EnviroInfo 2007 (Warschau)
Environmental Informatics and Systems Research
Copyright © Shaker Verlag, Aachen 2007. ISBN: 978-3-8322-6397-3

A Lattice-Theoretic Approach to Computing Averaged Ranks Illustrated on Pollution Data in Baden-Württemberg

Karel De Loof[1], Bernard De Baets[2], Hans De Meyer[1], Rainer Brüggemann[3]

Abstract

In the paper an algorithm for calculating the ranking distribution and the averaged ranking probability (averaged rank) of each element of a poset that avoids the enumeration of all linear extensions and relies on the lattice of ideals representation of the poset is presented.

1. Introduction

The theory of partially ordered sets (posets) has been successfully applied in several fields of environmental protection, for example in the analysis of monitoring data of chemical pollution in Baden-Württemberg (e.g. Brüggemann et al. 1999). The visualisation of a partial order by means of a Hasse diagram supports many important conclusions. However, the absence of a unique ranking is often seen as a disadvantage by decision makers. For that reason, the notion of ranking probabilities was developed, which is based on the set of linear extensions, i.e. the set of all possible rankings out of a poset (e.g. Brüggemann et al. 2001). It is well known that the number of all possible rankings of the elements of a poset, can be exponential in the number of its elements. Even the determination of the number of linear extensions of a poset is a so-called #P-complete counting problem (Brightwell and Winkler, 1991). Since presently no counting algorithm with a polynomial time complexity bound is known, it seems likely that the counting problem is no easier than the generation problem. An algorithm to generate all linear extensions of a poset with constant processing time for each linear extension is due to Pruesse and Ruskey (Pruesse and Ruskey, 1994). In this contribution, we introduce. Note that although approximation algorithms for these problems were suggested (see e.g. Lerche and Sørensen 2004, Brüggemann et al. 2004), here we are only interested in exact algorithms.

2. Methodology

In practice, the construction of the lattice of ideals could be done by the algorithm of Habib et al. (Habib, 2001), which is based on an intermediate ideal tree encoding of the lattice of ideals (Habib, 1996). This algorithm generates the lattice of ideals from the covering relation of the poset in constant amortized time, which means that it only needs constant time for each edge in the lattice of ideals, implying the algorithm is optimal up to a constant factor. As their algorithm is beyond the scope of the present contribution we will instead use a naive (non-optimal) algorithm for illustrative purposes.

Once the lattice of ideals is established, we use a two-pass counting algorithm that associates so-called LEF (as an abbreviation for LinearExtensionFilter) and LEI (LinearExtensionIdeal) numbers to each ideal. It is based on the well-known one-to-one correspondence between the linear extensions of a poset and the paths in the lattice of ideals from the source to the sink. Counting the number of linear extensions of a poset with a specific element on a specified rank then amounts to the problem of counting the number

[1] Department of Applied Mathematics and Computer Science, Ghent University, Krijgslaan 281-S9, 9000 Gent, Belgium, e-mail: Karel.DeLoof@ugent.be
[2] Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure Links 653, 9000 Gent, Belgium
[3] Wollgrasweg 30, 15566 Schoeneiche, Germany, e-mail: brg_home@web.de

of paths in the ideal lattice containing the edges labelled with the specific element at the specified height from the source. The first pass of the algorithm executes the counting algorithm from the sink to the source in order to calculate the LEF numbers, the second pass from the source to the sink calculates the LEI numbers. Additionally, to each edge in the lattice of ideals we attach the product of the LEF number of the covering vertex and the LEI number of the covered vertex. The numbers associated to each edge in the lattice then correspond exactly to the number of different paths from the source to the sink passing through this edge. Mutual ranking probabilities, expressing the probability that an element is ranked higher than another element in a linear extension, can also be derived in an analogous way (De Loof et al., 2006).

Since the number of ideals of posets with a limited width is in general much lower than the number of linear extensions, the suggested approach means a major gain in computing time. Not surprisingly, the number of ideals, i.e. the number of vertices in the lattice of ideals, can still be exponential in the number of elements of the poset. However, even in the extreme case of an antichain, the number of ideals equals 2^n, compared to its n! possible linear extensions. Although our approach clearly needs more memory since the lattice of ideals needs to be physically stored into memory, there is a huge class of posets for which the outlined approach makes the exact calculation of the ranking probabilities feasible.

3. Results

In this contribution we aim at ranking regions in Baden-Württemberg according to their pollution (Figure 1). Having the ranking distributions at our disposal, we can easily obtain the averaged rank of each element (Table 1). Mapping each element in our starting poset to its averaged rank induces a weak-order extension of this poset. Such a weak-order extension can be intuitively seen as a linear order on equivalence classes. Ties, i.e. multiple elements in an equivalence class, can always appear since different elements can give rise to an equal averaged rank. In our application one region (57) shows up to be ranked highest in the linear order induced by the averaged ranks. This region will form our region of highest concern. In order to determine how sharp we can state this, we analyse the ranking probability distribution of element 57 and we find that its lowest possible rank is 4 and its highest 9, indicating a quite large uncertainty.

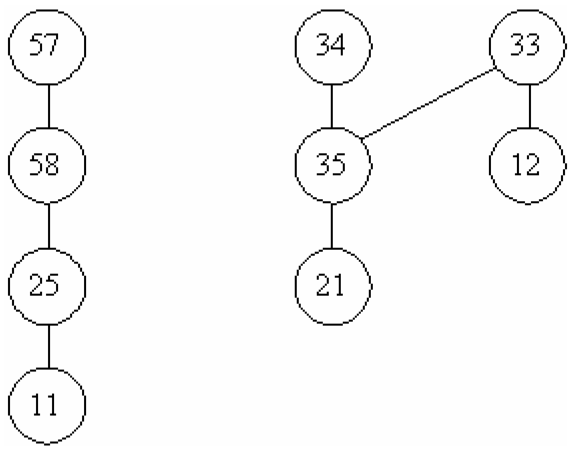

Figure 1: Partial order between granitic regions of Baden-Württemberg induced by their Lead (Pb) and Cadmium (Cd) pollution of the herb layer

Table 1:
Rank distributions and averaged ranks of the regions

rank region	1	2	3	4	5	6	7	8	9	r_{av}
11	0,44	0,28	0,16	0,08	0,03	0,01	0,00	0,00	0,00	2,00
25	0,00	0,17	0,24	0,24	0,19	0,12	0,05	0,00	0,00	4,00
58	0,00	0,00	0,05	0,12	0,19	0,24	0,24	0,17	0,00	6,00
57	0,00	0,00	0,00	0,01	0,03	0,08	0,16	0,28	0,44	8,00
21	0,40	0,28	0,18	0,10	0,04	0,01	0,00	0,00	0,00	2,14
35	0,00	0,12	0,20	0,24	0,22	0,15	0,07	0,00	0,00	4,29
34	0,00	0,00	0,02	0,05	0,10	0,15	0,20	0,24	0,24	7,14
12	0,16	0,16	0,16	0,15	0,14	0,11	0,08	0,04	0,00	3,81
33	0,00	0,00	0,00	0,02	0,06	0,12	0,20	0,28	0,32	7,62

4. Discussion

Although from an application point of view no new results are obtained by this approach, it is clear that the more efficient computation of the ranking distributions (as well as the mutual ranking distributions) of the elements of a large class of partially ordered sets is made feasible. Presently, work is undertaken on an efficient implementation of these algorithms in order to integrate them into existing software frameworks such as WHASSE (Brüggemann, 1995).

Bibliography

Brightwell G., Winkler P. (1991): Counting linear extensions, Order, **8**(3): 225-242.

Brüggemann R., Halfon E., Bücherl C. (1995): Theoretical base of the program „Hasse“, GSF-Bericht 20/95 Neuherberg.

Brüggemann, R., Pudenz, S., Voigt, K., Kaune, A., Kreimes, K. (1999): An algebraic/graphical tool to compare ecosystems with respect to their pollution. IV: Comparative regional analysis by Boolean arithmetics, Chemosphere, **38**: 2263-2279.

Brüggemann, R., Halfon, E., Welzl, G., Voigt, K., Steinberg, C. (2001): Applying the Concept of Partially Ordered Sets on the Ranking of Near-Shore Sediments by a Battery of Tests, Journal for Chemical Information and Computer Scientists, **41**: 918-925.

Brüggemann, R., Sørensen, P.B., Lerche, D., Carlsen, L. (2004): Estimation of averaged ranks by a local partial order model, Journal for Chemical Information and Computer Scientists, **44**: 618-625.

De Loof, K., De Meyer, H., De Baets, B. (2006) Exploiting the lattice of ideals representation of a poset, Fundamenta Informaticae, **71**: 309-321.

Habib, M., Medina, R., Nourine, L., Steiner, G. (2001): Efficient algorithms on distributive lattices and its applications, Discrete Applied Mathematics, **110**: 169-187.

Habib, M., Nourine, L. (1996): Tree structure for distributive lattices and its applications, Theoretical Computer Science, **165**: 391-405.

Lerche, D., Sørensen, P.B. (2004): Evaluation of the ranking probabilities for partial orders based on random linear extensions, Chemosphere, **53**: 981-992.

Pruesse, G., Ruskey, F. (1994): Generating linear extensions fast, SIAM Journal on Computing, **23**: 373-386.