Bases via Minimal Generators*

Pablo Cordero, Manuel Enciso, Angel Mora, Manuel Ojeda-Aciego
Universidad de Málaga, Spain
pcordero@uma.es, enciso@lcc.uma.es, amora@ctima.uma.es, aciego@uma.es

July 24, 2013

Abstract

The concept lattice corresponding to a context may be alternatively specified by means of attribute implications. One outstanding problem in formal concept analysis and other areas is the study of the equivalences between a given set of implications and its corresponding basis (notice that there exists a wide range of approaches to basis in the literature). In this work we introduce a method to provide a Duquenne-Guigues basis corresponding to the minimal generators and their closed sets from a context.

1 Introduction

The main goal of Formal Concept Analysis (FCA) is to identify the relationships between sets of objects and sets of attributes using information from a cross table. The derivation operators establish a Galois connection between the power sets of objects and attributes which generates a complete lattice, the so-called concept lattice.

One obvious goal in this framework is to remove redundancy and obtain a minimal basis. The most widely approach comes from the notion of *Duquenne-Guigues Basis* [4] also called *stem base*. This basis is minimal with respect to the number of implications, i.e. if some implication is removed from the basis, there exist valid and non-redundant implications which are valid in the dataset and cannot be inferred from the new reduced basis using Armstrong's Axioms.

In [2], the authors presented an algorithm to obtain *all* the minimal generators and their corresponding closures. From that information it is possible to build a set of implications which mimics exactly the underlying concept lattice, by using a minimal generator as antecedent and its corresponding closure as consequent. Obviously, this set can be somehow minimized, obtaining what is called a *basis*; in this paper, we will focus on the Duquennes-Guigues basis.

Specifically , a method is introduced to calculate a Duquenne-Guigues basis from all closed sets and their minimal generator.

*Partially supported by grants P09-FQM-5233 of the Junta de Andalucía, and TIN09-14562-C05-01, TIN2011-28084, and TIN12-39353-C04-01 of the Science and Innovation Ministry of Spain, co-funded by the European Regional Development Fund (ERDF).

2 Background

We assume as known the basic concepts of Formal Concept Analysis (FCA). [3, 11]

2.1 Simplification logic and closures

We summarize the axiomatic system of Simplification Logic for Functional Dependencies $\mathbf{SL}_{FD}$ equivalent to the well-know Armstrong's Axioms. It avoids the use of transitivity and is guided by the idea of simplifying the set of implications by efficiently removing redundant attributes inside the implications [1]. We define $\mathbf{SL}_{FD}$ as the pair $(\mathcal{L}_{FD}, \mathcal{S}_{FD})$ where the axiomatic system $\mathcal{S}_{FD}$ has the following axiom scheme and inference rules. The third rule is named *Simplification* rule and it is the core of $\mathbf{SL}_{FD}$:

$$\texttt{[Ref]} \quad \frac{A \supseteq B}{A \rightarrow B}$$

$$\texttt{[Frag]} \ \frac{A \rightarrow B \cup C}{A \rightarrow B} \qquad \texttt{[Comp]} \ \frac{A \rightarrow B,\ C \rightarrow D}{A \cup C \rightarrow B \cup D} \qquad \texttt{[Simp]} \ \frac{A \rightarrow B,\ C \rightarrow D}{A \cup (C \smallsetminus B) \rightarrow D}$$

3 Obtaining basis from minimal generators

As stated in the introduction, the goal of this position paper can be considered one step beyond the work presented in [2], where we illustrated the use of the Simplification paradigm to guide the search of all minimal generator sets.

Our main goal is studied here: a method to get a Duquenne-Guigues basis given the set of all the minimal generators and its corresponding closed sets.

Based on the properties of minimal generators - closed sets and in $\mathbf{SL}_{FD}$, we propose an operator that characterizes when it is possible to remove redundant attributes in a set of implications. The exhaustive application of this result produces a reduced implication set and, in some cases, with an empty right-hand side. These implications are removed from the output set, returning a Duquenne-Guigues basis (see Theorem 3.4).

Thus, summarizing our proposal, from a set of *(minimal generators, closed set)* the method returns a Duquenne-Guigues basis.

Theorem 3.1 *Let $\langle A \cup B, A \rangle, \langle C \cup D, C \rangle$ be two pairs obtained using* `MinGen` *algorithm [2]* [1] *where A, C are minimal generators and $A \cup B, C \cup D$ are closed sets. In this situation, the following implications are valid: $\{A \rightarrow B, C \rightarrow D\}$*

If $A \subseteq C$, then the following equivalence holds:

$$\{A \rightarrow B, C \rightarrow D\} \equiv \{A \rightarrow B, (C \cup B) \rightarrow (D \smallsetminus B)\} \tag{3.1}$$

Notice that $(C \cup B \cup D \smallsetminus B)$ is a closed set.

The above equivalence (3.1), infers the definition of an operator that reduces the set of implications if we apply it exhaustively.

[1] Closed sets - minimal generators can be calculated using others methods well known.

Definition 3.2 *Let $\langle A \cup B, A\rangle, \langle C \cup D, C\rangle$ be two pairs obtained using* `MinGen` *algorithm and let $\Gamma = \{A \to B, C \to D\}$ be the corresponding equivalent set of implications. We define the following operator:*

$$\Upsilon(A \to B, C \to D) = \{A \to B, C \cup B \to D \smallsetminus B\}$$

This operator is applied only when $A \subseteq C$. We traverse the set of implications and for any two implications we check whether $A \subseteq C$ or $C \subseteq A$, applying the operator if it is the case. We have developed in Prolog this operator.

In the following definition, we present the way in which the Υ operator will be applied to a set of implications. The way in which we check both inclusions of the left hand sides of the implications, reduces the traversing of the Γ set because we compare each implication only with all later implications in the Γ set.

Definition 3.3 *Let $\Gamma = \{A_1 \to B_1, A_2 \to B_2 \dots A_n \to B_n\}$ be a set of implications. We define the application of the operator Υ to a set of implications as its exhaustive application as follows: $\Upsilon(\Gamma) = \{\Upsilon(A_i, A_j), \quad i{=}1 \dots n{-}1, j{=}i{+}1 \dots n\}$*

Theorem 3.4 *Let $\Phi = (\langle C_1, mg(C_1)\rangle, \langle C_2, mg(C_2)\rangle, \dots)$ be a set where C_i is a closed set of attributes and $mg(C_i) = \{D\colon D$ is a mingen and $D^+ = C_i\}$. And let $\Gamma = \{A_1 \to B_1, \dots\}$ the set of implications deduced from Φ. The operator $\Upsilon(\Gamma)$ renders the Duquenne-Guigues basis equivalent to Γ.*

The proof aries from the transformation made with the Υ operator, which completes the left-hand side to be a pseudo-intent and remove implications from the original set to get minimal cardinality.

Example 3.1 *Let $\Gamma = \{b \to acef, ad \to ef, abd \to cef\}$ a set of implications equivalent to the following set of closed sets and their minimal generators $= \{\langle abcdef, \{abd\}\rangle, \langle adef, \{ad\}\rangle, \langle abcef, \{b\}\rangle, \langle \varnothing, \{\varnothing\}\rangle\}$. Applying exhaustively the Υ operator to any pair of implications of Γ we get the following set of implications, which conforms with a Duquenne-Guigues basis.*

$$\Gamma' = \{b \to acef, ad \to e\}$$

4 Conclusions and future works

In this paper we present an operator which allows the transformation of a set of implications builds over the set of all minimal generators into a Duquenne-Guigues basis. The first step is to use `MinGen` Algorithm introduced in [2] to compute all minimal generators corresponding to an arbitrary set of implications. A deep study about the soundness, completeness, and complexity of the algorithms proposed are the plan of work that we sketch in this proposal paper. In the future, our interest is to achieve a basis not only with minimal cardinality in the number of implications but also with minimal size in the attributes inside of the implications. The operator defined in this work is the first step in this direction.

References

[1] P. Cordero, M. Enciso, A. Mora, I.P, de Guzmán: SLFD logic: Elimination of data redundancy in knowledge representation, LNCS 2527: 141–150, 2002.

[2] P. Cordero, M. Enciso, A. Mora, M. Ojeda-Aciego: Computing minimal generators from implications: a logic-guided approach, Concept Lattice and Aplications - CLA 2012: 187-198, 2012.

[3] B. Ganter, Two basic algorithms in concept analysis. Technische Hochschule, Darmstadt, 1984.

[4] 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.

[5] M. Hermann and B. Sertkaya, On the Complexity of Computing Generators of Closed Sets. ICFCA 2008: 158–168

[6] R. Hill, Computational Intelligence and Emerging Data Technologies. 2nd Intl Conf on Intelligent Networking and Collaborative Systems (INCOS'10), pg 449–454, 2010.

[7] A. Mora, P. Cordero, M. Enciso, I. Fortes, Closure via functional dependence simplification, IJCM, 89(4): 510–526, 2012.

[8] L. Szathmary and A. Napoli and S. O. Kuznetsov, ZART: A Multifunctional Itemset Mining Algorithm , Proc. of the 6th Intl. Conf. on Concept Lattices and Their Applications (CLA '08): 47–58, 2008.

[9] L. Szathmary and P. Valtchev and A. Napoli and R. Godin, An Efficient Hybrid Algorithm for Mining Frequent Closures and Generators, Concept Lattices and Their Applications (CLA '07): 26–37, 2007.

[10] K. Nehmé, P. Valtchev, M. H. Rouane, R. Godin, On Computing the Minimal Generator Family for Concept Lattices and Icebergs, LNCS 3403: 192–207, 2005.

[11] R. Wille, Restructuring lattice theory: an approach based on hierarchies of concepts. In I. Rival (ed.), Ordered sets, pp. 445-470, 1982.