

ICCS-2001 International Workshop on Concept
Lattices-based Theory, Methods and Tools for
Knowledge Discovery in Databases
(CLKDD'01)

July 30, 2001

Stanford University, Palo Alto (CA)

edited by

Engelbert Mephu Nguifo & al.



Contents

Organisation	I
Preface	III
Call for Papers	VI
Workshop Schedule	X
Invited Talk	
Association Rules : Past, Present and Future <i>Ramakrishnan Srikant</i>	3
Theoretical Background and Applications	
Why Can Concept Lattices Support KDD ? <i>Rudolf Wille</i>	7
Structuration of phenotypes/genotypes through Galois lattices and implications <i>Vincent Duquenne, Caroline Chabert, Améziane Cherfouh, Jean-Maurice Delabar, Anne-Lise Doyen, Douglas Pickering</i>	21
Algorithms and Implications	
Comparing Performance of Algorithms for Generating Concept Lattices <i>Sergei O. Kuznetsov, Sergei A. Ob"edkov</i>	35
Computing Proper Implications <i>Rafik Taouil, Yves Bastide</i>	49
Lattices Compression & Temporal Analysis	
Building a pruned inheritance lattice for relational descriptions <i>Mélanie Courtine, Isabelle Bournaud</i>	65
A Lattice-based Data Structure for Information Retrieval and ML <i>FJ (Dean) van der Merwe, DG Kourie</i>	77
Temporal Concept Analysis <i>Karl Erich Wolff</i>	91

CLKDD 2001

Organizing Committee

Primary Chair :

Engelbert Mephu Nguifo (CRIL - IUT de Lens, France)

Workshop Co-Chairs :

Vincent Duquenne (CNRS Orleans, France)

Michel Liquiere (LIRMM Montpellier, France)

Program Committee

Claudio Carpineto (Fondazione Ugo Bordoni, Rome, Italy)

Edwin Diday (CEREMADE-LISE Paris Dauphine, France)

Vincent Duquenne (CNRS FRE 2134 Orleans, France)

Jean-Gabriel Ganascia (LIP6 Paris, France)

Bernhard Ganter (Technische Universität Dresden, Germany)

Robert Godin (University of Québec, Montréal, Canada)

Sergej Kuznetsov (Moscow VINITI Institute, Russia)

Lotfi Lakhal (LIM Marseille, France)

Michel Liquière (LIRMM Montpellier, France)

Engelbert Mephu Nguifo (CRIL Lens, France)

Sergei Ob'edkov (Russian State University for the humanities)

Ramakrishnan Srikant (IBM Research Center, San José CA)

Gerd Stumme (Universität Karlsruhe, Germany)

Rudolf Wille (Technische Universität Darmstadt, Germany)

Preface

Throughout the last decade, Knowledge Discovery in Databases (KDD) has become an increasingly important topic in research as well as in industrial applications, up to being now a well-established interdisciplinary research area, benefiting yet from diverse influences such as : Databases, Data Analysis and Machine Learning technology.

The main purpose of this workshop is to bring together members of these communities that are working on concept lattices (also termed 'galois lattices') structures, for considering the current state of art and identifying the strengths -and the potential weaknesses- of concept lattices for KDD. The workshop addresses the general methodologies and principles that are appropriate in the development of concept lattices-based datamining methods and tools.

The workshop notes contain an abstract of the invited talk and seven accepted papers divided into three parts.

The invited speaker, Dr. Srikant, will browse through association rules which have received considerable attention in KDD.

The first part concerns theoretical background and application of concept lattices, and contains two papers. In the first contribution, Wille demonstrates why concept lattices as mathematical abstraction of concept systems can support humans to discover information and then create knowledge. The paper of Duquenne & al. gives an illustration, with the GLAD software, of application of concept lattices in Biology that allows to obtain better interpretation.

The second part contains two papers, and is about algorithms and implications on concept lattices. Kuznetsov & Ob"edkov compare and discuss several algorithms that generate the set of all formal concepts and diagram graphs of concept lattices. Taouil & Bastide propose a new algorithm, Impec, which is designed to efficiently find proper implications given a set and a closure operator on this set.

The last part contains three papers that concerns lattices compression and temporal analysis. Courtine & Bournaud present an approach based on iterative reformulation of descriptions, to build a pruned inheritance concept lattices for relational descriptions. Van der Merwe & Kourie propose a generic data structure called compressed lattices, that is closely related to concept lattices. This compressed lattices allow to compress and expand the generated lattices. In the last contribution, Wolff's paper introduces two closely related types of conceptual time systems, in order to describe temporal phenomena with tools of Formal Concept Analysis.

Ample time has been scheduled for discussions during this one-day workshop. This will give an opportunity to push further the discussion upon the potentialities -and limitations- of concept lattices for KDD accross these communities,

but also to spread it furthermore among other trends of works in Conceptual Structures.

Thanks to the program committee members for their timely and thorough reviews of the papers, as well as for their many constructive suggestions to the authors.

Engelbert Mephu Nguifo
Workshop Primary Chair

Acknowledgements

We gratefully acknowledge the financial support of the Ganymède Project, the CRIL (Université d'Artois) and IUT de Lens. We thank Stanford University (Palo Alto) and ICCS organization for printing the workshop notes. We would also like to thank ICCS-2001 chairs for their flexibility and cooperation. Many thanks to Olivier Roussel for his help to format this document.

Call for Papers

Background (from the call for papers)

Throughout the last decade, Knowledge Discovery in Databases (KDD) has become an increasingly important topic in research as well as in industrial applications, up to being now a well-established interdisciplinary research area, benefiting yet from diverse influences such as : DataBases (DB), Data Analysis (DA) and Machine Learning (ML) technology. Among those communities, there are several sub-areas of research that can be unified by their common interest in "concept lattice structures" which, as a consequence, start to play an important role in DataMining (DM).

Over the last two decades, several trends of works have demonstrated how concept lattices formalize conceptual structures by coding any kind of dualities, and can be used to address a variety of problems in DB, DA and ML. These include :

- association rules or data dependencies in DB, searching frequent item sets, indexing documents for Information Retrieval (IR)...;
- exploring attributes in DA for binary as well as more complex symbolic data structures...;
- designing concept lattices-based ML Systems for conceptual clustering and classifications...

These studies have generated diverse methods and specific tools along their developments, which are now well implemented, commonly used and have shown the usefulness of this structural approach for each domain. Consequently, it may be the right time to exchange and share the researchers' experiences of these areas, in order to increase lattice-based contributions to KDD.

Workshop Goals

The main purpose of this workshop is to bring together the members of the DB, DA and ML communities that are working on such concept lattice structures, for considering the current state of art and identifying the strengths -and the potential weaknesses- of these approaches. The workshop will address the general methodologies and principles that are appropriate in the development of concept lattice based DM methods and tools.

Hence, we are seeking for papers devoted on the theory and practice of concept lattices for modelling conceptual structures and more specifically for KDD. The workshop will focus on the following questions :

- which tasks of KDD's processes can be improved with lattices ?
- which sampling data size could reasonably deal with lattices ?
- for which data (Boolean, symbolic, numerical, semistructured...) ?
- are there "fast" algorithms to design lattices from data ?
- how to revise and to incrementally build lattices ?
- how to appropriately visualize the generated lattices ?
- how to reduce computational complexity of lattices ?
- how to interpret the designed lattice and its structure ?
- are they "fast" algorithms to generate significant rules from lattices ?
- how to reduce the lattice size (approximations, irreducible nodes...) ?

These questions should also be demonstrated on real examples through applications of concept lattices for knowledge discovery (Biology, Genetics, Social Sciences, Banking, ...). Emphasis will be given on recent theory and algorithms of concept lattices for KDD, on systems based on concept lattices and their applications in Data Mining.

Intended Audience

This workshop is firstly aimed at bringing together members of different overlapping communities that share concept lattices as a kernel for developing Data Mining methods and tools :

- those coming from DB, which use lattices for generating association rules or indexing data ;
- those coming from DA, with an interest in formal concept analysis ;
- those coming from ML, developing lattice-based machine learning systems.

This will give an opportunity to push further the discussion upon the potentialities -and limitations- of concept lattices for KDD across these communities, but also to spread it furthermore among other trends of works in Conceptual Structures.

Workshop format

The workshop is planned to include three kinds of session :

Invited talks : The workshop will feature a number of invited presentations by people who have contributed to these areas.

Technical papers : Contributors to the workshop will be requested to highlight only key contributions when presenting their work, in order to set aside the maximum possible time for discussion and interaction.

Panel/discussion sessions : Key issues will be debated by panellists and the workshop participants as a whole.

How to contribute

We invite submissions of technical papers and short position papers (two pages). Authors of accepted technical and position papers will be invited to participate in the workshop. Format requirements for submissions of technical papers are :

- maximum 12 double-spaced pages, excluding title page and bibliography
- all submissions should be made electronically if possible, by email attachment and preferably in Postscript format. Only if electronic submission is impossible should you send three hardcopies.
- all submissions must be sent to the primary workshop contact, Engelbert Mephu Nguifo, at the address below.

Important Dates

8 May 2001 Submission of technical and position papers

8 June 2001 Notification of acceptance

1 July 2001 Camera-ready papers

Organization

CLKDD'2001 is organized in cooperation with the International Conference on Conceptual Structures (ICCS-2001).

Organizing Committee

Primary Chair : Engelbert Mephu Nguifo (CRIL Lens, France)

Workshop Co-Chairs : Vincent Duquenne (CNRS Orleans, France)
Michel Liquiere (LIRMM Montpellier, France)

Program Committee

Claudio Carpineto	Fondazione Ugo Bordoni, Rome, Italy
Edwin Diday	CEREMADE-LISE Paris Dauphine, France
Vincent Duquenne	CNRS FRE 2134 Orleans, France
Jean-Gabriel Ganascia	LIP6 Paris, France
Bernhard Ganter	Technische Universität Dresden, Germany
Robert Godin	University of Québec, Montréal, Canada
Sergei Kuznetsov	Moscow VINITI Institute, Russia
Lotfi Lakhal	LIM Marseille, France
Michel Liquière	LIRMM Montpellier, France
Engelbert Mephu Nguifo	CRIL Lens, France
Sergei Ob"edkov	Russian State University for the humanities
Ramakrishnan Srikant	IBM Almaden Research Center, San José, CA
Gerd Stumme	Universität Karlsruhe, Germany
Rudolf Wille	Technische Universität Darmstadt, Germany

Program Schedule

Workshop on Concept Lattices for KDD

Stanford University
Monday, July 30, 2001

9.00-9.15 Welcome to participants

Invited Talk

9.15-10.15 Association Rules : Past, Present and Future
Ramakrishnan Srikant

10.15-10.40 *Coffee Break*

Theoretical Background and Applications

10.40-11.10 Why Can Concept Lattices Support KDD ?
Rudolf Wille

11.10-11.40 Structuration of phenotypes/genotypes through Galois lattices and implications
Vincent Duquenne & al.

11.40-12.30 Discussion

12.30-14.00 *Lunch Break*

Algorithms and Implications

14.00-14.25 Comparing Performance of Algorithms for Generating Concept Lattices
Sergei O. Kuznetsov & Sergei A. Ob"edkov

14.25-14.50 Computing Proper Implications
Rafik Taouil & Yves Bastide

14.50-15.20 Discussion

15.20-15.45 *Coffee Break*

Lattices Compression & Temporal Analysis

15.45-16.10 Building a pruned inheritance lattice for relational descriptions
Mélanie Courtine & Isabelle Bournaud

16.10-16.35 A Lattice-based Data Structure for Information Retrieval and ML
FJ (Dean) van der Merwe & DG Kourie

16.35-17.00 Temporal Concept Analysis
Karl Erich Wolff

17.00-17.30 Discussion

General Discussion and Conclusion

17.30-17.50 General Discussion

17.50-18.00 Conclusion

Invited Talk

Theoretical Background and Applications

Algorithms and Implications

Lattices Compression and Temporal Analysis
