Computing the D-base and D-relation of finite closure systems

Kira Adaricheva,[1] Lhouari Nourine,[2] Simon Vilmin[3]

[1] Departments of Mathematics, Hofstra University, Hempstead, NY 11549, USA.
[2] Université Clermont Auvergne, Clermont Auvergne INP, CNRS, LIMOS, F-63000 Clermont–Ferrand, France.
[3] Aix-Marseille Université, CNRS, LIS, LIRICA, Marseille, France.
kira.adaricheva@hofstra.edu, lhouari.nourine@uca.fr, simon.vilmin@lis-lab.fr

Abstract

Implicational bases (IBs) are a common representation of finite closure systems and lattices, along with meet-irreducible elements. They appear in a wide variety of fields ranging from logic and databases to Knowledge Space Theory.

Different IBs can represent the same closure system. Therefore, several IBs have been studied, such as the canonical and canonical direct bases. In this paper, we investigate the D-base, a refinement of the canonical direct base. It is connected with the D-relation, a must-have tool in the study of free lattices. The D-base demonstrates desirable algorithmic properties, and together with the D-relation, it conveys essential properties of the underlying closure system. Hence, computing the D-base and the D-relation of a closure system from another representation is crucial to enjoy its benefits. However, complexity results for this task are lacking.

In this paper, we give algorithms and hardness results for the computation of the D-base and D-relation. Specifically, (1) we establish the $\mathbf{NP}$-completeness of finding the D-relation from an IB, even if the D-relation is acyclic; (2) we obtain a polynomial-delay algorithm computing the D-base from another IB, and (3) we give an output-quasipolynomial time algorithm to compute the D-base from meet-irreducible elements instead of an IB. We conclude the paper with a discussion regarding the E-base, a subset of the D-base.

Introduction

A *closure system* over a finite groundset X is a set system containing X and closed under taking intersections. The sets in a closure system are *closed sets*, and when ordered by inclusion they form a *lattice*. Lattices and closure systems are used in a number of fields of mathematics and computer science such as algebra (Gratzer 2011), databases (Mannila and Räihä 1992), logic (Hammer and Kogan 1995), or Knowledge Space Theory (Doignon and Falmagne 2012) to mention but a few.

Often, a closure system is implicitly given by one of the following two representations: *meet-irreducible elements* or *implications* and *implicational bases* (IBs). The family of meet-irreducible elements of a closure system is the unique minimal subset of closed sets from which the whole system can be rebuilt using intersections. An implication over groundset X is a statement $A \to c$ where A is a subset of X and c and element of X. In $A \to c$, A is the *premise* and c the *conclusion*. The implication $A \to c$ stands for "if a set includes A, it must also contain the element c". An implicational base (IB) over X is a collection of implications over X. An IB encodes a unique closure system. On the other hand, a closure system can be represented by several equivalent IBs. IBs are known as Horn CNFs in logic (Boros et al. 2009; Hammer and Kogan 1995), covers of functional dependencies in databases (Mannila and Räihä 1992), association rules in data mining (Agrawal et al. 1996) or entailments in Knowledge Space Theory (Doignon and Falmagne 2012).

This variety of applications brought a rich theory of implications still at the core of several works, as witnessed by surveys (Bertet et al. 2018; Wild 2017) and recent contributions (Bérczi, Boros, and Makino 2023a,b; Bichoupan 2022, 2023; Nourine and Vilmin 2023). Specifically, since two distinct IBs can represent the same closure system, several IBs have been studied (Adaricheva, Nation, and Rand 2013; Bertet and Monjardet 2010; Guigues and Duquenne 1986; Wild 1994). Within this galaxy, the *canonical base* and the *canonical direct base* are arguably the two most shining stars. They are unique, and they reflect two complementary ways of understanding IBs:

(1) The canonical base (Guigues and Duquenne 1986), also known as the *Duquenne-Guigues base*, puts emphasis on the premises of implications. Its implications are of the form $A \to C$ (rather than $A \to c$) where A is called *pseudo-closed* and C is the *closure* of A, i.e. the set of all elements derived from A. The canonical base has a minimum number of implications, and can be reached from any other IB in polynomial time (Wild 2017).

(2) The canonical direct base is defined by its conclusions (see the survey (Bertet and Monjardet 2010)). It describes, for each element c, the inclusion-wise minimal A sets deriving c in the closure system. Such minimal sets are the *minimal generators* of c. The canonical direct base then gathers all implications of the form $A \to c$ (known in logic as the prime implicates in Horn CNFs). It enjoys the property of being *direct*, meaning that the closure of a set can be computed with a single pass over the implications. Moreover, the canonical direct base naturally captures the δ-relation of lattice theory (Monjardet and Caspard 1997), where $b\delta a$ means that a belongs to some minimal generator of b. This relation and its transitive closure have been used implicitly in the context of Horn CNFs and implication-graphs for minimization (Boros, Čepek, and Kogan 1998) or to recognize

and optimize acyclic Horn functions and some of their generalizations (Boros et al. 2009; Hammer and Kogan 1995).

In this paper, we study a subset of the canonical direct base: the D-base (Adaricheva, Nation, and Rand 2013). It relies on particular minimal generators called *D-generators*. A minimal generator A of c is a *D-generator* of c if its closure with respect to binary implications—implications of the form $a \to c$—is minimal as compared to those of other minimal generators of c. The D-generators also appear under the name *minimal pairs* in the study of semimodular closure operators (Faigle and Herrmann 1981). The D-base then consists In general, the D-base is much smaller than the canonical direct base. Yet, it still enjoys directness as long as its implications are suitably ordered. This makes the D-base appealing for computational purposes. Several application projects were carried out recently that provide analysis of data by employing D-base (Nation et al. 2021; Adaricheva et al. 2023).

Besides, the D-base embeds the D-relation of a closure system: cDa holds if a belongs to a D-generator of c. The D-relation has played a major role in lattice theory since the 1970s (Jónsson and Nation 1975). It is crucial in the study of free lattices (Freese, Ježek, and Nation 1995) and for the doubling of convex sets (Day 1992). Also, the D-relation compactly convey structural information of the underlying closure system. The most striking examples are lower bounded closure systems that are precisely characterized by an acyclic D-relation (Freese, Ježek, and Nation 1995). Acyclic closure systems (i.e. acyclic Horn functions) (Adaricheva 2017) and closure systems generated by sub-semilattices of semilattices (Adaricheva 1991) are well-known examples of lower bounded systems.

Despite their importance, the algorithmic solutions for several questions related to D-base and D-relation are still lacking:

(1) Can we recover the D-relation from an arbitrary IB?

(2) Can we compute the D-base from an IB?

(3) How to find the D-base from the collection of meet-irreducible elements of a closure system?

Main Contributions. In this paper, we study the aforementioned questions. Moreover, we investigate the E-base, a refinement of the D-base. As the D-base may be of exponential size with respect to the input IB or meet-irreducible elements, we express the complexity of our algorithms in terms of the combined size of their input and their output. This is *output-sensitive complexity*[1] (Johnson, Yannakakis, and Papadimitriou 1988). We propose the following results:

(1) Given an IB and two elements a, c, we show that it is $\mathbf{NP}$-complete to decide whether cDa holds (problem D-RR). The problem is $\mathbf{NP}$-complete for both acyclic closure systems and D-acyclic closure systems with D-paths of constant sizes.

(2) Given an IB, we give a polynomial-delay algorithm to compute the corresponding D-base (problem DB-IB). Our algorithm uses the supergraph traversal method and previous results of (Ennaoui and Nourine 2016).

[1] the relevant notions are defined in the preliminaries

(3) Given a set of meet-irreducible elements (or closed sets), we show that computing the D-base (problem DB-M) can be achieved in output-quasipolynomial time. We obtain this result by establishing the equivalence between DB-M and dualization in distributive lattices (DLD). This latter problem admits an output-quasipolynomial time (Elbassioni 2022).

(4) In the final section, we discuss a subset of the D-base: the E-base. The E-base does not always faithfully represent the underlying closure system (Adaricheva, Nation, and Rand 2013). Henceforth, the following question becomes particularly interesting.

Question 1. For which closure systems is the E-base valid?

D-geometries are an example of such closure systems. These are anti-exchange closure systems, or convex geometries, without D-cycles. We show that D-geometries can be recognized in polynomial time from an IB. Then, we analyze closure systems with exchange properties, i.e. lattices of flats of matroids. We mention that the lattice of flats of binary matroids, projective geometries and atomistic modular lattices are systems with E-base.

Related work. We begin with complexity results regarding the D-relation. In (Adaricheva, Nation, and Rand 2013), the authors provide sub and supersets of the D-relation that can be obtained in polynomial time from any IB. On the one hand, it is possible to refine any IB to a subset of the D-base, which gives a subset of the D-relation. On the other hand, the transitive closure of the D-relation can be recovered from any IB. Similar results also exist for the δ-relation (Boros, Čepek, and Kogan 1998). The hardness of computing the δ-relation can be obtained as a corollary of the $\mathbf{NP}$-completeness of the prime attribute problem in databases (Lucchesi and Osborn 1978). Let us mention that when one is given meet-irreducible elements instead of an IB, both the D-relation and the δ-relation can be computed in polynomial time by means of lattice-theoretic characterizations (Monjardet and Caspard 1997; Freese, Ježek, and Nation 1995). Still, to our knowledge, none of these results formally settles the complexity of computing the D-relation from an IB, especially with regard to the structure of the underlying closure system as we do in this paper.

We move to the tasks of computing the D-base from an IB (DB-IB) or from meet-irreducible elements (DB-M). In (Rodríguez-Lorenzo et al. 2015, 2017), the authors use simplification logic to come up with two algorithms solving DB-IB. However, the complexity of these algorithms is not analyzed. As for DB-M, (Adaricheva and Nation 2017) give an algorithm based on *hypergraph dualization*. Yet, their algorithm will produce in general a proper superset of the D-base. Note that hypergraph dualization is a central open problem in algorithmic enumeration (Eiter, Makino, and Gottlob 2008). To date, the best algorithm for this task is due to Fredman and Khachiyan (Fredman and Khachiyan 1996) and runs in output-quasipolynomial time. We finally mention the of Freese et al. (Freese, Ježek, and Nation 1995) (Listing 11.12, p. 232) which computes the D-base from the whole closure system (in fact, a meet-join table of the lattice) as input. As there is usually an exponential gap between a closure and either of its representation, IB or meet-irreducible elements, we cannot afford using this algorithm in our context.

We also mention algorithms for finding the related canonical direct base. If the input representation is meet-irreducible elements, it is well-known that the problem is equivalent to hypergraph dualization (see e.g. (Mannila and Räihä 1992; Wild 2017)). When an IB is given, the algorithm of (Lucchesi and Osborn 1978) (see also (Bérczi, Boros, and Makino 2023a)) listing the minimal keys of a closure system can be adapted to list the minimal generators of a single element with polynomial-delay. Besides, the algorithm of (Boros, Crama, and Hammer 1990) computes the whole canonical direct base with a total time linearly dependent on the number of implications in the output. These techniques cannot be used out of the box to find the D-base as there may be an exponential gap between the two IBs. Instead, we adapt the procedure of (Ennaoui and Nourine 2016) to list the so called ideal-minimal keys of a closure system. Nevertheless, our approach naturally extends results on minimal generators, all the while taking care of the binary implications playing an important role in the D-base.

Paper Organization. First, we give necessary definitions regarding closure systems and their representations. Secondly, we establish the complexity of computing the D-relation. Thirdly, we give the algorithms to compute the D-base from different representations of a closure system. Then, we discuss the E-base and systems with valid E-base. We conclude the paper by recalling our results and some questions for further research.

Preliminaries

All the objects considered in this paper are finite. For notions not defined here, we refer the reader to (Gratzer 2011; Korte, Lovász, and Schrader 2012).

Closure operators, closure systems. Let X be a set. A *closure operator* over X is a map $\phi\colon \mathbf{2}^X \to \mathbf{2}^X$ such that, for every $A, B \subseteq X$: (1) $A \subseteq \phi(A)$; (2) $A \subseteq B$ implies $\phi(A) \subseteq \phi(B)$; and (3) $\phi(\phi(A)) = \phi(A)$. The pair (X, ϕ) is a *closure space*. A subset F of X is *closed* if $\phi(F) = F$. For brevity, we write $\phi(a)$ instead of $\phi(\{a\})$ for $a \in X$. The family of closed sets of ϕ is called $\mathcal{F}$. We have $\mathcal{F} = \{F \subseteq X : F = \phi(F)\} = \{\phi(A) : A \subseteq X\}$. We say that ϕ is *standard* if for every $a \in X$, $\phi(a) \setminus \{a\}$ is closed. Consequently, $\emptyset$ is closed and $\phi(a) \neq \phi(b)$ unless $a = b$. Let (X, ϕ) be a closure space and let $K \subseteq X$. We say that K is a *minimal spanning set* of its closure if $\phi(K') \subset \phi(K)$ for every $K' \subset K$. If moreover $\phi(K) = X$, K is a *minimal key* of the closure space.

We move to closure systems. A set system $(X, \mathcal{F})$ is a *closure system* if $X \in \mathcal{F}$ and $F_1 \cap F_2 \in \mathcal{F}$ whenever $F_1, F_2 \in \mathcal{F}$. There is a well-known correspondence between closure operators and closure systems. More precisely, if (X, ϕ) is a closure space, the set system $(X, \mathcal{F})$ is a closure system. Dually, a closure system $(X, \mathcal{F})$ induces a closure operator $\phi\colon \mathbf{2}^X \to \mathbf{2}^X$ given by $\phi(A) = \bigcap\{F \in \mathcal{F} : A \subseteq F\}$. This correspondence is one-to-one. The pair $\mathbf{L} = (\mathcal{F}, \subseteq)$ is the *(closure) lattice* of $(X, \mathcal{F})$. An *antichain* of $\mathbf{L}$ is a subset $\mathcal{B}$ of $\mathcal{F}$ consisting of pairwise incomparable closed sets.

Remark 1. In this paper, we only consider standard closure systems, without loss of generality. If $(X, \mathcal{F})$ is a closure system, we call ϕ the corresponding closure operator and $\mathbf{L}$ the corresponding closure lattice $(\mathcal{F}, \subseteq)$.

Let $(X, \mathcal{F})$ be a closure system and let $M \in \mathcal{F}$. The closed set M is a *meet-irreducible element* of $\mathcal{F}$ if for every $F_1, F_2 \in \mathcal{F}$, $F_1 \cap F_2 = M$ implies $F_1 = M$ or $F_2 = M$. We denote by $\mathrm{Mi}(\mathbf{L})$ the meet-irreducible elements of the closure lattice $\mathbf{L} = (\mathcal{F}, \subseteq)$. For every closed set F, we have $F = \bigcap\{M \in \mathrm{Mi}(\mathbf{L}) : F \subseteq M\}$. Given $a \in X$ and $M \in \mathrm{Mi}(\mathbf{L})$, we write $a \uparrow M$ if $a \notin M$ but $a \in F$ for every $F \in \mathcal{F}$ such that $M \subset F$.

Example 1. Let $X = \{1, 2, 3, 4, 5, 6\}$ and consider the closure system $(X, \mathcal{F})$ whose closure lattice $\mathbf{L} = (\mathcal{F}, \subseteq)$ is depicted in Figure 1. For convenience we write a set as the concatenation of its elements, e.g. 1234 stands for $\{1, 2, 3, 4\}$. For instance, $\phi(25) = 2456$. We have $\mathrm{Mi}(\mathbf{L}) = \{356, 13, 15, 1356, 1456, 124, 2456, 12456\}$.

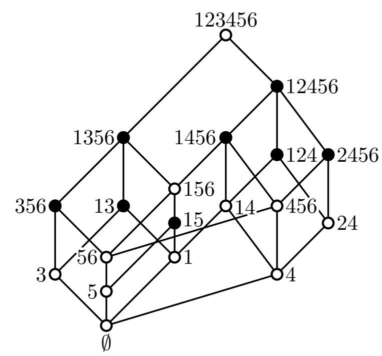

Figure 1: The closure lattice of Example 1. Meet-irreducible elements are marked by black dots.

Let $(X, \mathcal{F})$ be a standard closure system. We say that $(X, \mathcal{F})$ (and hence (X, ϕ)) is *atomistic* if $\phi(a) = \{a\}$ for every $a \in X$. The closure operator ϕ satisfies the *Anti-Exchange Property* (AEP), if $x, y \notin A = \phi(A)$, $x \neq y$ and $x \in \phi(A \cup \{y\})$ imply $y \notin \phi(A \cup \{x\})$. If, for the same assumptions, the conclusion is $y \in \phi(A \cup \{x\})$, the operator ϕ satisfies the *Exchange Property* (EP). The lattice of a closure system which statisfies EP is *geometric* and corresponds to the lattice of flats of a matroid (White 1986). A closure system whose closure operator satisfies AEP is a *convex geometry* (Edelman and Jamison 1985). In a convex geometry, each closed set F has a unique minimal spanning set K_F. If $\mathcal{F}$ is closed under union, $(X, \mathcal{F})$ is *distributive*. Distributive closure systems are convex geometries.

Implications, Implicational bases (IBs). We give some standard notations regarding implications and implicational bases (see also (Wild 2017)). Let X be a set. An *implication* over X is an expression $A \to c$ where $A \cup \{c\} \subseteq X$, also called *unit implication* in (Bertet and Monjardet 2010). In $A \to c$, A is the *premise* and c the *conclusion*. An implication is *binary* if A is a singleton. For simplicity, we write them $a \rightarrow c$ instead of $\{a\} \rightarrow c$. An *implicational base* (IB) over X is a pair (X, Σ) where Σ is a collection of implications over X. An IB induces a closure operator ϕ where a subset F of X is closed if for every implication $A \rightarrow c \in \Sigma$, $A \subseteq F$ implies $c \in F$. The closure of $C \subseteq X$ can be computed using the *forward chaining procedure*. This routine starts from C and construct a sequence $C = C_0, C_1, \ldots, C_m = \phi_\Sigma(C)$ of subsets of X such that, for every $1 \leq i \leq m$, $C_i = C_{i-1} \cup \{c \subseteq X : A \rightarrow c \in \Sigma, A \subseteq C_{i-1}\}$. Each closure system can be represented by several equivalent IBs. An implication $A \rightarrow c$ holds in a closure system if and only if $c \in \phi(A)$.

Example 2. An IB for the closure system of Example 1 is:

$$\Sigma = \{2 \rightarrow 4, 6 \rightarrow 5\} \cup \{245 \rightarrow 6, 34 \rightarrow 1, 34 \rightarrow 2, \\ 34 \rightarrow 5, 35 \rightarrow 6, 45 \rightarrow 6\}$$

We conclude this paragraph by recalling the well-known correspondence between IBs and (pure) Horn functions (see also (Wild 2017; Bertet et al. 2018)). More precisely, an implication $A \rightarrow c$ can be thought as a pure Horn clause $c \vee \bigvee_{a \in A} \neg a$ when the elements of X are seen as propositional variables. Thus, an IB (X, Σ) can be seen as a pure Horn CNF φ over variables X. Moreover, there is a direct one-to-one relationship between the closed sets associated to (X, Σ) and the models of the function described by ϕ.

Canonical direct and D-base. We focus on specific IBs: the *canonical direct base* and the *D-base*. In a later section, we also discuss a subset of D-base which forms the E-base in some closure systems. We delay the associated definitions to the appropriate place.

Let (X, ϕ) be a closure space and let $c \in X$. A subset A of X is a *(non-trivial) minimal generator of* c if $c \notin A$, $c \in \phi(A)$ but $c \notin \phi(A')$ for every $A' \subset A$. For $c \in X$, $\text{gen}_\delta(c)$ is the family of all minimal generators of c. The *canonical direct base* of the closure system $(X, \mathcal{F})$ is the IB (X, Σ_δ) where $\Sigma_\delta = \{A \rightarrow c : A \in \text{gen}_\delta(c), c \in X\}$. Implications of the canonical direct base are known as *prime implicates* in the corresponding Horn CNF. Hence, (X, Σ_δ) corresponds to the CNF with all possible prime implicates of the associated Horn function (Bertet and Monjardet 2010).

Example 3. IB of Example 2 is not (X, Σ_δ) of the closure system of Example 1. Indeed, $23 \in \text{gen}_\delta(1)$, but $23 \rightarrow 1$ is not in Σ. Also, $245 \rightarrow 6$ is in Σ, but $245 \notin \text{gen}_\delta(6)$ The following is the listing of Σ_δ from $\text{gen}_\delta(1)$ to $\text{gen}_\delta(6)$:

$$\Sigma_\delta = \left\{ \begin{array}{llll} 23 \rightarrow 1, & 34 \rightarrow 1, & 34 \rightarrow 2, & 2 \rightarrow 4, \\ 23 \rightarrow 5, & 34 \rightarrow 5, & 6 \rightarrow 5, & 23 \rightarrow 6, \\ 25 \rightarrow 6, & 34 \rightarrow 6, & 35 \rightarrow 6, & 45 \rightarrow 6 \end{array} \right\}$$

For the D-base, we need further definitions regarding closure systems. Let (X, ϕ) be a closure space. We consider a new closure operator ϕ^b over X defined by $\phi^b(A) = \bigcup_{a \in A} \phi(a)$. The corresponding closure system is $(X, \mathcal{F}^b)$ and the associated closure lattice is called $\mathbf{L}^b$. Note that $\mathcal{F}^b$ is closed under taking union, i.e. $F_1 \cup F_2 \in \mathcal{F}^b$ for every $F_1, F_2 \in \mathcal{F}^b$. Henceforth, the lattice $\mathbf{L}^b$ is *distributive*.

Let $c \in X$ and let $A \in \text{gen}(c)$. We say that A is a *D-generator* of c if $c \notin \phi^b(A)$ and for every $A' \in \text{gen}_\delta(c)$, $\phi^b(A') \subseteq \phi^b(A)$ implies $A' = A$. The family of D-generators of c is $\text{gen}_D(c)$. The D-base associated to the closure system $(X, \mathcal{F})$ is then (X, Σ_D) where $\Sigma_D = \{A \rightarrow c : A \in \text{gen}_D(c), c \in X\} \cup \Sigma^b$.

Example 4. Continuing Example 3, we can extract the IB (X, Σ^b) from (X, Σ): $\Sigma^b = \{2 \rightarrow 4, 6 \rightarrow 5\}$. The IB (X, Σ) in Example 2 is not the D-base of $(X, \mathcal{F})$. Indeed, $34 \in \text{gen}_D(6)$, yet $34 \rightarrow 6$ is not in Σ. Note that 25 is in $\text{gen}_\delta(6)$ but not in $\text{gen}_D(6)$ since $\phi^b(45) = 45 \subset 245 = \phi^b(25)$. Finally:

$$\Sigma_D = \Sigma^b \cup \{34 \rightarrow 1, 34 \rightarrow 2, 34 \rightarrow 5, \\ 34 \rightarrow 6, 35 \rightarrow 6, 45 \rightarrow 6\}$$

The D-base is always a subset of the canonical direct base. While the two IBs coincide in atomistic closure systems, there may be an exponential gap between them in the general case. This is demonstrated by the following example.

Example 5. Let $X = \{a_1, \ldots, a_n, b_1, \ldots, b_n, c\}$ and let $\Sigma = \{a_i \rightarrow b_i : 1 \leq i \leq n\} \cup \{b_1 \ldots b_n \rightarrow c\}$. The IB (X, Σ) is the D-base of its closure system, while we have $\Sigma_\delta = \{d_1 \ldots d_n \rightarrow c : d_i \in \{a_i, b_i\}, 1 \leq i \leq n\} \cup \{a_i \rightarrow b_i : 1 \leq i \leq n\}$. Thus, $|\Sigma| = n + 1$ and $|\Sigma_\delta| = 2^n + 1$.

As a consequence, one cannot straightforwardly apply the algorithms for finding the canonical direct base to find the D-base.

We conclude this paragraph by mentioning ϕ^b-minimal spanning sets, that will be useful in finding the D-base from an arbitrary IB. A minimal spanning set A of F is ϕ^b-minimal if for every (minimal) spanning set A' of F, $\phi^b(A') \subseteq \phi^b(A)$ implies $A' = A$. A ϕ^b-minimal spanning set of X is a ϕ^b-minimal key of the closure system.

Relations δ and D. Minimal and D-generators induce two binary relations over X. More precisely, we put $c\delta a$ if a belongs to a minimal generator of c. This is the δ-relation. Similarly, cDa means that a belongs to a D-generator of c.[2] The resulting relation is the D-relation. By definition, $D \subseteq \delta$, and the equality holds when $\Sigma^b = \emptyset$, i.e. when the underlying closure system is atomistic. However, in general, $D \subset \delta$ as illustrated in Figure 2.

There is a close connection between the relations δ and D and *implication-graphs* (Boros, Crama, and Hammer 1990; Hammer and Kogan 1995). The *implication-graph* of an IB (X, Σ) is the directed graph $G(\Sigma) = (X, E)$ where an arc (a, c) belongs to $G(\Sigma)$ if there exists an implication $A \rightarrow c$ in Σ such that $a \in A$. Now, the δ-relation precisely consists of the reversed arcs of $G(\Sigma_{cd})$. Similarly, the D-relation is obtained from $G(\Sigma_D)$ by removing the arcs coming from binary implications and reversing the remaining ones.

Example 6. Figure 2 illustrates the δ- and D-relations of the closure system of Example 1. Note that D is a proper subset of δ.

[2]the notations δ and D are standard terminology in lattice theory (Adaricheva, Nation, and Rand 2013; Bertet and Monjardet 2010; Freese, Ježek, and Nation 1995; Monjardet and Caspard 1997)

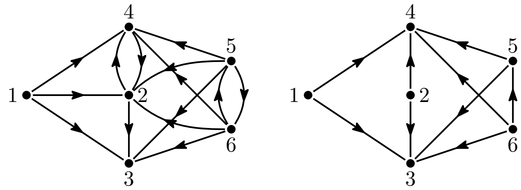

Figure 2: The δ-relation (left) and the D-relation (right) of Example 1 represented as directed graphs.

When the δ-relation of a closure system (X, Σ) does not have δ-cycles, we say that (X, Σ) is *acyclic*. Acyclic closure systems have been extensively studied as they correspond to acyclic Horn functions, poset type convex geometries and G-geometries (Boros et al. 2009; Hammer and Kogan 1995; Adaricheva 2017; Wild 1994). When the D-relation is acyclic the closure system is *lower bounded*, or simply without D-cycles (Freese, Ježek, and Nation 1995; Adaricheva, Nation, and Rand 2013). This class strictly generalizes acyclic closure systems. The closure system of Example 1 is without D-cycles but it contains δ-cycles. Hence, it is lower bounded but not acyclic.

Enumeration Complexity. We move to enumeration concepts (Johnson, Yannakakis, and Papadimitriou 1988). Let $\mathtt{A}$ be an enumeration algorithm with input x of size n and output a set of solutions $R(x)$ with size m. We assume that a solution in $R(x)$ has size poly(n). The algorithm $\mathtt{A}$ runs in *output-polynomial time* if its execution time is bounded by poly$(n + m)$. If the delay between two solutions output and after the last one is bounded by poly(n), $\mathtt{A}$ has *polynomial-delay*. Note that if $\mathtt{A}$ has polynomial-delay, it runs in output-polynomial time. We say that $\mathtt{A}$ runs in *output-quasipolynomial time* if its execution time is bounded by $2^{\text{polylog}(n+m)}$. An enumeration problem Π_1 is *harder than* than an enumeration problem Π_2 if there exists an output-polynomial time algorithm solving Π_2 whenever there is one solving Π_1. The problems Π_1 and Π_2 are *polynomially equivalent* if they are both harder than each other.

Computing the D-relation from an IB

In this section we establish the hardness of computing the D-relation of a closure system $(X, \mathcal{F})$ given by an arbitrary IB (X, Σ). The D-relation is a binary relation over elements of X. Hence, the problem of computing D can be reduced to the problem of checking whether cDa holds for each pair a, c of X:

D-Relation Recognition (D-RR)
Input: An IB (X, Σ), and $a, c \in X$.
Question: Does cDa hold?

The hardness of the corresponding problem for δ is a straightforward corollary of the hardness of the prime attribute problem in databases (Lucchesi and Osborn 1978). However, as demonstrated in Example 6, D is in general a proper subset of δ. This is due to the requirement on ϕ^b, the closure operator induced by the binary implications holding in $(X, \mathcal{F})$. Moreover, these existing results do not take care of the structure of the δ- and D-relations of the underlying closure system.

Here, we focus on the case where δ and D are acyclic. Note that since $D \subseteq \delta$, δ-acyclicity is stronger than D-acyclicity. As mentioned in the introduction, closure systems where δ or D are acyclic have been at the core of several positive results. However, we show that even when restricted to acyclic closure systems, D-RR is $\mathbf{NP}$-complete.

As a preliminary step, we show that D-RR belongs to $\mathbf{NP}$. On this purpose, we use a characterization of D-generators given in (Freese, Ježek, and Nation 1995). It is used in Algorithm 11.13 in combination with Equation (9) (p. 230). We rephrase it in our terminology.

Lemma 1. *Let $(X, \mathcal{F})$ be a closure system, $A \subseteq X$ and $c \in X$. Then, A is a D-generator of c if and only if for every $a \in A$, $c \notin \phi(\phi^b(A) \setminus \{a\})$.*

Corollary 2. *Let $(X, \mathcal{F})$ be a closure system, let $A \subseteq X$ and $c \in X$. Then, whether A is a D-generator of c can be checked with a number of calls to ϕ being polynomial in the size of X.*

One can compute ϕ in polynomial time from both an IB or meet-irreducible elements. We deduce:

Corollary 3. *DB-M belongs to* $\mathbf{NP}$.

We proceed to the hardness of D-RR in acyclic closure systems. In fact, the hardness result also holds under the assumption that the premises of the input IB have constant size. We give here the reduction, but the proof is omitted due to space limitations.

Theorem 4. *The problem D-RR is $\mathbf{NP}$-complete in acyclic closure systems represented by IBs with premises of size at most* 2.

Proof sketch. We first show that D-RR belongs to $\mathbf{NP}$. A certificate is a D-generator of A of c such that $a \in A$. Note that (X, Σ^b) can be computed in polynomial time from (X, Σ). Thus checking whether A is indeed a D-generator can be achieved in polynomial time. It follows that D-RR is in $\mathbf{NP}$.

To show hardness, we use a reduction from 1-in-3-SAT.

1-in-3 SAT
Input: A positive 3-CNF $\varphi = \{C_1, \dots, C_m\}$ over variables $V = \{v_1, \dots, v_n\}$
Question: Is there an assignment T of the variables in V such that $|T \cap C_i| = 1$ for each C_i? (T is seen as the set of variables set to 1)

Let $\varphi = \{C_1, \dots, C_m\}$ be a positive 3-CNF over variables $V = \{v_1, \dots, v_n\}$. We say that two variables v_i, v_j are in conflict if there exists a clause C_k in φ that contains both v_i and v_j. Let $X = \{c_1, \dots, c_m\} \cup \{x_{i,j} : 1 \leq i \leq m, 1 \leq j \leq 3\} \cup \{z\}$. An element c_i represents the clause C_i, while $x_{i,j}$ represents the j-th literal of the clause C_i. Now consider the following set of implications:

$$\Sigma = \Sigma_{var} \cup \Sigma_{mod} \cup \Sigma_{conf}, \text{ where}$$

- $\Sigma_{var} = \{x_{i,j} \to x_{k,\ell} : 1 \leq i \leq k \leq m, 1 \leq j, \ell \leq 3 \text{ and } x_{i,j} = x_{k,\ell} \text{ in } \varphi\}$,

- $\Sigma_{mod} = \{c_i x_{i,j} \to c_{i+1} : 1 \leq i < m, 1 \leq j \leq 3\} \cup \{c_m x_{m,j} \to z : 1 \leq j \leq 3\}$,
- $\Sigma_{conf} = \{x_{i,j} x_{k,\ell} \to z : 1 \leq i, k \leq m, 1 \leq j, \ell \leq 3$ and $x_{i,j}, x_{k,\ell}$ are in conflict in $\varphi\}$.

Informally, Σ_{var} connects the $x_{i,j}$'s that are the same variable in φ, Σ_{mod} models the fact that whenever the i-th clause is satisfied, we can proceed to the $(i+1)$-th clause or z, and Σ_{conf} represents the conflicts induced by the variables appearing together in a clause.

Considering the implicational base (X, Σ) the rest of the proof consists in showing that zDc_1 holds if and only if there is a 1-in-3 assignement of the variables in φ. Note that since $G(\Sigma)$ is acyclic, $(X, \mathcal{F})$ is also acyclic (Hammer and Kogan 1995), and that the premises of Σ have size at most 2. Moreover, as $\phi^b(c_1) = \{c_1\}$, $z\delta c_1$ is equivalent to zDc_1. For this reason, the analysis can be conducted over minimal generators instead of D-generators. ☐

Understanding what makes hard to decide whether cDa holds is an intriguing question. A first idea comes from the observation that in Theorem 4, c_1 and z are separated by a long D-path, i.e. a long directed path when the D-relation is seen as a directed graph. Namely, we have $zDc_mD \dots Dc_2Dc_1$. This suggests the difficulty of solving D-RR comes from the fact that two elements are separated by several implications. However, we prove that the problem remains $\mathbf{NP}$-complete even when the length of a D-path is constant. Interestingly though, this complexity comes at the cost of cycles in the δ-relation.

Theorem 5. *The problem D-RR remains* $\mathbf{NP}$*-complete even in closure systems without D-cycles where the length of a longest D-path between two elements is at most* 2.

Proof sketch. The reduction is another reduction from 1-in-3 SAT. Let $\varphi = \{C_1, \dots, C_m\}$ be a positive 3-CNF over variables $V = \{v_1, \dots, v_n\}$. Hence, let $V = \{x_1, \dots, x_n\}$ be a set of variables and let $\varphi = \{C_1, \dots, C_m\}$ be a positive 3-CNF over V. For convenience, we put $C_i = \{\ell_{i,1}, \ell_{i,2}, \ell_{i,3}\}$ for each $1 \leq i \leq m$. Let $X = V \cup \{c_1, \dots, c_m\} \cup \{a, b\}$ be a new groundset, and let Σ be the following set of implications:

$$\Sigma = \Sigma_{mod} \cup \Sigma_{conf} \cup \{c_i \to a : 1 \leq i \leq m\}$$

where:

- $\Sigma_{mod} = \{av_i \to c_j : 1 \leq i \leq n, 1 \leq j \leq m$ and $v_i \in C_j\} \cup \{c_1 \dots c_m \to b\}$
- $\Sigma_{conf} = \{v_i v_j \to b : 1 \leq i, j \leq n$ and v_i, v_j are in conflict in$\varphi\}$

Informally, Σ_{conf} models all the conflicts in between two variables appearing in a clause and Σ_{mod} models the fact that taking one element in a clause together with a for all the clauses yields a 1-in-3 assignment of φ. Finally, the implications $\{c_i \to a : 1 \leq i \leq m\}$ guarantee that a will never appear in a minimal generator together with one of the c_i's.

Note that (X, Σ) can be constructed in polynomial time from φ and V. Analyzing the D-relation we obtain relations of the form c_iDa, c_iDv_j if $v_j \in c_i$, bDc_i and bDv_j if v_j is in a conflict. Then, bDa will hold if and only if φ has a 1-in-3 assignment. ☐

Whether the D-relation can be computed in polynomial time when restricted to acyclic closure systems with D-paths of constant size remains an open question for future work.

Computing the D-base

In this section we give algorithms to compute the D-base from an IB or a set of meet-irreducible elements.

D-base from Meet-irreducible elements (DB-M)
Input: The family of meet-irreducible elements $\mathrm{Mi}(\mathbf{L})$ of a closure system $(X, \mathcal{F})$
Output: The D-base (X, Σ_D) associated to $(X, \mathcal{F})$

D-base from IB (DB-IB)
Input: An IB (X, Σ)
Output: The D-base (X, Σ_D) associated to (X, Σ)

For DB-M, we show that the problem is equivalent to the dualization of distributive lattices represented by implicational bases. A recent algorithm (Elbassioni 2022) solves this latter problem in output-quasipolynomial time. As for DB-IB we give a polynomial delay algorithm based on a procedure in (Ennaoui and Nourine 2016) for listing the so-called minimal key-ideals of a closed set.

Solving DB-M in output-quasipolynomial time

We show that the problem DB-M is polynomially equivalent to the problem of dualization a distributive lattice given by an IB. We deduce an algorithm to solve DB-M in output-quasipolynomial time.

We first give some definitions related to dualization. Let $\mathbf{L} = (\mathcal{F}, \subseteq)$ be the closure lattice associated to a closure system $(X, \mathcal{F})$ and let $\mathcal{B}^-$, $\mathcal{B}^+$ be two antichains of $\mathbf{L}$. We say that $\mathcal{B}^-$ and $\mathcal{B}^+$ are *dual* in $\mathbf{L}$ if $\uparrow \mathcal{B}^- \cup \downarrow \mathcal{B}^+ = \mathcal{F}$ and $\uparrow \mathcal{B}^- \cap \downarrow \mathcal{B}^+ = \emptyset$. For each antichain $\mathcal{B}^+$ in $\mathbf{L}$, there exists a unique dual antichain $\mathcal{B}^-$ to $\mathcal{B}^+$. In particular, we have $\mathcal{B}^- = \min_{\subseteq}(\{F \in \mathbf{L} : F \nsubseteq B, \text{ for all } B \in \mathcal{B}^+\})$. Symmetrically, each antichain $\mathcal{B}^-$ has a unique dual antichain $\mathcal{B}^+$. Now the problem of dualizing distributive lattices reads as

Distributive Lattice Dualization (DLD)
Input: An IB (X, Σ) of a distributive closure system $(X, \mathcal{F})$ with closure lattice $\mathbf{L}$, and an antichain $\mathcal{B}^+$ of $\mathbf{L}$.
Output: The dual antichain $\mathcal{B}^-$ of $\mathcal{B}^+$ in $\mathbf{L}$.

This problem is a generalization of hypergraph dualization. In hypergraph dualization, Σ is empty and the corresponding distributive lattice is in fact the Boolean lattice consisting in the powerset of X.

We proceed to establish that DB-M is harder than DLD.

Lemma 6. *There is an output-polynomial time algorithm solving DLD if there is one solving DB-M.*

We now show that DB-M reduces to solving $|X|$ instances of DLD. We first establish a correspondence between dualization in the distributive lattice $\mathbf{L}^b$ and the D-generators of $(X, \mathcal{F})$.

Lemma 7. *Let* $(X, \mathcal{F})$ *be a standard closure system. For* $c \in X$*, consider the two antichains of* $\mathbf{L}^b$*:*

- $\mathcal{B}^+ = \{M \in \mathrm{Mi}(\mathbf{L}) : c \uparrow M\}$
- $\mathcal{B}^- = \{\phi^b(A) : A \in \mathrm{gen}_D(c)\} \cup \{\phi^b(c)\}$

Then, the two antichains $\mathcal{B}^+$ and $\mathcal{B}^-$ are dual in $\mathbf{L}^b$. Moreover, ϕ^b is a one-to-one correspondence between $\mathrm{gen}_D(c)$ and $\{\phi^b(A) : A \in \mathrm{gen}_D(c)\}$.

An IB (X, Σ^b) can be computed in polynomial time from $\mathrm{Mi}(\mathbf{L})$ by setting $\Sigma^b = \{a \to c : a \in X, c \in \phi(a)\}$. The antichain $\mathcal{B}^+$ can also be computed in polynomial time by checking for each $M \in \mathrm{Mi}(\mathbf{L})$ whether or not $c \uparrow M$. Thus, listing the D-generators of some $c \in X$ reduces to DLD. By calling DLD for each $c \in X$ we deduce:

Lemma 8. *There exists an output-polynomial time algorithm solving DB-M if there is one solving DLD.*

Combining Lemmas 6 and 8 we obtain the desired result.

Theorem 9. *The problems DB-M and DLD are polynomially equivalent.*

It is shown in (Elbassioni 2022) that DLD can be solved in output-quasipolynomial time. We deduce:

Corollary 10. *The problem DB-M can be solved in output-quasipolynomial time.*

A polynomial-delay algorithm for DB-IB

Let (X, Σ) be an IB for $(X, \mathcal{F})$. We give an algorithm to solve DB-IB with polynomial delay. Our algorithm adapts the *supergraph traversal* procedure in (Ennaoui and Nourine 2016) for the enumeration of the ϕ^b-minimal keys of a closure system. Recall that a minimal key K of $(X, \mathcal{F})$ is ϕ^b-minimal if $\phi^b(K') \nsubseteq \phi^b(K)$ for each minimal key K' distinct from K.

As a first step, we show that given $c \in X$, we can list $\mathrm{gen}_D(c)$ with polynomial-delay. Define $\Sigma_c = \Sigma \cup \{c \to X\}$. We put $(X, \mathcal{F}_c)$ as the closure system associated to (X, Σ_c).

Lemma 11. *Let $(X, \mathcal{F})$ be a closure system, $c \in X$ and $A \subseteq X$. Then, $A \in \mathrm{gen}_D(c)$ if and only if A is a ϕ^b-minimal key in $(X, \mathcal{F}_c)$.*

The next statement describes a polynomial-delay algorithm which takes as input an IB (X, Σ) and computes the D-generators of an element in X.

Theorem 12. (Ennaoui and Nourine 2016, Proposition 8) *Let (X, Σ) be a standard IB, let $c \in X$ and let $\mathcal{S} \subseteq \mathrm{gen}_D(c)$ be such that $\mathcal{S} \neq \emptyset$. Then, $\mathcal{S} \neq \mathrm{gen}_D(c)$ iff there exists $A \in \mathcal{S}$ and $B \to d \in \Sigma$ such that $|B| \geq 2$ and $\phi^b((\phi^b(A) \setminus \phi^b(d)) \cup B)$ does not contain any set of $\mathcal{S}$.*

The *supergraph* $\mathcal{G}(c)$ of D-generators of c, is a graph whose vertices are the sets in $\mathrm{gen}_D(c)$, and an edge connects A_1 to A_2 if there exists $B \to c \in \Sigma$ such that $|B| \geq 2$ and $A_2 = \mathtt{Minimize}(\phi^b((\phi^b(A_1) \setminus \phi^b(c)) \cup B))$. The procedure `Minimize()` is a greedy algorithm that computes a D-generator from any generator, according to an ordering of the vertices of X: $a < d$, provided $\phi^b(a) \subset \phi^b(d)$. Such a procedure `Minimize` can be directly deduced from Lemma 1 and 2. Due to (Ennaoui and Nourine 2016), $\mathcal{G}(c)$ is strongly connected. Using a queue, a DFS-procedure lists all the D-generators with polynomial-delay and possibly exponential space.

Corollary 13. *Let (X, Σ) be an implicational base of a closure system $(X, \mathcal{F})$ and let $c \in X$. Then, there is a polynomial-delay algorithm which computes $\mathrm{gen}_D(c)$ from (X, Σ).*

We are interested to enumerate all D-generators of Σ, i.e. the set $\bigcup_{c \in X} \mathrm{gen}_D(c)$. Notice that a D-generator may belong to $\mathrm{gen}_D(a) \cap \mathrm{gen}_D(c)$ with $a \neq c$.

Example 7. In (X, Σ_D) of Example 4, 34 belongs to $\mathrm{gen}_D(1)$, $\mathrm{gen}_D(2)$, $\mathrm{gen}_D(5)$ and $\mathrm{gen}_D(6)$.

To avoid repetitions in the enumeration, we consider the supergraph $\mathcal{G}(X)$ being the union of the supergraphs $\mathcal{G}(c)$ for all $c \in X$. Note that $\mathcal{G}(X)$ may not be strongly connected, but it is a union of strongly connected graphs. Moreover, we can compute in polynomial time a D-generator in $\mathcal{G}(c)$ for every $c \in X$ by calling `Minimize` on $X \setminus \{c\}$. We denote by $\mathrm{root}(\mathcal{G}(X))$ these D-generators. Due to Corollary 13 we can traverse all vertices of $\mathcal{G}(X)$ starting from vertices in $\mathrm{root}(\mathcal{G}(X))$. Using a queue and DFS-search starting from any vertex in $\mathrm{root}(\mathcal{G}(X))$, we can list all D-generators without repetitions. When obtaining a solution A, we output the implications $A \to c$ for each c such that $A \in \mathrm{gen}_D(c)$. Since Σ^b can be computed in polynomial time from Σ at pre-proccesing time. We deduce

Theorem 14. *The problem DB-IB can be solved with polynomial-delay.*

E-base: a subset of the D-base

An important subset of the D-base is the *E-base* (Adaricheva, Nation, and Rand 2013). Terminology similar to D-relation follows. Let $A \subseteq X$ and $c \in X$. The set A is a *E-generator* of c if it is a D-generator of c whose closure is inclusion-wise minimal among closures of other minimal generators of E, i.e. if for every $A' \in \mathrm{gen}_D(c)$, $\phi(A') \subseteq \phi(A)$ entails $\phi(A) = \phi(A')$. Let $\mathrm{gen}_E(c)$ denote the E-generators of c. The E-base (X, Σ_E) of $(X, \mathcal{F})$ is $\Sigma_E = \{A \to c : A \in \mathrm{gen}_E(c), c \in X\} \cup \Sigma^b$.

Remark 2. There is connection between E-generators and *critical* generators known in closures systems with AEP (Korte, Lovász, and Schrader 2012). Moreover, closures of E-generators are *essential* closed sets (Wild 1994), which cannot be said about D-generators. These connections will be discussed in detail in (Vilmin 2023).

Unlike the canonical direct and D-bases, the E-base may not always form a valid IB of (X, ϕ) (Adaricheva, Nation, and Rand 2013). This is demonstrated by the next example.

Example 8. Let $X = \{1, 2, 3, 4\}$ and consider the closure system defined by its D-base (X, Σ_D) with $\Sigma_D = \{13 \to 2, 24 \to 3, 14 \to 2, 14 \to 3\}$. The implications $14 \to 2$ and $14 \to 3$ are not in the E-base since $\phi(13), \phi(24) \subset \phi(14) = X$. Hence, the E-base comprises only $13 \to 2$ and $24 \to 3$, which do not define the same closure system.

Still, in several important classes of closure systems the E-base is valid. One of them is the class of closure systems without D-cycles. The closure system of Example 4 fulfills this property.

Example 9 (Example 4 continued)**.** To see that the closure system is without D-cycles, remark that the directed graph to the right of Figure 2 representing its D-relation is acyclic.

Now, implication $34 \rightarrow 6$ is in the D-base, but not in the E-base, because there is $35 \rightarrow 6$, with $\phi(35) \subset \phi(34)$. Removing $34 \rightarrow 6$ will still provide the base of the same system, because $34 \rightarrow 6$ follows from $34 \rightarrow 5$ and $35 \rightarrow 6$. Thus, this gives an example of the valid E-base.

In the first part of this section we focus on the subclass of systems without D-cycles. We consider its subclass with (AEP) which is called *D-geometries*, and show that they can be recognized efficiently from an arbitrary IB. In a second part, we collect references and observations about E-base in systems satisfying (EP). To the best of our knowledge, little is known about E-base outside these two classes.

D-geometries. They are closure spaces (X, ϕ) without D-cycles satisfying (AEP). Hence they are convex geometries. While it is known that closure systems without D-cycles can be recognized in polynomial time from an IB (Adaricheva, Nation, and Rand 2013), it was shown in (Bichoupan 2022) that recognizing convex geometries from an arbitray IB is $\mathbf{coNP}$-complete. Nevertheless, (Adaricheva, Freese, and Nation 2022) state that convex geometries could be effectively identified from the D-base of a closure system whose closure lattice is *join semidistributive* (which generalizes both (AEP) and D-acyclicity). Here, we prove that under the assumption of D-acyclicity, (AEP) can be recognized in polynomial time from an arbitrary IB. Since D-acyclicity is easily verified, we deduce that the problem of recognizing D-geometries admits a polynomial time algorithm.

D-Geometry Recognition (DGR)
Input: An implicational base (X, Σ).
Question: Is $(X, \mathcal{F})$ a D-geometry?

Theorem 15. *A closure system without D-cycles is not a D-geometry if and only if its E-base contains implications of the form $Ac \rightarrow d$ and $d \rightarrow c$.*

In closure systems without D-cycles, the E-base can be retrieved in polynomial time from any IB (Adaricheva and Nation 2014). We obtain:

Corollary 16. DGR *can be solved in polynomial time.*

Systems with the Exchange Property Atomistic closure systems with (EP) of associated closure operator correspond to *lattices of flats of matroids*. The study of implicational bases of matroids was initiated in (Wild 1994) and continued in (Bérczi, Boros, and Makino 2023b). In particular, the case of binary matroids, thus, graphic matroids and partition lattices, is settled in the following statement. This section assumes familiarity with concepts from matroid theory and implicational systems. We refer the reader to (Wild 1994; White 1986) for the necessary definitions.

Theorem 17. (Wild 1994, Theorem 10) *In a simple (i.e. standard) binary matroid, a closed set is essential if and only if it is a closed circuit.*

Corollary 18. *Binary matroids have the valid E-base.*

In fact, we observe more generally that any closure system where essential sets are incomparable have valid E-base. It turns out that, besides simple binary matroids, all biatomic modular lattices fall into this scope. We note that every atomistic modular closure system is a matroid.

Proposition 1. (Vilmin 2023) *Let $(X, \mathcal{F})$ be an atomistic and modular closure system. Then its E-base is valid.*

Projective geometry is a rank-3 matroid, where every two rank-2 elements have a common atom.

Proposition 2. *Projective geometry has a valid E-base.*

Note that there exist projective geometries which are neither binary matroids, nor modular.

Example 10. With $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$ consider the lines $1234, 561, 572, 673, 178, 268, 358$. Any two lines have a common point. Since $1, 2, 3, 4$ form an interval isomorphic to $\mathbf{M}_4$, the 6-element lattice with four atoms, this matroid is not binary: in binary matroids every interval of height 2 has at most 5 elements. Moreover, the lattice is not *bi-atomic*, since $567 \rightarrow 4$, but $57 \rightarrow 4, 67 \rightarrow 4, 17 \rightarrow 4$ do not hold. As every atomistic modular lattice is be bi-atomic (Bennett 1987), we conclude it is not modular.

But not all matroids have valid E-base, as the following example demonstrates.

Example 11. Let $X = \{1, 2, 3, 4, 5\}$ and consider the atomistic semimodular (non-modular) closure system $(X, \mathcal{F})$ where $\mathcal{F}$ contains $\emptyset$, X, 123 and all singletons and pairs except 12, 13 and 23. The following implications holds: $12 \rightarrow 3$, $13 \rightarrow 2$, $23 \rightarrow 1$, and $abc \rightarrow X$, for all triples abc except 123. Hence, the implication $245 \rightarrow 1$ appears in the D-base, but not in the E-base. Indeed, 23 is a E-generator of 1 satisfying $\phi(23) \subset \phi(245)$. Yet, we cannot get $245 \rightarrow 1$ from E-covers. For this, one would need to use $23 \rightarrow 1$, the only E-cover available for 1, but one cannot get 3 from $2, 4, 5$ either, because $245 \rightarrow 3$ is also in D-base but 245 is not an E-generator of 3. Thus, $245 \rightarrow 1$ cannot be dropped and the E-base is not valid.

This prompts the following concluding questions.

Question 2. What are the matroids for which E-base is valid?

Question 3. Is it possible that, even though the E-base of a matroid is not valid, all essential sets are the closure of some critical circuits (E-generator)?

Conclusion

In this paper, we have investigated complexity aspects of the D-base and D-relation of a closure system. We have shown that computing the D-relation from an IB is $\mathbf{NP}$-complete even in the acyclic setup. Besides, we gave output-sensitive algorithms to compute the D-base from both an IB and meet-irreducible elements. Studying the structure of the D-relation in order to uncover other properties of closure systems is a topic of further research.

The E-base is also particularly intriguing. It does not always constitute a valid IB, and we highlighted some classes of closure systems (AEP or EP) where it is a faithful representation. Characterizing closure systems with valid E-base seems to be challenging and fascinating.

References

Adaricheva, K. 1991. A characterisation of finite lattices of subsemilattices. *Algebra i Logika*, 30: 385–404.

Adaricheva, K. 2017. Optimum Basis of Finite Convex Geometry. *Discrete Applied Mathematics*, 230: 11–20.

Adaricheva, K.; Bernhardt, J.; Liu, W.; and Schmidt, B. 2023. Importance of overnight parameters to predict sea breeze on Long Island. arxiv:2309.01803.

Adaricheva, K.; Freese, R.; and Nation, J. B. 2022. Notes on join semidistributive lattices. *International Journal of Algebra and Computation*, 32(02): 347–356.

Adaricheva, K.; and Nation, J. B. 2014. On Implicational Bases of Closure Systems with Unique Critical Sets. *Discrete Applied Mathematics*, 162: 51–69.

Adaricheva, K.; and Nation, J. B. 2017. Discovery of the D-Basis in Binary Tables Based on Hypergraph Dualization. *Theoretical Computer Science*, 658: 307–315.

Adaricheva, K.; Nation, J. B.; and Rand, R. 2013. Ordered Direct Implicational Basis of a Finite Closure System. *Discrete Applied Mathematics*, 161(6): 707–723.

Agrawal, R.; Mannila, H.; Srikant, R.; Toivonen, H.; Verkamo, A. I.; et al. 1996. Fast discovery of association rules. *Advances in knowledge discovery and data mining*, 12(1): 307–328.

Bennett, M. K. 1987. Biatomic lattices. *Algebra Universalis*, 24: 60–73.

Bérczi, K.; Boros, E.; and Makino, K. 2023a. Hypergraph Horn functions. *arXiv preprint arXiv:2301.05461*.

Bérczi, K.; Boros, E.; and Makino, K. 2023b. Matroid Horn functions. *arXiv preprint arXiv:2301.06642*.

Bertet, K.; Demko, C.; Viaud, J.-F.; and Guérin, C. 2018. Lattices, closures systems and implication bases: A survey of structural aspects and algorithms. *Theoretical Computer Science*, 743: 93–109.

Bertet, K.; and Monjardet, B. 2010. The multiple facets of the canonical direct unit implicational basis. *Theoretical Computer Science*, 411(22-24): 2155–2166.

Bichoupan, T. 2022. Complexity Results for Implication Bases of Convex Geometries. *arXiv preprint arXiv:2211.08524*.

Bichoupan, T. 2023. Independence of Essential Sets in Finite Implication Bases. *arXiv preprint arXiv:2304.06837*.

Boros, E.; Čepek, O.; and Kogan, A. 1998. Horn minimization by iterative decomposition. *Annals of Mathematics and Artificial Intelligence*, 23: 321–343.

Boros, E.; Čepek, O.; Kogan, A.; and Kučera, P. 2009. A subclass of Horn CNFs optimally compressible in polynomial time. *Annals of mathematics and artificial intelligence*, 57: 249–291.

Boros, E.; Crama, Y.; and Hammer, P. L. 1990. Polynomial-time inference of all valid implications for Horn and related formulae. *Annals of Mathematics and Artificial Intelligence*, 1: 21–32.

Day, A. 1992. Doubling constructions in lattice theory. *Canadian J. of Mathematics*, 44: 252–269.

Doignon, J.-P.; and Falmagne, J.-C. 2012. *Knowledge spaces*. Springer Science & Business Media.

Edelman, P. H.; and Jamison, R. E. 1985. The Theory of Convex Geometries. *Geometriae dedicata*, 19(3): 247–270.

Eiter, T.; Makino, K.; and Gottlob, G. 2008. Computational aspects of monotone dualization: A brief survey. *Discrete Applied Mathematics*, 156(11): 2035–2049.

Elbassioni, K. 2022. On dualization over distributive lattices. *Discrete Mathematics & Theoretical Computer Science*, 24(Discrete Algorithms).

Ennaoui, K.; and Nourine, L. 2016. Polynomial delay Hybrid algorithms to enumerate candidate keys for a relation. In *BDA 2016*.

Faigle, U.; and Herrmann, C. 1981. Projective geometry on partially ordered sets. *Transactions of the American Mathematical Society*, 266(1): 319–332.

Fredman, M. L.; and Khachiyan, L. 1996. On the complexity of dualization of monotone disjunctive normal forms. *Journal of Algorithms*, 21(3): 618–628.

Freese, R.; Ježek, J.; and Nation, J. B. 1995. *Free Lattices*, volume 42. American Mathematical Soc.

Gratzer, G. A. 2011. Lattice theory: foundation.

Guigues, J.-L.; and Duquenne, V. 1986. Familles Minimales d'implications Informatives Résultant d'un Tableau de Données Binaires. *Mathématiques et Sciences humaines*, 95: 5–18.

Hammer, P. L.; and Kogan, A. 1995. Quasi-acyclic propositional Horn knowledge bases: Optimal compression. *IEEE Transactions on knowledge and data engineering*, 7(5): 751–762.

Johnson, D. S.; Yannakakis, M.; and Papadimitriou, C. H. 1988. On generating all maximal independent sets. *Information Processing Letters*, 27(3): 119–123.

Jónsson, B.; and Nation, J. B. 1975. A report on sublattices of a free lattice. *Colloq. Math. Soc. Janos Bolyai*, 17: 223–257.

Korte, B.; Lovász, L.; and Schrader, R. 2012. *Greedoids*, volume 4. Springer Science & Business Media.

Lucchesi, C. L.; and Osborn, S. L. 1978. Candidate keys for relations. *Journal of Computer and System Sciences*, 17(2): 270–279.

Mannila, H.; and Räihä, K.-J. 1992. *The design of relational databases*. Addison-Wesley Longman Publishing Co., Inc.

Monjardet, B.; and Caspard, N. 1997. On a dependence relation in finite lattices. *Discrete Mathematics*, 165: 497–505.

Nation, J.; Cabot-Miller, J.; Segal, O.; Lucito, R.; and Adaricheva, K. 2021. Combining algorithms to find signatures that predict risk in early stage stomach cancer. *Journal of Computational Biology*, 28(310): 985–1006.

Nourine, L.; and Vilmin, S. 2023. Hierarchical decompositions of implicational bases for the enumeration of meet-irreducible elements. *Theoretical Computer Science*, 114030.

Rodríguez-Lorenzo, E.; Adaricheva, K.; Cordero, P.; Enciso, M.; and Mora, A. 2017. Formation of the d-basis from implicational systems using simplification logic. *International Journal of General Systems*, 46(5): 547–568.

Rodríguez-Lorenzo, E.; Adaricheva, K.; Cordero-Ortega, P.; Enciso-García-Oliveros, M.; Mora-Bonilla, A.; et al. 2015. From an implicational system to its corresponding D-basis.

Vilmin, S. 2023. Critical generators and pseudo-closed sets.

White, N. 1986. *Theory of matroids*. 26. Cambridge University Press.

Wild, M. 1994. A theory of finite closure spaces based on implications. *Advances in Mathematics*, 108(1): 118–139.

Wild, M. 2017. The joy of implications, aka pure Horn formulas: mainly a survey. *Theoretical Computer Science*, 658: 264–292.