Some Complexity Results about Essential Closed Sets

Felix Distel

Theoretical Computer Science, TU Dresden, Germany
felix@tcs.inf.tu-dresden.de

Abstract. We examine the enumeration problem for essential closed sets of a formal context. Essential closed sets are sets that can be written as the closure of a pseudo-intent. The results for enumeration of essential closed sets are similar to existing results for pseudo-intents, albeit some differences exist. For example, while it is possible to compute the lectically first pseudo-intent in polynomial time, we show that it is not possible to compute the lectically first essential closed set in polynomial time unless P = NP. This also proves that essential closed sets cannot be enumerated in the lectic order with polynomial delay unless P = NP. We also look at minimal essential closed sets and show that they cannot be enumerated in output polynomial time unless P = NP.

1 Introduction

The analysis of dependencies between attributes, so-called *implications*, is an important area of research within Formal Concept Analysis (FCA). Already in [7] it has been shown how a complete set of implications with minimal cardinality can be obtained from a formal context. This set is now commonly known as the *Duquenne-Guigues-Base* of a context. Since its discovery many results and algorithms in FCA, such as Attribute Exploration, have made use of the Duquenne-Guigues-Base.

Not surprisingly, a lot of effort has been directed at finding efficient algorithms to compute it. One of the earliest, and probably most well-known algorithms is *Next-Closure-Algorithm* [5]. It produces all concept intents and all pseudo-intents of a given formal context in a lexicographic order (called the *lectic order*). During the last decade, newer algorithms have been developed [9, 11].

It is known that the Duquenne-Guigues base cannot be computed in polynomial time in the size of the input, since the base itself can be exponentially large in the size of the input [8]. This leaves the question whether it can be enumerated in output-polynomial time. Until now, no output-polynomial algorithm has been found, and it is also not known whether such an algorithm exists. Recently, a lot of progress has been made with respect to this question. It has been shown that the implications from the Duquenne-Guigues Base cannot be enumerated in output-polynomial time unless the transversal hypergraph problem (cf. [4]) is in P [12, 13]. In [2] a connection between the boolean satisfiability problem

(SAT) and enumeration problems from FCA has been established. In particular, it has been shown using a reduction from SAT that the Duquenne-Guigues-Base cannot be enumerated with polynomial delay in the lectic order unless P = NP. A reduction from SAT has also been used in [1] to show that the problem of verifying whether a given set of attributes is a *pseudo-intent*, i. e. whether it occurs as the left-hand side of an implication in the Duquenne-Guigues Base, is coNP-complete. In the same paper it is also shown that pseudo-intents cannot be enumerated in the reverse lectical order with polynomial delay. Other works related to enumeration algorithms for pseudo-intents include [6] where optimizations based on hidden dependencies within the Duquenne-Guigues Base are considered. In [10] it is shown that the problem of counting pseudo-intents is #P-hard.

Previous work has mainly considered the pseudo-intents, i. e. the left-hand sides of the implications. In this paper we look at the right-hand sides, which are commonly called *essential closed sets*. In [1] it is shown that verifying whether a given set of attributes of a context is a pseudo-intent is as hard as verifying whether it is an essential closed set, i. e. it is coNP-complete. Unfortunately, a similar connection cannot be easily obtained for the decision problems considered in [2]. We therefore present yet another reduction from SAT which yields several complexity results about essential closed sets. Most of these results are similar to the ones for pseudo-intents. The main part of this paper is a reduction from SAT which proves that the problem of verifying whether a given set of attributes contains an essential closed set is NP-complete (Section 3). In Section 4 several other results are obtained using the same reduction. In particular, it is shown that the lectically first essential closed set cannot be computed in polynomial time unless P = NP, that essential closed sets cannot be enumerated in the lectic order with polynomial delay unless P = NP, and that minimal essential closed sets cannot be enumerated in output polynomial time unless P = NP.

2 Preliminaries

A *formal context* is a tuple (G, M, I) where G and M are finite sets and $I \subseteq G \times M$ is a binary relation. The elements of G are called *objects* and elements of M are called *attributes*. For a set of objects $A \subseteq G$ its *derivation* A' is defined as

$$A' = \{m \in M \mid \forall g \in A : gIm\}.$$

Analogously, for a set $B \subseteq M$ its *derivation* B' is defined as

$$B' = \{g \in G \mid \forall m \in B : gIm\}.$$

Applying the two derivation operators successively yields the closure operators $\cdot''$. Whenever we speak of a *closed set* in this work, we mean a set of attributes $B \subseteq M$ that is closed with respect to $\cdot''$, i. e. that satisfies $B'' = B$. Sets of attributes that can be written as $\{g\}'$ for some $g \in G$ are called *object intents*. The following result is common knowledge in FCA.

Proposition 1. *A set of attributes $B \subseteq M$ is closed if and only if it can be written as an intersection of object intents, i. e. there is a set $A \subseteq G$ such that*

$$B = \bigcap_{g \in A} \{g\}'.$$

A relevant research area in FCA are dependencies between sets of attributes. The simplest form of such a dependency is an implication $A \to B$, $A, B \subseteq M$. A set of attributes $D \subseteq M$ respects $A \to B$ if $A \not\subseteq D$ or $B \subseteq D$. $A \to B$ holds in the context (G, M, I) if all object intents respect $A \to B$.

Let $\mathcal{L}$ be a set of implications. We say that $A \to B$ follows semantically from $\mathcal{L}$ if and only if each subset $D \subseteq M$ that respects all implications from $\mathcal{L}$ also respects $A \to B$. $\mathcal{L}$ is an implicational base for (G, M, I) if it is

- *sound*, i. e. all implications from $\mathcal{L}$ hold in (G, M, I), and
- *complete*, i. e. all implications that hold in (G, M, I) follow from $\mathcal{L}$.

In [7] a minimum cardinality base, which is called the *Duquenne-Guigues-Base*, has been introduced. The left-hand sides of the implications in the Duquenne-Guigues-Base are called *pseudo-intents*. $P \subseteq M$ is a *pseudo-intent of* $\mathbb{K}$ if P is not closed and $Q'' \subseteq P$ holds for every pseudo-intent Q that is a proper subset of P. The Duquenne-Guigues-Base consists of all implications $P \to P''$, where P is a pseudo-intent. A set $R \subseteq M$ is an *essential closed set (of* $\mathbb{K}$*)* if there is a pseudo-intent P of $\mathbb{K}$ satisfying $P'' = Q$. Hence, the essential closed sets of a context are exactly those sets that occur as the right-hand side of an implication in the Duquenne-Guigues-Base. The following result is also common knowledge from FCA.

Proposition 2. *Let* $\mathbb{K}$ *be a formal context and let* $Q \subseteq M$ *be a set of attributes. If* Q *is not closed then* Q *contains a pseudo-intent of* $\mathbb{K}$.

The most well-known algorithm for computing the Duquenne-Guigues-Base is Next-Closure. It computes the set of all closed sets and all pseudo-intents of a context $\mathbb{K}$ in a special order, called the *lectic order*. Let $<$ be a total order on the elements of M. Then we say that $A \subseteq M$ is *lectically smaller* than $B \subseteq M$ if the smallest element with respect to $<$ that distinguishes A and B is contained in B. Formally, we write

$$A < B :\Leftrightarrow \exists x \in B \setminus A : \forall y < x : (y \in A \Leftrightarrow y \in B).$$

Notice that the lectic order extends the subset order, i. e. $A \subsetneq B$ implies $A < B$.

3 Main Reduction

In this section we prove that the following auxiliary problem is NP-hard. All other results will be based on this reduction and can be found in Section 4.

*Problem 1 (Essential Closed Subset (*ECS*)). Input:* A formal context $\mathbb{K} = (G, M, I)$ and a set $B \subseteq M$.
Question: Does an essential closed set $Q \subseteq B$ of $\mathbb{K}$ exist?

We prove NP-hardness using a reduction from SAT.

*Problem 2 (*SAT*). Input:* A boolean CNF-formula $f(p_1, \dots, p_n) = C_1 \wedge \cdots \wedge C_m$, where $C_i = (x_{i1} \vee \cdots \vee x_{il_i})$ and $x_{ij} \in \{p_1, \dots, p_n\} \cup \{\neg p_1, \dots, \neg p_n\}$ for all $i \in \{1, \dots, m\}$ and all $j \in \{1, \dots, l_i\}$.
Question: Is f satisfiable?

SAT remains NP-complete if we impose the additional condition that for every literal $l \in \{p_1, \dots, p_n\} \cup \{\neg p_1, \dots, \neg p_n\}$ there is a clause C_i in which l does not occur (otherwise we could simply add a new variable p_{n+1} and a new clause $C_{m+1} = (p_{n+1})$ without changing satisfiability of the formula).

Let an instance of SAT, i. e. a formula $f(p_1, \dots, p_n) = C_1 \wedge \cdots \wedge C_m$, where $C_i = (x_{i1} \vee \cdots \vee x_{il_i})$ and $x_{ij} \in \{p_1, \dots, p_n\} \cup \{\neg p_1, \dots, \neg p_n\}$ for all $i \in \{1, \dots, m\}$ and all $j \in \{1, \dots, l_i\}$, be given. We construct an instance of ECS. We define

$$M = \{\alpha_1, \dots, \alpha_n\} \cup \{p_1, \dots, p_n\} \cup \{\neg p_1, \dots, \neg p_n\} \cup \{\beta\} \tag{1}$$

For every $r \in \{1, \dots, n\}$ we define sets T_r and F_r.

$$\begin{aligned} T_r &= M \setminus \{\neg p_r, \alpha_r\} \\ F_r &= M \setminus \{p_r, \alpha_r\} \end{aligned} \tag{2}$$

Finally, for every $i \in \{1, \dots, m\}$ we define a set

$$\begin{aligned} A_i = M \setminus \{\beta\} & \\ & \setminus \{p_r \mid r \in \{1, \dots, n\},\ \text{the positive literal } p_r \text{ occurs in } C_i\} \\ & \setminus \{\neg p_r \mid r \in \{1, \dots, n\},\ \text{the negative literal } \neg p_r \text{ occurs in } C_i\} \\ & \setminus \{\alpha_r \mid r \in \{1, \dots, n\},\ p_r \text{ or } \neg p_r \text{ occurs in } C_i\} \end{aligned} \tag{3}$$

We construct a context $\mathbb{K}_f = (G, M, I)$ whose attribute set M is defined as in (1), whose set of objects is

$$G = \{g_{A_1}, \dots, g_{A_m}\} \cup \{g_{T_1}, \dots, g_{T_n}\} \cup \{g_{F_1}, \dots, g_{F_n}\} \cup \{g_{Q_1}, \dots, g_{Q_n}\},$$

and whose incidence relation I is such that

$$\{g_{A_i}\}' = A_i, \quad \{g_{T_r}\}' = T_r, \quad \{g_{F_r}\}' = F_r, \quad \{g_{Q_r}\}' = \{\alpha_r, p_r, \neg p_r\}$$

for all $r \in \{1, \dots, n\}$ and all $i \in \{1, \dots, m\}$. This context is shown in Table 1.

Our eventual objective is to reduce SAT to ECS by proving that f is satisfiable if and only if there exists a subset of

$$B = M \setminus \{\alpha_1, \dots, \alpha_n\}$$

Table 1. Context $\mathbb{K}_f$

	α_1	$\dots$	α_n	p_1	$\dots$	p_n	$\neg p_1$	$\dots$	$\neg p_n$	β
g_{A_1}			$\cdots$		A_1		$\cdots$			
$\vdots$					$\vdots$					
g_{A_m}			$\cdots$		A_m		$\cdots$			
g_{T_1}			$\cdots$		T_1		$\cdots$			
$\vdots$					$\vdots$					
g_{T_n}			$\cdots$		T_n		$\cdots$			
g_{F_1}			$\cdots$		F_1		$\cdots$			
$\vdots$					$\vdots$					
g_{F_n}			$\cdots$		F_n		$\cdots$			
g_{Q_1}	×			×			×			
$\vdots$		$\ddots$			$\ddots$			$\ddots$		
g_{Q_n}			×			×			×	

that is an essential closed set of $\mathbb{K}_f$. We need several technical results.

Let $\phi : \{p_1, \dots, p_n\} \to \{\texttt{true}, \texttt{false}\}$ be an assignment that assigns truth values to all variables. There is a natural correspondence between ϕ and a set of attributes S_ϕ. We define

$$S_\phi := \{p_r \mid \phi(p_r) = \texttt{true}\} \cup \{\neg p_r \mid \phi(p_r) = \texttt{false}\}. \tag{4}$$

The following result motivates our choice of the object intents $\{g_{A_i}\}' = A_i$, $i \in \{1, \dots, m\}$.

Lemma 1. *Let ϕ be an assignment of truth values. Then ϕ makes f true if and only if $S_\phi \nsubseteq A_i$ holds for all $i \in \{1, \dots, m\}$.*

Proof. Since f is in conjunctive normal form, ϕ makes f true if and only if ϕ makes every clause C_i, $i \in \{1, \dots, m\}$, of f true. For every $i \in \{1, \dots, m\}$ the assignment ϕ makes the clause C_i true if and only if one of the literals in C_i evaluates to true, i. e.

- there is some p_r satisfying $\phi(p_r) = \texttt{true}$, where the positive literal p_r occurs in C_i, or
- there is some p_r satisfying $\phi(p_r) = \texttt{false}$, where the negative literal $\neg p_r$ occurs in C_i.

According to (3) and (4) this is equivalent to saying that

- there is some $p_r \in S_\phi$, where $p_r \notin A_i$, or
- there is some $\neg p_r \in S_\phi$, where $\neg p_r \notin A_i$.

This is equivalent to $S_\phi \nsubseteq A_i$. Thus we have shown that ϕ makes f true if and only if $S_\phi \nsubseteq A_i$ holds for all $i \in \{1, \dots, m\}$.

The following proposition follows immediately from (2) and (4).

Proposition 3. *Let ϕ be an assignment of truth values and $X \subseteq S_\phi$ a set of attributes. Then*

$$X \cup \{\beta\} = \bigcap_{\neg p_r \notin X} T_r \cap \bigcap_{p_r \notin X} F_r \tag{5}$$

holds. Since for all $r \in \{1, \ldots, n\}$ the sets T_r and F_r are object intents this proves that $X \cup \{\beta\}$ is closed.

Proposition 4. *Let ϕ be an assignment of truth values and $X \subseteq S_\phi$ a set of attributes. X is closed if and only if there is some $i \in \{1, \ldots, m\}$ such that $X \subseteq A_i$ holds.*

Proof. ($\Leftarrow$) We already know from Proposition 3 that $X \cup \{\beta\} = \bigcap_{\neg p_r \notin X} T_r \cap \bigcap_{p_r \notin X} F_r$ holds for X. Since $X \subseteq A_i$ and $\beta \notin A_i$ it follows that

$$X = A_i \cap \big(X \cup \{\beta\}\big) = A_i \cap \bigcap_{\neg p_r \notin X} T_r \cap \bigcap_{p_r \notin X} F_r.$$

Since T_r and F_r, $r \in \{1, \ldots, n\}$, and A_i are object intents Proposition 1 proves that $X \cup \{\beta\}$ is closed.

($\Rightarrow$) The case where $X = \emptyset$ is trivial. Let $X = \{l\}$ be a singleton set. We have required that for every literal there is a clause in which it does not occur. Hence, there is a clause C_i in which l does not occur, and therefore $X = \{l\} \subseteq A_i$ holds. The case where X contains at least two elements remains. Since $X \subseteq S_\phi$ holds, X cannot contain $\{p_r, \neg p_r\}$ or $\{\alpha_r\}$ for any $r \in \{1, \ldots, n\}$. We obtain that if X has at least two elements then it cannot be a subset of $\{\alpha_r, p_r, \neg p_r\} = \{g_{Q_r}\}'$ for any $r \in \{1, \ldots, n\}$. Assume that $X \not\subseteq A_i = \{g_{A_i}\}'$ holds for all $i \in \{1, \ldots, m\}$. Then the only object intents that contain X are of the form $\{g_{T_r}\}'$ or $\{g_{F_r}\}'$. All object intents of the form $\{g_{T_r}\}'$ or $\{g_{F_r}\}'$ contain β, which yields $\beta \in X''$. Because $\beta \notin S_\phi \supseteq X$ this is a contradiction to the fact that X is closed. Therefore, the assumption that $X \not\subseteq A_i$ holds for all $i \in \{1, \ldots, m\}$ must be false, i. e. there must be some $i \in \{1, \ldots, m\}$ satisfying $X \subseteq A_i$.

Lemma 2. *For every $r \in \{1, \ldots, n\}$ it holds that*

$$\{\alpha_r\}'' = \{p_r, \neg p_r\}'' = \{\alpha_r, p_r, \neg p_r\}.$$

Proof. We have defined $\mathbb{K}_f$ in such a way that every object intent that contains $\{p_r, \neg p_r\}$ also contains $\{\alpha_r\}$. Conversely, every object intent that contains $\{\alpha_r\}$ also contains $\{p_r, \neg p_r\}$. This proves $\{\alpha_r, p_r, \neg p_r\} \subseteq \{p_r, \neg p_r\}''$ and $\{\alpha_r, p_r, \neg p_r\} \subseteq \{\alpha_r\}''$. On the other hand, we know that $\{\alpha_r, p_r, \neg p_r\} = \{g_{Q_r}\}'$ is closed since it is an object intent. This yields $\{\alpha_r, p_r, \neg p_r\} = \{p_r, \neg p_r\}'' = \{\alpha_r\}''$.

Theorem 1. *Define $B = \{p_1, \ldots, p_n\} \cup \{\neg p_1, \ldots, \neg p_n\} \cup \{\beta\}$. There is an essential closed set $Q \subseteq B$ if and only if f is satisfiable.*

Proof. (⇒) Assume that Q contains both p_r and $\neg p_r$ for some $r \in \{1, \ldots, n\}$. Lemma 2 yields $\alpha_r \in Q''$. This contradicts the fact that Q is a closed subset of B. Therefore, the assumption that Q contains both p_r and $\neg p_r$ for some $r \in \{1, \ldots, n\}$ must be false. Thus, Q must be a subset of $S_\phi \cup \{\beta\}$ for some assignment ϕ. Since Q is an essential closed set there must be a pseudo-intent $P \subseteq Q \subseteq S_\phi \cup \{\beta\}$. If P contains β then Proposition 3 yields that P is closed. This contradicts the fact that P is a pseudo-intent. Hence, P cannot contain β, i. e. $P \subseteq S_\phi$ holds. Since P is a pseudo-intent and therefore not closed we obtain from Proposition 4 that there is no $i \in \{1, \ldots, m\}$ such that $P \subseteq A_i$. $P \subseteq S_\phi$ yields that there is no $i \in \{1, \ldots, m\}$ such that $S_\phi \subseteq A_i$. It follows from Lemma 1 that ϕ makes f true.

(⇐) Let ϕ be an assignment that makes f true. Lemma 1 implies $S_\phi \not\subseteq A_i$ for all $i \in \{1, \ldots, m\}$. Proposition 4 shows that S_ϕ is not closed. Let X be minimal among all subsets of S_ϕ that are not closed. Then in particular all subsets of X are closed. Since X is not closed, but all of its subsets are closed, X must be a pseudo-intent of $\mathbb{K}_f$. Proposition 3 states that $X \cup \{\beta\}$ is closed, and therefore $X'' = X \cup \{\beta\}$ holds. This shows that $X \cup \{\beta\}$ is an essential closed set. Since $X \cup \{\beta\}$ is also a subset of B this proves the initial claim.

Corollary 1. ECS *is* NP-*hard.*

Proof. Every boolean formula f can be converted into an instance of ECS, namely the context $\mathbb{K}_f$ and the set $B = \{p_1, \ldots, p_n\} \cup \{\neg p_1, \ldots, \neg p_n\} \cup \{\beta\}$, in polynomial time. Theorem 1 states that f is a “Yes”-instance of SAT if and only if $\mathbb{K}_f$ and B are a “Yes”-instance of ECS.

We have thus shown that the problem of deciding whether a given set of attributes B in a context $\mathbb{K}$ contains an essential closed set, is NP-hard. Surprisingly, the problem becomes easier if we require B to be closed. If all subsets of B are closed then B cannot contain a pseudo-intent, and thus it does not contain an essential closed set. On the other hand, if B contains a set S that is not closed, then there must be a pseudo-intent $P \subseteq S$ because of Proposition 2. We obtain $P'' \subseteq S'' \subseteq B'' = B$. Therefore B contains the essential closed set P''. This proves that checking whether a closed set B contains an essential closed set is equivalent to checking whether all subsets of B are closed. It is well known that the latter can be done in polynomial time.

4 Further Results

Let $\mathbb{K} = (G, M, I)$ be a formal context. We call a set Q is a *minimal essential closed set (of* $\mathbb{K}$*)* if Q is minimal with respect to set inclusion among all essential closed sets of $\mathbb{K}$. It is known from [1] that the problem of deciding whether a given set of attributes is an essential closed set is coNP-hard. We first show that the problem becomes easier for minimal essential closed sets: it is possible to decide in polynomial time whether a given set is a minimal essential closed set. This result is required for later proofs.

Proposition 5. *Q is a minimal essential closed set if and only if*

- *Q is closed, and*
- *not every subset of Q is closed, and*
- *every closed set $R \subsetneq Q$ satisfies*

$$\forall S \subseteq R : S'' = S. \tag{6}$$

Proof. ($\Rightarrow$) As an essential closed set, Q is obviously closed. As an essential closed set, Q must contain a pseudo-intent P_1, which is not closed. Hence, not all subsets of Q are closed. Assume that there is a strict subset $R \subsetneq Q$ that is closed and a set $S \subseteq R$ that is not closed. By Proposition 2 S contains a pseudo-intent $P_2 \subseteq S \subseteq R$. Since R is closed it follows that $P_2'' \subseteq R \subsetneq Q$. Hence, P_2'' is an essential closed set and a strict subset of Q, which contradicts minimality of Q. Thus the assumption that such a set S exists must be false.

($\Leftarrow$) Since not all subsets of Q are closed there must be a pseudo-intent $P \subseteq Q$ by Proposition 2. Since Q is closed we obtain $P'' \subseteq Q$. P'' cannot be a strict subset of Q, because otherwise (6) would imply that P is closed. Therefore, $P'' = Q$ holds, and thus Q is an essential closed set. No strict subset of Q can be an essential closed set because of (6). Thus Q is a minimal essential closed set.

Notice that in order to decide whether a given set Q satisfies (6) for all closed sets $R \subsetneq Q$ it suffices to check whether (6) holds for all sets R that are maximal with respect to set inclusion among the closed strict subsets of Q. If Q is itself closed then these are of the form $Q \cap \{g\}'$, where $g \in G$ and $Q \not\subseteq \{g\}'$. Hence, it suffices to check (6) for at most $|G|$ strict subsets of Q. It has been established in previous works [2] that one can decide in polynomial time whether all subsets of a given set of attributes are closed. Hence, all conditions from Propostion 5 can be tested in polynomial time.

Corollary 2. *Let $\mathbb{K}$ be a formal context and $Q \subseteq M$ a set of attributes. It is possible to decide in time polynomial in the size of the context $\mathbb{K}$ and the size of Q whether Q is a minimal essential closed set.*

This gives us the containment result corresponding to the hardness result from Corollary 1. Clearly, for a formal context $\mathbb{K} = (G, M, I)$ and a set $B \subseteq M$ there is an essential closed set $Q \subseteq B$ if and only if there is a minimal essential closed set $R \subseteq B$. In order to decide in non-deterministic polynomial time whether B contains an essential closed set we can non-deterministically guess a subset of B and decide using Corollary 2 whether it is a minimal essential closed set. This proves that ECS is contained in NP. Together with the previous hardness result we obtain NP-completeness.

Theorem 2. ECS *is* NP*-complete.*

We want to take a closer look at the enumeration problem for essential closed sets. But first we consider the following decision problem.

*Problem 3 (Lectically Smaller Essential Closed Set (*LS-ECS*)). Input:* A formal context $\mathbb{K} = (G, M, I)$ and a set $B \subseteq M$.
Question: Is there an essential closed set Q of $\mathbb{K}$ which is lectically smaller that B?

Theorem 3. LS-ECS *is* NP*-complete.*

Proof. Containment: Since the lectic order extends the subset order there is an essential closed set that is lectically smaller than B if and only if there is a minimal essential closed set that is lectically smaller than B. This can be verified by non-deterministically guessing a subset of B and checking in polynomial time whether it is a minimal essential closed set.

Hardness: Given an instance f of SAT we can construct an instance of LS-ECS consisting of the context $\mathbb{K}_f$ and the set $B = \{p_1, \ldots, p_n\} \cup \{\neg p_1, \ldots, \neg p_n\} \cup \{\beta\}$ using the same reduction as in Section 3. We define the order on the set of attributes as follows

$$\alpha_1 < \cdots < \alpha_n < p_1 < \cdots < p_n < \neg p_1 < \cdots < \neg p_n < \beta.$$

Then the sets that are lectically smaller than B are exactly the subsets of B. The correctness of the reduction therefore follows from Theorem 1.

This result has consequences for the problem of enumeration of essential closed sets in the lectic order. If it were possible to compute the lectically first essential closed set of a context $\mathbb{K}$ in polynomial time then we could decide LS-ECS in polynomial time as follows. We would simply compute the lectically first essential closed set of $\mathbb{K}$ and check whether it is lectically smaller than B. Because of Theorem 3 it is not possible to decide LS-ECS in polynomial time unless P = NP.

Corollary 3. *Let $\mathbb{K}$ be a formal context. It is not possible to compute the lectically first essential closed set of $\mathbb{K}$ in polynomial time unless* P = NP*.*

In this respect, the computational behaviour of essential closed sets is worse than that of pseudo-intents, since the lectically first pseudo-intent can be computed in polynomial time [3]. Because not even the lectically first essential closed set can be computed in polynomial time it is obviously not possible to enumerate essential closed sets in the lectic order with polynomial delay.

Corollary 4. *It is not possible to enumerate essential closed sets in the lectic order with polynomial delay.*

Finally, we consider the problem of enumeration of minimal essential closed sets.

*Problem 4 (All Minimal Essential Closed Sets (*All-MECS*)). Input:* A formal context $\mathbb{K} = (G, M, I)$ and sets $Q_1, \ldots, Q_k \subseteq M$.
Question: Are $Q_1, \ldots, Q_k$ all minimal essential closed sets of $\mathbb{K}$?

Lemma 3. *Let f be a boolean CNF-formula and $\mathbb{K}_f$ the formal context constructed as in Section 3. Then $Q_1 = \{\alpha_1, p_1, \neg p_1\}$, …, $Q_n = \{\alpha_n, p_n, \neg p_n\}$ are all the minimal essential closed sets of $\mathbb{K}_f$ if and only if there is no essential closed set $Q \subseteq B = \{p_1, \ldots, p_n\} \cup \{\neg p_1, \ldots, \neg p_n\} \cup \{\beta\}$.*

Proof. ($\Rightarrow$) Assume that there is an essential closed set $Q \subseteq B$. Then Q contains a minimal essential closed set R. For all $r \in \{1, \ldots, n\}$ since B does not contain α_r the set R cannot contain α_r, either. Thus $R \neq Q_r$ holds for all $r \in \{1, \ldots, n\}$, which contradicts the fact that Q_1, …, Q_n are all the minimal essential closed sets of $\mathbb{K}$. Hence, the assumption must be false, i. e. an essential closed set $Q \subseteq B$ cannot exist.

($\Rightarrow$) By Lemma 2 every closed set that contains α_r for some $r \in \{1, \ldots, n\}$ must also contain $\{p_r, \neg p_r\}$. Therefore, $Q_r = \{\alpha_r, p_r, \neg p_r\}$ is the only minimal essential closed set of $\mathbb{K}_f$ that contains α_r. Thus, every essential closed set that is different from Q_1, …, Q_n must be a subset of B. The hypothesis states that such a set does not exist. Hence, Q_1, …, Q_n are all the minimal essential closed sets of $\mathbb{K}$.

Theorem 4. All-MECS *is* coNP*-complete.*

Proof. To prove hardness using a reduction from SAT from a given formula f we construct a context $\mathcal{K}_f$ as in Section 3 and sets Q_1, …, Q_n as in Lemma 3. Lemma 3 and Theorem 1 show that Q_1, …, Q_n are all the minimal essential closed sets of $\mathbb{K}$ if and only if f is not satisfiable. This proves that All-MECS is coNP-hard. Containment can be shown using Corollary 2: Given an instance of All-MECS consisting of a context $\mathbb{K} = (G, M, I)$ and sets Q_1, …, Q_n we can verify in polynomial time using Corollary 2 that all sets Q_1, …, Q_n are minimal essential closed sets. Subsequently, we non-deterministically guess a set $S \subseteq M$ and check in polynomial time whether it is a minimal essential closed set that is different from Q_1, …, Q_n.

If there were an algorithm $\mathcal{A}$ that enumerates the minimal essential closed sets of a context in output polynomial time, then we could construct an algorithm $\mathcal{A}'$ that decides All-MECS as follows: Since $\mathcal{A}$ can enumerate the minimal essential closed sets of a context $\mathbb{K}$ in output polynomial time there must be a polynomial $p(k, n)$ that serves as an upper bound for the runtime of $\mathcal{A}$, where k is the size of the input context $\mathbb{K}$ and n is the number of minimal essential closed sets of $\mathbb{K}$. To decide All-MECS for a context $\mathbb{K}$ and sets Q_1, …, Q_n we let $\mathcal{A}$ run on $\mathbb{K}$ and stop it after time $p(|\mathbb{K}|, n)$. Then we compare its output (if any) to $Q_1, \ldots, Q_n$. $Q_1, \ldots, Q_n$ are not all the minimal essential closed sets of $\mathbb{K}$ iff the outputs differ or $\mathcal{A}$ does not terminate within $p(|\mathbb{K}|, n)$ steps. Since All-MECS cannot be decided in polynomial time unless P = NP such an algorithm cannot exist unless P = NP.

Corollary 5. *Minimal essential closed sets cannot be enumerated in output polynomial time unless* P = NP*.*

5 Conclusion

Using a new reduction from SAT we have shown several complexity results about essential closed sets. Most of these results closely resemble those for pseudo-intents. Essential closed sets cannot be enumerated in the lectic order with polynomial delay unless P = NP, and minimal essential closed sets cannot be enumerated in output polynomial time unless P = NP. The same holds for pseudo-intents [2].

Essential closed sets differ from pseudo-intents computationally with respect to the following problems. For an arbitrary set of attributes it is NP-hard to verify whether it contains an essential closed set. By contrast, it is easy to check whether a given set of attributes contains a pseudo-intent, because this simply means checking for closedness. It is impossible to compute the lectically first essential closed set unless P = NP. The lectically first pseudo-intent can be computed in polynomial time [3].

These results are, of course, only a minor contribution to the question whether the Duquenne-Guigues Base can be enumerated efficiently. This important question remains open and should be part of future work.

References

1. Mikhail Babin and Sergei Kuznetsov. Recognizing pseudo-intents is coNP-complete. In Marzena Kryszkiewicz and Sergei Obiedkov, editors, *Proc. of the 7th Int. Conf. on Concept Lattices and Their Applications (CLA 2010)*, volume 672. CEUR Workshop Proceedings, 2010.
2. Felix Distel. Hardness of enumerating pseudo-intents in the lectic order. In Barış Sertkaya and Léonard Kwuida, editors, *Proc. of the 8th Int. Conf. on Formal Concept Analysis (ICFCA 2010)*, volume 5986 of *Lecture Notes in Artificial Intelligence*, pages 124–137. Springer, 2010.
3. Felix Distel and Barış Sertkaya. On the complexity of enumerating pseudo-intents. *Discrete Applied Mathematics*, to appear, 2011.
4. T. Eiter and G. Gottlob. Hypergraph transversal computation and related problems. In S. Flesca, S. Greco, N. Leone, and G. Ianni, editors, *Proceedings of JELIA 2002*, volume 2424 of *LNCS*, pages 549–564, Heidelberg, 2002. Springer.
5. Bernhard Ganter. Two basic algorithms in concept analysis. Preprint 831, Fachbereich Mathematik, TU Darmstadt, Darmstadt, Germany, 1984.
6. Alain Gély, Raoul Medina, Lhouari Nourine, and Yoan Renaud. Uncovering and reducing hidden combinatorics in guigues-duquenne bases. In Bernhard Ganter and Robert Godin, editors, *ICFCA*, volume 3403 of *Lecture Notes in Computer Science*, pages 235–248. Springer, 2005.
7. J.-L. Guigues and V. Duquenne. Familles minimales d'implications informatives résultant d'un tableau de données binaires. *Math. Sci. Humaines*, 95:5–18, 1986.
8. Sergei O. Kuznetsov. On the intractability of computing the Duquenne-Guigues base. *Journal of Universal Computer Science*, 10(8):927–933, 2004.
9. Sergei O. Kuznetsov and Sergei A. Obiedkov. Algorithms for the construction of concept lattices and their diagram graphs. In *Principles of Data Mining and Knowledge Discovery*, pages 289–300, 2001.

10. Sergei O. Kuznetsov and Sergei A. Obiedkov. Counting pseudo-intents and #p-completeness. In Rokia Missaoui and Jürg Schmid, editors, *ICFCA*, volume 3874 of *Lecture Notes in Computer Science*, pages 306–308. Springer, 2006.
11. S. Obiedkov and V. Duquenne. Attribute-incremental construction of the canonical implication basis. *Annals of Mathematics and Artificial Intelligence*, 49(1-4):77–99, 2007.
12. Barış Sertkaya. Some computational problems related to pseudo-intents. In Sébastien Ferré and Sebastian Rudolph, editors, *Proc. of the 7th Int. Conf. on Formal Concept Analysis, (ICFCA 2009)*, volume 5548 of *Lecture Notes in Artificial Intelligence*, pages 130–145. Springer Verlag, 2009.
13. Barış Sertkaya. Towards the complexity of recognizing pseudo-intents. In Frithjof Dau and Sebastian Rudolph, editors, *Proceedings of the 17th International Conference on Conceptual Structures, (ICCS 2009)*, volume 5662, pages 284–292, 2009.