

# ON IMPLICATIONAL BASES OF CLOSURE SYSTEMS WITH UNIQUE CRITICAL SETS

K. ADARICHEVA AND J. B. NATION

ABSTRACT. Presented results are inspired by the study of closure systems with unique critical sets, but many have the general nature. Among those is the statement that every optimum basis of a finite closure system, in D. Maier's sense, is also right-side optimum, which is a parameter of a minimum CNF representation of a Horn Boolean function. New parameters for the size of the binary part are also established. As a part of general studies, we introduce the K-basis of a closure system, which is a refinement of the canonical basis of V. Duquenne and J.L. Guigues, and discuss a polynomial algorithm to obtain it. The main part of the paper is devoted to the closure systems with unique critical sets, and some subclasses of these where the K-basis is unique. A further refinement in the form of the E-basis is possible for closure systems without D-cycles. There is a polynomial algorithm to recognize the D-relation from a K-basis. Thus, closure systems without D-cycles can be effectively recognized. While the E-basis achieves an optimum in one of its parts, the optimization of the others is an NP-complete problem.

## 1. INTRODUCTION

Closure system on a finite set is a unifying concept in logic programming, relational data bases and knowledge systems. Closure systems can be defined by a set of implications (a basis), and in such form they appear as Horn formulas in logic programming, dependencies in relational data bases, CNF representations of Horn Boolean functions and directed hypergraphs in discrete optimization.

Closure systems can also be presented in the terms of finite lattices, and the tools of economic description of a finite lattice have long existed in lattice theory. In this paper we continue the study of economic representation of a closure system based on the structure of its closure lattice, initiated in K. Adaricheva, J.B. Nation and R. Rand [2].

Since the seminal work of D. Maier [19], the main parameters of effective representation of a closure system have been (1) the number of implications in a basis, or (2) the total number of literals in all implications of the basis. It was shown by Maier, in a non-trivial argument, that the set of implications that achieves the minimum in the second parameter (an optimum basis) also achieves the minimum in the first (a minimum basis). The result of V. Duquenne and J.L. Guigues [13] is

---

2010 *Mathematics Subject Classification.* 05A05, 06B99, 52B05.

*Key words and phrases.* Closure systems, lattices of closed sets, canonical basis, stem basis, Duquenne-Guigues basis, unit basis, optimum basis, minimum basis, acyclic Horn formulas, shortest DNF-representation, shortest CNF-representation, shortest representations of acyclic hypergraphs, finite semidistributive lattices, lattices without  $D$ -cycles.

The first author was partially supported by AWM-NSF Mentor Travel grant N0839954.

that every closure system has a canonical minimum basis; moreover, it can be obtained from any given basis in time polynomial in the size of that basis, see A. Day [11]. Quite to the opposite, Maier showed that the problem of finding an optimum basis is NP-complete [19].

On the other hand, in the theory of Horn Boolean functions and directed hypergraphs, other parameters were developed, and no connection with Maier's parameters was so far realized. We establish such a connection in Theorem 10, for the minimum representation of a Horn Boolean function.

Then we introduce a  $K$ -basis in general closure systems following the idea of the minimal join representation of elements in finite lattices. This produces a refinement of the canonical basis of V. Duquenne and J.L. Guigues [13]. While the  $K$ -basis is not optimum, it does provide a reduction in size of the canonical basis and can be obtained in polynomial time from the canonical basis.

The  $K$ -basis allows us to establish an important link between the canonical basis and the  $D$ -relation in the closure lattice of a closure system: the latter plays an important role in the lattice theoretical literature, see R. Freese, J. Ježek and J.B. Nation [14]. In particular, the  $D$ -relation can be effectively recovered from the canonical basis via its refinement to any  $K$ -basis. This allows us to recognize closure systems without  $D$ -cycles (Theorem 43), which are generalizations of the quasi-acyclic systems defined in P. Hammer and A. Kogan[16].

We also suggest the general concept of *partial optimizations* based on the idea that any basis can always be divided into two natural parts: so-called *binary* and *non-binary*. Either of those parts can be optimized given various assumptions. Alternatively, a basis can be minimized with respect to the total size of all the *premises* of implications, which we call *left-side optimum*, or with respect to the total size of all the *conclusions*, which is called *right-side optimum*.

The essential part of the paper is devoted to the study of effective representations of closure systems with unique critical sets, or *UC-systems*. In fact, this class of closure systems was the main stimulus for the whole paper: the results of the first sections, which are proved for the general closure systems, grew out of the study of *UC-systems*. The definition of this class is based on the notion of *essential* and *critical* sets associated with a given closure system. In one important subclass of such systems, the  $K$ -basis is unique.

Further refinement in the form of the  $E$ -basis is possible in systems without  $D$ -cycles, which form a proper subclass of *UC-systems*. This basis is right-side optimized in its non-binary part. Still, we show that finding an optimum basis for such systems is an NP-complete problem.

The paper is organized as follows. We collect all the required definitions and recall important results in section 2. Then section 3 establishes the relationship between optimum and right-side optimum bases (Theorem 10). Section 4 deals with the problem of optimization of the binary part. It turns out, such optimization is independent of the form and size of the non-binary part (Theorem 20). We also introduce and discuss the concept of a *regular* basis (Definition 14) that becomes essential in section 6. In section 5 we introduce the  $K$ -basis (Definition 27) and describe a polynomial algorithm for retrieving it from the canonical basis (Proposition 24). This will be used in section 7, which discusses the  $K$ -basis in systems with unique critical sets. In section 6, a relationship between a canonical basis and the  $D$ -relation on the closure lattice is established (Theorem 32), which allows us

to investigate, in sections 8 and 9, closure systems without  $D$ -cycles. The latter form a proper subclass of the class of  $UC$ -systems (Definition 44). In particular, we prove that the  $E$ -basis that was introduced in [2] for closure systems without  $D$ -cycles, is optimized in one of three essential parts of the basis (Theorem 68). Nevertheless, two other parts of this basis cannot be effectively optimized, which is shown in section 10 (Corollaries 74 and 76).

## 2. PRELIMINARIES

Given a non-empty set  $S$  and the set  $\mathbf{2}^S$  of all its subsets, a *closure operator* is a map  $\phi : \mathbf{2}^S \rightarrow \mathbf{2}^S$  that satisfies the following, for all  $X, Y \in \mathbf{2}^S$ :

- (1) increasing:  $X \subseteq \phi(X)$ ;
- (2) isotone:  $X \subseteq Y$  implies  $\phi(X) \subseteq \phi(Y)$ ;
- (3) idempotent:  $\phi(\phi(X)) = \phi(X)$ .

It will be convenient for us to refer to the pair  $\langle S, \phi \rangle$  of a set  $S$  and a closure operator on it as a *closure system*.

A subset  $X \subseteq S$  is called *closed* if  $\phi(X) = X$ . The collection of closed subsets of closure operator  $\phi$  on  $S$  forms a lattice, which is usually called the *closure lattice* of the closure system and denoted  $\text{Cl}(S, \phi)$ .

The lattice operations are denoted  $\wedge$ , for the *meet*, and  $\vee$ , for the *join*. Simultaneously, every lattice is a partially ordered set in which every two elements have a least upper bound (which coincides with the join of those elements), and a greatest lower bound (the meet). We will use the notation  $0$  for the least element of a lattice, and  $1$  for its greatest element. If  $a \leq b$  in lattice  $L$ , then we denote by  $[a, b]$  the interval in  $L$ , consisting of all  $c$  satisfying  $a \leq c \leq b$ .

For every finite lattice  $L$ , let  $\text{Ji}(L)$  denote the set of *join irreducible* elements of  $L$ . An element  $j \in L$  is called join irreducible, if  $j \neq 0$ , and  $j = a \vee b$  implies  $a = j$  or  $b = j$ .

With every finite lattice  $L$ , we can associate a particular closure system  $\langle S, \phi \rangle$  in such a way that  $L$  is isomorphic to  $\text{Cl}(S, \phi)$ . Indeed, define a closure system with  $S = \text{Ji}(L)$  and the following closure operator:

$$\phi(X) = [0, \bigvee X] \cap J(L), X \subseteq S.$$

It is straightforward to check that the closure lattice of  $\phi$  is isomorphic to  $L$ .

There are infinitely many closure systems whose closure lattices are isomorphic to a given lattice  $L$ . On the other hand, the closure system just described is the unique one, up to one-to-one mappings of the base sets, that satisfies two additional properties:

- (1)  $\phi(\emptyset) = \emptyset$ ;
- (2)  $\phi(\{i\}) \setminus \{i\}$  is closed, for every  $i \in S$ .

Condition (2) just says that each  $\phi(\{i\})$  is join irreducible. Note that (1) is a special case of (2), and that (2) implies the property

- (3)  $\phi(\{i\}) = \phi(\{j\})$  implies  $i = j$ , for any  $i, j \in S$ .

We will call a closure system with properties (1), (2) above a *standard closure system*. It is straightforward to verify that the standard system is characterized by the property that the set  $S$  is of the smallest possible size. In other words, one cannot reduce  $S$  to define an equivalent closure system. There exists a standard

procedure to obtain a standard closure system equivalent to a given one, see [2]. We will assume that the closure systems in this paper are standard.

If  $y \in \phi(X)$ , then this relation between an element  $y \in S$  and a subset  $X \subseteq S$  in a closure system can be written in the form of an implication:  $X \rightarrow y$ . Thus, the closure system  $\langle S, \phi \rangle$  can be given by the set of implications:

$$\Sigma_\phi = \{X \rightarrow y : y \in S, X \subseteq S \text{ and } y \in \phi(X)\}.$$

Conversely, any set of implications  $\Sigma$  defines a closure system: the closed sets are exactly those subsets  $Y \subseteq S$  that *respect* the implications from  $\Sigma$ , i.e., if  $X \rightarrow x$  is in  $\Sigma$ , and  $X \subseteq Y$ , then  $x \in Y$ . There are numerous ways to represent the same closure system by sets of implications; those sets of implications with some minimality property are called *bases*. Thus we can speak of various sorts of bases.

It is convenient to define an implication  $X \rightarrow y$  as any ordered pair  $(X, y)$ , where  $X \subseteq S$  and  $y \in S$ , especially having in mind its interpretation as a propositional formula, as in the next paragraph below. On the other hand, from the point of view of closure systems, any single implication  $X \rightarrow x$ , with  $x \in X$ , defines a trivial closure system, where all subsets of  $S$  are closed. If such an implication is present in the set of implications  $\Sigma$ , then it can be removed without any change to the family of closed sets that  $\Sigma$  defines. We will assume throughout the paper that implications  $X \rightarrow x$ , where  $x \in X$ , are not included in the set of implications defining closure systems.

In general, implications  $X \rightarrow y$ , where  $X \subseteq S$  and  $y \in S$ , can be treated as the formulas of propositional logic over the set of variables  $S$ , equivalent to  $y \vee \bigvee_{x \in X} \neg x$ . Formulas of this form are also called *definite Horn clauses*. More generally, Horn clauses are disjunctions of several negative literals and at most one positive literal. The presence of a positive literal makes a Horn clause *definite*. A *Horn formula* is a conjunction of Horn clauses.

There is also a direct correspondence between Horn formulas and Horn Boolean functions: a Boolean function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  is called a (*pure or definite*) *Horn function*, if it has some CNF representation given by a (definite) Horn formula  $\Sigma$ .

Note that, in general, one can consider implications of the form  $X \rightarrow Y$ , where  $Y$  is not necessarily a one-element subset of  $S$ . The set  $X$  is called the *premise*, and  $Y$  the *conclusion* of an implication  $X \rightarrow Y$ . We will assume that any implication  $X \rightarrow Y$  is an ordered pair of non-empty subsets  $X, Y \subseteq S$ , and  $Y \cap X = \emptyset$ .

The following general observation about the computation of the closure via the basis is part of the folklore, see, for example, [20].

**Proposition 1.** *Let  $\langle S, \phi \rangle$  be a closure system, with a basis  $\Sigma$ . An implication  $A \rightarrow b$  holds in the closure system iff one can find a sequence  $\sigma_k = A_k \rightarrow B_k$ ,  $k \leq m$ , of implications from  $\Sigma$ , such that: (I)  $A_1 \subseteq A$ ; (II)  $b \in B_m$ ; (III)  $A_k \subseteq A \cup B_1 \cup \dots \cup B_{k-1}$  for  $k > 1$ .*

A sequence  $\sigma_k$ ,  $k \leq m$ , with properties (I)-(III) from Proposition 1 is called a  $\Sigma$ -*inference* of  $b$  from  $A$ . Thus, an implication  $A \rightarrow b$  holds in the closure system iff  $b$  has  $\Sigma$ -inference from  $A$ , for some basis  $\Sigma$ . We could also say that the implication  $A \rightarrow b$  *follows* from  $\Sigma$ .

Following K. Bertet and B. Monjardet [5], we will call the basis  $\Sigma$  a *unit implicational basis*, if  $|Y| = 1$  for all implications  $X \rightarrow Y$  in  $\Sigma$ .

Given a unit basis  $\Sigma$ , we can replace all implications  $X \rightarrow y$  with the same premise  $X$  by a single implication  $X \rightarrow Y$ , where  $Y$  is the union of all singletons

$y$  from the conclusions of these unit implications. Such a basis will be called the *aggregation* of  $\Sigma$ , and denoted by  $\Sigma^{ag}$ .

Vice versa, for every basis  $\Sigma$ , we may consider its *unit expansion*  $\Sigma_u$ , where  $X_i \rightarrow Y_i$  is replaced by  $\{X_i \rightarrow y : y \in Y_i\}$ . In particular,  $(\Sigma_u)^{ag} = \Sigma$ , for every aggregated basis  $\Sigma$ .

As in [2], we will call the subset  $\Sigma^b = \{(A \rightarrow B) \in \Sigma : |A| = 1\}$  of given basis  $\Sigma$  the *binary part* of the basis. The *non-binary* part of  $\Sigma$  is  $\Sigma^{nb} = \Sigma \setminus \Sigma^b$ , consisting of all implications  $A \rightarrow B$  in  $\Sigma$  with  $|A| > 1$ .

Assuming that the closure system  $\langle S, \phi \rangle$  defined by  $\Sigma$  is standard, we can claim that the binary relation  $\geq_\phi$  on  $X$  defined as:

$$a \geq_\phi b \text{ iff } b \in \phi(a)$$

is a partial order. This is exactly the partial order on the join irreducible elements in  $L = \text{Cl}(S, \phi)$ .

We write  $|\Sigma|$  for the number of implications in  $\Sigma$ . An aggregated basis  $\Sigma$  is called *minimum*, if  $|\Sigma| \leq |\Sigma^*|$ , for any other aggregated basis  $\Sigma^*$  of the same system.

The number  $s(\Sigma) = |X_1| + \dots + |X_n| + |Y_1| + \dots + |Y_n|$  is called the *size* of the basis  $\Sigma$ . A basis  $\Sigma$  is called *optimum* if  $s(\Sigma) \leq s(\Sigma^*)$ , for any other basis  $\Sigma^*$  of the system. Similarly, one can define  $s_L(\Sigma) = |X_1| + \dots + |X_n|$ , the *L-size*, and  $s_R(\Sigma) = |Y_1| + \dots + |Y_n|$ , the *R-size*, of a basis  $\Sigma$ . The basis will be called *left-optimum* (resp. *right-optimum*), if  $s_L(\Sigma) \leq s_L(\Sigma^*)$  (resp.  $s_R(\Sigma) \leq s_R(\Sigma^*)$ ), for any other basis  $\Sigma^*$ . Finally, the basis is called *non-redundant*, if removing any implication gives a set of implications that no longer defines the same closure system.

Now we recall the major theorem of V. Duquenne and J.L. Guigues [13] about the canonical basis; see also N. Caspard and B. Monjardet [9].

A set  $Q \subseteq S$  is called *quasi-closed* for  $\langle S, \phi \rangle$ , if

- (1)  $Q$  is not closed;
- (2)  $Q \cap X$  is closed, for every closed set  $X$  with  $Q \not\subseteq X$ .

In other words, adding  $Q$  to the family of  $\phi$ -closed sets, makes the new family stable under the set intersection; in particular, it is the family of closed sets of some closure operator.

A quasi-closed set  $C$  is called *critical*, if it is minimal, with respect to the containment order, among all quasi-closed sets with the same closure. Equivalently, if  $Q \subseteq C$  is another quasi-closed set and  $\phi(Q) = \phi(C)$ , then  $Q = C$ .

Let  $\mathcal{Q}$  be the set of all quasi-closed sets and  $\mathcal{C} \subseteq \mathcal{Q}$  be the set of critical sets of the closure system  $\langle S, \phi \rangle$ . Subsets of the form  $\phi(C)$ , where  $C \in \mathcal{C}$ , are called *essential*. It can be shown that by adding all quasi-closed sets to closed sets of  $\langle S, \phi \rangle$ , one obtains a family of subsets stable under set intersection, thus defining a new closure operator  $\sigma$ . This closure operator  $\sigma$  associated with  $\phi$  is called the *saturation* operator. In other words, for every  $Y \subseteq S$ ,  $\sigma(Y)$  is the smallest set containing  $Y$  which is either quasi-closed or closed.

**Theorem 2.** [13] *Consider the set of implications  $\Sigma_C = \{C \rightarrow (\phi(C) \setminus C) : C \in \mathcal{C}\}$ . Then*

- (1)  $\Sigma_C$  is a minimum basis.
- (2) For every other basis  $\Sigma$ , for every  $C \in \mathcal{C}$ , there exists  $(U \rightarrow V)$  in  $\Sigma$  such that  $\sigma(U) = C$ .

The basis  $\Sigma_C$  defined in Theorem 2 is called the *canonical basis* for the system  $\langle S, \phi \rangle$ .

We will make use of the following lemma regarding the saturation operator, due to M. Wild.

**Lemma 3.** [20] *Let  $\Sigma$  be a basis for the closure system  $\langle S, \phi \rangle$ , and let  $U \rightarrow V$  be an implication in  $\Sigma$ . Let  $\Sigma' = \{(X \rightarrow Y) \in \Sigma : \phi(X) = \phi(U)\}$ . For any subset  $W \subseteq S$  such that  $\phi(W) \subseteq \phi(U)$ , the implication  $W \rightarrow \sigma(W)$  follows from  $\Sigma \setminus \Sigma'$ .*

A big part of the current paper is devoted to the closure systems that we call *UC-systems*. In such a system every essential element  $X$  has exactly one critical set  $C \subseteq S$  with  $\phi(C) = X$ . The source of inspiration for *UC-systems* is its proper subclass of closure systems whose closure lattices satisfy the join-semidistributive law.

A lattice is called *join-semidistributive* if it satisfies the lattice law

$$(SD_{\vee}) \quad x \vee y = x \vee z \rightarrow x \vee y = x \vee (y \wedge z).$$

The join-semidistributive law plays an important role in lattice theory, for example in the study of free lattices, see [14].

An important subclass of finite join-semidistributive lattices are so-called lattices *without D-cycles*. First, we need to define the *D*-relation on  $\text{Ji}(L)$ . If  $x \leq \bigvee_{i \in I} x_i$  for  $x, x_i \in \text{Ji}(L)$ , and  $x \not\leq x_i$  for all  $i \in I$ , then  $X = \{x_i : i \in I\}$  is called a *non-trivial cover* of  $x$ . For any  $X, Y \subseteq L$ , we say that  $Y$  refines  $X$ , and write  $Y \ll X$ , if for every  $y \in Y$  there exists  $x \in X$  such that  $y \leq x$ . The set  $X \subseteq \text{Ji}(L)$  is called a *minimal cover* of  $x$ , if  $Y \ll X$  can be a cover of  $x$ , only if  $X \subseteq Y$ . In other words, no  $x_i \in X$  can be deleted, or replaced by a set of join irreducibles  $Z$  with  $z < x_i$  for all  $z \in Z$ , to obtain another cover for  $x$ . Finally, a binary relation  $D$  is defined on  $\text{Ji}(L)$ :  $x D y$  iff  $y \in X$  for some minimal cover  $X$  of  $x$ .

We say that a lattice  $L$  is *without D-cycles* if it does not have  $x_1, \dots, x_n \in \text{Ji}(L)$ , where  $n > 2$ , such that  $x_i D x_{i+1}$  and  $x_1 = x_n$ . It is well-known that every finite lattice without *D*-cycles is join-semidistributive. In lattice literature, the lattices without *D*-cycles are known as *lower bounded*. In section 10 we will also briefly mention *bounded* lattices: a lattice  $L$  is bounded if both  $L$  and the dual lattice  $L^\delta$  are without *D*-cycles.

### 3. MINIMUM UNIT BASIS VERSUS OPTIMUM BASIS

The following result of D. Maier [19] establishes the connections between different types of effective bases.

**Theorem 4.** *Every optimum basis is minimum, and every minimum basis is non-redundant.*

While the second implication is rather straightforward, the first one requires a non-trivial argument. Moreover, neither of these two statements can be reversed.

Given any basis  $\Sigma$ , we call it a *minimum unit* basis, if its unit expansion  $\Sigma_u$  has the minimum number of (unit) implications among all possible unit bases for this closure system. Being a minimum unit basis is equivalent to having right-side optimization:  $s_R = |Y_1| + \dots + |Y_n|$  is minimum among all bases for this system. The problem of determining a minimum unit basis is associated with the problem of finding the minimum CNF-representation of Boolean functions, or minimum representation of a directed hypergraph, see E. Boros et al. [6].

It turns out that minimum unit bases and optimum bases of any closure system have an intimate connection. Our goal in this section is to extend Maier's result

to include minimum unit bases into a hierarchy: *every optimum basis is right-side optimum.*

An important part of the statement is that, in every optimum basis, the left side of every implication has a *fixed size*  $k_C$ ,  $C \in \mathcal{C}$  (critical sets, refer to definition before Theorem 2), that does not depend on the choice of the optimum basis. This makes it into a parameter of the closure system itself.

**Theorem 5.** *Let  $\langle S, \phi \rangle$  be a closure system.*

- (I) *If  $\Sigma'$  is a non-redundant basis, then  $\{\sigma(U) : (U \rightarrow V) \in \Sigma'\} \subseteq \mathcal{Q}$ .*
- (II) *Let  $\Sigma_{\mathcal{O}}$  be an optimum basis. For any critical set  $C$ , let  $X_C \rightarrow Y_C$  be an implication from this basis with  $\sigma(X_C) = C$ . Then  $|X_C| = k_C := \min\{|U| : U \subseteq C, \phi(U) = \phi(C)\} = \min\{|U| : U \subseteq C, \sigma(U) = C\}$ .*

For the rest of this section we assume that  $\Sigma_u$  is a minimum unit basis of the closure system  $\langle S, \phi \rangle$ , and  $\Sigma = (\Sigma_u)^{ag}$  is the aggregation of  $\Sigma_u$ . The following two lemmas borrow from the argument of Theorem 5 in M. Wild [20].

**Lemma 6.** *There exists a minimum basis  $\Sigma^*$  such that  $|\Sigma_u^*| = |\Sigma_u|$ .*

*Proof.* We start from  $\Sigma$ . Since  $\Sigma_u$  is a minimum basis,  $\Sigma$  is non-redundant.

According to Theorem 2 (2), for each critical set  $C$ , we will be able to find an implication  $(X_C \rightarrow Y_C) \in \Sigma$ , such that  $\sigma(X_C) = C$ . Suppose  $\Sigma$  has another implication  $U \rightarrow V$  distinct from all  $X_C \rightarrow Y_C$ ,  $C \in \mathcal{C}$ . By Theorem 5 (I) we have  $\sigma(U) \in \mathcal{Q}$ , and thus we can find a critical set  $C_0 \in \mathcal{C}$  such that  $C_0 \subseteq \sigma(U)$  and  $\phi(C_0) = \phi(U)$ . The premise of the implication  $X_{C_0} \rightarrow Y_{C_0}$  from  $\Sigma$  also satisfies  $\phi(X_{C_0}) = \phi(C_0) = \phi(U)$ .

By Lemma 3, the implication  $U \rightarrow \sigma(U)$  follows from the set of implications  $\Sigma \setminus \{X_{C_0} \rightarrow Y_{C_0}, U \rightarrow V\}$ . Form  $\Sigma^\circ = (\Sigma \setminus \{X_{C_0} \rightarrow Y_{C_0}, U \rightarrow V\}) \cup \{X_{C_0} \rightarrow Y_{C_0} \cup V\}$ . Then  $U \rightarrow V$  follows from  $\Sigma^\circ$ . Moreover,  $|\Sigma_u^\circ| \leq |\Sigma_u|$ , but due to the minimality of  $\Sigma_u$ , we have  $|\Sigma_u^\circ| = |\Sigma_u|$ .

Repeat this procedure, replacing each implication  $U \rightarrow V$  in  $\Sigma^\circ$  distinct from all  $X_C \rightarrow Y_C$ , obtaining a basis  $\Sigma^*$ . It will have the same number of implications as the canonical basis  $\Sigma_C$ , whence it is minimum. Moreover,  $|\Sigma_u^*| = |\Sigma_u|$ .  $\square$

Lemma 6 suggests the following procedure for dealing with the problem of unit basis minimization.

**Corollary 7.** *Given a unit basis  $\Sigma_v$  with the non-redundant aggregation  $\Sigma$ , in time polynomial in  $s(\Sigma)$ , one can build a unit basis  $\Sigma_v^*$  such that  $|\Sigma_v^*| \leq |\Sigma_v|$ , while the aggregation  $\Sigma^*$  of  $\Sigma_v^*$  is a minimum basis.*

Indeed, it follows from the proof of Lemma 6 that there exists a partition of  $\Sigma = \bigcup_{C \in \mathcal{C}} \Sigma_C$  into blocks  $\Sigma_C$  with respect to critical sets  $C \in \mathcal{C}$ : if  $(A \rightarrow B) \in \Sigma_C$ , then  $C \subseteq \sigma(A)$ ,  $B \subseteq \phi(C)$ . Moreover, there exists at least one implication  $(D \rightarrow F) \in \Sigma_C$  such that  $C = \sigma(D)$ . Thus, the whole block can be replaced by  $D \rightarrow F'$ , where  $F'$  is the union of all right side sets of implications in  $\Sigma_C$ . Note that it takes polynomial time in  $s(\Sigma)$  to build the partition  $\Sigma = \bigcup_{C \in \mathcal{C}} \Sigma_C$ . In particular, every minimum unit basis  $\Sigma_v$  can be turned into a minimum unit  $\Sigma_v^*$ , whose aggregation is minimum.

**Lemma 8.** *There exists a minimum basis  $\Sigma^{**}$  such that, for every  $(X_C \rightarrow Y_C) \in \Sigma^{**}$ , we have  $|X_C| = k_C$ . Moreover,  $|\Sigma_u^{**}| = |\Sigma_u|$ .*

*Proof.* Start with  $\Sigma^*$  obtained in Lemma 6. Then  $\sigma(X_C) = C$ , and there exists  $X_C^*$  with  $\sigma(X_C^*) = C$  and  $|X_C^*| = k_C$ , as defined in Theorem 5(2). An implication  $X_C \rightarrow C$  follows from  $\Sigma^* \setminus \{X_C \rightarrow Y_C\}$ , due to Lemma 3. Hence,  $X_C \rightarrow Y_C$  follows from  $(\Sigma^* \setminus \{X_C \rightarrow Y_C\}) \cup \{X_C^* \rightarrow Y_C\}$ . Repeat this procedure for each implication in  $\Sigma^*$ , and obtain a new basis  $\Sigma^{**}$ . Since only the premise of each implication may change,  $|\Sigma_u^{**}| = |\Sigma_u^*| = |\Sigma_u|$ .  $\square$

The basis obtained as the result of previous two lemmas has a nice property.

**Lemma 9.**  $\Sigma^{**}$  is an optimum basis.

*Proof.* Suppose  $\Sigma^{**} = \{X_C \rightarrow Y_C : C \in \mathcal{C}\}$  is not optimum. Consider an optimum basis  $\Pi = \{A_C \rightarrow B_C : C \in \mathcal{C}\}$ , so that  $s(\Pi) < s(\Sigma^{**})$ . According to Theorem 5(II),  $|A_C| = k_C = |X_C|$ . It follows that  $\Sigma(|B_C| : C \in \mathcal{C}) < \Sigma(|Y_C| : C \in \mathcal{C})$ . Both sums represent the number of implications in the unit expansion of each basis. Thus,  $|\Pi_u| < |\Sigma_u^{**}| = |\Sigma_u|$ , which contradicts to the fact that  $\Sigma_u$  is a minimum unit basis.  $\square$

Now we can prove the main result of this section.

**Theorem 10.** There exists a minimum unit basis whose aggregation is an optimum basis. The unit expansion of every optimum basis is a minimum unit basis.

*Proof.* According to Lemma 9, the unit expansion of the basis  $\Sigma^{**} = \{X_C \rightarrow Y_C : C \in \mathcal{C}\}$  has the same number of implications as  $\Sigma_u$ , hence,  $\Sigma_u^{**}$  is also a minimum unit basis. Besides,  $(\Sigma_u^{**})^{ag} = \Sigma^{**}$ , which is optimum. This proves the first statement. Take another optimum basis  $\Pi = \{A_C \rightarrow B_C : C \in \mathcal{C}\}$ . Then  $s(\Pi) = s(\Sigma^{**})$ . Moreover,  $|A_C| = k_C = |X_C|$ , for each  $C \in \mathcal{C}$ , hence,  $\Sigma(|B_C| : C \in \mathcal{C}) = \Sigma(|Y_C| : C \in \mathcal{C})$ . The latter equality means that the unit expansions of both bases have the same number of implications. Hence,  $\Pi_u$  is also a minimum unit basis, and the second statement is also proved.  $\square$

**Corollary 11.** A basis  $\Sigma$  is optimum if and only if it is left-side optimum and right-side optimum.

*Proof.* It follows from Theorem 5 (II) that every optimum basis is left-side optimum. It was shown in Theorem 10 that every optimum basis is a minimum unit basis, equivalently, it is right-side optimum.

Vice versa, let  $\Sigma$  be both left-optimum and right-optimum and let  $\Sigma_O$  be any optimum basis. Recall that  $s(\Sigma_O) = s_L(\Sigma_O) + s_R(\Sigma_O)$ . It follows from left- and right-optimality of  $\Sigma$  that  $s_L(\Sigma) \leq s_L(\Sigma_O)$  and  $s_R(\Sigma) \leq s_R(\Sigma_O)$ . Hence,  $s(\Sigma) \leq s(\Sigma_O)$ , which implies  $s(\Sigma) = s(\Sigma_O)$  and  $\Sigma$  is optimum.  $\square$

It was proved in D. Maier [19] that finding an optimum basis of a given closure system is an NP-complete problem. A similar result about the minimum unit basis in the form of minimal directed hyper-graph representation was established later in G. Ausiello et al.[4]. It follows from Theorem 10 that the Maier's result can be obtained for free from the latter.

In [21], M. Wild proved that, given any basis  $\Sigma$  of a closure system whose lattice of closed sets is *modular*, one can obtain an optimum basis of this system in time  $O(s(\Sigma)^2)$ . Moreover, it was shown in C. Herrmann and M. Wild [15] that the test for the modularity of the closure system can also be achieved in polynomial time. This has the following consequence.

**Corollary 12.** *Let  $\langle S, \phi \rangle$  be a finite closure system, whose lattice of closed sets is modular. Then there is a polynomial time algorithm to obtain a minimum unit basis for this system.*

Indeed, according to Theorem 10, this basis is the unit expansion of the optimum basis built in Wild [21].

#### 4. OPTIMIZING BINARY PART OF A BASIS

In this section we touch upon the issue of optimization of the binary part of any basis. It turns out that the binary part is independent of the non-binary part, in the sense that there are parameters of optimization for the binary part pertinent to a closure system itself, and independent of the basis that is considered. These parameters are similar to the parameters  $k_C$  of left-side optimization considered in Theorem 5 (II).

For this section in particular, it is important that we consider a *standard* closure system  $\langle S, \phi \rangle$ , where it is assumed that  $\phi(\{a\}) \setminus \{a\}$  is a closed set, for every  $a \in S$ . This set will be denoted  $A_*$ .

First, we observe that singletons from  $S$  that are not closed are always represented in the binary part of any basis.

**Lemma 13.** *Let  $\Sigma$  be any aggregated basis of a standard closure system  $\langle S, \phi \rangle$ . Then  $(a \rightarrow B) \in \Sigma$  for some nonempty  $B \subseteq S$  iff  $\phi(\{a\}) \setminus \{a\} \neq \emptyset$ .*

*Proof.* If  $\phi(\{a\}) \setminus \{a\} \neq \emptyset$ , then  $\{a\}$  is a critical set for the essential set  $\phi(\{a\})$ . Thus  $(a \rightarrow B)$  must be in  $\Sigma$ , due to Theorem 2 (2). The converse statement is obvious, since  $\Sigma$  defines the closure system  $\langle S, \phi \rangle$  and  $B \subseteq \phi(\{a\})$ . Besides,  $B \neq \{a\}, \emptyset$ , according to the definition of implication.  $\square$

Secondly, we would like to introduce the bases with properly distinguished binary and non-binary parts. Some bases may have implications in  $\Sigma^{nb}$  that should truly belong to  $\Sigma^b$ .

**Definition 14.** Let  $\langle X, \phi \rangle$  be a closure system. An aggregated basis  $\Sigma$  is called *regular*, if  $\Sigma^{nb}$  does not have implications  $\{a\} \cup F \rightarrow D$  with  $F \subseteq \phi(\{a\})$ .

In other words, in a regular basis  $\Sigma$ , no implication  $A \rightarrow B \in \Sigma^{nb}$  is equivalent to  $a \rightarrow B$ , for any  $a \in A$ .

It is easy to observe that both the canonical basis and the aggregated  $D$ -basis of [2] are regular by the definition. The concept of the regular basis is highly important in the key Definition 31 of section 6.

**Proposition 15.** *Let  $\langle X, \phi \rangle$  be a standard closure system. If  $\Sigma$  is regular, then for every implication  $a \rightarrow B$  in  $\Sigma^b$  it holds that  $\phi(B) = \phi(\{a\}) \setminus \{a\}$ .*

*Proof.* Since  $\langle X, \phi \rangle$  is standard,  $\phi(\{a\}) \setminus \{a\}$  is closed. To simplify the notation, we will denote  $A_* = \phi(\{a\}) \setminus \{a\}$ .

If  $a \rightarrow B$  is in  $\Sigma$ , then  $B \subseteq \phi(\{a\})$  and  $a \notin B$ , whence  $B \subseteq A_*$ . Since  $A_*$  is closed, this implies  $\phi(B) \subseteq A_*$ . We want to verify that  $\phi(B) = A_*$ .

Consider any  $c \in A_*$ . Since  $c \in \phi(\{a\})$ , by Proposition 1, there should be a sequence of implications  $\sigma_1, \dots, \sigma_k \in \Sigma$ , say  $\sigma_i = A_i \rightarrow B_i$ , such that  $A_1 = \{a\}$ ,  $c \in B_k$ , and  $A_i \subseteq A_1 \cup B_1 \cup \dots \cup B_{i-1}$  for  $1 < i \leq k$ . Since  $\Sigma$  is aggregated, we have  $B_1 \subseteq \Phi(B)$ . As  $\Sigma$  is regular,  $a \notin A_2, \dots, A_k$ , whence by induction it follows that  $B_i \subseteq \phi(B)$ . For  $i = k$  this yields  $c \in B_k \subseteq \phi(B)$ , so that  $c \in \phi(B)$ , as desired.  $\square$

The next lemma shows that any basis can be modified to become regular, without increasing its size.

**Lemma 16.** *If  $\Sigma$  is any aggregated basis of a standard closure system, then one can find a new basis  $\Sigma_r$  that is regular,  $|\Sigma_r| \leq |\Sigma|$  and  $s_L(\Sigma_r) \leq s_L(\Sigma)$ ,  $s_R(\Sigma_r) \leq s_R(\Sigma)$ .*

*Proof.* Let  $\{a\} \cup F \rightarrow D$  in  $\Sigma$ , where  $F \subseteq \phi(\{a\})$ . By Proposition 1,  $\Sigma^b$  should have an implication  $a \rightarrow B$ , for some  $B$ . Consider the splitting of  $D$ :  $D = (D \setminus \phi(F)) \cup (D \cap \phi(F))$ . If both subsets in the union are non-empty, then one can replace this implication by  $a \rightarrow D \setminus \phi(F)$  and  $F \rightarrow (D \cap \phi(F))$  without changing  $s_L$  and  $s_R$ . If either of two subsets is empty, then the original implication can be replaced by one of the two new implications, correspondingly. Then  $a \rightarrow B$  can be aggregated with  $a \rightarrow D \setminus \phi(F)$ , reducing  $s_L$  and, possibly,  $s_R$  as well. Observe that this procedure will not increase the number of implications.  $\square$

**Corollary 17.** *Every optimum basis is regular.*

We note that the procedure in Lemma 16 does modify the non-binary part of a given basis  $\Sigma$ .

**Example 18.** Consider a closure system be given by the basis  $\Sigma = \{a \rightarrow b, ab \rightarrow c, bc \rightarrow d\}$ . The second implication has the premise  $ab \subseteq \phi(\{a\})$ , and it follows from  $a \rightarrow c$ . Hence, the regular basis obtained via procedure of Lemma 16 is  $\Sigma_r = \{a \rightarrow bc, bc \rightarrow d\}$ . We note that  $|\Sigma_2| = 2 < 3 = |\Sigma|$ ,  $s_L(\Sigma_r) = 3 < 5 = s_L(\Sigma)$ , and  $s_R(\Sigma_r) = s_R(\Sigma) = 3$ .

The binary part  $\Sigma^b$  of some standard basis will also be called *regular*, if it satisfies the condition of Proposition 15: for every implication  $a \rightarrow B$  in  $\Sigma^b$  it holds that  $\phi(B) = \phi(\{a\}) \setminus \{a\}$ .

The next statement shows that the binary part of any basis can be replaced by any regular binary part without any modification to the non-binary part.

**Lemma 19.** *Let  $\Sigma = \Sigma^b \cup \Sigma^{nb}$  be any basis of a standard closure system. One can replace any  $(a \rightarrow D) \in \Sigma^b$  by arbitrary  $a \rightarrow B$ , with  $\phi(B) = \phi(\{a\}) \setminus \{a\}$ , to obtain a new basis  $\Sigma_*$  that preserves the non-binary part:  $\Sigma_*^{nb} = \Sigma^{nb}$ .*

*Proof.* The argument proceeds by induction on the height of the element  $a$  in the partially ordered set  $(S, \geq_\phi)$ . By the *height* of an element  $x$  in a partial ordered set  $(S, \geq)$ , we mean the length  $k$  of the longest chain  $x = x_0 > x_1 \cdots > x_k$  in  $(S, \geq)$ .

The argument for  $k = 1$  and the induction step are similar, so we combine both cases. Assume that  $a \rightarrow B$  is in  $\Sigma_*^b$  and that  $a \rightarrow D$  is in  $\Sigma^b$ , where either  $a$  is of height 1, or of height  $n + 1$ , and it is already proved that any implication  $(b \rightarrow F) \in \Sigma^b$  with height at most  $n$  follows from  $\Sigma_*$ . If  $a$  is of the height 1, then every  $c \in \phi(\{a\})$  is an *atom*, i.e.,  $\{c\} = \phi(\{c\})$ . In particular,  $c$  cannot be a premise of any binary implication.

We want to show that  $a \rightarrow D$  follows from  $\Sigma_*$ . Let use again the notation  $A_* = \phi(\{a\}) \setminus \{a\}$ .

Assume there exists  $d \in D \setminus B$ . Since  $d \in A_* = \phi(B)$ , the implication  $B \rightarrow d$  should follow from  $\Sigma$ . Then one can find a sequence of implications  $\sigma_1, \dots, \sigma_k \in \Sigma$ ,  $\sigma_i = A_i \rightarrow B_i$ , such that  $B \subseteq A_1$ ,  $d \in B_k$ , and  $A_i \subseteq A_1 \cup B_1 \cup \dots \cup B_{i-1} \subseteq A_*$ . If  $a$  is of height 1, then none of these implications is binary, since all elements in  $A_*$  are atoms. If  $a$  is of height  $n + 1$ , then all binary implications are of the form

$(b \rightarrow F) \in \Sigma^b$ , where  $b \in A_*$ , whence  $b$  is of height  $\leq n$ , and such implications follow from  $\Sigma_*$  by the inductive hypothesis. Place  $\sigma_0 = a \rightarrow B$  prior to the first implication in inference  $\sigma_1, \dots, \sigma_k$ , and replace each binary implication among  $\sigma_i$  by its inference from  $\Sigma_*$ . This will give the inference of  $a \rightarrow d$  from  $\Sigma_*$ .  $\square$

We note that similar statement is proved in Corollary 5.7 of E. Boros et al. [7], where an exclusive set of implications may, in particular, be any subset of the basis, whose left and right sides are contained in some  $\phi(Y)$ . In our case, it is  $\phi(\{a\})$ .

The consequence of two lemmas is the following statement about the right-side optimization of the binary part.

**Theorem 20.** *Let  $\Sigma_C$  be the canonical basis of a standard closure system  $\langle S, \phi \rangle$ , and let  $x_C \rightarrow Y_C$  be any binary implication from  $\Sigma_C$ . Every (regular right-side) optimum basis  $\Sigma$  will contain an implication  $x_C \rightarrow B$ , where  $|B| = b_C = \min\{|Y| : \phi(Y) = \phi(\{x_C\}) \setminus \{x_C\}\}$ .*

*Proof.* Suppose we are given a (regular right-side) optimum basis  $\Sigma$ . For each  $(x_C \rightarrow Y_C) \in \Sigma_C^b$ , there should be  $(x_C \rightarrow D) \in \Sigma$ ; besides,  $\phi(D) = \phi(\{x_C\}) \setminus \{x_C\}$ . Suppose there exists  $B$  such that  $\phi(B) = \phi(\{x_C\}) \setminus \{x_C\}$  and  $|B| < |D|$ . Then, by Lemma 19, it would be possible to replace  $x_C \rightarrow D$  by  $x_C \rightarrow B$ , without changing other implications of the basis. This would reduce  $s_R$ , a contradiction with the right-side optimality of  $\Sigma$ . Hence,  $|D|$  is minimal among subsets whose closure is  $\phi(\{x_C\}) \setminus \{x_C\}$ .  $\square$

We will call a basis *B-optimum*, if the size of its binary part is minimum among all possible *regular* bases of the closure system  $\langle S, \phi \rangle$ .

**Corollary 21.** *Every optimum basis is B-optimum. Moreover, every B-optimum basis  $\Sigma$  has the size of its binary part  $s(\Sigma^b) = |\Sigma_C^b| + \sum_{|C|=1} b_C$ .*

*Proof.* According to Theorem 20, every optimum basis has the smallest right size of its binary part, among the same measurement of all other regular bases. The premises in the binary parts are just singletons, so the total size of the premises in every aggregated basis  $\Sigma^b$  is  $|\Sigma^b|$ .  $\square$

We can point some easy computable lower bound for parameters  $b_C$ . Recall that by an *extreme point* of a closed set  $X$  one calls an element  $x \in X$  such that  $x \notin \phi(X \setminus \{x\})$ . The set of extreme points of  $X$  is denoted by  $Ex(X)$ .

**Corollary 22.** *Let  $x_C \rightarrow B$  be an implication from any optimum basis of a closure system  $\langle S, \phi \rangle$ , and let  $X_*$  denote  $\phi(\{x_C\}) \setminus \{x_C\}$ . Then  $Ex(X_*) \subseteq B$ . In particular,  $|B| = b_C \geq |Ex(X_*)|$ .*

*Proof.* Since an optimum basis is regular,  $\phi(B) = X_*$ . Suppose  $y \in Ex(X_*) \setminus B$ . Then  $B \subseteq X_* \setminus \{y\}$ , whence  $\phi(B) \subseteq \phi(X_* \setminus \{y\})$  and  $y \notin \phi(B)$ , a contradiction. Therefore,  $Ex(X_*) \subseteq B$ .  $\square$

In general, it is not true that  $\phi(Ex(Y)) = Y$ , for any  $Y \subseteq S$ . Thus, it is possible that  $b_C > |Ex(X_*)|$ . It is easy to observe that this lower bound is attained in closure systems called *convex geometries*, so that the optimum binary part is tractable in such closure systems. This is treated with more detail in K. Adaricheva [1].

One natural choice for the binary implications in a basis  $\Sigma$  is the *cover relation*  $\succ_\phi$  of  $\geq_\phi$ . (By the definition,  $a \succ_\phi b$  means that  $a \geq_\phi b$ ,  $a \neq b$  and  $a \geq_\phi c \geq_\phi b$

implies  $c = a$  or  $c = b$ .) Namely,  $a \rightarrow A_{\succ}$  is included into  $\Sigma^b$ , if  $A_{\succ} = \{b \in \phi(\{a\}) : a \succ_{\phi} b\}$ . It is straightforward to verify that  $\phi(A_{\succ}) = \phi(\{a\}) \setminus \{a\}$ . Hence, such a binary part can be included into any basis, irrelevant to its non-binary part. It is algorithmically easy to compute  $A_{\succ}$ , see [2, Proposition 16]. Also, the advantage of this choice is that every binary implication  $a \rightarrow b$  that follows from  $\Sigma$ , follows from binary part only. Of course, it is by no means guaranteed that  $A_{\succ}$  has the minimal cardinality required in an optimum basis. It can even be reducible, i.e., some proper subset of  $A_{\succ}$  can still have the same closure as  $A_{\succ}$ .

**Example 23.** Consider the closure system given by the canonical basis  $\Sigma_C = \{a \rightarrow bcd, bcdy \rightarrow a, bc \rightarrow d\}$ . Apparently, the only binary implication  $a \rightarrow bcd$  is presented by  $\succ_{\phi}$ , i.e.,  $A_{\succ} = \{b, c, d\}$ . On the other hand, we can reduce this implication to  $a \rightarrow bc$ , since  $\phi(\{b, c\}) = \phi(\{a\}) \setminus \{a\}$ .

We will consider in section 7 another form of the basis, for the  $UC$ -closure systems, whose binary part will be non-redundant.

## 5. K-BASIS IN GENERAL CLOSURE SYSTEMS

As follows from the results of sections 3 and 4, an optimum basis is the most desirable among all bases of the closure system, since it implies other forms of minimality.

On the other hand, finding an optimum basis is generally a computationally hard problem. In this section we describe a new notion of a  $K$ -basis which can be computed from the canonical basis in polynomial time and provides reduction in size, while maintaining the minimum number of implications.

Let  $\Sigma_C = \{C \rightarrow Y_C : C \in \mathcal{C}\}$  be a canonical basis. We will call any basis  $\Sigma^* = \{X_C^* \rightarrow Y_C^* : C \in \mathcal{C}\}$  a *refinement* of the canonical basis if  $X_C^* \subseteq C$ ,  $Y_C^* \subseteq Y_C$ , and  $\phi(X_C^*) = \phi(C)$  for all  $C \in \mathcal{C}$ . Every refinement is a minimum basis, while  $s(\Sigma^*) \leq s(\Sigma_C)$ .

According to Theorem 2 (2) and Theorem 5 (II), every optimum basis is a refinement of the canonical one. Besides, for every critical set  $C$ , an optimum basis would have an implication  $X^* \rightarrow Y^*$  with  $|X^*| = k_C$ , the minimum cardinality of a subset  $X \subseteq C$  for which  $\phi(X) = \phi(C)$  (equivalently,  $\sigma(X) = C$ ). Such set  $X$  is called a *minimal generator* for the critical set  $C$ .

The  $K$ -basis is built on the idea of minimizing the left side of implications of the canonical basis with respect to all special *order ideals* contained in critical sets. Thus, the left sides of the  $K$ -basis are not necessarily of the smallest cardinality. On the other hand, the algorithm producing a  $K$ -basis is fast and easy.

We observe that, according to the definition of a critical set, every  $C$  is an  $\geq_{\phi}$ -order ideal, i.e., if  $a \in C$  and  $a \geq_{\phi} b$ , then  $b \in C$ . Among all order ideals  $Y \subseteq C$  with the property  $\phi(Y) = \phi(C)$ , we can find a minimal one with respect to containment. Finally, for every order ideal  $Y$ , it is clear that  $\phi(\max(Y)) = \phi(Y)$ , where  $\max(Y)$  is the subset of  $\geq_{\phi}$ -maximal elements of  $Y$ .

We will call a set  $X^*$  a *minimal order generator* of the critical set  $C$  if  $X^* = \max(X)$  for some minimal (with respect to containment) order ideal  $X \subseteq C$  such that  $\phi(X) = \phi(C)$ .

**Proposition 24.** *Let  $C$  be a critical set for the closure system  $\langle S, \phi \rangle$ . Build a sequence  $X_0 = C, X_1, \dots, X_n$  of subsets of  $C$  as follows: (1)  $X_{k+1} = X_k \setminus \{x_k\}$ , where  $x_k$  is any  $\geq_{\phi}$ -maximal element of  $X_k$ ; (2)  $\phi(X_k) = \phi(C)$ ; (3) none of the*

maximal elements of  $X_n$  can be removed to obtain a set with the same closure as  $C$ . Then  $C_K := \max(X_n)$ , the set of  $\geq_\phi$ -maximal elements of  $X_n$ , is a minimal order generator, and every minimal order generator may be obtained via such a sequence.

*Proof.* One can reach any minimal order ideal  $X_n$  contained in  $C$  and with  $\phi(X_n) = \phi(C)$ , through the series of steps, when a maximal element of an order ideal is removed, producing another order ideal with the same closure, which is exactly the procedure described in the proposition.  $\square$

To align the notion of a minimal order generator with existing lattice terminology we include the following tautological statement.

**Proposition 25.** *Every minimal order generator  $C_K \subseteq C$  is a  $\ll$ -minimal join representation of the element  $\phi(C)$  in the closure lattice of  $\langle S, \phi \rangle$ .*

A critical set might have several minimal order generators, and the procedure described in Proposition 24 may lead to any of them, depending on the order in which the maximal elements are removed. See Example 29 for an illustration.

**Remark 26.** Not every minimal generator for a critical set  $C$  will be simultaneously a minimal order generator. Both are sets satisfying  $\phi(X) = \phi(C)$ , but the former type is minimal with respect to  $\subseteq$ , while the latter is minimal with respect to  $\ll$ . A concrete example is the lattice in Example 72.

**Definition 27.** A set of implications  $\Sigma_K$  is called a *K-basis* if it is obtained from the canonical basis  $\Sigma_C$  by replacing each implication  $(C \rightarrow Y_C) \in \Sigma_C$  by  $C_K \rightarrow Y_K$ , where  $C_K$  is a minimal order generator of  $C$ , and  $Y_K = \max(Y_C)$ .

We summarize easy facts about *K*-bases.

**Lemma 28.** *Let  $\Sigma_K$  be a *K*-basis for  $\langle S, \phi \rangle$ .*

- (1)  $\Sigma_K^b$  represents the cover relation  $\succ_\phi$  of  $\geq_\phi$ ;
- (2)  $\Sigma_K$  is a basis of  $\langle S, \phi \rangle$ ;
- (3)  $\Sigma_K$  is a minimum basis and  $s(\Sigma_K) \leq s(\Sigma_C)$ .

*Proof.* First, note that if  $x \geq_\phi y$  and  $y \neq x$ , then  $(x \rightarrow Y) \in \Sigma_C$  and  $y \in Y$ , due to definition of  $\Sigma_C$ . Secondly,  $\{x\}$  will be a minimal order generator for itself and  $Y$  comprises all elements  $y \in S$  with  $x \geq_\phi y$ . Choosing only  $\geq_\phi$ -maximal elements  $Y_K \subseteq Y$  will give exactly the lower covers of  $x$ , which proves (1).

Since every  $a \geq_\phi b$  is in the transitive closure of cover sub-relation of  $\geq_\phi$ , every  $(x \rightarrow Y) \in \Sigma_C^b$  follows from  $\Sigma_K^b$ . Also,  $C_K \rightarrow Y_C$  follows from  $C_K \rightarrow Y_K$  and  $\Sigma_K^b$ . Finally,  $(C \rightarrow Y_C) \in \Sigma_C$  follows from  $C_K \rightarrow Y_C$ , since  $C_K \subseteq C$ . Hence,  $\Sigma_C$  follows from  $\Sigma_K$ .

Vice versa,  $C_K \rightarrow Y_C$  follows from  $\Sigma_C$ , due to the definition of  $C_K$ . Moreover,  $C_K \rightarrow Y_K$  follows from  $C_K \rightarrow Y_C$ , since  $Y_K \subseteq Y_C$ , which finishes the proof of (2).

Part (3) easily follows from the definition of  $\Sigma_K$ , since  $C_K \subseteq C$  and  $Y_K \subseteq Y_C$ .  $\square$

The following example demonstrates the advantage of considering *K*-g due to possible size reduction of the canonical basis, while preserving its property of the minimality. The other purpose of this example is to show that there are possibly several *K*-bases associated with a closure system.

**Example 29.** Consider the standard closure system on  $S = \{x, y, z, e, d, u\}$  given by its closure lattice in Figure 1. Consider the critical set  $C = S \setminus \{e\}$ , for which

$\phi(C) = S$ . Following procedure in Proposition 24, we would have two maximal elements  $x, y \in C$  that can be removed:  $X_1 = C \setminus \{x\}$ ,  $X_2 = C \setminus \{y\}$ ,  $\phi(X_1) = \phi(X_2) = \phi(C)$ .

Thus,  $(C \rightarrow e) \in \Sigma_C$  can be refined to either  $yd \rightarrow e$  ( $\max X_1 = \{y, d\}$ ) or  $xd \rightarrow e$  ( $\max X_2 = \{x, d\}$ ), and this system has two  $K$ -bases.

Here is the first of  $K$ -bases:

$y \rightarrow u, z \rightarrow u, d \rightarrow z, e \rightarrow d, yd \rightarrow e, xu \rightarrow y, zy \rightarrow x$ , with  $s(\Sigma_K) = 17$ .

In the second one, the implication  $yd \rightarrow e$  is replaced by  $xd \rightarrow e$ , which does not change the size of the basis.

Compare both with  $\Sigma_C$ :

$y \rightarrow u, z \rightarrow u, d \rightarrow zu, e \rightarrow dzu, xyzdu \rightarrow e, xu \rightarrow y, zyu \rightarrow x$ , with  $s(\Sigma_C) = 24$ .

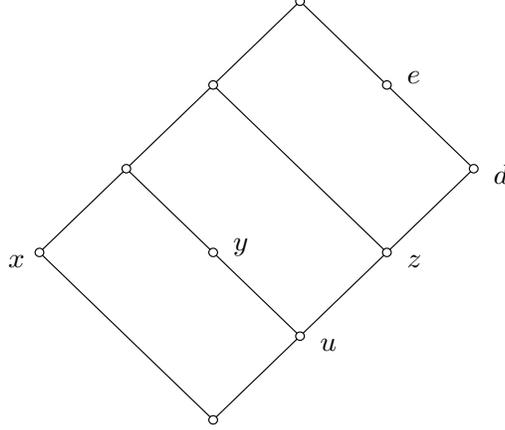


FIGURE 1. Example 29

While searching for a minimal order generator of smallest cardinality, for each critical set  $C$ , could be a long process, finding *some* minimal order generator is a fast quadratic algorithm.

Let  $\Sigma_C$  be a canonical basis of a standard closure system  $\langle S, \phi \rangle$ . Assume that  $|X| = n$ ,  $|\Sigma_C| = m$ ,  $s(\Sigma_C) = k$  and  $s(\Sigma_C^b) = k_b$ . It was shown in Theorem 11.3 of [14] that it will take time  $O(nk_b + n^2)$  to produce the cover relation of the partially ordered set  $(S, \geq_\phi)$ . Moreover, an upper bound for the number of pairs in the cover relation is  $k_b$ .

**Proposition 30.** *Given the canonical basis  $\Sigma_C$  of a standard closure system with  $s(\Sigma_C) = k$ , it will take time  $O(k^2)$  to produce one minimal order generator for each  $C \in \mathcal{C}$ . Thus, it requires time  $O(k^2)$  to produce a  $K$ -basis.*

*Proof.* For each critical set  $C$ , one would need at maximum  $k_b$  steps to recognize a maximal element  $x \in C$ . Then, the time  $O(k)$  is needed to check whether  $\phi(C \setminus \{x\}) = \phi(C)$  (using the forward chaining procedure, for example). If this is true, then  $X_1 = C \setminus \{x\}$ , and one proceeds with  $X_1$  in place of  $C$ . Otherwise, we would search during time  $O(k_b)$  for another maximal element of  $C$ . The number of maximal elements checked overall cannot be more than  $|C|$ , and the time spent on

each of them is  $O(k)$ . The overall time is the summation over all critical sets, i.e.,  $O(k^2)$ .  $\square$

## 6. $D$ -RELATION FROM THE CANONICAL BASIS

The goal of this section is to show that the  $D$ -relation defined on  $S$ , for the standard closure system  $\langle S, \phi \rangle$ , via the concept of minimal covers, can be recovered from the canonical basis. The main result is achieved in Theorem 32 by the series of Lemmas 33-40. This allows us to recognize effectively the systems without  $D$ -cycles that will be treated further in section 8.

Recall that  $bDa$ , for  $a, b \in S$  in a standard closure system  $\langle S, \phi \rangle$ , iff there is a  $\ll$ -minimal cover  $A$  for  $b$  such that  $a \in A$ . In particular, the closure system satisfies the implication  $A \rightarrow b$ . We note that such an implication belongs to the non-binary part of the  $D$ -basis  $\Sigma_D$ , for the closure system  $\langle S, \phi \rangle$ , see [2]. Vice versa, if  $A \rightarrow b$  is a non-binary implication of  $\Sigma_D$ , then  $A \cup \{b\} \subseteq \text{Ji}(L)$  and  $A$  is a minimal cover for  $b$ . The binary part of  $\Sigma_D$  can be given by  $\succ_\phi$ .

We will denote  $D^\delta$  the dual of  $D$ , i.e.,  $(a, b) \in D^\delta$  iff  $(b, a) \in D$ . Thus,  $a$  is the first entry in the pair  $(a, b) \in D^\delta$  iff  $a \in A$  appears on the left in a *non-binary* implication  $A \rightarrow b$  from the  $D$ -basis.

Recall that in the standard closure system  $\langle S, \phi \rangle$ , the binary relation  $\geq_\phi$  on  $S$  defined as:

$$a \geq_\phi b \text{ iff } b \in \phi(a)$$

is a partial order. A subset  $J \subseteq S$  will be called a  $\geq_\phi$ -ideal, if  $a \in J$  and  $a \geq_\phi b$  imply  $b \in J$ . We will denote by  $\text{Id}_\phi(A)$  an  $\geq_\phi$ -ideal generated by  $A \subseteq S$ .

**Definition 31.** Given any *regular* basis  $\Sigma$ , we define a binary relation  $\Delta_\Sigma$  on  $S$  as follows:  $(a, b) \in \Delta_\Sigma$  iff there is  $(A \rightarrow B) \in \Sigma^{nb}$  such that  $a \in A, b \in B$ . By  $\Delta_\Sigma^{tr}$  we will denote the transitive closure of the relation  $\Delta_\Sigma$ .

Finally, we introduce an important modification of the canonical basis  $\Sigma_C$  that will prove useful. For each  $A_C \rightarrow B_C$  in  $\Sigma_C^{nb}$  let  $A_K \rightarrow B_K$  be a corresponding implication from any  $K$ -basis of the closure system. Replace each implication  $A_C \rightarrow B_C$  in  $\Sigma_C^{nb}$  by  $A_K \rightarrow B_K$ . Denote by  $\Sigma^*$  this new set of implications which, apparently, also forms a basis. Note that  $\Sigma^*$  is a refinement of the canonical basis, with no change to the right sides of implications.

We note that although  $\Sigma_C$  is uniquely defined for a closure system,  $\Sigma^*$  is not, reflected in the fact that there might be several  $K$ -bases for the closure system.

Our main goal now is to prove the following statement.

**Theorem 32.**  $(D^\delta)^{tr} = \Delta_{\Sigma^*}^{tr}$

The meaning of Theorem 32 is that the basis  $\Sigma^*$  recovers an important sub-relation  $\Delta_{\Sigma^*}$  of the  $D$ -relation, so that any pair in the  $D$ -relation can be obtained via transitive closure of this sub-relation.

The theorem directly follows from Lemmas 33 and 40.

**Lemma 33.**  $\Delta_{\Sigma^*} \subseteq D^\delta$ .

*Proof.* We need to show that, for any  $A_C \rightarrow B_C$  from the canonical basis, if  $A_K \subseteq A_C$  is a  $\ll$ -minimal representation of  $U = \phi(A_C)$ , then  $(a, b) \in D^\delta$ , for every  $a \in A_K$  and  $b \in B_C$ . Apparently,  $A_K$  is a non-trivial cover of  $b$ , for every  $b \in B_C$ . Suppose that  $A_K$  can be properly refined to another cover  $A''$  of  $b$ , i.e.  $A'' \ll A_K$ ,

$A''$  is a cover for  $b$  and  $A_K \not\subseteq A''$ . Since  $A_K$  is a  $\ll$ -minimal representation for  $U = \phi(A_C)$ , we cannot have  $\phi(A'') = U$ . Hence,  $\phi(A'') \subset \phi(A_K) = U$ . Then, since  $A_C$  is a quasi-closed set and  $A'' \subseteq A_C$ , we should have  $b \in \phi(A'') \subseteq A_C$ , a contradiction. Hence,  $A_K$  is a minimal cover for  $b \in B_C$  and  $(a, b) \in D^\delta$  follows, for every  $a \in A_K$ .  $\square$

This lemma allows us to establish a connection between any  $K$ -basis and the  $D$ -basis  $\Sigma_D$ .

**Corollary 34.** *For any  $K$ -basis  $\Sigma_K$ , its unit expansion  $\Sigma_K^u$  is contained in  $\Sigma_D$ . In particular,  $|\Sigma_K^u| \leq |\Sigma_D|$ .*

*Proof.* Indeed, as shown in Lemma 33, if  $A_K \rightarrow B_C$  is in  $\Sigma^*$ , then  $A_K$  is a minimal cover for every  $b \in B_C$ . Recall that  $A_K$  is taken from implication  $A_K \rightarrow B_K$  of some  $K$ -basis, and  $B_K = \max(B_C) \subseteq B_C$ . Hence every implication  $A_K \rightarrow b$ , with  $b \in B_K$ , is in the  $D$ -basis.  $\square$

We observe that Theorem 32 does not hold when  $\Sigma^*$  is replaced by  $\Sigma_C$ . This is due to the fact that an even weaker form of Lemma 33 fails for  $\Sigma_C$ .

**Example 35.** It is possible that  $\Delta_{\Sigma_C} \not\subseteq (D^\delta)^{tr}$ . Consider the closure system  $\langle \{1, 2, 3, 4, 5, 6\}, \phi \rangle$  with  $\text{Cl}(X, \phi)$  on Fig. 35, whose canonical basis  $\Sigma_C$  is

$$2 \rightarrow 1, 3 \rightarrow 1, 5 \rightarrow 4, 6 \rightarrow 13, 14 \rightarrow 3, 123 \rightarrow 6, 1345 \rightarrow 6, 12346 \rightarrow 5.$$

Then  $\Sigma^{*nb}$  is

$$14 \rightarrow 3, 23 \rightarrow 6, 15 \rightarrow 6, 24 \rightarrow 5$$

and  $\Sigma_D^{nb}$  is

$$14 \rightarrow 3, 23 \rightarrow 6, 15 \rightarrow 6, 24 \rightarrow 5, 24 \rightarrow 6$$

We observe that  $(6, 5) \notin (D^\delta)^{tr}$ , since 6 does not occur on the left side of the implications in  $\Sigma_D^{nb}$ , while  $(6, 5) \in \Delta_{\Sigma_C}$  due to the implication  $12346 \rightarrow 5$  in  $\Sigma_C$ .

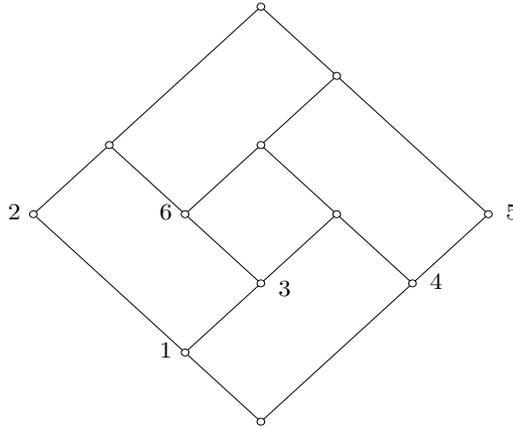


FIGURE 2. Example 35

The reverse inclusion needed in Theorem 32 will come in Lemma 40, which will follow from a series of preliminary lemmas.

We will call a basis  $\Sigma$  *round*, if  $A \cup B = \phi(A)$  for every  $(A \rightarrow B) \in \Sigma^{nb}$ . In particular, the canonical basis is round: by the definition, the implications in the canonical basis are of the form  $A \rightarrow (\phi(A) \setminus A)$ , see Theorem 2.

**Lemma 36.** *Let  $\Sigma$  be any round basis of a closure system  $\langle S, \phi \rangle$ . If an element  $b \in S$  has a  $\Sigma$ -inference from  $A \subseteq S$ , then it has a  $\Sigma^{nb}$ -inference from  $\text{Id}_\phi A$ . The converse statement is also true for an arbitrary basis  $\Sigma$ .*

*Proof.* Let  $\sigma_k = A_k \rightarrow B_k$ ,  $k \leq m$ , be an inference of  $A \rightarrow b$  from  $\Sigma$ .

We want to show that one can remove binary implications from this sequence at the cost of replacing  $A$  in properties (I) and (III) of Proposition 1 by  $\text{Id}_\phi A$ .

For this we show by induction on  $k$  that set  $X_k = (\text{Id}_\phi A) \cup B_1 \cup \dots \cup B_{k-1}$  is an  $\geq_\phi$ -ideal.  $X_1 = \text{Id}_\phi A$  is an ideal by the definition. Suppose  $X_{k-1}$  is an ideal and  $\sigma_k \in \Sigma^{nb}$ . Recall that  $A_k \cup B_k = \phi(A_k)$ , a closed set in  $\langle S, \phi \rangle$ , in particular, it is an  $\geq_\phi$ -ideal. As  $A_k \subseteq X_{k-1}$ , it follows that  $X_k = X_{k-1} \cup B_k$  is an  $\geq_\phi$ -ideal as well.

If  $\sigma_k \in \Sigma^b$ , then  $A_k = \{a\}$  and  $a \in X_{k-1}$ , but then  $B_k \subseteq X_{k-1}$  by the induction hypothesis. In particular, we see from the above argument that  $X_{k-1} = X_k$ , when  $\sigma_k \in \Sigma^b$ , and any such  $\sigma_k$  can be removed from the sequence.

The converse statement follows from the observation that  $A \rightarrow \text{Id}_\phi A$  holds in the closure system, whence by Proposition 1, there exists a  $\Sigma$ -inference of every element of  $\text{Id}_\phi A$  from  $A$ . One can append an  $\Sigma^{nb}$ -inference of  $b$  from  $\text{Id}_\phi A$ , to obtain an inference of  $b$  from  $A$ .  $\square$

**Lemma 37.** *Let  $\Sigma$  be any basis of a closure system  $\langle S, \phi \rangle$ . Consider  $aD^\delta b$ , and let  $A \rightarrow b$  be the corresponding implication, where  $A$  is a  $\ll$ -minimal cover for  $b$  and  $a \in A$ . If  $\sigma_1, \dots, \sigma_n$  is a  $\Sigma^{nb}$ -inference of  $b$  from  $\text{Id}_\phi A$ , then  $a \in A_k$  for some  $k \leq n$ .*

*Proof.* Since  $A$  is a  $\ll$ -minimal cover for  $b$ , every  $a \in A$  is a maximal element in  $\text{Id}_\phi A$ . Therefore,  $J_a = \text{Id}_\phi A \setminus \{a\}$  is also an  $\geq_\phi$ -ideal. If some  $a \in A$  does not appear in any  $A_k$ ,  $k \leq n$ , then sequence of  $\sigma_k$ ,  $k \leq n$ , is, in fact, a  $\Sigma^{nb}$ -inference of  $b$  from  $J_a$ . Let  $A'$  be the set of maximal elements of  $J_a$ . Then  $A' \ll A$  and  $J_a = \text{Id}_\phi A'$ . By Lemma 36, there exists a  $\Sigma$ -inference of  $b$  from  $A'$ , and thus  $A' \rightarrow b$  holds in the system. Since  $A \not\ll A'$ , this contradicts the  $\ll$ -minimality of cover  $A$  for  $b$ .  $\square$

**Lemma 38.** *For any regular basis  $\Sigma$ , let  $\sigma_k \in \Sigma^{nb}$ ,  $k \leq n$ , be a non-redundant inference of  $b$  from  $\text{Id}_\phi A$ . Let  $a \in A_s$ , for some  $s \leq n$ . Then  $(a, b) \in \Delta_\Sigma^{tr}$ .*

*Proof.* Proceed by induction on  $t = n - k$ . If  $t = 0$ , i.e.,  $k = n$ , then  $a' \in A_n$ ,  $b \in B_n$ , whence  $(a', b) \in \Delta_\Sigma$ . Suppose we have shown this for all  $t < n - k$ , and consider  $\sigma_k = A_k \rightarrow B_k$  and some  $a' \in A_k$ . If none of elements in  $B_k$  belongs to any  $A_s$ ,  $s > k$ , then this implication could be removed from the sequence, contradicting the assumption that the sequence is non-redundant. Hence, there is  $b' \in B_k \cap A_s$ ,  $s > k$ . Since  $n - s < n - k$ , by the inductive assumption,  $(b', b) \in \Delta_\Sigma^{tr}$ . Besides,  $(a', b') \in \Delta_\Sigma$ . Hence,  $(a', b) \in \Delta_\Sigma^{tr}$ .  $\square$

Note that the following statement now follows from Lemmas 36-38.

**Corollary 39.** *Let  $\Sigma$  be any regular round basis of a closure system  $\langle S, \phi \rangle$ . Then  $D^\delta \subseteq \Delta_\Sigma^{tr}$ . In particular, this holds for  $\Sigma = \Sigma_C$ .*

It remains to show that the containment of Corollary 39 will still hold when we replace basis  $\Sigma_C$  by  $\Sigma^*$  defined prior to Theorem 32.

**Lemma 40.**  $D^\delta \subseteq \Delta_{\Sigma^*}^{tr}$ .

*Proof.* Consider  $(a, b) \in D^\delta$ . Let  $A \rightarrow b$  be an implication corresponding to a minimal cover  $A$  of  $b$ , where  $a \in A$ . Let  $\sigma_1, \dots, \sigma_n$  be a  $\Sigma_C^{nb}$ -inference of  $b$  from  $\text{Id}_\phi A$ . Replace  $A_t$  in each  $\sigma_t = A_t \rightarrow B_t$ ,  $t \leq n$ , by the corresponding  $\sigma^* = A_t^* \rightarrow B_t$  in  $\Sigma^*$ . Since  $A_t^* \subseteq A_t$ , and conclusions of all implications are still the same, the sequence  $\sigma_1^*, \dots, \sigma_n^*$  is an  $\Sigma^{*nb}$ -inference of  $b$  from  $\text{Id}_\phi A$ . Remove unnecessary implications, if needed, making it non-redundant. By Lemma 37,  $a \in A_k$  for some  $k \leq n$ , and by Lemma 38,  $(a, b) \in \Delta_{\Sigma^*}^{tr}$ .  $\square$

Note that Lemma 40 does not hold when  $\Delta_{\Sigma^*}^{tr}$  is replaced by  $\Delta_{\Sigma^*}$ .

**Example 41.** In the closure system of Example 35, we have  $(4, 6) \in D^\delta$ , but  $(4, 6) \notin \Delta_{\Sigma^*}$ .

The consequence is that the  $D$ -relation can be fully recovered from the transitive closure of the modified canonical basis  $\Sigma^*$ , whose premises are taken from any  $K$ -basis.

**Corollary 42.** *A closure system is without  $D$ -cycles iff  $\Delta_{\Sigma^*}^{tr}$  does not have cycles.*

This follows from Theorem 32.

**Theorem 43.** *There exists a polynomial algorithm, in the size of  $s(\Sigma_C)$ , to recognize a closure system without  $D$ -cycles.*

*Proof.* By Proposition 30, it takes polynomial time to recover  $\Delta_{\Sigma^*}$ . Then it takes polynomial time to check whether this relation contains cycles.  $\square$

Note that the systems without  $D$ -cycles include, as a subclass, so-called *quasi-acyclic* closure systems that are studied in the form of Horn Boolean functions in P. Hammer and A. Kogan [16]. For every basis  $\Sigma$  one can define binary relation  $\square_\Sigma$  similar to  $\Delta_\Sigma$ :  $(a, b) \in \square_\Sigma$  iff there exists  $(A \rightarrow B) \in \Sigma$  such that  $a \in A, b \in B$ . The difference is that  $\Delta_\Sigma$  is built only on the non-binary part of (regular)  $\Sigma$ . A quasi-acyclic closure system can be defined as a closure system for which there exists a basis  $\Sigma$  such that  $\square_\Sigma^{tr}$  does not have cycles, unless a cycle is fully within the binary part  $\square_{\Sigma_b}^{tr}$ . The latter cycles never occur in the standard closure systems considered in the current paper. In particular, if a closure system is quasi-acyclic, then it does not have  $D$ -cycles. See K. Adaricheva [1] for further discussion and examples.

## 7. $K$ -BASIS IN CLOSURE SYSTEMS WITH UNIQUE CRITICALS

Starting with this section, we will consider the closure systems which we call  $UC$ -systems.

**Definition 44.** A closure system  $\langle S, \phi \rangle$  is called a system with *unique criticals*, or  $UC$ -system, if, for any critical sets  $C_1, C_2 \in \mathcal{C}$ ,  $\phi(C_1) = \phi(C_2)$  implies  $C_1 = C_2$ .

An immediate consequence of this definition is that, for every essential element  $Y = \phi(Y)$ , there exists a unique implication  $C \rightarrow \phi(C) \setminus C$  in the canonical basis, for which  $\phi(C) = Y$ .

**Proposition 45.** *For every closure system  $\langle S, \phi \rangle$  given by some basis  $\Sigma$ , it requires time  $O(s(\Sigma)^2 \cdot |\Sigma|)$  to determine whether it is a UC-system.*

*Proof.* Indeed, it will require  $O(s(\Sigma)^2)$  time to build the canonical basis. With  $\Sigma_C$  available, one needs to check whether  $\phi(C_1) \neq \phi(C_2)$ , for any two premises  $C_1, C_2$  of the canonical basis. It takes time  $O(s(\Sigma_C))$  to find the closure of any given set, so verifying all the pairs requires time  $O(s(\Sigma_C) \cdot n^2)$ . Thus, overall time should not exceed  $O(s(\Sigma)^2 \cdot |\Sigma|)$ .  $\square$

One important advantage of UC-closure systems is that one can establish a lower bound for  $s_R(\Sigma^{nb})$ , among all minimum bases  $\Sigma$ .

Let  $\mathcal{C}_{>1} \subseteq \mathcal{C}$  be the subset of all critical sets with more than one element. For every  $C \in \mathcal{C}_{>1}$ , define  $M(C) = \{y \in X : y \text{ is a } \geq_\phi \text{-maximal in } \phi(C) \setminus C, \text{ and } y \notin \phi(C') \text{ when } \phi(C') \subset \phi(C), C' \in \mathcal{C}_{>1}\}$ .

Recall from Theorem 2 (2) that any minimum basis of a closure system has the form  $\Sigma = \{X_C \rightarrow Y_C : C \in \mathcal{C}, \sigma(X_C) = C\}$ .

**Lemma 46.** *Let  $\langle S, \phi \rangle$  be a UC-closure system, and  $\Sigma$  be any minimum basis. If  $(X_C \rightarrow Y_C) \in \Sigma$  is an implication in  $\Sigma$  corresponding to  $C \in \mathcal{C}$ , then  $M(C) \subseteq Y_C$ .*

*Proof.* We have  $y_m \in \phi(X_C)$ , so the implication  $X_C \rightarrow y_m$  follows from  $\Sigma$ . Let  $\sigma_1, \dots, \sigma_k \in \Sigma$ ,  $\sigma_i = A_i \rightarrow B_i$ , be a  $\Sigma$ -inference of  $y_m$  from  $X_C$ . By inductive argument, one can show that  $A_i, B_i \subseteq \phi(X_C)$ . Since  $y_m$  is maximal in  $\phi(X_C) \setminus C = \phi(C) \setminus C$ , the last implication  $\sigma_k$  is not binary. If  $\phi(A_k) \subset \phi(C)$ , then  $A_k \subseteq C'$  for a critical set  $C' \neq C$ ,  $\phi(C') = \phi(A_k) \subset \phi(C)$ . Thus,  $y_m$  is also the maximal in  $\phi(C') \setminus C'$ , which is not possible by assumption. Hence,  $\phi(A_k) = \phi(X_C) = \phi(C)$ . But in closure systems without  $D$ -cycles, there exists only one critical set for an essential element  $\phi(C)$ , i.e.,  $X_C = A_k$  and  $y_m \in Y_C$ .  $\square$

**Corollary 47.** *If  $\mathcal{C}_{>1} = \{C_1, \dots, C_k\}$ , then, for every minimum basis  $\Sigma$  of an UC-system,  $s_R(\Sigma^{nb}) \geq |M(C_1)| + \dots + |M(C_k)|$ .*

We will see in section 9 that this lower bound can be attained in the  $E$ -basis of a closure system without  $D$ -cycles.

One important subclass of UC-systems is the class of closure systems whose closure lattice satisfies  $(SD_\vee)$ . We will call an element  $t \in \text{Cl}(S, \phi)$  *join-semidistributive*, if  $(SD_\vee)$  holds at  $t$ :  $t = x \vee y = x \vee z \rightarrow t = x \vee (y \wedge z)$ . We first recall a well-known theorem of B. Jónsson and J.E. Kiefer.

**Theorem 48.** [17] *An element  $t$  of a finite lattice is join-semidistributive iff  $t$  has a unique  $\ll$ -minimal representation.*

We will call closure system  $\langle S, \phi \rangle$  *join-semidistributive* if all elements of  $\text{Cl}(S, \phi)$  are such. As was pointed in section 5, the premises of implications in  $K$ -basis, in its non-binary part, are associated with the  $\ll$ -minimal join-representation of essential elements in the closure lattice.

**Proposition 49.** *Let standard closure system  $\langle S, \phi \rangle$  be join-semidistributive. Then*

- (1)  $\langle S, \phi \rangle$  is an UC-closure system;

(2)  $\langle S, \phi \rangle$  has unique  $K$ -basis.

*Proof.* (1) Fix any essential set  $Y$ , and let  $C$  be any critical subset with  $\phi(C) = Y$ . Then  $C$  should contain some minimal order generator  $C_K$ , and  $\sigma(C_K) = C$ . On the other hand,  $C_K$  gives a  $\ll$ -minimal representation of  $\phi(C) = Y$ , and this representation is unique in closure systems with  $(SD_\vee)$ . Hence,  $C = \sigma(C_K)$  is unique critical set associated with the essential element  $Y$ .

(2) Since, by Theorem 48, every element of  $\text{Cl}(S, \phi)$  has a unique  $\ll$ -minimal join representation, apply Proposition 25 to conclude that the  $K$ -basis of the associated standard closure system must be unique.  $\square$

For both statements in Proposition 49,  $(SD_\vee)$  is a sufficient condition. It is easy to show that it is not necessary. Example 29 provides an  $UC$ -system which has two  $K$ -bases, and hence is not join-semidistributive. Even if the closure system is both  $UC$  and has a unique  $K$ -basis, which guarantees that  $(SD_\vee)$  holds at all essential elements, it does not imply that  $(SD_\vee)$  holds at non-essential elements.

**Example 50.** Take a closure system on  $S = \{a, b, c, d\}$  defined by implications  $\Sigma_C = \{ac \rightarrow b, bd \rightarrow c\}$ . There are two essential elements  $\phi(ac), \phi(bd)$ , and  $SD_\vee$  holds at both. In particular, the  $K$ -basis is unique and coincides with  $\Sigma_C$ . Nevertheless,  $(SD_\vee)$  fails at the non-essential element  $\phi(abcd) = 1$ :  $(b \vee c) \vee a = (b \vee c) \vee d = 1$ , but  $(b \vee c) \vee (a \wedge d) = b \vee c$ .

This leaves us with the following open questions.

**Problem 51.**

- (A) Is there an algorithm that allows us to recognize in polynomial time whether the system has a unique  $K$ -basis?
- (B) Is there an effective algorithm that allows us to recognize whether a closure system  $\langle S, \phi \rangle$  is join-semidistributive, given its canonical basis?

A positive answer to question (A), even in  $UC$ -system, does not automatically imply the positive answer to (B), as Example 50 demonstrates. We will address lattice theoretical aspects of  $UC$ -systems in a separate paper [3].

At the end of this section, we discuss a new binary part for join-semidistributive closure systems.

Recall from Lemma 19 that any implication  $(x \rightarrow A) \in \Sigma^b$  can be replaced by  $x \rightarrow B$  with  $\phi(B) = \phi(\{x\}) \setminus \{x\}$ .

In the most closure systems we treated so far, the binary part was built out of the cover relation of the order  $\geq_\phi$ . Precisely,  $B = \max_{\geq_\phi}(\phi(\{x\}) \setminus \{x\})$ , for every  $(x \rightarrow B) \in \Sigma^b$ . In join-semidistributive closure systems, there is another natural choice for  $B$ .

**Definition 52.** We will call set of implications  $\Sigma_F$  the  $F$ -basis of a join-semidistributive standard closure system, if

- (1)  $\Sigma_F^{nb} = \Sigma_K^{nb}$ ;
- (2)  $\Sigma_F^b = \{x \rightarrow B : B \text{ is a minimal order generator of } \phi(\{x\}) \setminus \{x\}\}$ .

We note that  $F$ -basis of any join-semidistributive closure system is unique. We saw in Example 23 that binary part based on the cover relation  $\succ_\phi$  might be

redundant. The advantage of the binary part of the  $F$ -basis is that it possesses a weak form of minimality. Recall that  $\Sigma_u^b$  denotes the unit expansion of the implications  $\Sigma^b$ .

**Proposition 53.**  $\Sigma_u^b$  of the  $F$ -basis  $\Sigma_F$  for a join-semidistributive closure system  $\langle S, \phi \rangle$  is non-redundant.

*Proof.* We need to show that none of the implications in  $\Sigma_u^b$  of the  $F$ -basis can be removed.

Suppose that  $a \rightarrow B$  is a binary implication in  $\Sigma_F$ , where  $B = \{b_0, \dots, b_n\}$  is a minimal order generator of  $\phi(\{a\}) \setminus \{a\}$ . Consider a closure system  $\Sigma = \Sigma_F \setminus \{a \rightarrow b_0\}$ . Let  $\phi_\Sigma$  be an associated closure operator.

**Claim.**  $\phi_\Sigma(\{a\}) = \phi(\{b_1, \dots, b_n\}) \cup \{a\}$ .

Indeed, to show  $\supseteq$ , we note that  $b_1, \dots, b_n, a \in \phi_\Sigma(\{a\})$ . All the implications of  $\Sigma_F$  that can be applied to elements of  $\phi(\{b_1, \dots, b_n\})$  are also in  $\Sigma$ , due to the fact that  $a \notin \phi(\{b_1, \dots, b_n\})$ .

Now take any  $w \in \phi_\Sigma(\{a\})$ , with  $w \neq a$ . Let  $\sigma_1, \dots, \sigma_k \in \Sigma$  be a  $\Sigma$ -inference of  $w$  from  $\{a\}$ . We may assume that  $\sigma_1 = (a \rightarrow \{b_1, \dots, b_n\})$  and  $a$  does not appear in a premise of any implication  $\sigma_2, \dots, \sigma_k$ , since  $\Sigma_F$  is regular. Hence,  $A_i, B_i \subseteq \phi(\{b_1, \dots, b_n\})$ , for all  $i \geq 2$ , by the inductive argument. This implies  $w \in \phi(\{b_1, \dots, b_n\})$ , which proves the Claim.

It immediately follows from the Claim that  $\phi_\Sigma(\{a\}) \subset \phi(\{a\})$ , since  $b_0 \notin \phi(\{b_1, \dots, b_n\})$ , due to the  $\ll$ -minimality of  $B$  for  $\phi(\{a\}) \setminus \{a\}$ .  $\square$

We note that there is no guarantee that  $|B|$  in an implication  $(a \rightarrow B) \in \Sigma_F$  for some join-semidistributive closure system is minimal with the property  $\phi(B) = \phi(\{a\}) \setminus \{a\}$ . Examples of such systems will be shown in section 10.

## 8. $E$ -BASIS AND ITS CONNECTION TO THE CANONICAL BASIS

We introduced the  $E$ -basis in [2] as a finer version of the  $D$ -basis for closure systems without  $D$ -cycles. In section 6 we showed that one can recognize whether a closure system is without  $D$ -cycles from its canonical basis, in time polynomial of the size of that basis.

In this section, we will establish the connection between the  $E$ -basis and the canonical basis. Since every finite lattice without  $D$ -cycles is join-semidistributive, the results of section 7 are applicable here. In particular, the  $K$ -basis is unique for closure systems without  $D$ -cycles.

We will proceed as follows. First, we prove that the  $E$ -basis in its aggregated form is a refinement of the canonical basis. Secondly, we establish that the left sides of the  $E$ -basis and the  $K$ -basis are the same. Thirdly, we will analyze possible differences in the right sides of the non-binary parts of the  $K$ -basis and the  $E$ -basis. This will allow us to create a blend of the two, which we will call the optimized  $E$ -basis. There is an effective algorithm to produce the optimized  $E$ -basis from the canonical basis.

First, recall the definition of the  $E$ -basis from [2].

Let  $\langle S, \phi \rangle$  be a standard closure system without  $D$ -cycles. For every  $x \in S$ , let  $M(x) = \{Y \subseteq S : Y \text{ is a minimal cover of } x\}$ . The family  $\phi(M(x)) = \{\phi(Y) : Y \in M(x)\}$  is ordered by set containment, so we can consider its  $\subseteq$ -minimal elements. Let  $M^*(x) = \{Y \in M(x) : \phi(Y) \text{ is } \subseteq\text{-minimal in } \phi(M(x))\}$ .

The set of implications  $\Sigma_E = \Sigma_E^b \cup \Sigma_E^n$ :

- (1)  $\Sigma_E^b = \{y \rightarrow x : y \succ_\phi x\}$ ,
- (2)  $\Sigma_E^n = \{X \rightarrow x : X \in M^*(x)\}$

is called the *E-basis* of the closure system. Recalling the definition of the *D-basis* given in section 6, we note that  $\Sigma_E \subseteq \Sigma_D$ .

It was noted in [2] that  $\Sigma_E$  might not form a basis of a closure system in general, so the requirement of no *D-cycles* is essential.

**Example 54.** Consider a closure system given by the basis  $\Sigma = \{ac \rightarrow b, bd \rightarrow c, ad \rightarrow bc\}$ , which is the canonical basis, at the same time the *K-basis* of the system. It is easy to check that this system is join-semidistributive, but it has the *D-cycle*  $bDcDb$ . We have  $\{a, c\}, \{a, d\} \in M(b)$ , and  $\phi(\{a, c\}) \subset \phi(\{a, d\})$ , so that only  $\{a, c\}$  belongs to  $M^*(b)$ . Symmetrically, only  $\{b, d\}$  belongs to  $M^*(c)$ . Hence, the set of implications  $\Sigma_E$  for this system is  $\{ac \rightarrow b, bd \rightarrow c\}$ , which is not a basis, since  $\{a, d\}$  would be a closed set with respect to  $\Sigma_E$ .

Let  $\Sigma_E^{ag}$  be the aggregated *E-basis* and  $\Sigma_C$  be the canonical basis of  $\langle S, \phi \rangle$ . As usual,  $\sigma$  denotes the saturation operator associated with the closure operator  $\phi$ .

**Theorem 55.** *Let  $\langle S, \phi \rangle$  be a standard closure system without *D-cycles*. Then there is a one-to-one mapping  $f : \Sigma_E^{ag} \rightarrow \Sigma_C$  such that  $f(A \rightarrow A') = (\sigma(A) \rightarrow \phi(A))$  with  $A' \subseteq \phi(A) \setminus \sigma(A)$ . In particular, the aggregated *E-basis* is a refinement of  $\Sigma_C$ :  $s(\Sigma_E^{ag}) \leq s(\Sigma)$ .*

In fact, the result follows, if we argue over  $\Sigma_E$  in its original unit form, and the unit expansion  $\Sigma_u$  of  $\Sigma_C$ , and show that there is one-to-one mapping  $f^* : \Sigma_E \rightarrow \Sigma_u$  such that  $f^*(A \rightarrow x) = (\sigma(A) \rightarrow x)$ .

We now describe a well-known algorithm of producing the canonical basis due to A. Day [11].

Given any basis  $\Sigma = \{A_i \rightarrow B_i : i \leq k\}$  of a closure system  $\langle S, \phi \rangle$ , the procedure of obtaining the canonical basis  $\Sigma_C$  consists of two steps:

- (1) replace each  $A_i \rightarrow B_i$  by  $\sigma(A_i) \rightarrow \phi(A_i)$ ;
- (2) if there are two implications  $C \rightarrow F$  and  $D \rightarrow F$  obtained in the first step, and  $C \subseteq D$ , then remove  $D \rightarrow F$ ; also, remove all implications of the form  $F \rightarrow F$ .

The unit expansion  $\Sigma_u$  of  $\Sigma_C$  can now be written from all remaining implications and will consist of  $\sigma(A_i) \rightarrow x$ , where  $A_i$  is some premise in the original basis and  $x \in \phi(A_i) \setminus \sigma(A_i)$ .

In the proof of Theorem 55 below, we will apply this algorithm to a given basis  $\Sigma_E$  of some closure system without *D-cycles*. First, we prove several auxiliary statements. We need to recall that the saturation operator  $\sigma$  associated with a closure operator  $\phi$  is defined as  $\sigma(X) = \bigcup_{k \geq 1} q^k(X)$ , where  $q(X) = X \cup \bigcup \{\phi(Y) : Y \subseteq X \text{ and } \phi(Y) \subset \phi(X)\}$ .

**Lemma 56.** *If  $(A \rightarrow x) \in \Sigma_E^n$ , then  $x \in \phi(A) \setminus \sigma(A)$ .*

*Proof.* It is enough to show that  $x \notin \sigma(A)$ . We have  $x \notin A$ . Also,  $x \notin \phi(Y)$ , for any  $Y \subseteq A$  with  $\phi(Y) \subset \phi(A)$ . Indeed, otherwise,  $Y$  would be a cover for  $x$ , so we would be able to find a minimal cover  $Y'$  with  $Y' \ll Y$ , hence,  $\phi(Y') \subset \phi(A)$ . This implies that  $A \notin M^*(x)$ , hence,  $A \rightarrow x$  cannot be included in *E-basis*. Thus,  $x \notin q(A)$ . Apparently,  $x \notin q(A)$  with  $\phi(q(A)) = \phi(A)$ , so we can apply the same

argument, replacing  $A$  by  $q(A)$ , to show that  $x \notin q^2(A)$ . Proceeding with this inductive argument, we conclude that  $x \notin q^k(A)$ , for all  $x \geq 1$ .  $\square$

**Lemma 57.** *If  $(A \rightarrow x) \in \Sigma_E$ , then  $A$  is a  $\ll$ -minimal join representation of  $\phi(A)$ .*

*Proof.* We may assume that  $|A| > 1$ , since for binary implication the statement is trivial.

According to the definition of the  $E$ -basis,  $A$  is a minimal cover of  $x$ . If there is a set  $F \ll A$  such that  $\phi(F) = \phi(A)$  and  $A \not\subseteq F$ , then  $F$  will also be a cover of  $x$ , which contradicts the  $\ll$ -minimality of the cover  $A$ .  $\square$

*Proof of Theorem.* Apply the algorithm for obtaining the canonical basis to  $\Sigma_E$ . Apparently,  $\Sigma_E^b \subseteq \Sigma_u^b$ . So we need to make sure the one-to-one mapping  $f^*$  exists for non-binary implications.

According to Lemma 57, if  $(A \rightarrow x), (B \rightarrow y) \in \Sigma_E$  and  $A \neq B$ , then  $\phi(A) \neq \phi(B)$ . Thus, after applying the first step of the algorithm, we will not obtain any two implications  $C \rightarrow F, D \rightarrow F$ , with  $C \neq D$ . Moreover, by Lemma 56, we will not have implications of the form  $F \rightarrow F$  after the first step. Hence, the second step of the algorithm may be applied only to remove repeating implications, and they may occur here only when  $A = B$ . In this case  $x \neq y$ , and  $x, y \in \phi(A) \setminus \sigma(A)$  by Lemma 56. Therefore,  $\sigma(A) \rightarrow x$  and  $\sigma(A) \rightarrow y$  are in the unit expansion of the implication  $\sigma(A) \rightarrow \phi(A)$  from the canonical basis. Hence, the mapping  $f^*(A \rightarrow x) = (\sigma(A) \rightarrow x)$  from  $\Sigma_E$  to  $\Sigma_u$  is one-to-one.

*End of Proof.*

Now we show that the premises of implications in the  $E$ -basis and  $K$ -basis for a closure system without  $D$ -cycles are the same.

**Lemma 58.** *Let  $\langle S, \phi \rangle$  be a standard closure system without  $D$ -cycles with  $E$ -basis  $\Sigma_E$  and  $K$ -basis  $\Sigma_K$ . For each critical set  $C$ , if  $(C_E \rightarrow Y_E) \in \Sigma_E^{ag}$  and  $(C_K \rightarrow Y_K) \in \Sigma_K$  with  $\sigma(C_E) = \sigma(C_K) = C$ , then  $C_E = C_K$ .*

*Proof.* According to Proposition 25,  $C_K$  is a  $\ll$ -minimal join representation of  $\phi(C)$ , which is unique in join-semidistributive closure systems. Due to Lemma 57,  $C_E = C_K$ .  $\square$

The following examples show that the conclusions of non-binary parts of the  $E$ -basis and  $K$ -basis might be different.

**Example 59.** The  $K$ -basis may have smaller size than the  $E$ -basis.

Consider the closure system on  $S = \{a, b, c, d\}$  given by its canonical basis  $\Sigma_C = \{d \rightarrow cb, c \rightarrow b, ab \rightarrow dc\}$ . This is also the  $E$ -basis. On the other hand, the  $K$ -basis optimizes the last implication to  $ab \rightarrow d$ , since  $d$  is the maximal element of  $\{d, c\}$  with respect to the  $\geq_\phi$ -order.

**Example 60.** The  $E$ -basis may have smaller size than the  $K$ -basis.

Consider the closure system on  $S = \{2, 3, 4, 5\}$  defined by its canonical basis  $\Sigma_C = \{2 \rightarrow 5, 45 \rightarrow 23, 35 \rightarrow 2\}$ . This is also the  $K$ -basis of the system. On the other hand, 2 appears on the right side of two non-binary implications, in particular, both  $\{4, 5\}$  and  $\{3, 5\}$  are the minimal covers for 2. Since  $\phi(45) > \phi(35)$ ,

only  $35 \rightarrow 2$  appears in the  $E$ -basis. Thus, the  $E$ -basis has smaller size than the  $K$ -basis:  $\Sigma_E = \{2 \rightarrow 5, 45 \rightarrow 3, 35 \rightarrow 2\}$ .

**Definition 61.** We call  $\Sigma_{OE}$  an optimized  $E$ -basis if every implication  $A_E \rightarrow B_E$  from the non-binary part of  $E$ -basis is replaced by  $A_E \rightarrow B_{OE}$ , where  $B_{OE} = \max_{\geq \phi}(B_E)$ .

Note that the  $E$ -basis is *ordered direct*, as was shown in Theorem 23 [2]. The optimized  $E$ -basis might not longer have this property. On the other hand, applying the ordered sequence of the  $E$ -basis to the optimized  $E$ -basis and concatenating the binary part again at the end will produce an ordered sequence. See the definition of the ordered sequence and further details in section 8 of [2].

Let us demonstrate this on Example 59. The optimized  $E$ -basis is the same as the  $K$ -basis:  $\Sigma_{OE} = \{d \rightarrow c, c \rightarrow b, ab \rightarrow d\}$ .  $\Sigma_{OE}$  is no longer direct, with this order of implications that is inherited from the order of the  $E$ -basis, since  $\rho(ab) = abd \neq abcd = \phi(ab)$ . If we concatenate the binary part at the end, we obtain the following sequence:  $d \rightarrow c, c \rightarrow b, ab \rightarrow d, d \rightarrow c, c \rightarrow b$ . It is easy to verify that it is an ordered direct sequence.

**Lemma 62.** *Given a closure system without  $D$ -cycles, it will take time  $O(s^2(\Sigma_C))$  to obtain optimized  $E$ -basis from the canonical basis.*

*Proof.* Given the canonical basis, it takes time  $O(s^2(\Sigma_C))$  to verify that there is no  $D$ -cycles and to obtain the unique  $K$ -basis  $\Sigma_K$ . For each  $x \in S$ , choose all non-binary implications  $(X_K \rightarrow Y_K) \in \Sigma_K$  such that  $x \in Y_K$ . We know that  $X_K$  is a  $\ll$ -minimal cover for  $x$ . Compute  $\phi(X_K)$  for all such implications and choose minimal ones, with the respect to the containment order. Keep  $x$  in  $Y_K$  only if  $\phi(X_K)$  is  $\subseteq$ -minimal; otherwise, remove  $x$  from  $Y_K$ . There are no more than  $|X|$  elements to check, and each may appear in no more than  $|\Sigma_C|$  implications. Computation of  $\phi(X_K)$  takes time  $O(s(\Sigma_C))$ . Assuming that  $|X| \cdot |\Sigma_C| \approx s(\Sigma_C)$ , it will take time  $O(s^2(\Sigma_C))$  to build the optimized  $E$ -basis.  $\square$

## 9. ASPECTS OF OPTIMALITY OF THE $E$ -BASIS AND ITS FURTHER MODIFICATIONS

As was discussed in sections 3 and 4, every optimum basis of a standard closure system has parameters fully determined by the closure operator: the right size of each binary implication (associated with one-element critical sets) and the left size of each non-binary implication (associated with critical sets of more than one element).

In section 10 we will show that finding any of those optimal parts in closure systems without  $D$ -cycles is an NP-complete problem. On the other hand, the optimum right size of the non-binary part can be achieved.

First, we summarize from the previous results.

**Proposition 63.** *Let  $\langle S, \phi \rangle$  be a standard closure system. There exists a constant  $s_R^{nb}(S, \phi)$  such that  $s_R(\Sigma^{nb}) = s_R^{nb}(S, \phi)$  for every optimum basis  $\Sigma$ .*

*Proof.* Indeed,  $s(\Sigma) = s_L(\Sigma^b) + s_R(\Sigma^b) + s_L(\Sigma^{nb}) + s_R(\Sigma^{nb})$ , for every basis  $\Sigma$ . Now assume that  $\Sigma$  is optimum. The number of implications in  $\Sigma^b$  is fully determined by the number of one-element critical sets, so that  $s_L(\Sigma^b)$  is the same for all optimum bases. By Theorem 20,  $s_R(\Sigma^b)$  is the sum of parameters  $b_C$ , where  $C$  runs over all

singletons in  $\mathcal{C}$ , and  $s_L(\Sigma^n)$  is the sum of parameters  $k_C$ , where  $C$  runs over all non-singletons in  $\mathcal{C}$ , as shown in Theorem 2 (2). Therefore, the last components in the sum,  $s_R(\Sigma^{nb})$ , must be the same for all optimum bases.  $\square$

**Corollary 64.** *The inequality  $s_R^{nb}(S, \phi) \leq s_R(\Sigma_*^{nb})$  holds for any basis  $\Sigma_*$  of a standard closure system  $\langle S, \phi \rangle$ .*

*Proof.* By Lemma 19 we may assume that the binary part of  $\Sigma_*$  matches the binary part of some optimum basis  $\Sigma$ . Now  $\Sigma$  is right-side optimum, due to Corollary 11, whence  $s_R(\Sigma) \leq s_R(\Sigma_*)$ . Since these bases have identical binary parts, we also have  $s_R^{nb}(S, \phi) = s_R(\Sigma^{nb}) \leq s_R(\Sigma_*^{nb})$ .  $\square$

**Remark 65.** Unlike the parameters  $b_C$  and  $k_C$  of parts of optimum basis used in argument of Proposition 63, there is no fixed value for the size of the *conclusion* in the *non-binary* implication  $A_C \rightarrow B_C$  of any optimum basis.

**Example 66.** Let the closure system on  $S = \{a, b, c, z\}$  be given by the canonical basis  $\Sigma_C = \{z \rightarrow a, ab \rightarrow cz, ac \rightarrow bz\}$ . There are two optimum bases:  $\Sigma_1 = \{z \rightarrow a, ab \rightarrow c, ac \rightarrow bz\}$  and  $\Sigma_2 = \{z \rightarrow a, ab \rightarrow cz, ac \rightarrow b\}$ . While  $s_R(\Sigma_1^{nb}) = s_R(\Sigma_2^{nb}) = 3$ , implications in  $\Sigma_1$  and  $\Sigma_2$  with premise  $ab$  have conclusions of different sizes, and the same holds for implications with  $ac$  in the premise.

On the other hand, we notice that the closure system in Example 66 is not *UC*, since essential element  $\phi(ab) = \phi(ac)$  contains two critical subsets.

**Conjecture 67.** If  $A_1 \rightarrow B_1, \dots, A_k \rightarrow B_k$  are *all* implications of any optimum basis, corresponding to critical sets with the same closure, i.e.,  $\phi(A_1) = \dots = \phi(A_k)$ , then  $s = |B_1| + \dots + |B_k|$  does not depend on the choice of the optimum basis. In particular, in *UC*-closure systems, for every critical set  $C$ , and corresponding implication  $A_C \rightarrow B_C$  in any optimum basis,  $|B_C|$  does not depend on the choice of the optimum basis.

The main statement of this section is that the optimized *E*-basis of any closure system without *D*-cycles has the optimum right size in its non-binary part. For this, we recall that every closure system without *D*-cycles is join-semidistributive, and thus it is also a *UC*-system.

**Theorem 68.** *Let  $\langle S, \phi \rangle$  be any closure system without *D*-cycles. Then  $s_R(\Sigma_{OE}^{nb}) = s_R^{nb}(S, \phi)$ .*

*Proof.* Consider any optimum basis  $\Sigma$ . Since it is minimum, Lemma 46 is applicable to  $\Sigma$ . Take any  $(X_C \rightarrow Y_C) \in \Sigma^{nb}$  and corresponding  $(A_C \rightarrow B_C) \in \Sigma_{OE}^{nb}$ ,  $C \in \mathcal{C}$ . By the definition of  $\Sigma_{OE}$ ,  $B_C$  consists of maximal elements  $y_m$  in  $\phi(A_C) \setminus \sigma(A_C) = \phi(C) \setminus C$  such that  $y_m \notin \phi(C') \setminus C'$ , for every  $C' \in \mathcal{C}$  with  $\phi(C') \subset \phi(C)$ . According to Lemma 46, this implies  $B_C \subseteq Y_C$ , hence,  $s_R(\Sigma_{OE}^{nb}) \leq s_R(\Sigma^{nb})$ . On the other hand,  $s_R(\Sigma_{OE}^{nb}) \geq s_R^{nb}(S, \phi) = s_R(\Sigma^{nb})$ , by Corollary 64. Thus,  $s_R(\Sigma_{OE}^{nb}) = s_R^{nb}(S, \phi)$ .  $\square$

**Remark 69.** We note that the result of Theorem 68 cannot be extended to join-semidistributive closure systems, when replacing the *E*-basis by a *K*-basis. We saw in Example 54 that the conclusions of non-binary implications in a *K*-basis cannot be reduced by taking away some  $y_m \in \phi(C) \setminus C$ , even if  $y_m \in \phi(C') \setminus C'$ , for some critical sets  $C, C'$ ,  $\phi(C') \subset \phi(C)$ . The same example shows that a *K*-basis may not reach  $s_R^{nb}(S, \phi)$  in the size of its non-binary right side. Indeed, the implication

$ad \rightarrow bc$  can be reduced to either  $ad \rightarrow b$  or  $ad \rightarrow c$ , to obtain a right-side optimum basis.

At the end of this section we mention the binary modification for basis  $\Sigma_{OE}$  that is available for all join-semidistributive closure systems discussed in section 7. It combines the binary part of the  $F$ -basis and non-binary part of the optimized  $E$ -basis.

**Definition 70.**  $\Sigma_{FOE} = \Sigma_F^b \cup \Sigma_{OE}^{nb}$ .

**Proposition 71.** *The unit expansion of  $\Sigma_{FOE}$  is non-redundant.*

Indeed, this follows from Theorem 68 and Proposition 53.

#### 10. FINDING OPTIMUM OR MINIMUM UNIT BASIS FOR BOUNDED LATTICES IS NP-COMPLETE

As was noted in section 3, finding an optimum basis for an arbitrary closure system is an NP-complete problem, while it could be done effectively in closure systems with modular closure lattices.

The effective procedure that allows us to build the optimized  $E$ -basis for systems without  $D$ -cycles and reaches the optimum in its non-binary right side, could suggest that such systems also have an effective procedure for finding an optimum basis. The goal of this section is to show, that, to the contrary, finding an optimum basis even in the strict subclass of systems without  $D$ -cycles (closure systems with *bounded* closure lattices) is NP-complete.

In this section we show that each of the problems is NP-complete, for the closure systems without  $D$ -cycles:

- (1) find a basis  $\Sigma$  that reaches the minimum in  $s_L(\Sigma^{nb})$ ;
- (2) find a basis  $\Sigma$  that reaches the minimum in  $s_R(\Sigma^b)$ .

For this aim, we need to identify a known NP-complete problem that can be reduced to either of the problems above.

The following problem known as *the set cover problem* is included into original ‘‘Karp’s21’’, the list of 21 NP-complete problems in Karp [18].

Given a finite set  $Q$  and a family of its subsets  $\mathcal{Q} = \{Q_i : i \leq k\}$  that *covers*  $Q$ , i.e.,  $Q \subseteq \bigcup Q_i$ , the set cover problem is to identify the smallest subfamily of  $\mathcal{Q}$  that still covers  $Q$ .

Now we will develop the path to reduce the set cover problem to the problem of finding a *minimal generator* for some critical set in closure system without  $D$ -cycles (see section 5 to recall definitions). This is equivalent to identifying a premise for one non-binary implication in an optimum basis. First, we illustrate it on a simple example.

**Example 72.** Let  $Q = \{q_1, q_2, q_3, q_4\}$  be a finite set with the set cover  $\mathcal{Q} = \{\{q_1\}, \{q_2\}, \{q_3\}, \{q_4\}, \{q_1, q_2\}\}$ . Apparently, the solution to the set cover problem will be given by family  $\mathcal{B} = \{\{q_3\}, \{q_4\}, \{q_1, q_2\}\}$ .

We will construct a closure system without  $D$ -cycles on an extension  $S$  of  $Q$  so that  $Q$  will be the minimal *order* generator for some critical set. At the same time,  $Q$  will not be a minimal generator, and the latter will be rather associated with the subfamily  $\mathcal{B}$ .

The procedure will use a lattice theoretical construction to build  $L = \text{Cl}(S, \phi)$ , for a standard system  $\langle S, \phi \rangle$ . For every subset  $Y$  of family  $\mathcal{Q}$  which is a not a

singleton, add a new element  $z(Y)$  to  $Q$ , also add another element  $w$ . Thus, with notation  $z = z(\{q_1 q_2\})$ , we have in our case  $S = \{q_1, q_2, q_3, q_4, z, w\}$ .

Start from  $L_0 = \mathbf{2}^Q$ . Then use the *doubling construction* introduced in A. Day [10] to double the element  $b = \{q_1, q_2\} \in L_0$ . This will replace  $b$  by a 2-element interval  $[b, z]$  and extend lattice operations so that  $b \vee x = z \vee x$ , for every  $x \not\leq b$ , and  $b \wedge x = z \wedge x$ , for every  $x \not\geq b$ ,  $x \in L_0$ . Let  $L_1$  be the lattice after this doubling. Then  $z$  is a join irreducible element in  $L_1$ ; moreover,  $z \geq q_1, q_2$ .

Finally, for  $w$ , choose any maximal proper subset in  $Q$ , say,  $\{q_1, q_2, q_3\}$ , and let  $t \in L_1$  be the element inherited from  $\{q_1, q_2, q_3\} \in L_0$ . Apply the doubling construction to  $t$  to obtain a 2-element interval  $[t, w]$ , so a new lattice  $L_2$  after this second doubling will have  $w$  as a new join irreducible element.

Let  $\langle S, \phi \rangle$  be a closure system with  $\text{Cl}(S, \phi) = L_2$ . It is known that every lattice obtained from a Boolean lattice by the series of doublings of intervals is a *bounded* lattice, so that in particular,  $\langle S, \phi \rangle$  is without  $D$ -cycles, see A. Day [10].

It is straightforward to check that  $C = \{q_1, q_2, q_3, q_4, z\}$  is a critical set in this closure system with  $\phi(C) = S$ . Indeed,  $z \in \phi(q_1, q_2, q_3) \subset \phi(C)$  and  $\phi(C) \neq C$ . Besides,  $Q$  is a minimal *order* generator for  $C$ .

In particular, the implication  $Q \rightarrow w$  is in the  $E$ -basis (and  $K$ -basis). Nevertheless, one can find a generator  $B = \{z, q_3, q_4\}$  of smaller size. Indeed, since  $z \rightarrow q_1, z \rightarrow q_2$  must be in the binary part of any basis for  $\langle S, \phi \rangle$ , we have  $\sigma(B) = C$ . Thus, an optimum basis for  $\langle S, \phi \rangle$  will have  $B \rightarrow w$ , not  $Q \rightarrow w$ .

Note that every element in  $B$  corresponds to an element of the family  $\mathcal{B}$  that solves the original set cover problem.

This example should clarify the general procedure described in the next statement. We note that the set cover problem has an effective solution when  $\mathcal{Q}$  contains  $Q$ , because the minimal cover will be just a single set. Similarly, it has an effective solution, if  $\mathcal{Q}$  does not have  $Q$ , but contains the set  $Q_1 = Q \setminus q$ , for one of  $q \in Q$ . In this case the solution is two-element family  $Q_1, Q_2$ , where  $q \in Q_2$ . These instances of set cover problem will be considered *trivial*.

**Lemma 73.** *Suppose  $(Q, \mathcal{Q})$  is a non-trivial instance of the set cover problem. One can effectively find a closure system, whose closure lattice is bounded, such that a particular non-binary implication  $U \rightarrow V$  in its optimum basis translates into the solution of this set cover problem.*

*Proof.* Due to assumption,  $\mathcal{Q}$  does not have  $Q$  or any  $Q \setminus \{q\}$ . Suppose  $Q = \{q_1, \dots, q_n\}$  and  $\mathcal{Q}$  has  $k$  elements of cardinality  $> 1$ .

First, we make a slight modification to the instance of the set cover problem. Extend the family  $\mathcal{Q}$  to  $\mathcal{Q}^* = \mathcal{Q} \cup \{\{q_i\} : i \leq n\}$ . A solution to the instance  $(Q, \mathcal{Q}^*)$  should be of cardinality at most as large as the solution to original instance  $(Q, \mathcal{Q})$ , simply because we might have more available covering subfamilies in second instance. On the other hand, if the solution to the second instance contains any one-element subsets which are not in  $\mathcal{Q}$ , we can replace them by subsets in  $\mathcal{Q}$  that contain those singletons, obtaining the solution at most as large as the solution to the second instance.

Thus, we may replace original instance by the instance  $(Q, \mathcal{Q}^*)$ .

To build a finite bounded lattice  $L$ , start from  $L_0 = \mathbf{2}^Q$ , then double every element  $b_i \in \mathbf{2}^Q$  that corresponds to  $Q_i \in \mathcal{Q}$ ,  $|Q_i| > 1$ ,  $i \leq k$ . This will add new

join-irreducible elements  $z_i$ ,  $i \leq k$ . It implies, in particular, that  $z_i \geq q$ , for every  $q \in Q_i$ .

Also, choose an arbitrary element  $t$  corresponding to some  $Q \setminus \{q\} \in L_0$ , and also double it, adding a new join-irreducible element  $w$ . The resulting lattice  $L$  is the closure lattice of a standard closure system on its set of join-irreducible elements  $S = \{q_1, \dots, q_n, z_1, \dots, z_k, w\}$ . Due to the nature of the construction,  $L$  is a bounded lattice, in particular, the standard closure system  $\langle S, \phi \rangle$  corresponding to  $L$  is without  $D$ -cycles.

It is straightforward to verify that  $C = \{q_1, \dots, q_n, z_1, \dots, z_k\}$  is a (unique) critical set of this closure system with  $\phi(C) = S$ . Hence, every optimum basis for  $\langle S, \phi \rangle$  should have an implication  $U \rightarrow w$ , where  $U \subseteq C$  is a minimal generator for  $C$ .

Also,  $Q$  is a unique  $\ll$ -minimal representation of  $S$ . Hence, every  $W \subseteq S$  with  $\phi(W) = S$  should satisfy  $Q \ll W$ . Since it should hold for  $U$ , we should have for every  $q \in Q$  some  $u \in U$  such that  $u \geq_\phi q$ , or simply  $u \geq q$  in  $L$ .

Every element in  $U$  can be interpreted as an element of family  $\mathcal{Q}^*$ : if  $u = q_i$ , then it is  $\{q_i\}$ , and if  $u = z_i$ , then it is  $Q_i = \phi(z_i) \cap Q$ . Thus,  $U$  can be interpreted as a covering subfamily of  $\mathcal{Q}^*$ .

If we find a minimal generator  $U$  for  $C$ , then it will serve as a solution to the instance  $(Q, \mathcal{Q}^*)$  of the set cover problem. □

**Corollary 74.** *The problem of finding a basis  $\Sigma$ , for a closure system without  $D$ -cycles (and even closure systems with bounded closure lattice), which reaches minimum in  $s_L(\Sigma^{nb})$ , is NP-complete.*

*Proof.* Suppose, we can find a polynomial algorithm of building such a basis for a closure systems without  $D$ -cycles, given its  $E$ -basis. Then the set cover problem can be effectively solved as well. Indeed, start from any instance  $(Q, \mathcal{Q})$  of the set cover problem. If it is trivial, then it is solved effectively. If it is not trivial, then we can effectively write the  $E$ -basis of a bounded lattice constructed in Lemma 73 from  $(Q, \mathcal{Q})$ . Since, according to assumption, this allows us to write effectively an optimum basis for this closure system, the solution to the given set cover problem can be recovered from one of its implications in polynomial time. This contradicts the NP-completeness of the set cover problem. □

Now we turn to second problem on our list.

**Lemma 75.** *Suppose  $(Q, \mathcal{Q})$  is a non-trivial instance of the set cover problem. One can effectively find a closure system, whose closure lattice is bounded, such that a particular binary implication  $w \rightarrow B$  of its optimum basis translates into a solution of this set cover problem.*

*Proof.* As in the proof of Lemma 73, we assume that  $Q = \{q_1, \dots, q_n\}$ ,  $\mathcal{Q} = \{Q_i : i \in I\}$  is some set cover of  $Q$ , and  $\mathcal{Q}$  has  $k$  elements of cardinality  $> 1$ . We also can replace the instance of set cover problem by  $(Q, \mathcal{Q}^*)$  with  $\mathcal{Q}^* = \mathcal{Q} \cup \{\{q_i\} : i \leq n\}$ .

We start from  $L_0 = \mathbf{2}^Q \cup \{w\}$ , where  $w > x$ , for all  $x \in \mathbf{2}^Q$ . For each  $Q_i \in \mathcal{Q}$ ,  $|Q_i| > 1$ , double element  $b_i = Q_i \in L_0$ ,  $i \leq k$ , adding new join-irreducible elements  $z_i$ .

The resulting bounded lattice  $L$  is built on the set of join-irreducible elements  $S = \{q_1, \dots, q_n, z_1, \dots, z_k, w\}$ . Let  $\langle S, \phi \rangle$  be a closure system with  $L = \text{Cl}(S, \phi)$ .

According to Theorem 20, every (right-side) optimum basis should have an implication  $w_0 \rightarrow B$ , where  $\phi(B) = S \setminus \{w_0\} = \phi(\{w_0\}) \setminus \{w_0\}$  and  $|B|$  is minimal among subsets of  $S \setminus \{w_0\}$  with this property. Recall that  $Q$  is a unique  $\ll$ -minimal representation of  $S \setminus \{w_0\}$ . Hence,  $Q \ll B$ . This means that for every  $q \in Q$  there exists  $b \in B$  such that  $b \geq_\phi q$ , or simply  $b \geq q$  in  $L$ . Thus,  $B$  can be interpreted as a set cover for  $Q$ , where each  $b \in B$  is thought as an element of covering family  $\mathcal{Q}^*$ . The rest of the argument is similar to the proof of Lemma 73.  $\square$

**Corollary 76.** *The problem of finding a basis  $\Sigma$ , for systems without  $D$ -cycles (and even in closure systems with bounded closure lattices), that reaches minimum in  $s_R(\Sigma^b)$ , is NP-complete. In particular, the problem of finding an optimum basis for such systems is NP-complete.*

*Proof.* Indeed, starting from an instance of the cover set problem, build a finite bounded lattice  $L$  from Lemma 75. If the (right side) optimum basis for  $L$  can be effectively found, then taking a binary implication corresponding to the top element of  $L$ , we would get a solution to the set cover problem. The last sentence follows from Corollaries 11 and 21.  $\square$

*Acknowledgments.* The first draft of the paper was written during the first author's visit to University of Hawai'i, supported by AWM-NSF Travel Mentor Grant. The warm and encouraging atmosphere of the Department of Mathematics of UofH is highly appreciated. The work on the paper was also inspired by the communication with V. Duquenne and E. Boros.

#### REFERENCES

- [1] K. Adaricheva, *Optimum basis of finite convex geometry*, available at <http://arxiv.org/abs/1205.3236>
- [2] K. Adaricheva, J.B. Nation and R. Rand, *Ordered direct implicational basis of a finite closure system*, Disc. Appl. Math. **161**(2013), 707–723.
- [3] K. Adaricheva and J.B. Nation, *On closure systems with unique critical sets*, a manuscript.
- [4] G. Ausiello, A. D'Atri and D. Saccá, *Minimal representation of directed hypergraphs*, Journal on computing **15** (1986), 418–431.
- [5] K. Bertet and B. Monjardet, *The multiple facets of the canonical direct unit implicational basis*, Theor. Comp. Sci. **411**(2010), 2155–2166.
- [6] E. Boros, Horn Functions, in Boolean Functions: Theory, Algorithms, and Applications (Y. Crama and P.L. Hammer eds.), Cambridge University Press, New York, 2011.
- [7] E. Boros, O. Čepek, A. Kogan and P. Kucěra, *Exclusive and essential sets of implicates of Boolean functions*, Disc. Appl. Math. **158** (2010), 81–96.
- [8] E. Boros, O. Čepek, A. Kogan and P. Kucěra, *A subclass of Horn CNFs optimally compressible in polynomial time*, Annals of Math. and Artif. Intelligence **57** (2012), 249–291.
- [9] N. Caspard and B. Monjardet, *The lattices of closure systems, closure operators, and implicational systems on a finite set: a survey*, Disc. Appl. Math. **127** (2003), 241–269.
- [10] A. Day, *Characterizations of finite lattices that are bounded homomorphic images or sublattices of free lattices*, Canad. J. Math. **31** (1979), 69–78.
- [11] A. Day, *The lattice theory of functional dependencies and normal decompositions*, Int.J.Alg. Comp. **2**(1992), 409–431.
- [12] V. Duquenne, *On the core of finite lattices*, Discr. Math. **88** (1991), 133–147.
- [13] J. L. Guigues and V. Duquenne, *Familles minimales d'implications informatives résultant d'une tables de données binaires*, Math. Sci. Hum. **95** (1986), 5–18.
- [14] R. Freese, J. Ježek and J. B. Nation, *Free Lattices*, Mathematical Surveys and Monographs **42**, Amer. Math. Soc., Providence, 1995.

- [15] C. Herrmann and M. Wild, *A polynomial algorithm for testing congruence modularity*, Int. J. Alg. Comp. **6** (1996), 379–387.
- [16] P.L. Hammer and A. Kogan, *Quasi-acyclic propositional Horn knowledge bases: optimal compression*, IEEE Transactions on knowledge and data engineering **7** (1995), 751–762.
- [17] B. Jónsson and J. E. Kiefer, *Finite sublattices of a free lattice*, Canad. J. Math. **14** (1962), 487–497.
- [18] R. Karp, *Reducibility among combinatorial problems*, in Complexity of computer computations (R.E. Miller and J.W. Thatcher, eds.), Plenum, NY, 1972, 85–103.
- [19] D. Maier, *Minimum covers in the relational database model*, JACM **27** (1980), 664–674.
- [20] M. Wild, *A theory of finite closure spaces based on implications*, Adv. Math. **108**(1994), 118–139.
- [21] M. Wild, *Optimal implicational bases for finite modular lattices*, Quaestiones Mathematicae **23** (2000), 153–161.

DEPARTMENT OF MATHEMATICAL SCIENCES, YESHIVA UNIVERSITY, 245 LEXINGTON AVE., NEW YORK, NY 10016, USA

*E-mail address:* `adariche@yu.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HAWAII, HONOLULU, HI 96822, USA

*E-mail address:* `jb@math.hawaii.edu`