Three Generalisations of Lattice Distributivity: An FCA Perspective

Dissertation

zur Erlangung des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt

der Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

von

Dipl.-Math. Heiko Reppe
geboren am 24. Dezember 1976 in Räckelwitz

Gutachter:	Prof. Dr. Bernhard Ganter, Technische Universität Dresden
	Prof. Dr. Marcus Greferath, University College Dublin
Eingereicht am:	12. Oktober 2010
Tag der Disputation:	27. Januar 2011

Berichte aus der Mathematik

Heiko Reppe

Three Generalisations of Lattice Distributivity: An FCA Perspective

Shaker Verlag
Aachen 2011

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Zugl.: Dresden, Techn. Univ., Diss., 2011

Printed in Germany.

ISBN 978-3-8440-0037-5
ISSN 0945-0882

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • e-mail: info@shaker.de

Preface

In the present study we investigate concept lattices and their corrsponding formal contexts $\mathbb{K} = (G, M, I)$ that are usually given by a cross-table and from wich we infer a set of implications. The triple $\mathbb{K}$ consists of two sets, namely the set of objects G, which refers to the rows of the table, and the set of attributes M, which refers to the names of the columns of the cross-table. The triple is complete with an incidence relation I that states that an object g possesses an attribute m if and only if in row g there is a cross in the column m. The concept lattice $\underline{\mathfrak{B}}(\mathbb{K})$ arises from $\mathbb{K} = (G, M, I)$ by collecting all maximal rectangles, which are full of crosses (regardless of succession of rows and columns). Such a maximal rectangle is characterised by the collection of its objects $A \subseteq G$ and its attributes $B \subseteq M$. The pair (A, B) is called formal concept. All formal concepts can naturally be ordered and thereby form a complete lattice.

Moreover, such a concept lattice $\underline{\mathfrak{B}}(\mathbb{K})$ is compounded by two closure systems, namely $\mathrm{Ext}(\mathbb{K}) \subseteq \mathfrak{P}(G)$, the system of extents, and $\mathrm{Int}(\mathbb{K}) \subseteq \mathfrak{P}(M)$, the system of intents. When ordered by inclusion, these two form dually isomorphic lattices. Instead of studying the closure systems, we investigate their closure operators, which are generated from some set of implications $\mathcal{L}$. Of particular interest for applications are small sets of implications that characterise the closure operator. Likewise, implications are of interest, the premises of which are small in size. To obtain the closure of an attribute set, the implications have to be applied iteratively. Thus, it is preferable to deal with sets of implications that do not need to be applied frequently.

Concept lattices are complete lattices. Moreover, every complete lattice is isomorphic to a concept lattice. A complete lattice can be regarded as an algebraic structure $\underline{L} = (L, \vee, \wedge)$ of the type $(2, 2)$ as well as a relational structure $(L, \leq)$.

The three subjects (cross-table, concept lattice, set of implications) find their opponent in data management. The cross-table is an abstraction of a data-base. The concept lattice is a representation of the information within the data-base. It organises objects in sets with common attribute sets and allows navigation within the data. Finally, the implications are inferences and dependencies of the data-base.

The theoretical background for our analysis is founded in the algebraic nature of a concept lattice $\underline{\mathfrak{B}}(\mathbb{K})$. We use contributions from two mathematical fields, namely *order theory* and *universal algebra*. A lattice and a concept lattice in particular possesses an order relation induced by the inclusion of extents (and dually the superset relation of intents). At the same time it allows to determine the greatest lower bound and the least upper bound for every set of concepts. Thus, a concept lattice is an algebraic structure with two binary operations (which, in addition to it, extends to operations of arbitrary arity). It is the field of universal algebra that investigates those structures in general.

Application	Data-Base	Data-Representation	Data-Analysis
Abstraction	*Pattern*	*Structure*	*Inference*
3 Subjects of Study	Formal Context	Concept Lattice Closure Systen Complete Lattice	Closure Operator
Theoretic Background	Order Theory →	$(L, \leq) \cong (L, \vee, \wedge)$ Lattice Theory	← Universal Algebra

Since each of the three fully characterises the other two, the question arises what influence has the algebraic structure of $\underline{\mathfrak{B}}(\mathbb{K})$ on the formal context and the set of implications. Exemplary, we study three degrees of lattice distributivity that are of interest to other authors as well. These lattice properties are n-distributivity, introduced by Huhn in [H72], n-modularity studied by Grätzer and Wehrung in [GW99a, GW99b] and k-join-semidistributivity, proposed by Geyer [G92].

This study is divided into five chapters. Chapter one focuses on n-distributivity. Let n be a positive integer. A lattice $\underline{L}$ is called *n-distributive* if for all x, y_0, …, $y_n \in L$:

$$x \vee \Big(\bigwedge_{i=0}^{n} y_i\Big) = \bigwedge_{j=0}^{n} \Big(x \vee \bigwedge_{\substack{i=0 \\ i \neq j}}^{n} y_i\Big).$$

We achieve by Theorem 1.39 characterisations of this lattice property in the formal context as well as in the closure system of intents. Moreover, we obtain a characterisation of n-distributivity by the help of substructures of the lattice, see Theorem 1.23. We investigate how close the characterisation of n-distributivity by mean of implications is related to the concept of the character of a closure system, cf. Section 1.2.1.

The second chapter is concerned with the n-modularity. Here, a partial success in the direction of further characterisations was also achieved. Let $\underline{L}$ be a lattice, $x, y, z \in L$, and $n \in \mathbb{N}$. We define the following polynomials:

$$\begin{array}{lll} p_0 := x, & q_0 := y, & r_0 := z, \\ p_1 := x \vee (y \wedge z), & q_1 := y \vee (x \wedge z), & r_1 := z \vee (x \wedge y), \\ \vdots & \vdots & \vdots \\ p_{n+1} := p_n \vee (q_n \wedge r_n), & q_{n+1} := q_n \vee (p_n \wedge r_n), & r_{n+1} := r_n \vee (p_n \wedge q_n). \end{array}$$

Additionally, if $n > 0$, then the lattice identity μ_n is defined as $p_n = p_{n+1}$. A lattice satisfying μ_n is called *n-modular*.

Concerning the determination of the least integer $n \in \mathbb{N}$, for which a lattice is n-modular, three-element antichains of the lattice are important that do not form a *balanced triple*. On the other hand, the set of balanced triples of a finite lattice forms a lattice. For that lattice, we determined the reduced formal context, cf. Proposition 2.13. The comprehension of that helped to perceive the difference between the tensor product of concept lattices and the lattice of bonds for their formal contexts. Consequently, we define a tensor product for concept lattices in a way that the result is isomorphic to the lattice of bonds for their formal contexts, cf. Theorem 2.23.

Moreover, the second chapter considers the question of the free lattice, generated by three elements, which is 2-distributive and dually 2-distributive. It turns out that this lattice is infinite. For the free lattice, generated by three elements, which is 2-distributive and dually 2-distributive as well as 2-modular and dually 2-modular, we obtain a list of subdirectly irreducible lattices which must be contained as sublattice. It is not known yet whether this list is complete. However, the subdirect product of those lattices turns out to be very large.

Let n be a positive integer. A complete lattice $\underline{L}$ is called *n-join-semi-*

distributive if it holds for all $x, y_0, y_1, \dots, y_n \in L$:

$$x \vee y_0 = \cdots = x \vee y_n \text{ implies } x \vee y_0 = \bigvee_{\substack{i,j=0 \\ i \neq j}}^{n} x \vee (y_i \wedge y_j).$$

These lattice properties generalise join-semidistributivity ($n = 1$). This property is related to a lattice construction, called local doubling. We characterise the subsets of a lattice, which can be doubled such that the result is a lattice again, cf. Theorems 3.31 and 3.33. This generalises results for order-convex subsets of lattices. We show that the collection of all sets, which can be doubled, also forms a complete lattice.

Additionally, the third chapter of the thesis provides three families of lattices which demonstrate that the three properties n-distributivity, n-modularity and n-join-semidistributivity are independent to a considerable degree. Furthermore, we provide a formal context for one particular join-semidistributive lattice, namely the lattice of all suborder relations of a finite linear order, cf. Theorem 3.37, and, likewise, for the lattice of all quasiorders of a finite set. Both constructions are recursive.

The fourth chapter deals with the lattice of all closure systems on a finite set. Here, too, we provide a formal context of this lattice by using a recursive construction, cf. Proposition 4.3. In addition, we provide the reduced formal context for the lattice of all closure systems of a complete lattice, the reduced context for the lattice of all complete sublattices of a distributive lattice, and we show that the latter construction is generalisable, cf. Section 4.3. In the sequence of formal contexts we also obtain the reduced formal context of lattice refinements of a closure system.

Finally, we demonstrate an attribute exploration algorithm, which is based on implications with proper premise, cf. Section 4.5. The advantages of this may be summarised as follows: When dealing with large attribute sets, in the most cases, a user is not interested in all dependencies between attributes, but only in those with a small size of a premise, since they seem more easy to manage. In contrast to the usual method—using the stem base—the algorithm used here accomplishes that strategy. However, it has not been implemented yet, but is here provided in a pseudo-code.

In an appendix, cf. Appendix A, we turn to left-clearings and present them in a new light. A left-clearing of an ordered set is a subset of the order relation equipped with an additional property. The set of left-clearings of an ordered set of finite length ordered by inclusion forms a complete lattice. Lattices that arise in this way generalise Tamari-lattices and can be explained purely by order theory.

The present study is based on three publications for conferences in the neighbourhood of the Formal Concept Analysis community, including the ICFCA and the ICCS-series. The main results of the publications [GR07, R07, R08] are presented here in a different light.

Acknowledgements

First and foremost, I would like to express my special thanks to my supervisor, Bernhard Ganter. Most of all, I appreciate his intuition of what is correct and what is not in the world of mathematics, and, of course, his humour. I also want to express my sincere thanks to Leonard Kwuida: His comments have always been more than helpful and have highly encouraged me throughout. I also want to acknowledge Christian Zschalig and Mike Behrisch, my office-mates in the *Willersbau*, for countless, spontaneous discussions. Some of them were fruitful, all of them were helpful.

I would also like to thank all my colleagues of the *Institut für Algebra*. At all times I felt comfortable in this friendly and supportive environment—some might even say that I felt too comfortable.

I would like to express my gratitude to my family, to my wife Daniela Saaro for her incessant love, affection and support, and to my parents Karin and Jürgen Reppe for their patience and encouragement throughout. Furthermore, I express my deep gratitude to Anna-Maria Gramatté for improving my English and for her persistent enquiry about the deadline of my thesis.

For final revisions I express my gratitude to Ursula and Matthias Diestel as well as Christian Zschalig.

Affirmation

(a) Hereby I affirm that I wrote the present thesis without any inadmissible help by a third party and without using any other means than indicated. Thoughts that were taken directly or indirectly from other sources are indicated as such. This thesis has not been presented to any other examination board in this or a similar form, neither in this nor in any other country.

(b) The present thesis has been produced since November 2003 at the Institut für Algebra, Department of Mathematics, Faculty of Science, TU Dresden under the supervision of Prof. Bernhard Ganter.

(c) There have been no prior attempts to obtain a PhD at any university.

(d) I accept the requirements for obtaining a PhD (Promotionsordnung) of the Faculty of Science of the TU Dresden, issued March 20, 2000 with the changes in effect since April 16, 2003 and October 1, 2008.

Versicherung

(a) Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

(b) Die vorliegende Dissertation wurde im November 2003 am Institut für Algebra, Fachrichtung Mathematik, Fakultät Mathematik und Naturwissenschaften, Technische Universität Dresden unter der Betreuung von Prof. Bernhard Ganter angefertigt.

(c) Es wurden zuvor keine Promotionsvorhaben unternommen.

(d) Ich erkenne die Promotionsordnung der Fakultät Mathematik und Naturwissenschaften der TU Dresden vom 20. März 2000, in der geänderten Fassung mit Gültigkeit vom 16. April 2003 und vom 1. Oktober 2008 an.

Dresden, den 12. Oktober 2010

Contents