The prime stems of rooted circuits of closure spaces and minimum implicational bases

Kenji Kashiwabara and Masataka Nakamura

Department of System Sciences, University of Tokyo
Komaba 3-8-1, Meguro, Tokyo 153-8902, JAPAN
kashiwa@idea.c.u-tokyo.ac.jp   *nakamura@klee.c.u-tokyo.ac.jp*

Abstract

A rooted circuit is firstly introduced for convex geometries (antimatroids). We generalize it for closure systems or equivalently for closure operators. A rooted circuit is a specific type of a pair (X, e) of a subset X, called a stem, and an element $e \notin X$, called a root. We introduce a notion called a 'prime stem,' which plays the key role in this article. Every prime stem is shown to be a pseudo-closed set of an implicational system. If the sizes of stems are all the same, the stems are all pseudo-closed sets, and they give rise to a canonical minimum implicational basis. For an affine convex geometry, the prime stems determine a canonical minimum basis, and furthermore gives rise to an optimal basis. A 'critical rooted circuit' is a special case of a rooted circuit defined for an antimatroid. As a precedence structure, 'critical rooted circuits' are necessary and sufficient to fix an antimatroid whereas critical rooted circuits are not necessarily sufficient to restore the original antimatroid as an implicational system. It is shown through an example.

Keywords: closure system, Armstrong's axioms, pseudo-closed set, convex geometry

1 Introduction

A family $\mathcal{K} \subseteq 2^E$ on a nonempty finite set E is a closure system if it contains E and closed under intersection. Closure systems appear in many areas such as implicational system (functional dependencies in relational database theory), formal concept analysis [7], knowledge space [5], logic [14], and so on. An implicational system is called a set of functional dependencies in relational database theory. As is well known, a closure system and a closure operator $\mu : 2^E \to 2^E$ are equivalent, and in one-to-one correspondence with each other. We shall call a pair (μ, E) a closure space.

For an element $e \in E$ and a subset $X \subseteq E - e$, a pair (X, e) is a rooted circuit when X is minimal with respect to inclusion relation under satisfying $e \in \mu(X)$. $X \subseteq E - e$ is called a *stem*, and $e \in E$ is called a *root*. A rooted circuit (X, e) is equivalent to the left-minimal implication $X \to e$ defined in Bertet and Monjardet [2]. A specific type of a stem, called a *prime stem*, is introduced. In short, a stem is a prime stem if its closure includes no boundary element. A prime stem plays a key role in this article.

Firstly, every prime stem is shown to be a pseudo-closed set (Theorem 11). In particular, if the stems of a closure space are all of the same size, they are all pseudo-closed sets, and give rise to the canonical implicational basis (Theorem 15).

Convex geometries are special cases of closure spaces. In case of an affine convex geometry, the prime stems give rise to the canonical minimum basis (Theorem 18), and even more an optimal basis (Theorem 20).

We reveal the difference between the set of critical rooted circuits of an antimatroid (convex geometry) and an implicational basis. The so-called critical rooted circuits are necessary and sufficient to determine the original antimatroid (convex geometry) as a 'transversal' precedence structure, but not enough to fix the convex geometry (antimatroid) as an implicational basis. It will be shown through an example in Section 5.

2 Preliminaries

2.1 Closure spaces and rooted circuits

Let E be a nonempty finite set. A map $\mu : 2^E \to 2^E$ is a *closure operator* if it satisfies

$$(1)\ A \subseteq \mu(A), \qquad (2)\ A \subseteq B \implies \mu(A) \subseteq \mu(B), \qquad (3)\ \mu(\mu(A)) = \mu(A).$$

We shall call the pair (μ, E) a *closure space*.

$\mathcal{K} \subseteq 2^E$ is a *closure system* (or a *Moore family*) if it satisfies the following.

$$(1)\ E \in \mathcal{K}, \qquad (2)\ X, Y \in \mathcal{K} \implies X \cap Y \in \mathcal{K}.$$

We call an element of $\mathcal{K}$ a *closed set*. Since $\mathcal{K}$ is a finite lower semilattice having the maximum element E, it is actually a lattice with respect inclusion relation.

For a closure system $\mathcal{K}$, $\mu : 2^E \to 2^E$ below is a closure operator.

$$\mu(A) = \bigcap_{X \in \mathcal{K}, A \subseteq X} X. \tag{1}$$

Conversely, $\mathcal{K} = \{X \subseteq E : \mu(X) = X\}$ holds. Hence there is a one-to-one correspondence between the closure systems and the closure operators.

A closure system $\mathcal{K}$ arising from a closure operator which enjoys the exchange property is said to be a *matroid*. When a closure operator μ satisfies the anti-exchange property, the closure system (μ, E) is a *convex geometry*.

We can define independent sets and dependent sets for a closure space (μ, E) in the following way. A map $\mathrm{ex} : 2^E \to 2^E$ defined by

$$\mathrm{ex}(A) = \{x \in A : x \notin \mu(A - x)\} = \{x \in A : \mu(A - x) \neq \mu(A)\} \qquad (A \subseteq E) \qquad (2)$$

is said to be the *extreme-point operator*. An element in $\mathrm{ex}(A)$ is called an *extreme point* of A, and otherwise it is a *non-extreme point* of A. For $A \subseteq E$, if $\mathrm{ex}(A) = A$, then A is said to be an *independent set*. In other words, a set A is independent if it contains no non-extreme point. Any subset of an independent set is again an independent set [15]. A set which is not independent is a *dependent set*. A minimal dependent set is called a *circuit*. For a matroid, the independent sets and the circuits defined here agree with the ordinary definitions of matroid theory.

A pair (X, e) of a set X and an element e is a *rooted set* if $e \notin X$. In a closure space (μ, E), a rooted set (X, e) is a *rooted circuit* if $X \subseteq E - e$ is a minimal set satisfying $e \in \mu(X)$ with respect to inclusion relation. X is called the *stem* and e is called the *root*. If C is a circuit of a closure space, C contains at least one non-extreme point, say $e \in C$. Then $(C - e, e)$ is necessarily a rooted circuit of the closure space.

Let $\mathcal{C}$ denote the set of the rooted circuits of a closure space (μ, E), and $Stem(\mu)$ be the collection of the stems of rooted circuits, i.e. $Stem(\mu) = \{X : (X, e) \in \mathcal{C}$ for some $e \in E\}$. For a stem X, put $int(X) = \{f \in E : (X, f) \in \mathcal{C}\}$ and $bd(X) = \mu(X) - (X \cup int(X))$. Clearly $\mu(X) = X \cup int(X) \cup bd(X)$ is a partition. If $bd(X) = \emptyset$, we shall call X a *prime stem*.

For a matroid M, take any circuit C of M and any element e in C. Then $(C - \{e\}, e)$ is a rooted circuit, and vice versa. For a graphic matroid, $C - \{e\}$ is a prime stem if and only if C is a chordless cycle. In case of an affine convex geometry in an affine space, a stem is the set of vertices of a simplex. If the convex closure of the stem includes no boundary point except the vertices, it is a prime stem.

Compared with our notation of a rooted circuit (X, e), in the former literature such as [4, 11], they call $(X \cup e, e)$ a rooted circuit, and an 'independent set' in this article is called a 'free set' in them. We shall call a 'closed independent set' a free set [15]. The independent sets of a closure space as well as the free sets constitute a simplicial complex, in other words, an independence system.

2.2 Implicational systems

We shall describe that an implicational system determines a closure system, and hence a closure space.

Let E be a nonempty finite set. An *implicational system* S on E is a relation on 2^E, that is, it is a collection of ordered pairs $(A, B) \in 2^E \times 2^E$. When $(A, B) \in S$, we write $A \to B \in S$, and call it an *implication*. For an implication $A \to B \in S$, A is called a *premise*, and B is called a *conclusion*.

A set $X \subseteq E$ is said to *respect* $A \to B$ if $A \subseteq X$ implies $B \subseteq X$, or equivalently if either $A \nsubseteq X$ or $B \subseteq X$. For an arbitrary implicational system S, we shall define

$$\mathcal{K}(S) = \{X \subseteq E : X \text{ respects every implication } A \to B \in S\}.$$

That is, a set $Y \subseteq E$ respects all the implications in S if and only if $Y \in \mathcal{K}(S)$. Let us denote the closure operator of $\mathcal{K}(S)$ by μ_S. For a subset A of E, let $\mathcal{K}(S)_A = \{X \in \mathcal{K}(S) : A \subseteq X\}$. For an implicational system S and implication $A \to B$, if $S' = S \cup \{A \to B\}$ is equivalent to S, we say that $A \to B$ is *compatible* with S.

Theorem 1 (folklore). For an arbitrary implicational system S, $\mathcal{K}(S)$ is a closure system.

An implicational system S is said to be a *generating set* of a closure system $\mathcal{K}(S)$. If two implicational systems give rise to the same closure system, they are said to be *equivalent*.

If an implication $A \to B$ can be derived from an implicational system S by applying Armstrong's axioms, we say that $A \to B$ is *induced* from S. Armstrong's axioms are known to be robust in the sense that $C \to D$ is induced from S if and only if $C \to D$ is compatible with S. (Bertet and Monjardet [2] as well as Caspard and Monjardet [3] is a rich source for closure systems and implicational systems.)

If an implicational system S is invariant under applying the three inference rules (A1), (A2) and (A3) below, we shall call it a *full implicational system*.

(A1) If $A \supseteq B$, then $A \to B \in S$.   (reflexivity)

(A2) If $A \to B \in S$ and $C \to D \in S$, then $A \cup C \to B \cup D \in S$.   (augmentation)

(A3) If $A \to B \in S$ and $B \to C \in S$, then $A \to C \in S$.   (transitivity)

These inference rules are called *Armstrong's axioms*. The following is an extension of (A3). The triple of (A1), (A2) and (A3') is equivalent to Armstrong's axioms.

(A3') If $A \to B \in S$ and $B \cup D \to C \in S$, then $A \cup D \to C \in S$.   (overlap)

For an implicational system S on E, suppose $\mathcal{K}(S)$ and μ_S to be the associated closure system and the closure operator, respectively. It is easy to see that for $A, B \subseteq E$, every closed set $X \in \mathcal{K}(S)$ respects $A \to B$ if and only if $B \subseteq \mu_S(A)$. Since the set of Armstrong's axioms is known to be sound and complete [13], the full implicational system equivalent to S is $\bar{S} = \{A \to B : A, B \subseteq E, \ B \subseteq \mu_S(A)\}$. Hence

Proposition 2. For an implicational system S on E and $A, B \subseteq E$, the following are equivalent.

(1) $A \to B$ is compatible with S.

(2) $B \subseteq \mu_S(A)$.

(3) $\mathcal{K}(S)_A \subseteq \mathcal{K}(S)_B$.

An implicational system giving rise to a closure space (μ, E) is called a *generating implicational system* of (μ, E). An implicational system S is *non-redundant* if for any $X \to Y \in S$, $S' = S \setminus \{X \to Y\}$ is not equivalent to S any more. An *implicational basis*

of a closure space (μ, E) is a generating implicational system which is non-redundant. An implicational basis S is *minimum* if $|S| \leqslant |S'|$ for any S' that is equivalent to S, and S is called *optimal* if $s(S) \leqslant s(S')$ for any S' which is equivalent to S where $s(S) = \sum_{X \to Y \in S}(|X| + |Y|)$. It is stated in [12] that a minimum implicational basis can be found in time $O(np)$, where $n = \Sigma_{X \to Y \in S}(|X| + |Y|)$ and $p = |S|$, and an optimal basis problem is NP-complete.

For any $A \subseteq E$, we put

$$A^\circ = A \cup \bigcup \{\mu(X) : X \subsetneq A, \mu(X) \subsetneq \mu(A)\}, \tag{3}$$

$$A^\bullet = A^\circ \cup A^{\circ\circ} \cup A^{\circ\circ\circ} \cup \cdots . \tag{4}$$

Then $A \mapsto A^\bullet$ is a closure operator, and $A^\bullet \subseteq \mu(A)$ holds [17].

$W \subseteq E$ is *quasi-closed* if $W = W^\bullet$ and $W \neq \mu(W)$. A quasi-closed set W is *pseudo-closed* if W is a minimal quasi-closed set with respect to inclusion among the quasi-closed sets W' such that $\mu(W) = \mu(W')$. Then $S_{pseu} = \{P \to (\mu(P) - P) : P \text{ is a pseudo-closed set}\}$ is actually an implicational basis of (μ, E) [3, 8]. This is called the *canonical basis* or the *Duquenne-Guigues basis*.

Example 3. Let us show two examples in which quasi-closed sets are not necessarily pseudo-closed. Consider a uniform matroid $U_{1,3}$. Each set of cardinality two is quasi-closed, but not a pseudo-closed set whereas every singleton is a pseudo-closed set. Similarly, in a uniform matroid $U_{2,4}$, each set of size three is quasi-closed, but not pseudo-closed. All the sets of size two are pseudo-closed sets.

Theorem 4 ([2, 8, 16, 17]). For a closure space (μ, E), the following hold.

(1) The canonical basis $S_{pseu} = \{P \to (\mu(P) - P) : P \text{ is a pseudo-closed set}\}$ is a minimum basis.

(2) An optimal basis is always a minimum basis.

(3) For each pseudo-closed set P, every optimal basis contains an implication $A_P \to B_P$ such that $A_P \subseteq P$, $\mu(A_P) = \mu(P)$ and $|A_P| = \min\{|A| : A \subseteq P, \mu(A) = \mu(P)\} = \min\{|A| : X \subseteq P, A^\bullet = P\}$.

3   Stems of closure spaces and pseudo-closed sets

The rooted circuits of a closure space give rise to an implicational generating set of the original closure system. We write $X \to e$ to denote $X \to \{e\}$ for simplicity.

Lemma 5. Let X be an independent set of a closure space (μ, E). Then X is quasi-closed if and only if it is pseudo-closed.

Proof. A pseudo-closed set is necessarily a quasi-closed set. Conversely, suppose X is quasi-closed, but not pseudo-closed. Then there is a proper subset $Y \subsetneq X$ such that Y is pseudo-closed and $\mu(Y) = \mu(X)$. Take an element $e \in X - Y$. Since X is independent, $e \notin \mu(X - e)$ while $\mu(Y) \subseteq \mu(X - e)$. Hence $e \notin \mu(Y)$. This contradicts the assumption that $e \in \mu(X) = \mu(Y)$. □

Lemma 6. Every stem of a closure space is an independent set.

Proof. Let $X \in Stem(\mu)$ be a stem. If X contains a non-extreme point $f \in X$, then $\mu(X) = \mu(X - f)$. This contradicts the minimality of stems. □

Lemma 7. In a closure space (μ, E), if A is an independent set, and $A' \subsetneq A$, then $\mu(A') \subsetneq \mu(A)$.

Proof. By definition, for any $f \in A - A'$, $f \notin \mu(A - f)$ must hold. Hence $\mu(A') \subsetneq \mu(A)$ follows. □

Lemma 8. For an independent set X of a closure space (μ, E), if there exists $(Y, f) \in \mathcal{C}$ such that $Y \subsetneq X$, then $f \notin X$.

Proof. Suppose contrarily $f \in X$. Set $X' = X - (Y \cup f)$. Then $X = X' \cup Y \cup f$. From the assumption, $Y \cup f \subseteq \mu(Y)$. Hence $X \subseteq \mu(X') \cup \mu(Y)$. Now $\mu(X) \supseteq \mu(X' \cup Y) \supseteq \mu(X') \cup \mu(Y) \supseteq X$. Thus $\mu(X) = \mu(X' \cup Y) = \mu(X - f)$ holds. This implies that X is not independent, a contradiction. □

Lemma 9. For a stem $X \in Stem(\mu)$ of a closure space (μ, E), there is no stem $Y \in Stem(\mu)$ such that $Y \subsetneq X$ if and only if $X = X°$. Hence a minimal stem among $Stem(\mu)$ with respect to inclusion relation is a pseudo-closed set.

Proof. (if part) Suppose there exists $Y \subsetneq X$ with $(Y, f) \in \mathcal{C}$. By Lemma 8, $f \notin X$. If $\mu(Y) = \mu(X)$, it contradicts the minimality of the stem X. Hence $\mu(Y) \subsetneq \mu(X)$, which implies $\mu(Y) \subseteq X°$ and $f \in X°$. Thus $X° \supseteq X \cup f \supsetneq X$, and $X° \neq X$.
(only if part) Suppose contrarily $X \neq X°$. Take an element $f \in X° - X$. By definition, there exists a rooted circuit $(Y, f) \in \mathcal{C}$ such that $Y \subsetneq X$ and $\mu(Y) \subsetneq \mu(X)$. This contradicts the assumption. Hence $X = X°$ occurs.

For the last statement, from the above proof, a minimal stem X is a quasi-closed set since $\mu(X) \neq X$ holds by definition. On the other hand, the minimality of X among $Stem(\mu)$ implies the minimality of $\mu(X)$ from Lemmas 5 and 6. □

Proposition 10. For a stem $X \in Stem(\mu)$ of a closure space (μ, E), the following are equivalent.

    (1) X is a prime stem.

    (2) X is minimal among $Stem(\mu)$ with respect to inclusion relation.

    (3) X is a pseudo-closed set.

    (4) X is a quasi-closed set.

Proof. $(1) \Rightarrow (2)$: Suppose contrarily that X is not minimal. Then there exists a rooted circuit $(Y, f) \in \mathcal{C}$ such that $Y \subsetneq X$. By Lemma 6, X is an independent set. Hence $f \notin X$ follows from Lemma 8. Thus $f \notin int(X)$ and $f \in \mu(X)$ holds, which implies $\mu(X) \neq X \cup int(X)$, a contradiction.

$(2) \Rightarrow (3)$ follows from Lemma 9.

$(3) \Rightarrow (4)$: This is immediate from the definitions.

$(4) \Rightarrow (1)$: X is a quasi-closed set, but not a prime stem. Then there exists an element $f \in bd(X)$ such that (X, f) is not a rooted circuit. Hence there exists a subset $Y \subsetneq X$ such that $f \in \mu(Y)$. Since X and Y are independent sets, Lemma 7 yields $\mu(Y) \subsetneq \mu(X)$. Now $f \in X°$ and $f \notin X$ hold. This implies $X \neq X°$, and X is not a quasi-closed, a contradiction. □

Theorem 11. For a closure space (μ, E), every prime stem is a pseudo-closed set.

Proof. This directly follows from Proposition 10. □

For an implicational system S on a nonempty finite set E, let $\mathcal{K}(S)$ be the associated closure system, and (μ_S, E) be the corresponding closure space. For an element e in E, let $\mathcal{K}(S)_{\{e\}}$ denote $\{X \in \mathcal{K} : e \in X\}$. Then

Proposition 12. Suppose that $\mathcal{K}(S)$ contains the empty set $\emptyset$, and $\mathcal{K}(S)_{\{a\}}$ is a subset of $\mathcal{K}(S)_{\{b\}}$ for two distinct elements $a, b \in E$. Then $\{a\} \to \{b\}$ is an implication compatible with S, and $\{a\}$ is a pseudo-closed set of the closure space (μ_S, E).

Proof. The first statement and $\mu_S(\{a\}) \supseteq \{a, b\}$ follow from Proposition 2. The assumption implies $\mu_S(\emptyset) = \emptyset$. Hence $\{a\}° = \{a\} \cup \{\mu_S(A) : A \subsetneq \{a\}, \mu_S(A) \subsetneq \mu_S(\{a\})\}$ is equal to $\{a\}$, and so $\{a\}$ is a pseudo-closed set. □

Example 13. The converse of Theorem 11 is not true in general, i.e. a pseudo-closed set is not necessarily a (prime) stem. Let us consider an implicational system $S = \{\{1, 4\} \to \{2\}, \{3, 4\} \to \{1\}\}$ on $E = \{1, 2, 3, 4\}$. Fig. 1 describes the closure system arising from S. Now $\{4\}$ and $\{1, 4\}$ are all the pseudo-closed sets whereas $\{4\}$ and $\{3, 4\}$ are all the stems. So $\{1, 4\}$ is a pseudo-closed set, but not a stem. The canonical implicational basis is $\{\{4\} \to \{2\}, \{1, 4\} \to \{2\}\}$. In this example, $K(S)_4 \subsetneq K(S)_2$ holds. Proposition 12 applies to this case. In fact $\{4\}$ is a pseudo-closed set, and $\{4\} \to \{2\}$ is a compatible implication.

Theorem 14. For a closure space (μ, E), if every pseudo-closed set X is a prime stem and $|\mu(X) - X| = 1$, then $S_{prime} = \{X \to (\mu(X) - X) : X \text{ is a prime stem}\}$ is the unique optimal basis.

Proof. First note that every prime stem is a pseudo-closed set by Theorem 11. Hence a set is pseudo-closed if and only if it is a prime stem. From (3) of Theorem 4, for the premise A of any implication $A \to B$ of an arbitrary optimal basis, there exists a pseudo-closed set P such that $A \subseteq P$ and $\mu(A) = \mu(P)$. From the remark above, P is a prime
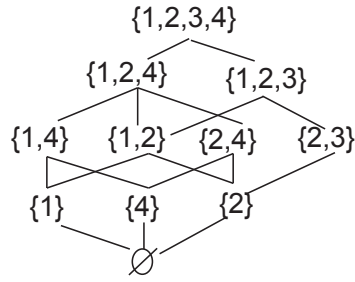
Figure 1: A closure system $\mathcal{K}(S)$ of Example 13

stem. Since a stem is necessarily an independent set by Lemma 6, a prime stem P is an independent set. Hence from Lemma 6, it is a minimal generating set of $\mu(P)$. This implies $A = P$, and that every premise of an implication in S_{prime} is of the minimum size. By assumption, the conclusions of implications are all singletons, so that they are of the minimum size. Thus S_{prime} is shown to be an optimal basis. $\qquad\square$

Theorem 15. Let (μ, E) be a closure space, and suppose that the sizes of the stems are all the same. Then for a subset $X \subseteq E$, the following are equivalent.

(1) X is a stem.

(2) X is a prime stem.

(3) X is a pseudo-closed set.

Proof. (1)$\Rightarrow$(2) : We shall show that every stem is a prime stem. If a stem $X \in Stem(\mu)$ is not prime, $bd(X) \neq \emptyset$. Take any $f \in bd(X)$. Then there exists a subset $Y \subsetneq X$ such that $(Y, f) \in \mathcal{C}$ is a rooted circuit. This contradicts the assumption that the stem sizes are all the same. Hence X is a prime stem.
(2)$\Rightarrow$(3) : By Theorem 11, a prime stem X is a pseudo-closed set.
(3)$\Rightarrow$(1) : From the assumption, $X = X^\circ$ and $X \neq \mu(X)$ hold. Take an element e in $\mu(X) - X$. Then there is a subset Y of X such that (Y, e) is a rooted circuit. If $X = Y$, then the proof is done. Hence suppose Y is a proper subset of X. If $\mu(Y) \subsetneq \mu(X)$ holds, then e is in X°, and $X \neq X^\circ$ happens, which contradicts the assumption that X is a pseudo-closed set. Hence $\mu(Y) = \mu(X)$ holds.

We shall show Y is a quasi-closed set. Now suppose contrarily $Y \neq Y^\circ$. By the definition of a quasi-closed set, there is a proper subset A of Y such that $\mu(A) \subsetneq \mu(Y)$ and $\mu(A) \not\subseteq Y$. Then there exists an element f in $\mu(A) \backslash Y$ and a rooted circuit (A', f) such that $A' \subseteq A$. A' and Y are both stems, and A' is a proper subset of Y, which contradicts the assumption that the stem sizes are all the same, and so $Y = Y^\circ$ is established. Since f is in $\mu(Y) - Y$, $\mu(Y) \neq Y$ is obvious So Y is a quasi-closed set. By the minimality of a pseudo-closed set, $X = Y$ should occur because of $\mu(Y) = \mu(X)$. Now X is proven to be a stem. $\qquad\square$

The stem sizes are all the same for some classes of convex geometries, which are presented in Example 17.

4 Convex geometries

In this article we say that a closure space (μ, E) is a *convex geometry* on E if μ meets the *anti-exchange property* (Ant).

(Ant) If $x \neq y$, $x, y \notin \mu(A)$ and $y \in \mu(A \cup x)$, then $x \notin \mu(A \cup y)$. ($x, y \in E, A \subseteq E$)

The closure system $\mathcal{K} = \{X : \mu(X) = X, X \subseteq E\}$ is called a convex geometry. $\mathcal{A} = \{E - X : X \in \mathcal{K}\}$ is called an *antimatroid*. For more details of convex geometries and antimatroids, we refer to [6, 11].

By our definition, a circuit includes at least one non-extreme point. In case of a convex geometry, a non-extreme element e in a circuit C is unique, and it gives a unique partition $C = (C - e) \cup e$. $(C - e, e)$ is defined to be a rooted circuit in [11]. This is the origin of rooted circuits. We have generalized it to a closure space in general. For a convex geometry, a rooted circuit presented in [11] and our definition are equivalent.

Proposition 16 ([15]). Let (μ, E) be a convex geometry. Then the following hold.

(1) If (X, e) is a rooted circuit, then $X \cup e$ is a circuit.

(2) Let C be a circuit, and e be the unique non-extreme element in C. Then $(C - e, e)$ is a rooted circuit.

Example 17. As a result of Theorem 15, for the convex geometries described below, a set is a pseudo-closed set if and only if it is a stem. For a poset P, we can define a convex geometry as $\mathcal{K} = \{X \subseteq P : \text{if } a, b \in X \text{ and } a \leqslant c \leqslant b, \text{ then } c \in X\}$ forms a convex geometry, called a *double shelling convex geometry* of a poset P. If $a < c < b$ in P, then $(\{a, b\}, c)$ is a rooted circuit, and vice versa. Hence the sizes of stems of a double shelling poset convex geometry are all two. The following are the classes of convex geometries for which the stem sizes are all two: *edge shelling convex geometry* of a tree, *node shelling convex geometry* of a tree, *simplicial shelling convex geometry* of a chordal graph, a convex geometry arising from the *edge shelling of a transitively closed digraph*. For more details, we refer to [6, 11].

We shall present another important class of convex geometries. Let $E \subseteq \mathbb{R}^n$ be a nonempty finite set, which is called an *affine point configuration*, and *conv.hull* denote the ordinary convex hull operator in $\mathbb{R}^n$. Then

$$\mu_E(A) = conv.hull(A) \cap E \qquad (A \subseteq E)$$

is a closure operator on E satisfying the anti-exchange property, and we shall call the convex geometry arising from μ_E the *affine convex geometry* on E. Although the stems can be of different sizes for an affine convex geometry, we can determine the canonical basis.

Theorem 18. Let (μ, E) be an affine convex geometry on a nonempty finite set $E \subseteq \mathbb{R}^n$. Then for $X \subseteq E$, X is a prime stem if and only if it is a pseudo-closed set. Hence

$$S = \{X \to (\mu(X) - X) : X \text{ is a prime stem}\}$$

is the canonical minimum basis of (μ, E).

Proof. (only if part) From Theorem 11, a prime stem is necessarily a pseudo-closed set. (if part) Suppose that X is a pseudo-closed set. Let $P = conv.hull(X)$ be the polytope of the convex hull of X, and V be the set of the vertices of P. We shall show that P is a simplex. If it is not a simplex, P can be decomposed to the union of multiple simplexes, say $Y_1, \ldots, Y_k$ ($k \geqslant 2$). Then $V = Y_1 \cup \cdots \cup Y_k$, and $\mu(V) = \mu(Y_1) \cup \cdots \cup \mu(Y_k) = \mu(X)$. Since $X \neq \mu(X)$ from the assumption, there exists an element $f \in \mu(X) - X$. Since $f \in \mu(X) = \cup_{i=1}^{k} \mu(Y_i)$, $f \in \mu(Y_j)$ for some j. Then $f \in X^\circ$ and $f \notin X$, which leads to $X \neq X^\circ$ and $X \neq X^\bullet$. This contradicts the assumption. Hence P is a simplex. If $bd(X) \neq \emptyset$, $X \neq X^\circ$ holds, contradicting the assumption. Hence $bd(X) = \emptyset$, and $\mu(X) = X \cup int(X)$, while $\mu(V) = \mu(X)$. Then $V \subseteq X$ is a quasi-closed set and $\mu(V) = \mu(X)$. Hence X should be equal to V since X is a pseudo-closed set. Now X is proved to be a prime stem. $\qquad\square$

An equivalent result to Theorem 18 is obtained in [1]. It is worthwhile to compare Theorem 18 and the problem mentioned in Example 4 of [17].

We shall present an optimal implicational basis of an affine convex geometry. Let (μ, E) be the affine convex geometry of E in $\mathbb{R}^n$. For each prime stem $X \in Stem(\mu)$, choose arbitrarily an element in $int(X)$, which we denote by e_X. Then

Lemma 19. Let $S_P = \{X \to e_X : X \text{ is a prime stem}\}$ be the set of implications, and S_A be the collection of all the implications induced from S_P by applying Armstrong's axioms. Then for every rooted circuit $(X, e) \in \mathcal{C}$, $X \to e$ is in S_A.

Proof. For any $(X, e) \in \mathcal{C}$, we shall show that $X \to e$ is in S_A. We use induction on $d = |X|$. Since it is an affine convex geometry, d is at least two.
(1) Suppose $d = 2$: For a given $(X, e) \in \mathcal{C}$, let $k = |\mu(X)|$. We use induction on k. By definition, k is at least 3. If $k = 3$, $|int(X)| = 1$ and $e = e_X$ should occur. Hence $X \to e \in S_A$ holds.

Suppose $k \geqslant 4$. Let $X = \{a, b\}$. Then k elements, $a = a_1, a_2, \ldots, a_k = b$, are on a line in the affine space in this order. By definition, there exists $2 \leqslant j \leqslant k-1$ such that $a_j = e_X$. If $e = a_j = e_X$, the proof is done. Otherwise $e = a_m$, and either $2 \leqslant m \leqslant j-1$ or $j+1 \leqslant m \leqslant k-1$ occurs. In the former case, $(\{a, e_X\}, e)$ is in $\mathcal{C}$, and $\mu(\{a, e_X\}) = j < k = \mu(\{a, b\})$ holds. Hence by induction hypothesis, $\{a, e_X\} \to e$ is in S_A, while $\{a, b\} \to e_X$ is in S_A. The overlap rule gives $\{a, b\} \to e \in S_A$. In the latter case, the proof is completely similar. Hence $X \to e \in S_A$ is verified.
(2) Suppose $d \geqslant 3$: We shall use induction on $k = |\mu(X)|$. By definition, k is at least $d + 1$.

If $k = d + 1$, then $|int(X)| = 1$, and $int(X) = \{e\}$. Hence $e = e_X$ should hold, and the proof is done.

Suppose $k > d + 1$. Let $\{Y \subseteq X : |Y| = d - 1\} = \{Y_1, \ldots, Y_d\}$. Then $\mu(X) = \cup\{\mu(Y_i \cup e_X) : i = 1, \ldots, d\}$ since (μ, E) is an affine convex geometry. (This can be actually considered to be a simplicial decomposition of the simplex X.) Hence $e \in \mu(Y_j \cup e_X)$ for some $1 \leqslant j \leqslant d$. (In the affine space, $Y_j \cup e_X$ is the set of vertices of a simplex in $\mathbb{R}^n$.) By definition, there exists such a subset $W \subseteq Y_j \cup e_X$ that $(W, e) \in \mathcal{C}$.

If $|W| < d$, then by induction hypothesis on d, $W \to e \in S_A$ holds.

In case that $|W| = d$, it implies that $W = Y_j \cup e_X$ and $(Y_j \cup e_X, e)$ is in $\mathcal{C}$. Since $\mu(Y_j \cup e_X) < \mu(X)$, by induction hypothesis on k, $Y_j \cup e_X \to e$ is in S_A, while we have $X \to e_X \in S_A$. By the overlapping rule, we obtain $X \cup Y_j \to e \in S_A$. Since $Y_j \subset X$, this is equal to $X \to e \in S_A$. Now the proof is completed. $\qquad\square$

Theorem 20. For an affine convex geometry (μ, E), $S_P = \{X \to e_X : X$ is a prime stem$\}$ is an optimal implicational basis.

Proof. By Lemma 19, S_P and S_A are both implicational generating sets of (μ, E), and for every rooted circuit $(X, e) \in \mathcal{C}$, $X \to e$ is in S_A. Since X is a stem, it is an independent set by Lemma 6, and $X = \text{ex}(X)$ is the unique minimum generating set of $\mu(X)$ because (μ, E) is a convex geometry. Hence by (3) of Theorem 4, each premise of the implication in S_P is of the minimum cardinality. Furthermore, every conclusion of S_P is a singleton. Hence the conclusions of S_P is of the minimum size, and S_P is an optimal basis. $\qquad\square$

5 Critical rooted circuits of precedence structures and implicational bases

For the closure system $\mathcal{K} \subseteq 2^E$ of a convex geometry (μ, E), $\mathcal{A} = \{E - X : X \in \mathcal{K}\}$ is called an *antimatroid*. Originally, the notion of critical rooted circuits is defined for an antimatroid $(\mathcal{A}, E)$. Let $e \in E$, $X \subseteq E - e$, and B be the maximum element of $\mathcal{A}$ in $E - (X \cup e)$. In [11], (X, e) is defined to be a *critical rooted circuit* of an antimatroid $\mathcal{A}$ if $B \cup e \notin \mathcal{A}$ and $B \cup e \cup x \in \mathcal{A}$ for every $x \in X$. It is not immediate, but a critical rooted circuit is a rooted circuit of the closure operator of the convex geometry. This notion is made clearer when it is generalized to a closure space (μ, E) [15]. A rooted circuit (X, e) of (μ, E) is defined to be a *critical rooted circuit* if the closure $\mu(X)$ is minimum among the closures of the rooted circuits of the same root e with respect to inclusion relation. Actually, these two definitions are equal in case of a convex geometry which is a special case of closure space.

An antimatroid is first called a *shelling greedoid* in [9], and next a *shelling structure* in Korte and Lovász [10] which is constructed by an *alternative precedence structure*. There is another precedence structure hidden in the proof of Theorem 3.11 of Korte, Lovász and Schrader [11], which is called here a *transversal precedence structure*. We shall describe it explicitly.

A rooted clutter $\mathcal{F}$ is the collection of clutters $\mathcal{F}(e)$ in 2^{E-e} for $e \in E$. Taking a rooted clutter $\mathcal{F} = \{\mathcal{F}(e) : e \in E\}$ as a transversal precedence structure, an antimatroid language is determined recursively as follows: the null word ϵ is a feasible word. Suppose

$f_1 \cdots f_k$ to be a feasible word. $f_1 \cdots f_k f_{k+1}$ is a feasible word if $f_{k+1} \notin \{f_1, \ldots, f_k\}$ and $\{f_1, \ldots, f_k\} \cap X \neq \emptyset$ for every $X \in \mathcal{F}(f_{k+1})$. For a word $\alpha = a_1 a_2 \cdots a_k$, let $[\alpha]$ denote $\{a_1, a_2, \ldots, a_k\}$. Then $\mathcal{A} = \{[\alpha] : \alpha$ is a feasible word of $\mathcal{L}\}$ is an antimatroid.

Let $\mathcal{C}_c = \{\mathcal{C}_r(e) : e \in E\}$ be defined by $\mathcal{C}_c(e) = \{X : (X, e)$ is a critical rooted circuit of $\mathcal{A}\}$. As a transversal precedence structure, the critical rooted clutter $\mathcal{C}_c$ of an antimatroid $\mathcal{A}$ restores $\mathcal{A}$. Conversely, if a transversal precedence structure of a rooted clutter $\mathcal{F}$ gives rise to an antimatroid $\mathcal{A}$, then $\mathcal{F}$ necessarily includes $\mathcal{C}_c$, i.e. $\mathcal{C}_c(e) \subseteq \mathcal{F}(e)$ for each $e \in E$. That is, as a transversal precedence structure, the critical rooted circuits are necessary and sufficient to determine the original antimatroid. As an implicational generating set, the set of critical rooted circuits is not enough to reconstruct the original antimatroid (convex geometry). The implicational system $\{X \to e : e \in E, X \in \mathcal{C}_r(e)\}$ may not be an implicational generating set of the convex geometry. Example 21 below shows it.

Example 21. In an affine space, suppose that four points $1, 2, 3, 4$ are collinear on a line in this order. See Fig. 2. When applying Theorem 4, we know that $S = \{\{1, 3\} \to \{2\}, \{2, 4\} \to \{3\}, \{1, 4\} \to \{2, 3\}\}$ is the canonical basis, while Theorem 20 implies that $S' = \{\{1, 3\} \to \{2\}, \{2, 4\} \to \{3\}, \{1, 4\} \to \{2\}\}$ is an optimal basis. $\mathcal{C}_r = \{(\{1, 3\}, 2), (\{2, 4\}, 3)\}$ is the set of critical rooted circuits whereas $S_r = \{\{1, 3\} \to \{2\}, \{2, 4\} \to \{3\}\}$ is not an implicational generating set of the closure system of the affine convex geometry of Fig. 2.

Figure 2: Four collinear points

When taking the rooted cocircuits of an antimatroid as an alternative precedence structure, it restores the original antimatroid. For the details, we refer to [11].

Acknowledgements

References

[1] K. Adaricheva, and J. B. Nation. Ordered direct implicational basis of a finite closure system. *Disc. Appl. Math.*, in press.

[2] K. Bertet, and B. Monjardet. The multiple facets of the canonical direct implicational basis. *Theoretical Computer Science*, 411:2155–2166, 2010.

[3] N. Caspard, and B. Monjardet. The lattice of closure systems, closure operators, and implicational systems on a finite set: A survey. *Disc. Appl. Math.* 127:241–269, 2003.

[4] B. L. Dietrich. A Circuit Set Characterization of Antimatroids. *J. Combinatorial Theory Ser. B* 43:314–321, 1987.

[5] J-P. Doignon, and J-C. Falmagne. *Knowledge Spaces.* Springer, 1999.

[6] P. H. Edelman and R. E. Jamison. The theory of convex-geometries. *Geometricae dedicata* 19:247–270, 1985.

[7] B. Ganter, and R. Wille. *Formal Concept Analysis.* Springer, 1999.

[8] J. L. Guigues, and V. Duquenne. Families minimales d'implications informatives résultant d'une table dedonnées binaires. *Mathematiques & Sciences Humanites* 95:5–18, 1986.

[9] B. Korte, and L. Lovász. Mathematical structures underlying the greedy algorithms. In: F. Gécseg (ed.). *Foundations of Computation Theory*, volume 117 of *Lecture Notes in Computer Science*, pages 205–209. Springer, 1981.

[10] B. Korte, and L. Lovász. Shelling structure, convexity, and a happy end. In: B. Bollobas, ed. *Graph Theory and Combinatorics: Proceedings of the Cambridge Combinatorial Conference in Honour of Paul Erdős*, pages 219–232. Academic Press, 1984.

[11] B. Korte, L. Lovász, and R. Schrader. *Greedoids.* Springer, 1991.

[12] D. Maier. Minimum covers in the relational database model. *J. ACM* 27:664–674, 1980.

[13] D. Maier. *The Theory of Relational Database.* Computer Science Press, 1983.

[14] N. M. Martin, and S. Pollard Closure Spaces and Logic. Kluwer Academic Publisher, Dordrecht, 1996.

[15] M. Nakamura. On the NBC-complexes and β–invariants of abstract convex geometries. *Disc. Appl. Math.* 157:1799–1805, 2009.

[16] R. C. Shock. Computing the minimum cover of functional dependencies. *Information Processing Letters* 22:157–159, 1986.

[17] M. Wild. A theory of finite closure spaces based on implications. *Advances in Math.* 108:118–139, 1994.