EnviroInfo 2008 (Lüneburg)
Environmental Informatics and Industrial Ecology
Copyright © Shaker Verlag, Aachen 2008. ISBN: 978-3-8322-7313-2

PYHASSE a New Software Tool for Partially Ordered Sets Example: Ranking of Contaminated Regions in the Alps

Kristina Voigt[1], Rainer Brüggemann[2], Karl-Werner Schramm[3] and Manfred Kirchner[4]

Abstract

In complex environmental data-sets it is often necessary to compare different sets of criteria (attributes). In the similarity analysis which is part of theory of partially ordered sets (Hasse Diagram Technique) we intend to calculate the similarity of different posets. This similarity analysis is an important feature of the newly developed software packages called PYHASSE developed by the second author. As an example we use a data-matrix which was generated in an international project named MONARPOP (Monitoring Network in the Alpine Region for Persistent and other Organic Pollutants) in which selected chemicals in environmental media in the mountain area of the Alps were analyzed in the years 2004 and 2005 (MONARPOP, 2008). 17 pesticides were chosen and analyzed in soil samples in Germany, Austria, Switzerland, Italy, and Slovenia. The samples were taken at different heights. In this approach we took a look at soil samples at different heights in the five mentioned countries. Two soil types were analyzed: humus and mineral soil layer. So we perform a similarity analysis of the two different soils (humus and mineral soil) in order to get an idea if the type of soil has a great impact on the ranking of the pesticides and hence on the contamination. It is demonstrated that the different soil layers have some kind of influence but not a great impact on the structure of the Hasse Diagrams.

1. Introduction: Hasse Diagrams and its Software

Partial order is a discipline of Discrete Mathematics and one may consider partial order as an example of mathematics without arithmetic. A good overview can be found in a book edited by the second author (Brüggemann, 2006). The graphical representation of partial orders is laid down in so-called Hasse Diagrams. The first software was already written in the nineteen eighties and was developed under MS-DOS. In the nineteen ninetieth the software was adapted to the MS-Windows platform. The functionalities were constantly enhanced and improved in the following years (Brüggemann, 2005). The innovative tool called METEOR (Method of Evaluation by Order Theory) which attempts to resolve the incomparabilities among objects by inclusion of external knowledge is incorporated in the WHASSE program (Brüggemann, 2008, Voigt, 2008). This software named WHASSE written in Delphi has always been available for scientific purposes from the second author free of charge. However, new features (e.g. like the similarity analysis or the fuzzy Hasse) made it necessary to develop a new software written in PYTHON (http://www.python.org/) being more flexible than the older software.

In complex environmental data-sets it is often necessary to compare different sets of criteria (attributes). In the similarity analysis we intend to calculate the similarity of different posets. This similarity analysis is an important feature of the newly developed software package called PYHASSE. PYHASSE is elaborated

1 Helmholtz Zentrum Muenchen - German Research Center for Environmental Health (GmbH), IBB, Neuherberg, Germany, kvoigt@helmholtz-muenchen.de
2 Leibniz- Institute of Fresh Water Ecology and Inland Fisheries, Berlin, Germany
3 Helmholtz Zentrum Muenchen - German Research Center for Environmental Health (GmbH), IÖC, Neuherberg, Germany

by the second author and is still under development. Test versions are, however, available from Rainer Brüggemann.

2. Similarity Analysis in PYHASSE

2.1 Evaluation by Hasse Diagram Technique

Consider the two Hasse diagrams, shown in Figure 1.

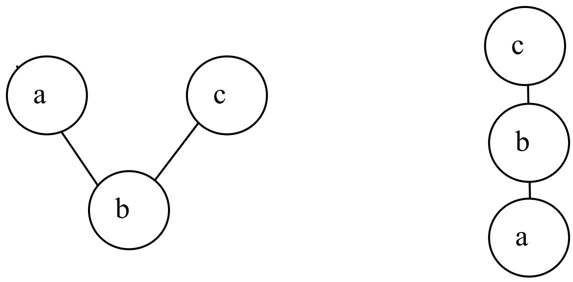

Fig. 1: Two Hasse diagrams to demonstrate the similarity analysis

There is one contradiction, namely a ≤ b in the Hasse diagram (rhs) and b ≤ a in the other Hasse diagram (lhs). There is one coincidence, namely c ≥ b. A matrix-like scheme gives a comprehensive view:

$$
\begin{array}{cc}
 & \begin{array}{ccc} a & b & c \end{array} \\
\begin{array}{c} a \\ b \\ c \end{array} & \left[\begin{array}{ccc} - & >< & \|< \\ <> & - & << \\ \|> & >> & - \end{array}\right]
\end{array}
$$

Beside the diagonal which is not of interest, the entries are found by comparing the row defining element with the column defining one. The first symbol belongs to the lhs-, the second to the rhs-Hasse diagram. Entries like >> or << are contributions, which are counting the ‘isotone’ character of both partial orders; entries like ><, <> contribute to the ‘antitone’ character, i.e. to the conflicts between the two partial orders, <||, >||, ||>,||< , =||, ||= or || || are considered as indifferent, combinations like > =, < =, = <, = > are called weak isotone. Finally the entry of type = = contributes to equivalence relations. There are inherent symmetries, so that it is sufficient to examine only the upper or lower triangular part of the matrix.

In the example above we count: 1 isotone, 1 antitone, 0 weak isotone, 1 indifferent and 0 equivalence relations. One may also be interested, which object is mainly contributing to conflicts, i.e. to the antitone character of the two partial orders. In the case above the elements a and b contribute once, whereas c does not have any counter current entry.

Copyright © Shaker Verlag, Aachen 2008. ISBN: 978-3-8322-7313-2

2.2 Similarity Analysis of Humus Soil Versus Mineral Soil

As an example we use a data-matrix which was generated in an international project named MONARPOP (Monitoring Network in the Alpine Region for Persistent and other Organic Pollutants) in which selected chemicals in environmental media in the mountain area of the Alps were analyzed in the years 2004 and 2005 (MONARPOP, 2008). 17 pesticides were chosen and analyzed in soil samples in Germany, Austria, Switzerland, Italy, and Slovenia. The samples were taken at different heights. In this approach we took a look at soil samples at different heights in the five mentioned countries. Two soil types were analyzed: humus and mineral soil. Our aim is to get an idea if the type of soil has a great impact on the ranking of the pesticides and hence on the contamination.

Table 1 lists the objects to be analyzed. Most of the chosen chemicals belong to the group of Persistent Organic Pollutants (POPs). The initial twelve POPs chemicals, known as 'poisons without passports', pose particular hazards because of their common characteristics. They are toxic to humans and wildlife. They are persistent and resist breaking down. POPs are semivolatile and mobile, travelling great distances on wind and water currents and are widely distributed throughout the environment. Through global distillation, they travel from temperate and tropical regions to be deposited in the colder regions of the poles (IPEN, 2005) and as demonstrated in the MONAPPOP project in the Alps. Being aware of the immense damage these chemicals are performing the Stockholm Convention on Persistent Organic Pollutants was established and came into force in 2004. The objective of the Stockholm Convention on Persistent Organic Pollutants is to protect human health and the environment from POPs. Unlike other chemical treaties that rely primarily on notification requirements or end-of-life management controls, the POPs Convention aims to eliminate the production, use and emissions of POPs. It also aims to ensure the environmentally sound destruction of POPs waste stockpiles, as well as preventing the introduction of new chemicals with POP-like characteristics.

Common Name	Trivial Name	Acronym
Alpha Hexachlorocyclohexane	alpha-HCH	AHCH
Beta Hexachlorocyclohexane	beta-HCH	BHCH
Gamma Hexachlorocyclohexane	gamma-HCH	GHCH
Delta Hexachlorocyclohexane	delta-HCH	DHCH
Epsilon Hexachlorocyclohexane	epsilon-HCH	EHCH
p,p’ Dichloro Diphenyl Trichloroethane	p,p’-DDT	PPDT
o,p’ Dichloro Diphenyl Trichloroethane	o,p’-DDT	OPDT
p,p’-Dichloro Diphenyl Dichloroethane	p,p’-DDD	PPDD
o,p’-Dichloro Diphenyl Dichloroethane	o,p’-DDD	OPDD
p,p’-Dichloro Diphenyl Dichloroethane	p,p’-DDE	PPDE
o,p’-Dichloro Diphenyl Dichloroethane	o,p’-DDE	OPDE
Aldrin	Aldrin	ALDR
Dieldrin	Dieldrin	DIEL
Endrin	Endrin	ENDR
Mirex	Mirex	MIRE
Hexachlorobenzene	HCB	HCBE
Heptachlor	Heptachlor	HECL

Tab. 1: Pesticides Analyzed (Trivial Name, Acronym)

Copyright © Shaker Verlag, Aachen 2008. ISBN: 978-3-8322-7313-2

It is obvious that performing sound data-analysis strategies will support reaching the aims of the convention. The ranking method applied here in the first place identifies the comparabilities and incomparabilities among chemicals as well as shows which chemicals are better or worst than the others.

Two different data-sets are analyzed: 17 pollutants in humus soil data in all five countries (17 x 56) and in mineral soils (17 x 25). The question to answer is the following: Are the two different resulting Hasse Diagrams similar and to which degree? The Hasse Diagrams are given in Figure 2a (17x56) and Figure 2b (17x25).

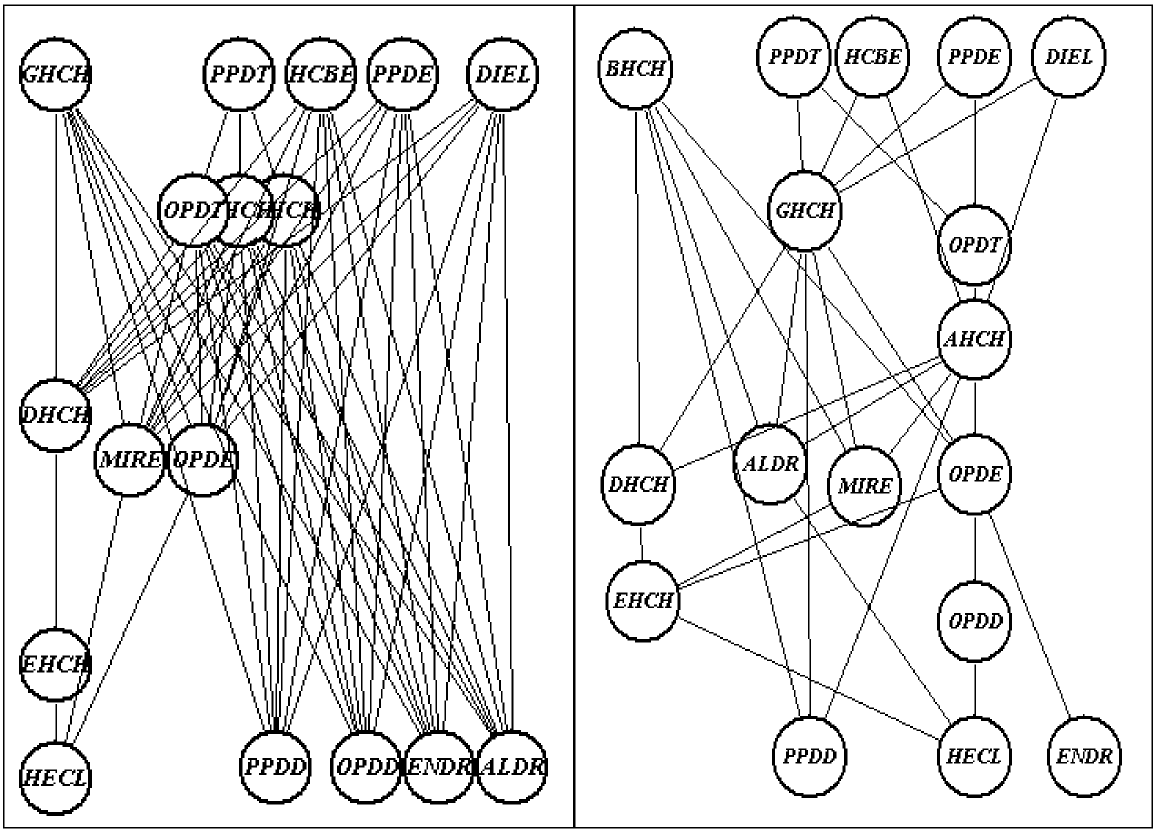

Fig. 2: HD 17 Pesticides x 56 Humus Soils (lhs) HD 17 Pesticides x 25 Mineral Soils (rhs)

Both Hasse Diagrams show a lot of similarities but also quite a few differences. The diagram of the 17x56 data-matrix (lhs) has five levels whereas the HD of the 17x25 data-matrix shows six levels. Concerning the maximal objects (on top of the diagrams) both analyses give the same number, namely 5. Only one chemical is different. Taking a look at the minimal objects (at the bottom of each diagram) we encounter more differences not only in the number but also the chemicals themselves. The differences are also given in the comparabilities and incomparabilities.

The similarity analysis should give some more insight into these differences. So the two data-matrices 17x56 (humus soil) and 17x25 (mineral soil) are now subject to the similarity analysis.

Figure 3 gives a screenshot of the User interface of the Python-program "similarity" performing the similarity analysis. Two files (17x56) and (17x25) have to be uploaded. Both Hasse Diagrams (see Figure 2a and b) can be displayed as well as their initial data-matrices. In Figure 4 the results of the similarity analysis are given.

Copyright © Shaker Verlag, Aachen 2008. ISBN: 978-3-8322-7313-2

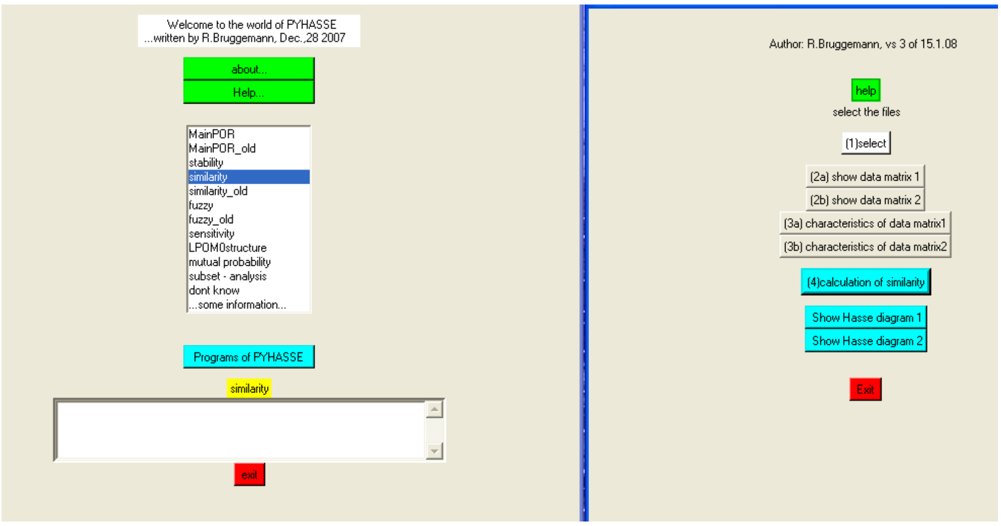

Fig. 3: PyHasse Similarity Analysis

Only isotone and indifferent and no antitione relations are calculated. 158 isotone relations are given and 114 indifferent relations. This means that the different soil layer has some kind of influence but not a great impact on the structure of the Hasse Diagrams. Isotone relations demonstrate a high degree of similarity of the two data-sets whereas indifferent relations mean incomparability among the relations.

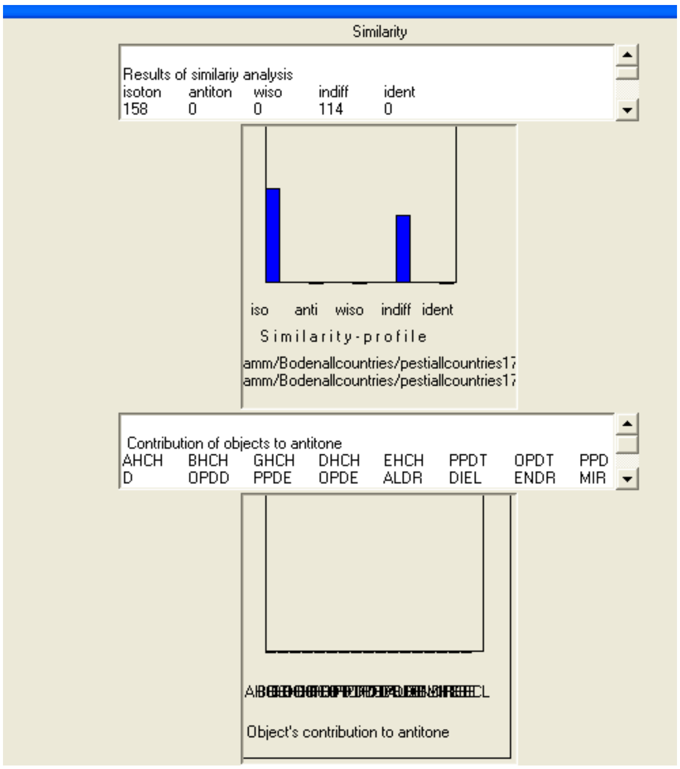

Fig. 4: Results of Similarity Analysis of two Data-sets (17x56 and 17x25) Humus / Mineral Soil

Copyright © Shaker Verlag, Aachen 2008. ISBN: 978-3-8322-7313-2

3. Discussion and Outlook

A PYTHON program allows a deepened graph theoretical analysis of the directed graphs, generated by the partial order relation. For the programming language PYTHON, see for example (Weigend, 2006) or (Lutz, 2003). In the current data-analysis the test-sets of 17 pesticides in 56 humus soil samples and the same 17 pesticides in 25 mineral soil samples were looked upon using the newly developed feature of similarity analysis. It exhibited that the type of soil only showed some indifferent relations but did not have a great impact on the Hasse Diagram. This means that the pesticides are found not only at the first soil layer the humus layer but also in considerably comparable concentrations in the mineral soil layer which is situated underneath the humus layer. The performed similarity analysis could give some support to stakeholders where further research should or could be initiated and which kind of samples could be taken and for which targets. As the same chemicals have already been tested in needles the similarity analysis of soil and needles will be performed in the near future in order to see if the target is of great importance concerning the concentration of chemicals in the Alps.

A direct influence by different stakeholders into the ranking procedure is given by the Method of Evaluation by Order Theory only available in the WHASSE program. The concepts behind this are the so-called stability fields and hot spots. The basics of this are given in two recent publications by the first and second authors (Brüggemann, 2008; Voigt, 2007, 2008).

References

Brüggemann, R., Simon, U., Mey, S. (2005): Estimation of averaged ranks by extended local partial order models. MATCH Communications in Mathematical and Computer Chemistry, 54, pp. 489-518.

Brüggemann, R., Carlson, L. (Eds.) (2006): Partial Orders in Environmental Sciences and Chemistry, Springer-Verlag, Berlin, pp. 1-406.

Brüggemann, R., Voigt, K., Restrepo, G., Simon, U. (2008): Concept of Stability Fields and Hot Spots in Ranking of Environmental Chemicals, Environmental Modelling & Software, 23, pp. 1000-1012.

Lutz, M., Ascher, D. (2003): Learning Python. O'Reilly Publisher, Beijing (CH), pp. 1-591.

IPEN (2005): IPEN Body Burden Community Monitoring Handbook, http://www.oztoxics.org/cmwg/chemicals/pop%20chem.html.

MONARPOP, Monitoring Network in the Alpine Region for Persistent and other Organic Pollutants http://www.monarpop.at/, 2008.

Voigt, K., Brüggemann, R., Pudenz, S. (2006): A multi-criteria evaluation of environmental databases using the Hasse Diagram Technique (ProRank) software, Environmental Modelling & Software, 21, pp. 1587-1597.

Voigt, K., Brüggemann, R. (2007): Data-Availability of Pharmaceuticals detected in water: An evaluation study by order theory (METEOR), in: Hryniewicz Olgierd, Studzinski Jan, Szediw Anna, Enviro-Info Warsaw 2007, Environmental Informatics and Systems Research, Volume 2: Workshop and application papers, pp. 201-208, Shaker Verlag, Aachen.

Voigt K., Brüggemann R. (2008 accepted): Ranking of Pharmaceuticals Detected in the Environment: Aggregation and Weighting Procedures, submitted to Combinatorial Chemistry & High Throughput Screening.

Weigend, M. (2006): Objektorientierte Programmierung mit Python. mitp-Verlag, Bonn (GE), pp. 1-700.

Copyright © Shaker Verlag, Aachen 2008. ISBN: 978-3-8322-7313-2