d \\ \hline a & 1 & 0 & 0 & 0 \\ b & 1 & 1 & 0 & 0 \\ c & 1 & 0 & 1 & 0 \\ d & 1 & 1 & 1 & 1 \end{array},$$

while that of Poset B is

$$\zeta_A = \begin{array}{c|cccccc} & a & b & c & d & e & f \\ \hline a & 1 & 0 & 0 & 0 & 0 & 0 \\ b & 0 & 1 & 0 & 0 & 0 & 0 \\ c & 1 & 0 & 1 & 0 & 0 & 0 \\ d & 0 & 1 & 0 & 1 & 0 & 0 \\ e & 1 & 0 & 1 & 0 & 1 & 0 \\ f & 1 & 1 & 1 & 1 & 0 & 1 \end{array}$$

The entries in these matrices are most easily filled in column by column. For example, consider column  $a$  of  $\zeta_A$ . A glance at the Hasse diagram (Fig. 5) shows that the list of elements less than or equal to  $a$  is  $\{a, b, c, d\}$  so that column  $a$  has a 1 in every entry. Similarly, the list of elements less than or equal to  $c$  is  $\{c, d\}$  so that column  $c$  has a 1 in row  $c$  and row  $d$ .

A useful fact to notice is that an element is maximal if and only if its row is zero except for a single 1 in the diagonal position. Similarly, an element is minimal if and only if its column is zero except for a single 1 in the diagonal position. Thus, for Poset B, we can see directly from the zeta matrix that the maximal elements are  $a$  and  $b$ , while the minimal elements are  $e$  and  $f$ .

When a poset is not connected, it is best to group together the members of each connected component before writing down the row and column labels. With this convention, the zeta matrix for Poset D assumes a block diagonal form:

$$\zeta_D = \begin{array}{c|cccc} & a & c & e & b & d \\ \hline a & 1 & 0 & 0 & 0 & 0 \\ c & 1 & 1 & 0 & 0 & 0 \\ e & 1 & 1 & 1 & 0 & 0 \\ \hline b & 0 & 0 & 0 & 1 & 0 \\ d & 0 & 0 & 0 & 1 & 1 \end{array}$$

Suppose we have a square matrix  $\zeta$  whose entries are either 0 or 1. What conditions must  $\zeta$  satisfy in order that it defines a partial order (on its row/column labels). Reflexivity means that  $\zeta$  must have 1s along its diagonal. Antisymmetry means that  $\zeta_{a,b}$  and  $\zeta_{b,a}$  cannot both be 1 when  $a \neq b$ . Transitivity is a little more complicated, but it means that  $\zeta_{a,b} = 1$  whenever there is an element  $x$  such that  $\zeta_{a,x}$  and  $\zeta_{x,b}$  are both equal to 1. This is the same as saying that  $\zeta_{a,b}$  is nonzero whenever  $\sum_x \zeta_{a,x} \zeta_{x,b}$  is nonzero.

Now, let us digress briefly. If  $\mathbf{A}$  is a matrix with nonnegative entries, we write  $\mathcal{L}(\mathbf{A})$  for the matrix that results when every positive entry of  $\mathbf{A}$  is replaced with the value 1;  $\mathcal{L}(\mathbf{A})$  is the logical form of  $\mathbf{A}$ . If  $\mathbf{A}$  and  $\mathbf{B}$  are (conformable) matrices with nonnegative entries, we write  $\mathbf{A} \circ \mathbf{B}$  for  $\mathcal{L}(\mathbf{AB})$ . Thus,  $\mathbf{A} \circ \mathbf{B}$  is obtained by computing the ordinary matrix product and then replacing every entry in the result by its logical equivalent. Finally, we write  $\mathbf{A}^* \mathbf{B}$  for the component-wise product of two matrices of the same size.

With these notations in hand, we can now say that  $\zeta$  satisfies the transitivity condition if and only if  $\zeta \circ \zeta \leq \zeta$ , where matrix comparisons are component-wise. In summary, an  $n \times n$  matrix  $\zeta$  whose entries are either 0 or 1 defines a partial order if and only if the following conditions are all satisfied:

1.  $\zeta$  has all 1s along its diagonal;
2.  $\zeta * \zeta = \mathbf{I}_{n \times n}$ ; and
3.  $\zeta \circ \zeta \leq \zeta$ .

A computer can check these conditions in a rather mechanical fashion. Note that condition (2) implies condition (1).

It is sometimes more convenient to work with the matrix  $\eta$  that is obtained by putting the diagonal entries of  $\zeta$  equal to 0. Thus,  $\zeta = \mathbf{I}_{n \times n} + \eta$ . We can also characterize the conditions for a partial order in terms of the matrix  $\eta$ . In fact,  $\zeta = \mathbf{I}_{n \times n} + \eta$  defines a partial order if and only if the following conditions are all satisfied:

1.  $\eta$  has all 0s along its diagonal;
2.  $\eta * \eta = \mathbf{0}_{n \times n}$ ; and
3.  $\eta \circ \eta \leq \eta$ .

Once again, condition (2) implies condition (1).

### 3.1.3 Cover matrix

Like the zeta matrix, the cover matrix,  $\gamma$ , of a poset is a square matrix whose rows and whose columns are labeled by the members of the poset. The entries of the cover matrix are either 0 or 1 as follows:

$$\gamma_{a,b} = 1, \quad \text{if } a < b; \quad \text{otherwise } \gamma_{a,b} = 0.$$

Thus, nonzero entries of the cover matrix indicate the existence of a cover relationship. All the diagonal entries of the cover matrix must vanish. According to the definition,  $a < b$  means that  $a < b$  and there is no  $x$  such that  $a < x$  and  $x < b$ . But, this is the same as saying that  $\eta_{a,b} = 1$  but  $\sum_x \eta_{a,x} \eta_{x,b} = 0$ . Thus, the cover matrix can be computed from the zeta matrix by the formula

$$\gamma = \eta - \eta \circ \eta. \quad (1)$$

The converse is also true: the zeta matrix can be computed from the cover matrix. To see this, we need some more notation. If  $\mathbf{A}$  is a square matrix with nonnegative entries and  $m$  is a positive integer, we write

$$\mathbf{A}^{\circ m} = \mathcal{L}(\mathbf{A}^m) = \mathbf{A} \circ \mathbf{A} \circ \cdots \circ \mathbf{A} \text{ (} m \text{ factors)}.$$

Now,  $\eta_{a,b} = 1$  if and only if  $a < b$  if and only if there is a sequence

$$a = x_0 < x_1 < \cdots < x_m = b, \quad (2)$$

where  $1 \leq m < n$  with  $n$  as the number of elements in the poset. The existence of such a sequence is equivalent to requiring the  $a, b$  component of  $\gamma^m$  to be nonzero. Thus,  $\eta_{a,b} = 1$  if and only if  $\gamma^m = 1$  for some  $m$  with  $1 \leq m < n$ . We can therefore conclude that

$$\eta = \mathcal{L}(\gamma + \gamma^{\circ 2} + \gamma^{\circ 3} + \cdots + \gamma^{\circ(n-1)}) = \mathcal{L}(\gamma + \gamma^2 + \gamma^3 + \cdots + \gamma^{n-1}),$$

and

$$\begin{aligned}\zeta &= \mathcal{L}(\mathbf{I}_{n \times n} + \gamma + \gamma^{\circ 2} + \gamma^{\circ 3} + \cdots + \gamma^{\circ(n-1)}) \\ &= \mathcal{L}(\mathbf{I}_{n \times n} + \gamma + \gamma^2 + \gamma^3 + \cdots + \gamma^{n-1}).\end{aligned}$$

A simpler expression results by using the relationship  $a \prec' b \Leftrightarrow a \prec b$  or  $a = b$ . The indicator matrix for  $\prec'$  is  $\bar{\gamma} = \mathbf{I}_{n \times n} + \gamma$ . By inserting enough equalities in the sequence (2), we can say that  $\zeta_{a,b} = 1$  if and only if  $a \leq b$  if and only if there is a sequence

$$a = x_0 \prec' x_1 \prec' \cdots \prec' x_m = b,$$

for some sufficiently large value of  $m$  ( $m \geq n - 1$  is always large enough). But then

$$\zeta = \mathcal{L}(\bar{\gamma}^m) = \mathcal{L}(\bar{\gamma}^m), \quad \text{whenever } m \geq n - 1.$$

Notice that the required high power of  $\bar{\gamma}$  can be obtained by successively squaring matrices:

$$\bar{\gamma}^2, \bar{\gamma}^4, \bar{\gamma}^8, \bar{\gamma}^{16}, \dots$$

The logical operator  $\mathcal{L}(\cdot)$  can be applied at any point(s) along the way in order to avoid numerical overflow in computing these matrix powers.

### 3.2 Obtaining the Hasse diagram from the zeta matrix

Starting from the zeta matrix of a poset, we want to extract the information needed to draw the Hasse diagram. There are three steps involved in implementing this program:

- *Step 1:* Determine if the poset is connected and, if it is disconnected, identify all the connected components.
- *Step 2:* For each connected component, determine the top-down levels for its Hasse diagram.
- *Step 3:* For each connected component, determine its cover matrix (which is needed to know where to insert edges in the Hasse diagram).

Technically, it is not necessary to carry out the first step; but if you do not do so and if the poset is disconnected, the connected components of the Hasse diagram (as it is drawn) tend to interweave in such a way that they are not visually obvious.

We take up these three steps in reverse order. Step 3 is solved by Equation (1), which shows how to calculate the cover matrix from the zeta matrix. The key ideas for Step 2 were summarized in Section 3.2 where we pointed out that the top-down levels are obtained by successively identifying and removing maximal elements from the poset. Furthermore, an element is maximal exactly when its row sum in the zeta matrix is unity. Therefore, we augment the zeta matrix by additional columns on the right that will hold the row sums. For Poset B, the augmented zeta matrix is

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	L1	L2	L3
<i>a</i>	1	0	0	0	0	0	1	0	0
<i>b</i>	0	1	0	0	0	0	1	0	0
$\zeta_B = c$	1	0	1	0	0	0	2	1	0
<i>d</i>	0	1	0	1	0	0	2	1	0
<i>e</i>	1	0	1	0	1	0	3	2	1
<i>f</i>	1	1	1	1	0	1	5	3	1

The row sums of  $\zeta_B$  are stored in the first augmentation column (headed with the label L1 for Level 1). The value 1 in each of the first two rows of augmentation column L1 indicates that *a* and *b* are maximal elements. The next step would be to remove *a* and *b* from the poset. We don't want to actually effect this removal because it would be very inconvenient in computer memory. Instead, we will adjust the row sums for the virtual removal of *a* and *b*. This is accomplished by subtracting columns *a* and *b* from column L1 and storing the results in column L2. The two values of zero in column L2 serve to remind that *a* and *b* have already been removed, and the two values of unity tell us that the next set of maximal elements consists of *c* and *d*. These two elements comprise Level 2. Finally, we subtract columns *c* and *d* from column L2 and store the results in column L3 to discover that Level 3 is the final level and that it consists of elements *e* and *f*. This algorithm for finding the top-down levels is surprisingly simple and can be easily implemented on a computer. The bottom-up levels can be found in much the same way except that the zeta matrix is augmented with column sums.

### 3.2.1 Determining the connected components of a poset from its zeta matrix

Thinking in terms of the Hasse diagram, two elements *a* and *b* are in the same connected component if there is a path through the diagram from *a* to *b*; this path can go up or down or both. The indicator matrix  $\kappa$  for the connected components is defined by  $\kappa_{a,b} = 1$  if *a* and *b* are in the same connected component; otherwise,  $\kappa_{a,b} = 0$ . Note that  $\kappa$  is symmetric and has 1s along the diagonal. We want to be able to calculate  $\kappa$  from the zeta matrix. To this end, we say that two elements *x* and *y* are comparable, written as  $x \hat{=} y$ , if either  $x \leq y$  or  $y \leq x$ ; note that the first possibility corresponds to moving upward from *x* to *y* in the Hasse diagram and the second possibility to moving downward from *x* to *y*. Clearly,  $\kappa_{a,b} = 1$  if and only if there is a finite sequence

$$a = x_0 \hat{=} x_1 \hat{=} \dots \hat{=} x_m = b.$$

By deleting cycles from and/or inserting equalities into this sequence, we may suppose that  $m = n - 1$ ; then by inserting more equalities we can suppose that *m* takes any given value bigger than  $n - 1$ . The indicator matrix,  $\tilde{\zeta}$ , of the (symmetric) relation  $\hat{=}$  is the symmetrized form of the zeta matrix, namely,

$$\tilde{\zeta} = \mathcal{L}(\zeta + \zeta^T) = \zeta + \zeta^T - \mathcal{I}_{n \times n}.$$

Examining the above sequence, we see that  $\kappa_{a,b} = 1$  if and only if the *a, b* component of  $\tilde{\zeta}^m$  is nonzero whenever  $m \geq n - 1$ . Thus, we obtain that

$$\kappa = \tilde{\zeta}^{om} = \mathcal{L}(\tilde{\zeta}^m),$$

whenever  $m \geq n - 1$ . As before, the needed matrix power  $\tilde{\zeta}^m$  can be calculated by iterated squares. Once  $\kappa$  is calculated, the connected components of the poset are obtained as follows. Fix any element  $a$  of the poset and scan across row  $a$  of  $\kappa$ . Whenever a value of 1 is encountered, it means that the corresponding column label is in the same connected component as  $a$ ; otherwise, the column label is in a different connected component. These scans also tell us how to group and permute the elements of the poset in order to change the zeta matrix into block diagonal form. See the discussion of  $\zeta_D$  in Section 3.1.2. We can also conclude that a poset is connected if and only if every entry of  $\kappa$  equals unity.

The foregoing method for identifying connected components is easy to implement (since it involves only matrix multiplication) and reasonably efficient provided there are no more than a few hundred elements in the poset. More efficient (but more complicated) methods are available. See Knuth (1973) or Cormen *et al.* (2001).

### 3.3 Example: The human-environment interface

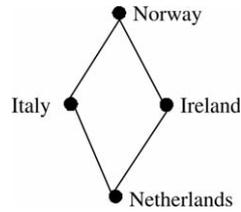
UNEP has compiled data to compare and rank the countries of the world according to environmental quality. Three indicator values are available for each of 106 countries. The indicators are intended to assess three different components of the human-environment interface (land, air, and water). Each indicator takes values between 0 and 1 with large values representing ‘‘better’’ conditions. We refer to the original report (Singh and Shansieve-Cohen, 2001) for data tabulations and details concerning construction of the indicators. Patil *et al.* (2001) gives a statistical analysis and summarization of the data.

Table 1 shows the indicators for four European countries: Norway, Italy, Ireland, and The Netherlands. We see that Norway has a higher indicator value than Italy for each of the three indicators so that Norway should be ranked higher than Italy for any reasonable ranking method (based on these indicators). In other words, Norway is intrinsically better than Italy according to these data.

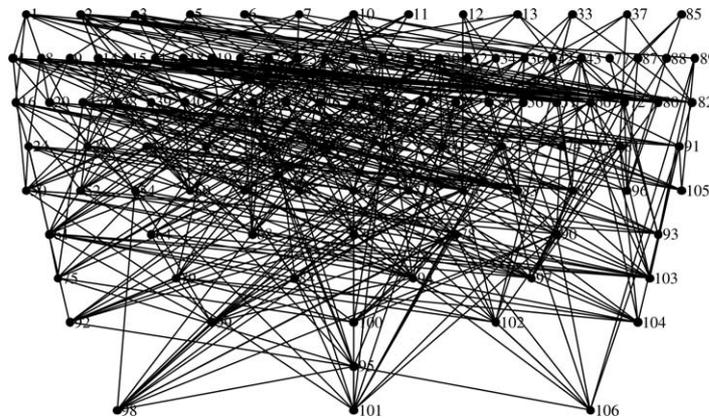
A different picture emerges when we compare Italy with Ireland. Italy is better with respect to the Air and Land indicators but Ireland is better with respect to the Water indicator. In this case, we would say that Italy and Ireland are not intrinsically comparable with respect to the HE indicators. This means that different investigators might rank Italy and Ireland differently, depending upon the relative weight or importance to be attached to the different indicators. The human-environment index (HEI) proposed by Singh and

**Table 1.** Human-environment indicator values for four countries. The average of the three indicators is the HEI. The last column gives the country’s rank when all 106 countries are ranked according to the HEI.

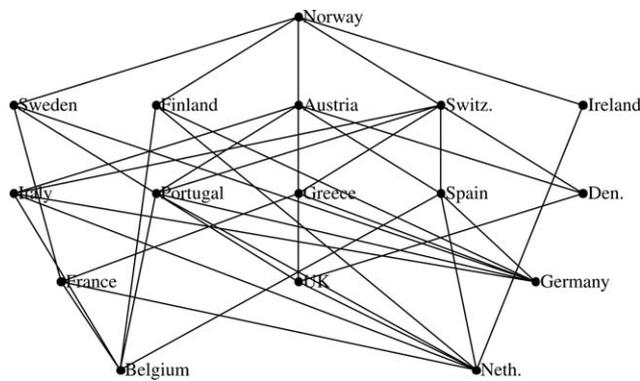
Country	Air	Water	Land	HEI	HEI Rank
Norway	0.43	1.00	1.00	0.81	2
Italy	0.37	0.82	0.58	0.59	63
Ireland	0.30	0.99	0.22	0.50	84
The Netherlands	0.22	0.61	0.16	0.33	104



**Figure 7.** Hasse diagram for the four countries of Table 1. Note that it has the same structure as Poset A in Fig. 5.

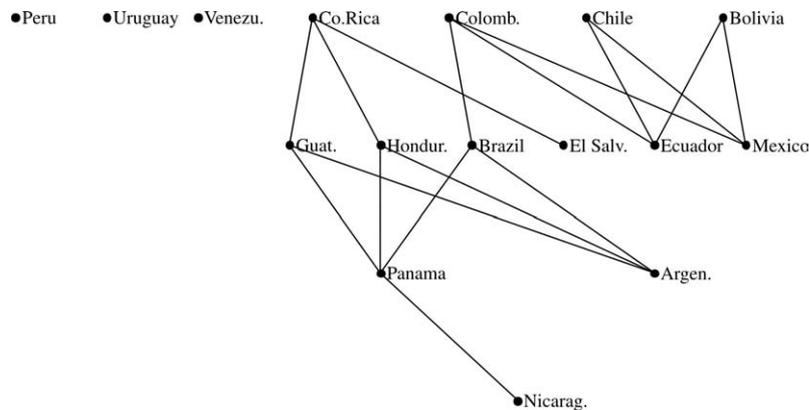


**Figure 8.** Hasse diagram for all 106 countries. Labels are the HEI ranks. The diagram is connected.



**Figure 9.** Hasse diagram for the countries of Western Europe. The diagram is connected.

Shansieve-Cohen (2001) takes the average of the three indicator values and accordingly gives equal weight to the indicators. According to HEI, Italy ranks higher than Ireland. A different way of combining the three indicators might have Ireland outranking Italy. Fig. 7 gives the Hasse diagram for the four countries of Table 1. This diagram shows that Italy and Ireland are not intrinsically comparable, while Norway is intrinsically better than each of the three other countries. Figs. 8–10 show Hasse diagrams for all 106 countries, for



**Figure 10.** Hasse diagram for Latin America. There are four connected components. Three of these components are isolates; the remaining component contains 13 countries.

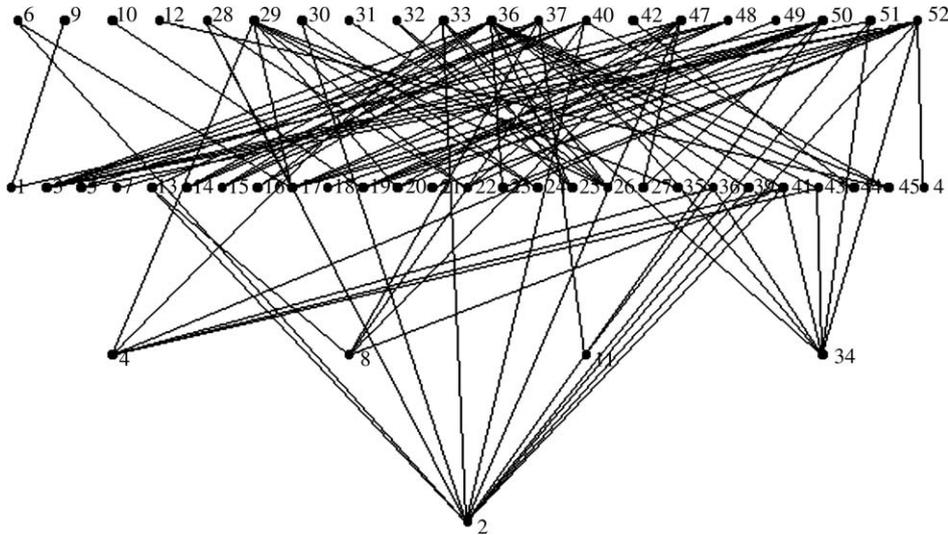
Western Europe, and for Latin America, respectively. The latter is of interest because it is disconnected.

### 3.4 Example: Using landscape metrics to prioritize watersheds

The United States Environmental Protection Agency (USEPA, 1997) has developed a suite of indicators to evaluate and rank environmental impact across 114 watersheds of the Mid-Atlantic region of the eastern United States. This region runs from Pennsylvania in the north down to Virginia in the south and from West Virginia eastward to the Atlantic Ocean. The “watersheds” are actually hydrologic units with eight-digit hydrologic unit codes (HUCs). The suite includes several dozen indicators, many of which are strongly correlated. We have selected nine indicators for this Hasse analysis (Table 2). Indicator directions were adjusted so that large indicator values correspond to heavy environmental impact.

**Table 2.** The nine indicators used in the Hasse diagram analysis. Indicator directions were adjusted so that large indicator values correspond to heavy environmental impact.

Indicator	Description
POPDENS	1990 population density
POPCHG	1970–1990 population change
RDDENS	Road density
SO4DEP	Sulfate deposition
RIPFOR	Proportion forested stream-length
STRD	Proportion stream-length with nearby roads
DAMS	Impoundments per 1000 km stream-length
CROPSL	Proportion of WS with crops on > 3% slope
INTALL	Proportion of WS with interior forest habitat



**Figure 11.** Hasse diagram for the 52 watersheds in the primary component. Labels are (arbitrary) row numbers in the data matrix.

The Hasse diagram is highly disconnected with 60 connected components:

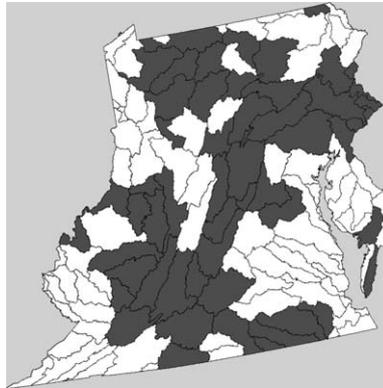
- 58 of the components are isolates.
- 1 component (*secondary component*) contains four watersheds.
- 1 component (*primary component*) contains 52 watersheds.

Over half of the watersheds are isolates showing the great difficulty of comparing environmental impact across this region. The nine indicators present such a multi-faceted view of environmental condition that comparison of the watersheds becomes problematic. The Hasse diagram for the primary component is shown in Fig. 11. The diagram has four levels and all but five watersheds fall into the top two levels. There is no logical connection between connectivity in the Hasse diagram and geographic connectivity, but mapping the watersheds reveals that the primary component tends to be geographically connected (Fig. 12).

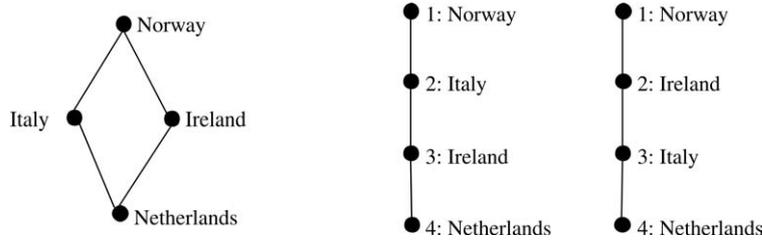
#### 4. Rank-intervals: Quantifying rank ambiguity

The examples in Sections 3.3 and 3.4 emphasize that considerable ambiguity can occur in attempting to rank the members of a poset  $S$ . We quantify this ambiguity by considering all the possible rankings of  $S$  that are consistent with its partial order. Across all these rankings, a given member  $a$  of  $S$  may be assigned several different numerical ranks. The set of all possible ranks turns out to be an interval (of integers), which we call the rank-interval of  $a$ .

Throughout, we follow the “man-in-the-street” convention of assigning small numerical ranks to top elements and large ranks to bottom elements. This is opposite to the statistical convention of assigning rank 1 to the smallest value. Our convention has the



**Figure 12.** Map of the Mid-Atlantic region showing the primary Hasse component (shaded). Geographically, there are three connected components of which two are small and located near the periphery of the region.



**Figure 13.** Hasse diagrams (right) of the two possible rankings for the poset on the left.

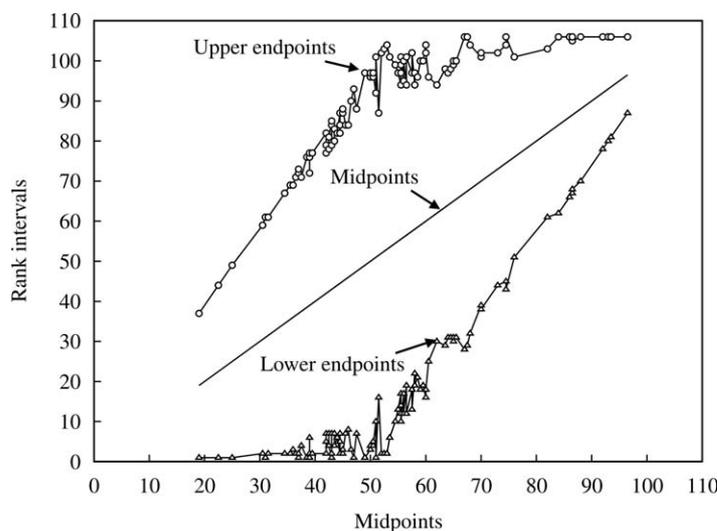
consequence that the numerical ranks tend to be negatively correlated with the indicators—which can lead to confusion if the convention is not kept in mind. We occasionally use the phrase “our ranking convention” to remind the reader.

We first illustrate with the simple example of the four countries of Table 1, whose Hasse diagram is repeated in Fig. 13. Any consistent ranking of the four countries must assign rank 1 to Norway and rank 4 to The Netherlands. Ranks 2 and 3 can be assigned indifferently to Italy and Ireland. This gives two possible rankings:

- Norway > Italy > Ireland > The Netherlands.
- Norway > Ireland > Italy > The Netherlands.

Hasse diagrams for these rankings appear in Fig. 13. The possible ranks assigned to each of the four countries are given in the following tabulation:

Country	Possible Ranks
Norway	1
Italy	2, 3
Ireland	2, 3
The Netherlands	4



**Figure 14.** Rank-intervals for all 106 countries. The intervals (countries) are labeled by their midpoints as shown along the horizontal axis. For each interval, the lower endpoint and the upper endpoint are shown vertically. The length of each interval corresponds to the ambiguity inherent in attempting to rank that country among all 106 countries.

For each country, the list of possible ranks is a set of consecutive integers, which we are calling the rank-interval for that country. The length of the rank-interval measures the ambiguity involved in attempting to rank that country against all other countries in the poset. We have computed rank-intervals for all 106 countries. Results are shown visually in Figs 14 and 15. Observe that the rank-intervals are generally quite wide indicating substantial ambiguity in any attempt to rank the human-environment poset.

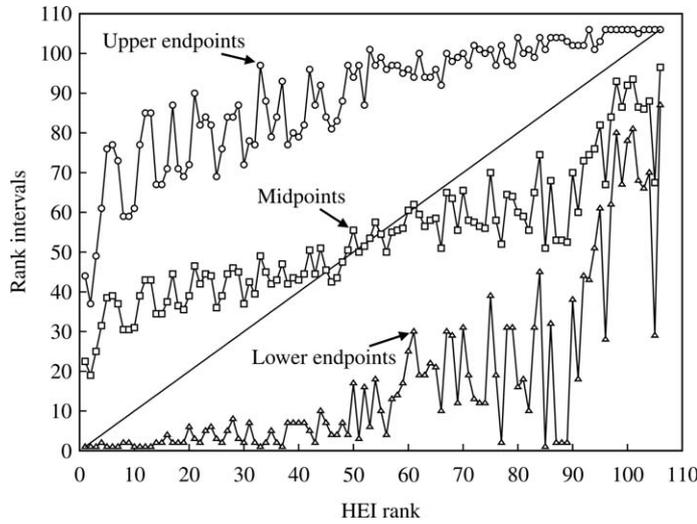
In Fig. 14, we have used the rank-interval midpoints to arrange the countries along the horizontal axis. This suggests using the midpoints themselves as a kind of index to rank the countries. More generally, for this purpose, we might use the upper endpoints, or the lower endpoints, or any other “average” computed across each rank interval. We will return to this notion of a rank-index in Section 0 after we prove that each rank-interval is indeed an interval.

#### 4.1 Rank-intervals are intervals

A (consistent) ranking of poset  $S$  is essentially a listing or enumeration,  $a_1, a_2, \dots, a_n$ , of its elements with the requirement that

$$a_i > a_j \Rightarrow i < j. \quad (3)$$

The inequality sign on the left refers to the partial order in  $S$  while that on the right refers to the ordering of real numbers—integers in this case. If condition (3) appears mysterious,



**Figure 15.** Rank-intervals for all 106 countries, plotted against their HEI rank. The HEI rank appears as the 45-degree line. The HEI tends to be optimistic (closer to the lower endpoint) for better-ranked countries and pessimistic (closer to the upper endpoint) for poorer-ranked countries.

think of the listing  $a_1, a_2, \dots, a_n$  as an integer-valued function,  $R$ , defined on  $S$  and mapping element  $a_i$  to its subscript  $i$ . Condition (3) requires  $R$  to be strictly anti-isotonic (“anti” because of our ranking convention):

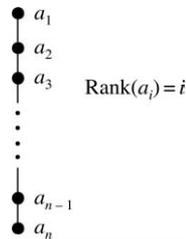
$$a_i > a_j \Rightarrow R(i) < R(j).$$

The listing  $a_1, a_2, \dots, a_n$  is best visualized as a linear Hasse diagram with  $a_1$  at the top and  $a_2$  at the bottom (Fig. 16).

In the poset literature, rankings as depicted by the Hasse diagram of Fig. 16 are called linear extensions of the poset  $S$ . In general, if  $\leq$  and  $\leq^*$  are partial orders on a set  $S$ , then  $\leq^*$  is an extension of  $\leq$  provided

$$a \leq b \Rightarrow a \leq^* b, \quad \text{for all } a, b \in S.$$

The partial order  $\leq^*$  is linear if every pair of elements of  $S$  is comparable under  $\leq^*$ , i.e., if the Hasse diagram is linear as in Fig. 16.



**Figure 16.** A ranking of a poset determines a linear Hasse diagram. The numerical rank assigned to each element is that element’s depth in the Hasse diagram.

Each element  $a \in S$  determines three subsets of  $S$ :

- The upper set  $U_a$  contains all members of  $S$  that strictly exceed  $a$

$$U_a = \{x \in S : x > a\}.$$

This set corresponds to the first quadrant of Fig. 1.

- The lower set  $L_a$  contains all members of  $S$  that are strictly less than  $a$

$$L_a = \{x \in S : x < a\}.$$

This set corresponds to the third quadrant of Fig. 1.

- The ambiguous set  $A_a$  contains all members of  $S$  that are not comparable with  $a$

$$A_a = \{x \in S : x \leq a \text{ and } x \geq a \text{ are both false}\}.$$

This set corresponds to the union of the third and fourth quadrants of Fig. 1.

These three sets together with the singleton  $\{a\}$  comprise a partition of  $S$ ,

$$S = \{a\} \cup U_a \cup L_a \cup A_a.$$

Using  $|A|$  to stand for the cardinality of set  $A$ , the above partitioning implies that

$$|A_a| + |U_a| + 1 = n - |L_a|, \quad (4)$$

where  $n$  is the cardinality of  $S$ . We are now ready to show that rank-intervals really are intervals.

**Theorem 1** *Let  $a$  be a given member of a partially ordered set  $S$ . If a ranking of  $S$  assigns rank  $r$  to  $a$  then*

$$|U_a| + 1 \leq r \leq |A_a| + |U_a| + 1. \quad (5)$$

*Equivalently,*

$$|U_a| + 1 \leq r \leq n - |L_a|. \quad (6)$$

*Conversely, if a positive integer  $r$  satisfies condition (5) then there is a ranking of  $S$  that assigns rank  $r$  to  $a$ . The ambiguity of  $a$  (i.e., the length of its rank-interval) equals  $|A_a|$ .*

**Proof:** Fix a ranking of  $S$  and think of it as a linear Hasse diagram as in Fig. 16. Since the ranking is consistent with the partial order on  $S$  and since all the members of  $U_a$  exceed  $a$  in that partial order, these members must all lie (strictly) above  $a$  in the linear Hasse diagram. This implies that the depth of  $a$  in the linear Hasse diagram is at least  $|U_a| + 1$ . Similarly, all the members of  $L_a$  must lie below  $a$  in the linear Hasse diagram implying that the depth of  $a$  in that diagram cannot exceed  $n - |L_a|$ . To prove the converse, we describe a universal technique for constructing linear extensions of a (finite) poset. Select an arbitrary maximal element of the poset, remove that element, and enter it as the first member of the list. Continue by selecting an arbitrary maximal element from the remaining poset, remove that element, and enter it as the second member of the list. Continue in this manner until the poset is exhausted. The resulting list forms a linear extension of the original poset. Use this technique to construct a linear extension  $x$  of  $U_a$ , a linear extension  $y$  of  $A_a$ , and a

linear extension  $z$  of  $L_a$ . Form a linear Hasse diagram by stacking  $x$  above  $y$  above  $z$ . We claim that this gives a linear extension of  $S - \{a\}$ . To see this, note that  $y \geq x$  can never be true if  $x \in U_a$  and  $y \in A_a$ ; otherwise, we would have  $y \geq a$  (because  $x > a$  by the definition of  $U_a$ ), which would contradict the definition of  $A_a$ . Thus, our linear Hasse diagram is consistent with the order relations (in  $S$ ) between members of  $U_a$  and members of  $A_a$ . Similarly, there is cross-consistency for  $(A_a, L_a)$  and for  $(U_a, L_a)$ . Now we ask: Where can the element  $a$  be inserted into the linear Hasse diagram to obtain a linear extension of  $S$ ? Clearly,  $x$  must lie above and  $z$  must lie below that location. But the insertion can occur anywhere within  $y$ , immediately above  $y$ , or immediately below  $y$  without violating any of the order relationships in  $S$ . These choices cover the interval described by condition (6). This completes the proof.  $\square$

One might think that it is necessary to examine all the linear extensions of  $S$  in order to determine the rank-intervals. Fortunately, this is not the case since the bounds in (6) can be calculated directly from the zeta matrix.

**Theorem 2** *Let  $a$  be a member of a poset  $S$ . The sum of the entries in row  $a$  of the zeta matrix is  $|U_a| + 1$  and the sum of the entries in column  $a$  is  $|L_a| + 1$ .*

**Proof:** This follows directly from the definition of the zeta matrix. The extra summand of 1 in these expressions is due to diagonal entries in the zeta matrix.  $\square$

## 4.2 Rank-indexes

By a rank-index, we mean a function,  $R(a, S)$ , defined for every (finite) poset  $S$  and every element  $a$  in  $S$  and which is anti-isotonic with respect to  $a$ , i.e.,

$$a \geq b \Rightarrow R(a, S) \leq R(b, S), \quad \text{whenever } a, b \in S. \quad (7)$$

The ‘‘anti’’ is unnatural and results from our ranking convention. The dependence on  $S$  emphasizes that a rank-index is a general procedure that is available for all finite posets and not just some particular poset. More importantly, if  $S$  is expanded to a larger poset  $S'$ , there is no requirement that  $R(a, S)$  and  $R(a, S')$  be the same. This permits the index value assigned to element  $a$  to reflect the structure of the containing poset and not be an inherent property of  $a$  by itself. The composite indexes discussed in Section 2 depend only on the indicator values and do not change their values if more elements are added to the poset. In this sense, rank-indexes are more general than composite indexes. In Section 4, we suggested that rank-indexes could be constructed by calculating a generalized mean across each rank interval. We want to show that the resulting indexes are indeed anti-isotonic as specified by condition (7).

**Lemma** *Let  $[A, B]$  and  $[A', B']$  be non-degenerate intervals of real numbers such that  $A \leq A'$  and  $B \leq B'$ . Then, the uniform distribution on  $[A, B]$  is stochastically smaller than the uniform distribution on  $[A', B']$ . The same conclusion holds for the discrete uniform distribution provided  $A, B, A', B'$  are integers.*

**Proof:** The result can be proved algebraically but we give a simple geometric argument. Let  $F(x)$  and  $F'(x)$  be the cumulative distribution functions for the uniform distribution on the intervals  $[A, B]$  and  $[A', B']$ , respectively. We need to show that  $F(x) \geq F'(x)$  for all  $x$ . But, this is trivially true for  $x < A$ , so we restrict attention to  $A \leq x$ . The graph of  $F(x)$  consists of a linear segment rising from  $(A, 0)$  to  $(B, 1)$  followed by a horizontal straight line at height 1. The region below this graph and to the right of  $x = A$  is convex. Consider the rising linear segment portion of the graph of  $F'(x)$ . Its endpoints,  $(A', 0)$  and  $(B', 1)$ , lie within the above-mentioned convex region and so also for the straight line joining these points.  $\square$

**Corollary** *Let  $S$  be a poset and write  $A(a)$  and  $B(a)$  for the respective upper and lower endpoints of the rank-interval of  $a \in S$ . If  $a, a' \in S$  with  $a' \leq a$ , then  $A(a) \leq A(a')$  and  $B(a) \leq B(a')$ . In particular,  $a' \leq a$  implies that the uniform distribution (discrete or continuous) on the rank-interval of  $a$  is stochastically smaller than that of  $a'$ . (The order-reversal is due to our ranking convention.)*

**Proof:** This is obvious from the explicit expressions given in Equation (6) for the rank-interval end-points.  $\square$

Let  $G(x)$  be a real-valued function defined over the positive real axis and suppose that  $G(x)$  is either strictly increasing or strictly decreasing. The class of functions we particularly have in mind is  $G(x) = x^p$  if  $p \neq 0$  and  $G(x) = \ln(x)$  if  $p = 0$ . Let  $\mu$  be a probability measure defined across a bounded subinterval of the positive real-axis. The generalized mean of  $\mu$  induced by  $G(x)$  is

$$\mu_G = G^{-1}(E_\mu[G(\cdot)]).$$

From well-known results on stochastic ordering (Lehmann, 1986), it follows that  $\mu_G \leq \mu'_G$  whenever  $\mu$  is stochastically smaller than  $\mu'$ . Taking  $\mu$  as uniform across rank-intervals, we can conclude that the rank-index induced by  $G(x)$  is anti-isotonic.

## 5. Rank-frequency distributions

In Section 4.2, we considered the uniform distribution across each rank-interval. This ignores the fact that ranks near the ends of the interval are usually less frequent than ranks near the middle of the interval. We can get a better assessment of rank-ambiguity if we regard each linear extension as a ‘‘voter’’ or ‘‘judge’’ and count the number of times a given rank is assigned to a given element  $a$  of the poset. This gives a unimodal (discrete) frequency distribution across the rank-interval of  $a$ , which we call the rank-frequency distribution of  $a$ .

We need some notation before proceeding. Throughout,  $S$  is a given poset and we let  $\Omega = \Omega_S$  stand for the collection of all linear extensions of  $S$ ; the set  $\Omega$  is finite but generally very large. Members of  $\Omega$  are denoted generically by the symbol  $\omega$ , and the rank which  $\omega$  assigns to  $a \in S$  is written as  $\omega(a)$ . With these notations, the (unnormalized) rank-frequency distribution of  $a \in S$  is given by

$$f_a(r) = \#\{\omega \in \Omega : \omega(a) = r\}, \quad r \text{ in the rank-interval of } a, \quad (8)$$

and the corresponding cumulative rank-frequency distribution becomes

$$\begin{aligned}
 F_a(r) &= f_a(1) + f_a(2) + \dots + f_a(r) \\
 &= \#\{\omega \in \Omega : \omega(a) \leq r\}.
 \end{aligned}
 \tag{9}$$

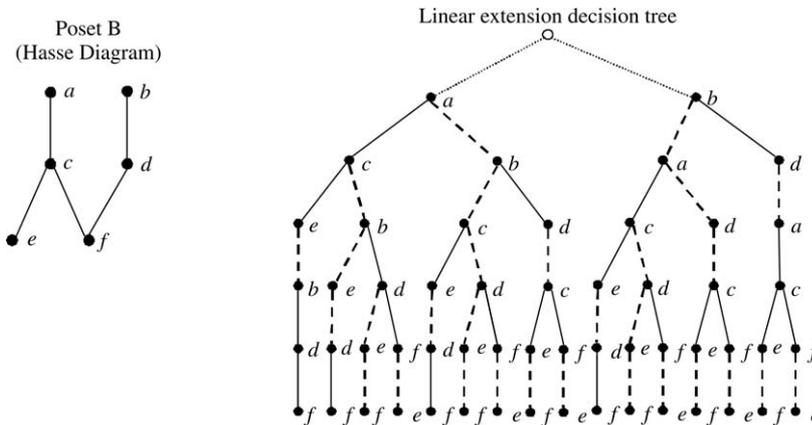
From (8), it is clear that  $\sum_r f_a(r) = |\Omega|$  is the same for every  $a \in S$ . Thus, it matters little whether one uses normalized or unnormalized rank-frequency distributions.

**Theorem 3** *If  $a, a' \in S$  and  $a \geq a'$ , then  $F_a(r) \geq F_{a'}(r)$  for all positive integers  $r$ . In other words, the rank-frequency distribution of  $a$  is stochastically smaller than that of  $a'$ . Again, the order reversal is a result of our ranking convention.*

**Proof:** Think of linear extensions  $\omega$  as linear Hasse diagrams. The condition  $a \geq a'$  implies that  $a'$  appears lower in the diagram than does  $a$ . Recall that  $\omega(\cdot)$  is the depth in the linear Hasse diagram. The result is then clear from (9) since

$$\{\omega \in \Omega : \omega(a) \leq r\} \supseteq \{\omega \in \Omega : \omega(a') \leq r\}.$$

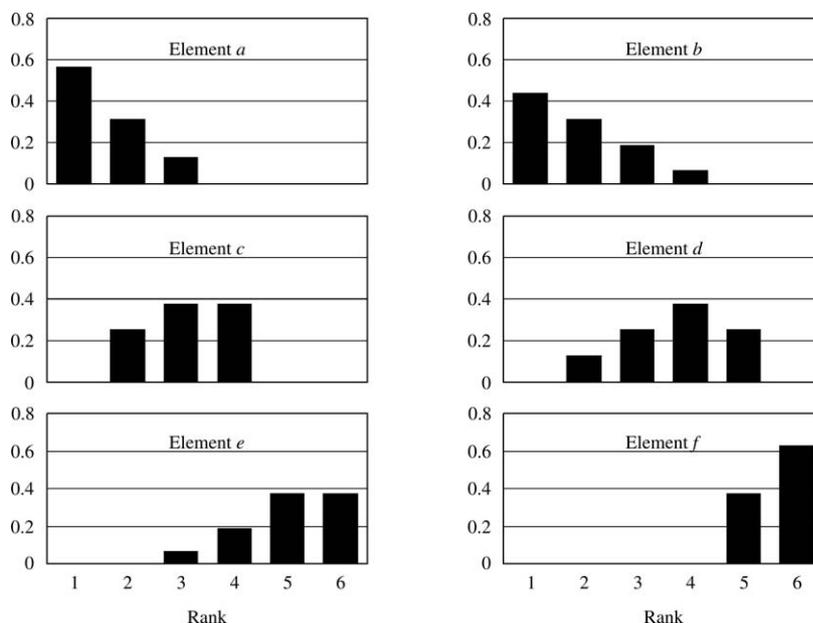
We illustrate these ideas with Poset B in Fig. 5. We need to enumerate all the linear extensions of the poset. The proof of Theorem 1 described an algorithm for doing this by successively removing maximal elements from the poset. The algorithm can be laid out in the form of a decision tree whose branches correspond to the choices of maximal elements available at each stage. The decision tree for Poset B is shown in Fig. 17. In all, there are 16 possible linear extensions compared with the  $6! = 720$  permutations of the six objects in the poset. Fix attention on element  $a$ . Examining the tree, we find that nine linear extensions assign rank 1 to  $a$ . Therefore,  $f_a(1) = 9$ . Similarly,  $f_a(2) = 5$  and  $f_a(3) = 2$ .



**Figure 17.** Hasse diagram of Poset B (left) and a decision tree enumerating all possible linear extensions of the poset (right). Every downward path through the decision tree determines a linear extension. Dashed links in the decision tree are not implied by the partial order and are called jumps. If one tried to trace the linear extension in the original Hasse diagram, a ‘‘jump’’ would be required at each dashed link. Note that there is a pure-jump linear extension (path  $a, b, c, d, e, f$ ) in which every link is a jump.

**Table 3.** Rank-frequency table for the poset of Fig. 17. Each row gives the rank-frequency distribution for the corresponding element of the poset.

Element	Rank						Totals
	1	2	3	4	5	6	
<i>a</i>	9	5	2	0	0	0	16
<i>b</i>	7	5	3	1	0	0	16
<i>c</i>	0	4	6	6	0	0	16
<i>d</i>	0	2	4	6	4	0	16
<i>e</i>	0	0	1	3	6	6	16
<i>f</i>	0	0	0	0	6	10	16
Totals	16	16	16	16	16	16	



**Figure 18.** Histograms of the rank-frequency distributions for Poset B.

The rank-frequency distributions for the other members of the poset are given in Table 3.

Note that each of the row sums is equal to 16—the total number of linear extensions. Histograms of the rank-frequency distributions appear in Fig. 18. Notice that these histograms are all unimodal and are often markedly non-uniform. It is part of the folklore of poset theory that rank-frequency distributions are always unimodal—in fact log-concave. We do not have a ready reference for this result, however. □

## 6. Linearizing a poset: Cumulative rank-frequency operator

This section explores the question of whether there is some canonical procedure for ranking a poset without combining indicators. A hint in this direction is provided by Theorem 3, which suggests that we use stochastic ordering of the rank frequency distributions to define a new partial order on  $S$ . We call this new partial order as the CRF ordering and symbolize it by  $\geq_{\text{CRF}}$ . Specifically, if  $a, a' \in S$ , then

$$a \geq_{\text{CRF}} a' \Leftrightarrow F_a(r) \geq F_{a'}(r), \quad \text{for all } r.$$

Theorem 3 now tells us that the CRF ordering is an extension of the original partial ordering on  $S$ , i.e., that  $a \geq a' \Rightarrow a \geq_{\text{CRF}} a'$ . This procedure for transforming the original partial order into the CRF order is called the CRF operator.

The CRF distributions for Poset B are plotted in Fig. 19. The curves are stacked one above the other, giving a linear ordering of  $S$ :

$$a >_{\text{CRF}} b >_{\text{CRF}} c >_{\text{CRF}} d >_{\text{CRF}} e >_{\text{CRF}} f.$$

It would be very exciting if the CRF ordering were always a linear extension of the original partial order on  $S$ . Unfortunately, this is not the case—Poset B is quite unusual in this respect. Fig. 20 displays a poset for which two iterations of the CRF operator are required to produce a linear ordering. We conjecture that repeated application of the CRF operator eventually terminates in a linear ordering. The process must eventually terminate and the terminal ordering must be a fixed point of the CRF operator. Thus, our conjecture is that linear orders are the only fixed points of the CRF operator. The CRF operator can also produce ties, which reflect symmetries in the structure of the original poset (Fig. 20).

### 6.1 Non-uniform (weighed) distributions

The approach described above treats each linear extension as an equal ‘‘voter’’ in arriving at a final ranking. It may sometimes be preferable to give more weight to certain linear extensions. A simple example might weight each linear extension according to its number of jumps (see Fig. 17). At first, it might seem more natural to weight according to the

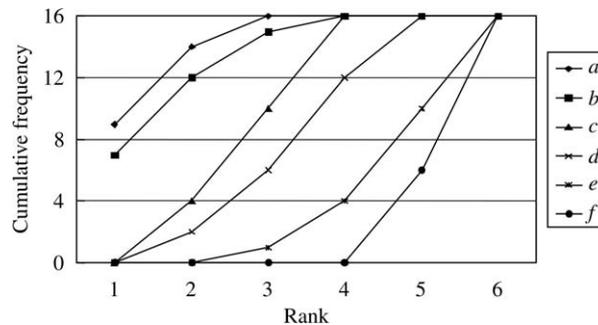
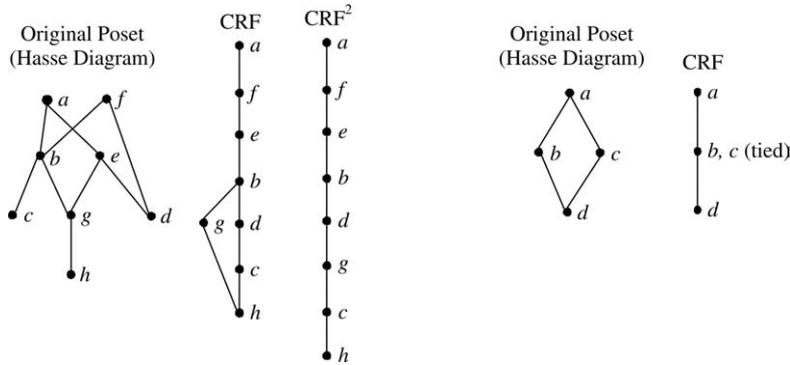


Figure 19. CRF distributions for Poset B.



**Figure 20.** (Left) Two iterations of the CRF operator are required to transform this poset into a linear ordering. (Right) A poset for which the CRF operator produces ties in the final linear ordering.

number of non-jumps since linear extensions with few jumps replicate large portions of the original poset. However, this also has the result that maximal elements from the original poset can appear rather low in the linear extension—consider element  $b$  in the first linear extension of the decision tree in Fig. 17. Thus, weighting according to the number of jumps gives preference to linear extensions for which elements that are high in the original Hasse diagram also tend to be high in the linear extension. Interesting weights can also be constructed using the original suite of indicators. For example, we might scan a linear extension and weight in proportion to the number of links that are consistent with an indicator that is deemed to be particularly important.

In general, suppose we have a nonnegative weight  $w(\omega)$  defined for each linear extension  $\omega \in \Omega$ . This determines a probability distribution,

$$\Pr(\omega) = \frac{w(\omega)}{\sum_{\omega' \in \Omega} w(\omega')}, \tag{10}$$

over the set  $\Omega$  of linear extensions. From this we obtain a (normalized) rank frequency distribution for each element  $a \in S$  as follows:

$$f_a(r) = \Pr\{\omega \in \Omega | \omega(a) = r\}$$

We can compute these weighted rank-frequency distributions in much the same way as was done for the unweighted case in the Table 3. We do need an enumeration of the linear extensions. Then, for each linear extension  $\omega$ , we add the weight  $w(\omega)$  to each cell  $(a, r)$  in the table for which  $\omega(a) = r$ . Note that  $w(\omega)$  will be added to exactly one cell in each row of the table and to exactly one cell in each column of the table. Just as in Table 3, then, the row sums and the column sums are all equal to one another and their common value is the total weight

$$\sum_{\omega \in \Omega} w(\omega). \tag{11}$$

### 6.2 Markov chain Monte Carlo (MCMC) sampling

Except for very small posets, it is computationally impossible to enumerate all possible linear extensions. Their number is simply too large. For instance, an earlier UNEP human

environment poset had 141 members but the number of linear extensions exceeded  $8 \times 10^{105}$ . As an alternative to full enumeration, one can use MCMC methods to estimate the (row-normalized) rank-frequency table. This will entail sampling from the uniform distribution on the set  $\Omega$  of all linear extensions of a given poset. If  $\omega \in \Omega$  is the current linear extension, then transition to the next (proposed) linear extension is accomplished by randomly selecting a jump (see Fig. 17) from  $\omega$  and interchanging its two endpoints. Computer implementation of this procedure is planned for the near future. See Aldous (1987), Brightwell and Winkler (1991), and Haggstrom (2002) for elaboration of MCMC methods applied to discrete data structures.

MCMC methods can also be used to estimate rank-frequency distributions when non-uniform weights are employed. We need to sample from  $\Omega$  according to the probability distribution given in Equation (10). An advantage of MCMC in this context is that it not necessary to know the normalizer (11); MCMC only requires efficient computation of  $w(\omega)$  for any given  $\omega$ .

### 6.3 Handling ties

If two (or more) members of  $S$  have exactly the same set of indicator values, then they are represented by the same point in indicator space and are said to be tied. In case of ties, the anti-symmetry condition fails and the intrinsic ordering of hotspots is a pre-order instead of a partial order. This is a common occurrence in the mathematical poset literature, and the standard solution is to identify all members of a tied set, which produces a partial order on the quotient space. However, in applications, each member of  $S$  has its own unique identity and it is unacceptable to identify different objects simply because they have the same set of indicator values. Our solution starts with the quotient space and its Hasse diagram, just as in the mathematical literature. But, we also attach to each node in the Hasse diagram the integer count of the number of hotspots represented by that node. The count is called the node's ramification index. When there are no ties, each ramification index equals unity. It is essential that the ramification index values be taken into account in doing the MCMC sampling and in compiling the rank-frequency table. Drawing an analogy with football rankings, if two teams are tied for number one then they collectively consume two ranks and the next team receives rank 3. Note that the CRF operator can produce ties even when there are no ties according to the original suite of indicators (Fig. 20). Accordingly, the CRF methodology must address the tie-handling issue.

### 6.4 Measurement and estimation error

To put this issue into perspective, suppose there are two indicators and we wish to compare elements  $a$  and  $b$ . Also, suppose  $I_1(a) = 10$  and  $I_1(b) = 18$ , while  $I_2(a) = 5$  and  $I_2(b) = 4.99$ . Then,  $b$  is better according to  $I_1$ , whereas  $a$  is better according to  $I_2$ . Strict application of the intrinsic ordering says that  $a$  and  $b$  are not comparable. Nonetheless, their  $I_2$ -values are so close (possibly differing only by measurement error) that one might be inclined to order the hotspots according to just  $I_1$ . A similar issue arises in applying the CRF operator where the columns of CRF are the effective indicators. When MCMC is used, the rank-frequencies must be estimated and are therefore subject to estimation error.

In making comparisons, should raw estimates be used or should one use only statistically significant differences? At this stage, we have no prescription for settling these issues but we plan to explore the multiple comparison and fuzzy comparison literature to develop appropriate procedures. The methods of data envelopment analysis and stochastic frontier analysis may also be helpful in this regard; see Charnes *et al.* (1994), Filar and Ross (2001), and Kumbhakar and Knox Lovell (2002).

## Acknowledgments

Prepared with partial support from U.S. EPA Star Grant for Atlantic Slope Consortium, Cooperative Agreement Number R-82868401, and the National Science Foundation Digital Government Program Award Number EIA-0307010. The contents have not been subjected to Agency review and therefore do not necessarily reflect the views of the Agencies and no official endorsement should be inferred.

## References

- Aldous, D. (1987) On the Markov chain simulation method for uniform combinatorial distributions and simulated annealing. *Probability in the Engineering and Informational Sciences*, **1**, 33–46.
- Brightwell, G. and Winkler, P. (1991) Counting linear extensions. *Order*, **8**, 225–42.
- Budach, L., Graw, B., Meinel, C., and Waack, S. (1988) *Algebraic and Topological Properties of Finite Partially Ordered Sets*, Teubner-Texte zur Mathematik, Band 109, Leipzig.
- Charnes, A., Cooper, A.Y., Lewin, A.Y., and Seiford, L.M. (1994) *Data Envelopment Analysis: Theory, Methodology, and Applications*, Kluwer, Boston.
- Cormen, T.H., Leiserson, C.E., Rivest, R.L., and Stein, C. (2001) *Introduction to Algorithms* (second edition), MIT Press, Cambridge, Massachusetts.
- Di Battista, G., Eades, P., Tamassia, R., and Tollis, I.G. (1999) *Graph Drawing: Algorithms for the Visualization of Graphs*, Prentice Hall, Upper Saddle River, New Jersey.
- Filar, J.A. and Ross, N.P. (2001) Generalized data envelopment analysis, and environmental indicators. Invited Paper. Plenary Forum on Environmental Indicators and their Integration for Quality of Life, Index 2001 Congress, Rome, Italy.
- Fishburn, P.C. (1985) *Interval Orders and Interval Graphs: A Study of Partially Ordered Sets*, Wiley, New York.
- Haggstrom, O. (2002) *Finite Markov Chains and Algorithmic Applications*, Cambridge University Press, Cambridge.
- Knuth, D.E. (1973) *The Art of Computer Programming: Volume 1, Fundamental Algorithms* (second edition), Addison-Wesley, Reading, Massachusetts.
- Kumbhakar, S.C. and Knox Lovell, C.A. (2002) *Stochastic Frontier Analysis*, Cambridge University Press, Cambridge.
- Lehmann, E.L. (1986) *Testing Statistical Hypotheses* (second edition), Wiley, New York.
- Neggers, J. and Kim, H.S. (1998) *Basic Posets*, World Scientific, Singapore.
- Sugiyama, K. (2002) *Graph Drawing and Applications*, World Scientific, Singapore.
- Trotter, W.T. (1992) *Combinatorics and Partially Ordered Sets*, Johns Hopkins University Press, Baltimore.
- USEPA (1997) *An Ecological Assessment of the United States Mid-Atlantic Region*, EPA/600/R-97/130, United States Department of Environmental Protection, Office of Research and Development, Washington, DC.

## Biographical Sketches

G.P. Patil is Distinguished Professor of Mathematical Statistics and Director of the Center for Statistical Ecology and Environmental Statistics in the Department of Statistics at The Pennsylvania State University.

C. Taillie is a Senior Research Associate in the Center for Statistical Ecology and Environmental Statistics in the Department of Statistics at The Pennsylvania State University.