Formal Concept Analysis

Springer
Berlin
Heidelberg
New York
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

Bernhard Ganter • Rudolf Wille

Formal Concept Analysis

Mathematical Foundations

With 105 Figures

Springer

Prof. Dr. Bernhard Ganter
Institut für Algebra
Fakultät für Mathematik
und Naturwissenschaften
Technische Universität Dresden
D-01062 Dresden, Germany

Prof. Dr. Rudolf Wille
Arbeitsgruppe Allgemeine Algebra
Fachbereich Mathematik
Technische Universität Darmstadt
D-64289 Darmstadt, Germany

Translated from the German by Cornelia Franzke
Title of the Original German Edition:
Formale Begriffsanalyse – Mathematische Grundlagen

Library of Congress Cataloging-in-Publication Data

Ganter, Bernhard.
[Formale Begriffsanalyse. English]
Formal concept analysis: mathematical foundations/Bernhard Ganter, Rudolf Wille; [translated from the German by Cornelia Franzke].
p. cm.
Includes bibliographical references and index.

1. Lattice theory. 2. Comprehension (Theory of knowledge) – Mathematical models. 3. Information theory. 4. Artificial intelligence – Mathematical models. 5. Logic, Symbolic and mathematical. I. Wille, Rudolf. II. Title.
QA171.5.G3513 1999
511.3'3–ddc21
98-47620
CIP

ISBN 978-3-540-62771-5 ISBN 978-3-642-59830-2 (eBook)
DOI 10.1007/978-3-642-59830-2

Typesetting: Camera ready pages by the authors
Cover Design: Künkel + Lopka, Werbeagentur, Heidelberg
Printed on acid-free paper SPIN 10552245 33/3142 – 5 4 3 2 1 0

Garrett Birkhoff
with his application-oriented view of lattice theory[1] and
Hartmut von Hentig
with his critical yet constructive understanding of science[2]
have had a decisive influence on the genesis of Formal Concept Analysis.

---

[1] G. Birkhoff: *Lattice Theory.* Amer. Math. Soc., Providence. 1st edition 1940, 2nd (revised) edition 1948, 3rd (new) edition 1967.

[2] H. von Hentig: *Magier oder Magister? Über die Einheit der Wissenschaft im Verständigungsprozeß.* Klett, Stuttgart 1972.

Preface

Formal Concept Analysis is a field of applied mathematics based on the mathematization of *concept* and *conceptual hierarchy*. It thereby activates mathematical thinking for conceptual data analysis and knowledge processing.

The underlying notion of "concept" evolved early in the philosophical theory of concepts and still has effects today. For example, it has left its mark in the German standards DIN 2330 and DIN 2331. In mathematics it played a special role during the emergence of mathematical logic in the 19th century. Subsequently, however, it had virtually no impact on mathematical thinking. It was not until 1979 that the topic was revisited and treated more thoroughly. Since then, through a large number of contributions, Formal Concept Analysis has obtained such breadth that a systematic presentation is urgently needed, but can no longer be realized in one volume.

Therefore, the present book focuses on the mathematical foundations of Formal Concept Analysis, which can be regarded chiefly as a branch of applied lattice theory. A series of examples serves to demonstrate the utility of the mathematical definitions and results; in particular, to show how Formal Concept Analysis can be used for the conceptual unfolding of data contexts. These examples do not play the role of case studies in data analysis. A separate volume is intended for a comprehensive treatment of methods of conceptual data and knowledge processing. The general foundations of Formal Concept Analysis will also be treated separately.

It is perfectly possible to use Formal Concept Analysis when examining human conceptual thinking. However, this would be an application of the mathematical method and a matter for the experts in the respective science, for example psychology. The adjective "formal" in the name of the theory has a delimiting effect: we are dealing with a mathematical field of work, that derives its comprehensibility and meaning from its connection with well-established notions of "concept", but which does not strive to explain conceptual thinking in turn.

The mathematical foundations of Formal Concept Analysis are treated in seven chapters. By way of introduction, elements of mathematical order and lattice theory which will be used in the following chapters have been compiled in a *chapter "zero"*. However, all difficult notation and results from this chapter will be introduced anew later on. A reader who knows what is understood by a lattice in mathematics may skip this chapter.

The *first chapter* describes the basic step in the formalization: An elementary form of the representation of data (the "cross table") is defined mathematically ("formal context"). A formal concept of such a data context is then explained. The totality of all such concepts of a context in their hierarchy can be interpreted as a mathematical structure ("concept lattice"). It is also possible to allow more complex data types ("many-valued contexts"). These are then reduced to the basic type by a method of interpretation called "conceptual scaling".

The *second chapter* examines the question of how all concepts of a data context can be determined and represented in an easily readable diagram. In addition, implications and dependencies between attributes are dealt with. The *third chapter* supplies the basic notions of a structure theory for concept lattices, namely part- and factor structures as well as tolerance relations. In each case the extent to which these can be elaborated directly within the contexts is studied.

These mathematical tools are then used in the *fourth* and *fifth chapter*, in order to describe more complex concept lattices by means of decomposition and construction methods. Thus, the concept lattice can be split up into (possibly overlapping) parts, but it is also possible to use the direct product of lattices or of contexts as a decomposition principle. A further approach is that of substitution. In accordance with the same principles, it is possible to construct contexts and concept lattices. As an additional construction principle, we shall describe a method of doubling parts of a concept lattice.

The structural properties examined in mathematical lattice theory, for example the distributive law and its generalizations or notions of dimension, play a role in Formal Concept Analysis as well. This shall be treated in the *sixth chapter.* The *seventh chapter* finally deals with structure-comparing maps, examining various kinds of morphisms. Particular attention is given to the scale measures, occuring in the context of conceptual scaling.

We limit ourselves to a concise presentation of ideas for reasons of space. Therefore, we endeavour to give a complete reference to further results and the respective literature at the end of each chapter. However, we have only taken into account such contributions closely connected with the topic of the book, i.e., with the mathematical foundations of Formal Concept Analysis. The index contains all technical terms defined in this book, and in addition some particularly important keywords. The bibliography also serves as an author index.

The genesis of this book has been aided by the numerous lectures and activities of the "Forschungsgruppe Begriffsanalyse" (Research Group on Concept Analysis) at Darmstadt University of Technology. It is difficult to state in detail which kind of support was due to whom. Therefore, we can here only express our gratitude to all those who contributed to the work presented in this book.

Two years after the German edition, this English translation has been finished. In its content there are only a few minor changes. Although there is ongoing active work in the field, the mathematical foundations of Formal Concept Analysis have been stable over the last years.

The authors are extremely grateful to Cornelia Franzke for her precise and cooperative work when translating the book. They would also like to thank K.A. Baker, P. Eklund and R.J. Cole, M.F. Janowitz, and D. Petroff for their careful proofreading.

Contents