Formal concept analysis – a method for exploring complex responses of tourist surveys

Karen Ziener

Abstract — The Formal concept analysis goes back to the 1940s and is advanced by Rudolf Wille in the eighties. Based on the algebraic lattice theory so-called concept lattices will be used for visualisation of hierarchical structures in a line diagram. The method needs as input a context consisting of sets of objects and attributes and a binary relation between these objects and attributes. In a concept lattice quantitative and qualitative data can be combined and relations between the objects as well as between the attributes can be analysed. Therefore concept lattices are suitable for analysing the results of surveys. The respondents and their answer patterns constitute the context. Precondition is a hierarchical structure of the answers (e.g. ordinal ordered data). In this presentation will be demonstrated how to use the line graph of the concept lattice for analysing complex questions of tourist surveys. As an example the main activities of about 600 tourists and daily visitors in the Lake Neusiedl Region are analysed. With the procedure of Formal Concept Analysis the different answer patterns of respondents are arranged as nodes in a line diagram. This diagram is constructed by means of two partial concept lattices that will be integrated later. The interpretation of the line graph will start at the biggest nodes, consider the whole structure of the graph and include additional attributes in order to describe groups of tourists with the same activity pattern.

Index Terms — Formal Concept Analysis, hierarchical structures, Lake Neusiedl Region, tourist surveys

1 Formal concept analysis as a method for analysing complex responses of tourist surveys

The Formal concept analysis goes back to the 1940s and is advanced by Rudolf Wille in the eighties. Based on the algebraic lattice theory so-called concept lattices will be used for visualisation of hierarchical structures in a line diagram. The method needs as input a formal context (matrix) consisting of sets of objects and attributes and a binary relation between these objects and attributes. In a concept lattice quantitative and qualitative data can be combined and relations between the objects as well as between the attributes can be analysed. The line diagram shows the order of objects, on the one hand, and the structure of the whole issue, on the other. In contrast to the creation of types the attributes of each object are unmodified visible.

Therefore concept lattices are suitable for analysing the results of surveys with mostly binary and nominal scale data. The respondents and their answer patterns constitute the context. Precondition is a hierarchical structure of the answers, e.g. ordinal scale data, attribute is not existent / is existent, less or more attributes are existent. Since the beginning of the nineties at the Geographical Institute of the Potsdam University Formal Concept Analysis has been used for analysing different results of surveys, see [1], pp. 80-88, [2], pp. 43-49, 127-129.

Karen Ziener, Dr. habil, geographer, currently self-employed, Klagenfurt, Austria. E-mail: kziener@web.de.

2 Activities of tourists and one day visitors in the Lake Neusiedl Region – construction and analysis of line diagrams

The mathematical foundations, which can be found in Ganter & Wille [3], are not subject of this paper. In this presentation should be demonstrated how to use the line graph of the concept lattice for analysing complex questions of tourist surveys.

TABLE 1

TOURIST ACTIVITIES IN LAKE NEUSIEDL REGION 2006

Activities	Percent
To relax	86,0
Being in nature	85,5
(Wine)tavern	64,8
Cycling	61,1
Bird watching	48,1
Bathing	46,4
Culture / Museums	39,1
Water sports	21,4
Sports	14,3

Source: questioning in the MaB-Project "Redesigning Biosphere Reserve Neusiedler See" in 2006, 588 persons

As an example the activities of 588 tourists and daily visitors in the Lake Neusiedl Region (questioning in the framework of the MaB-Project "Redesigning the Biosphere Reserve Neusiedler See", financed by Austrian Academy of Sciences, MAB Programme, in 2006) are analysed. The respondents were given a list of nine activities to say "yes" or "no", the result you can see in Table 1. But the aim was to get more information about the combination of activities, the activity patterns.

The combination of all activities shows 168 patterns of answer, most of them said by only one, two or three persons. In order to reduce this variety and to take only the main activity patterns into account the context include exclusively those answer combinations as objects which were mentioned by not less than 6 respondents (one percent). The formal context consists of 30 objects (respondents with the same activity pattern) and the 9 activities as attributes. It represents 337 persons (57 percent off all respondents). Now the concept lattice with 38 concepts can be created by means of Formal Concept Analysis software. The line diagram is hand-made because it should be human readable.

For a better construction of such complex line diagram a second step of preparation is required: segmentation of the formal context into two contexts with less attributes. In this case the first context includes the main activities (4) and all 30 objects and the created concept lattice has 7 concepts. The second context covers the other activities (5) and also all objects. The outcome of this is a concept lattice with 13 concepts. Later the two line diagrams of these concept lattices will be combined to the line diagram of the whole concept lattice (38 concepts respectively nodes).

In the line diagrams the different answer patterns of respondents are arranged as nodes. The lines show the relations between the objects representing through the change of attributes. The size of the nodes shows the number of persons with the same activity pattern. Going along the lines to the top of the graph the attributes of each node can be read.

2.1 First partial concept lattice – main activities

The analysis of the line graph of the first concept lattice (Fig. 1) starts at the biggest node. This node (173 persons) on the bottom of the graph combines all main activities. The opposite is the node on the top (6 persons) without any activities. Excepting these six persons all respondents make activities in nature and most of them like to relax. The attribute "nature" is standing at the second node an all other nodes are below it. But you can see a dichotomy – tourists making cycling-tours but they don't go to a tavern, in the left part of the line diagram, and tourist who visit restaurants and taverns and don't like cycling, in the right part.

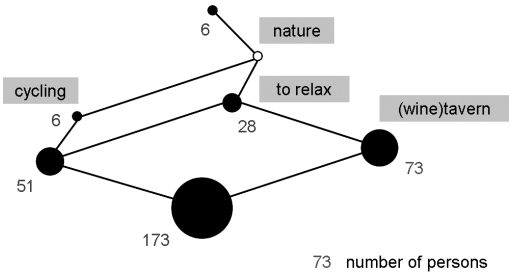

Fig. 1 First partial concept lattice (main activities)
Source: own compilation, survey in the MaB-Project "Redesigning Biosphere Reserve Neusiedler See" 2006

2.2 Second partial concept lattice – the other activities

The line diagram of the second concept lattice shows three activities – "bathing", "bird watching" and "museums" – in three different directions forming a cuboid. In addition water sports and sports are included (see Fig. 2).

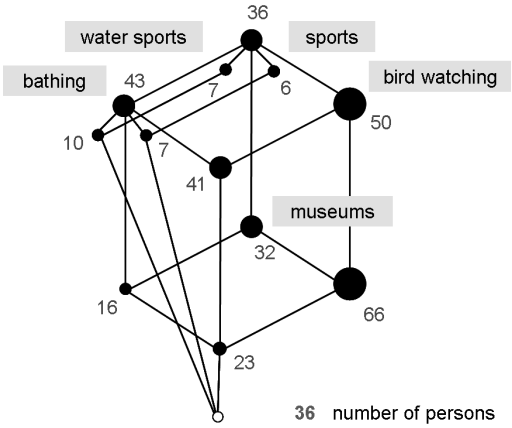

Fig. 2 Second partial concept lattice (the other activities)
Source: own compilation, survey in the MaB-Project "Redesigning Biosphere Reserve Neusiedler See" 2006

In this graph you can explore the various activity patterns. On the right below the node with the label "bird watching" you can see all activity patterns with bird watching. Most of them combine only bird watching and partly museums, but not bathing, sports or water sports. On the opposite side combinations of bathing, sports or water sports are located. The size of nodes shows the different number of respondents having these activity patterns. Of course there is also a relation between "bathing" and "bird watching" (nodes in front of the graph with 43 and 23 persons). The lower level of line diagram shows the interests of culture and museum. Here you can find the biggest node (66 respondents) with the activities "museums" and "bird watching".

2.3 Whole concept lattice – all activities

The construction of the line diagram for the whole concept lattice of all nine activities is based on these two small line diagrams. Initially the first graph is drawn to a larger scale. Afterwards the second line diagram is integrating in each node. However, this graph is only in the bottom node completely developed. Finally the lines between the nodes are supplemented.

Fig. 3 presents the complex line diagram developed in this manner. If you can read it you will get a lot of information in a very short time. The rules for analysing are the same. The graph combines two different structures – the main four activities and the other five activities (see 2.1 and 2.2). I want to explain only some aspects:

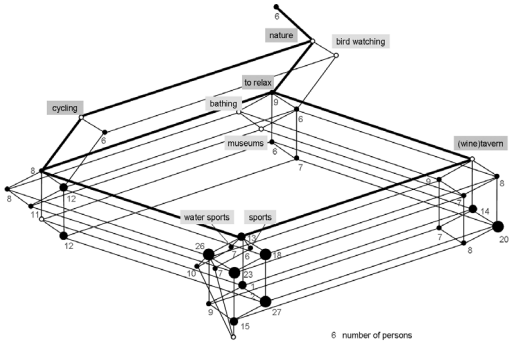

Fig. 3 Activity patterns in the Lake Neusiedl Region 2006,
Source: own compilation, survey in the MaB-Project "Redesigning Biosphere Reserve Neusiedler See" 2006

Most of the bigger nodes you can see in the lower part of the line diagram. These activity patterns combine a broad range of interests. Water sports and sports appear primary

in this bottom part, but their importance is very low.

The biggest node on the right side indicates the interesting relation between culture, nature and bird watching as well as the visit of taverns or wine taverns (so-called Heurige). This is a potential for an enlargement of the very short summer season.

The importance of "bathing" decreases not only quantitative and relative but also as sole activity. Bathing alone or with one or two activities doesn't exist in the concept lattice. The biggest nodes with bathing are situated on the bottom of the graph. This means that bathing is combined with several other activities.

2.4 Inclusion of additional attributes

A further possibility to explore complex responses of touristic surveys is the integration of additional attributes. In this case the time of questioning – spring, summer, autumn – should be correlated with the activities of tourist and one day visitors. In Fig 4 is shown in which season the different activity patterns dominate (residual smallest 1). The thick lines show all nodes with bathing in order to point up the relation to the summer tourism. In spring activity pattern are especially oriented to cycling and bird watching and in autumn we can find more complex interests with museums and wine taverns. Besides, some of the complex interests occur evenly in all seasons.

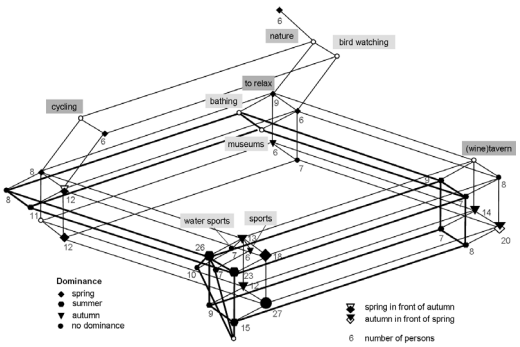

Fig. 4 Activity patterns in the Lake Neusiedl Region 2006 – time of survey, Source: own compilation, survey in the MaB-Project "Redesigning Biosphere Reserve Neusiedler See" 2006

Conclusion

Formal Concept Analysis is a world-wide used method of data analysis which identifies hierarchical structures between different objects. It can be applied in decision and modelling processes. For the analysis and interpretation of the results of questioning this method offers many possibilities. By means of concept lattices complex questions can be structured and relations between statements, opinions or perceptions as well as with other attributes of the persons or groups can be analysed. The line diagram of a concept lattice shows the whole structure of the answer patterns and the relations among it. The visual analysis of the graph provides the formation of groups and types if the structure of data is suitable. Until now the author of this paper is short on experience of attribute implications in concept lattices which can help to explore the relations between different answers and within complex answer patterns for a sample of tourists or visitors in protected areas.

References

[1] G. Saupe and K. Ziener, "Begriffsverbände zur Analyse von Nachfragestrukturen in Erholungs- und Freizeitgebieten", *Beiträge des 10. Kolloquiums für Theorie und quantitative Methoden in der Geographie (Göhren auf Rügen, 23. - 26.2.1994)*, pp. 80-88, Greifswald, 1994

[2] K. Ziener, „Das Bild des Touristen in Nationalparken und Biosphärenreservaten im Spiegel von Befragungen", pp. 43-49, 127-129, Potsdam

[3] B. Ganter and R. Wille, "Formal Concept Analysis: Mathematical Foundations", Berlin, Heidelberg, 1999

Karen Ziener, Dr. phil (1990) Dr. rer. nat. habil. (2003), Scientific Assistant at the University of Potsdam (Germany), Institute of Geography (1987-2000), and at the Alps-Adriatic-University of Klagenfurt (Austria), Department of Geography and Regional Studies (2003-2007), currently self-employed geographer, last project with research grants "Redesigning the Biosphere Reserve Neusiedler See", Books: Grenzüberschreitende Wirtschaftskooperationen zwischen Kärnten, Slowenien und Friaul-Julisch Venetien (2007, collective with P. Hössl), Auf dem Weg zu einer regionalen und vorausschauenden Konfliktstrategie

für Nationalparke und Biosphärenreservate: Sechs Regionen unter die Lupe genommen (2005), Das Konfliktfeld Erholungsnutzung – Naturschutz in Nationalparken und Biosphärenreservaten (2003, Habilitation Thesis), Das Bild des Touristen in Nationalparken und Biosphärenreservaten im Spiegel von Befragungen (2001), Current research interests: conflicts in national parks and biosphere reserves, regional development and spatial planning, cross-border economic cooperation, member of Austrian Geographic Society (ÖGG), Society for German Research (GfD), German Society for Tourism Science (DGT)