Theoretical Computer Science ••• (••••) •••–•••

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Lattices, closures systems and implication bases: A survey of structural aspects and algorithms

Karell Bertet *, Christophe Demko, Jean-François Viaud, Clément Guérin

Laboratory L3I, University of La Rochelle, av Michel Crépeau, 17042 La Rochelle, France

A R T I C L E I N F O

Article history:
Received 9 April 2015
Received in revised form 15 November 2016
Accepted 16 November 2016
Available online xxxx

Keywords:
Lattice
Closed set lattice
Concept lattice
Galois lattice
Implicational system
Closure operator
Closure system
Canonical direct basis
Canonical basis
Dependence graph
Minimal generators

A B S T R A C T

Concept lattices and closed set lattices are graphs with the lattice property. They have been increasingly used this last decade in various domains of computer science, such as data mining, knowledge representation, databases or information retrieval. A fundamental result of lattice theory establishes that any lattice is the concept lattice of its binary table. A consequence is the existence of a bijective link between lattices, contexts (via the table) and a set of implicational rules (via the canonical (direct) basis). The possible transformations between these objects give rise to relevant tools for data analysis.

In this paper, we present a survey of lattice theory, from the algebraic definition of a lattice, to that of a concept lattice, through closure systems and implicational rules; including the exploration of fundamental bijective links between lattices, reduced contexts and bases of implicational rules; and concluding with the presentation of the main generation algorithms of these objects.

1. Introduction

Birkhoff's book [1] in 1940 is the first reference book on lattice theory, in which the lattice structure is introduced as an algebraic structure provided with two operators: the lower and upper bound. Lattice theory is also the subject of the books of Grätzer [2,3] and Davey and Priestley [4]. However, in recent works [5,6], Monjardet establishes that the algebraic lattice structure appeared for the first time in the works of Dedekind in 1900 as the *dualgruppe* term and, later, with several forms and terminologies between 1928 and 1936 in the works of Merge, Klein, Blind, Birkhoff, Öre or von Neuman. Monjardet also mentions that lattices appeared for the first time in 1847, as a Boolean form, in the works of Boole on the algebra of logic, then in the works of Peirce in 1880. The *lattice* term was proposed by Birkhoff during the first symposium on lattices in 1938, while Merge and Öre respectively proposed the *System von Dingen* and *structure* terms.

In other reference books [4,7], a lattice is defined as an ordered set with particular elements called upper and lower bounds. The notion of irreducible elements [7,8] of a lattice allows the conception of compact representations of lattices from which their reconstruction is possible. The *Galois lattice* term was introduced in 1970 [7], while the *Galois connection* term was introduced by Öre in 1944 [9]. Since the composition of the two mappings of a Galois connection is a closure

* Corresponding author.
E-mail address: kbertet@univ-lr.fr (K. Bertet).

http://dx.doi.org/10.1016/j.tcs.2016.11.021
0304-3975/© 2016 Published by Elsevier B.V.

operator, the fundamental bijection can be extended to *closed set lattices* [10] and to *closure systems*, i.e. set systems provided with a closure operator.

The fundamental result of lattice theory [7] establishes the bijection between any lattice and the Galois lattice of its *irreducible binary table*. A direct consequence is the existence of bijective links between lattices, reduced contexts (via the table) and a set of implication rules (via the canonical (direct) basis).

Concept lattices were introduced by Wille in 1982 [11]. The 1990s reference book of B. Ganter and R. Wille [12] is often cited as a foundation of Formal Concept Analysis (FCA). A concept lattice is a graph defined from a context – i.e. a binary relation between a set of objects and a set of attributes. The lattice, composed of concepts connected by the generalization/specialization relation, supplies a very intuitive representation of the data. FCA allows to analyze and extract information from a context where implicit knowledge regarding a dataset can be represented, either by a concept lattice or by implications.

In data mining, various symbolic methods stemming from FCA have been studied for (un)supervised [13,14] classification and frequent pattern extraction [15]. Most of the literature on the subject using concept lattices relies on selection-based strategies, which consists in selecting/choosing the concepts which encode the most relevant information from the huge amount of available data. In opposition to these selection-based strategies are navigation-based approaches which perform the classification stage by navigating through the complete lattice (similar to the navigation in a classification tree) [16]. Pattern mining and association rules extraction is another important issue in data mining and knowledge extraction. This problem was first addressed in an exhaustive way with the well-known APriori algorithm [17]. Closed and frequent patterns [15] and bases of association rules have then been introduced. This problem turned out to be equivalent to the generation of minimal generators and bases of implications.

Some information retrieval methods from structured data propose a navigation in a conceptual space where places are concepts connected by navigation links, similar to the definition of a concept lattice. Let us cite the Lattice Based Information Retrieval (LBIR) approach introduced by U. Priss in [18], and Abstract Conceptual Navigation (ACN) introduced by Ferré in [19,20]. In ACN, "concepts are characterized by a formal expression (e.g., a query, an update), and are made of two parts: the extension is made of concrete objects (e.g., entities, values) while the intension is made of formal expressions. The conceptual space is not static but is induced by concrete data, and evolves with it. Navigation promotes usability by guiding users, and freeing them from the burden of writing formal expressions".

Therefore, FCA is a knowledge representation approach the use of which increased during the last two decades in various domains of computer science, such as data mining, knowledge representation, databases or information retrieval. Indeed, the technological improvements allow the use of these structures for large data in these domains though they are exponential in space/time (worst case). They make the development of a large number of applications possible. The need for appropriate and efficient algorithms to manipulate these structures remains a major challenge.

Equivalent notions such as lattices (or closed set lattices), closure systems, closure operators (or dual closure operators), (pure) Horn functions have been studied by different authors in different domains (topology, lattice theory, hypergraph theory, choice functions, relational data bases, data mining and concept analysis, artificial intelligence and expert systems, knowledge spaces, logic and logic programming, theorem proving...). In the area of logic, the representation of a formula by Horn clauses is linked to implication systems, where the closed set lattice is composed of all the true subsets of variables. In graph theory, a hyper-directed edge between two subsets of nodes can also be interpreted as an implication, and some path problems are linked to closure computation or forward chaining.

It is not surprising that one finds the same notions, results or algorithms under various names. One can also find many original results or algorithms only known in a specific domain. It would be very profitable to increase (or create) the communication between the various domains that use the same (or equivalent) notions and tools. The Dargsthul seminar "Horn formulas, directed hypergraphs, lattices and closure systems: related formalisms and applications" organized by Kira V. Adaricheva and Giuseppe F. Italiano and Hans Kleine Büning and György Turán in 2014 [21] is a huge step in this direction.

The first part of our review (Section 2) provides an overview of existing definitions stemming from lattice theory, some structural results and the fundamental bijective links between lattices, reduced contexts and bases of implicational rules. In a second part (Section 3), we address the generation problems issued from these bijections, and we present the main generation algorithms that are mainly implemented in our Java library, named `Galactic`.[1]

2. Basis of lattice theory

Section 2.1 provides an overview of existing definitions stemming from the lattice theory in the finite case: algebraic and ordinal definitions, irreducible elements, table and dependence graph. Section 2.2 defines a closed set lattice, a closure operator and a closure system. All sets, and especially lattices, are considered to be finite.

2.1. Algebraic and ordinal lattice

In the literature we find two definitions of a lattice: the algebraic definition and the ordinal definition. These two definitions introduce both notions of *upper bound* (or *join*) and *lower bound* (or *meet*) as binary operators in the algebraic definition, and particular elements in the ordinal definition:

---

[1] http://lattices.univ-lr.fr.