

A Lattice View of Functional Dependencies in Incomplete Relations *

Mark Levene
Department of Computer Science
University College London
Gower Street
London WC1E 6BT, U.K.
Email: mlevene@cs.ucl.ac.uk

Abstract

Functional Dependencies (or simply FDs) are by far the most common integrity constraint in the real world. When relations are incomplete and thus contain null values the problem of whether satisfaction is additive arises. Additivity is the property of the equivalence of the satisfaction of a set of functional dependencies (FDs), F , with the individual satisfaction of each member of F in an incomplete relation. It is well known that, in general, satisfaction of FDs is not additive. Previously we have shown that satisfaction is additive if and only if the set of FDs is monodependent. Thus monodependence of a set of FDs is a desirable property when relations may be incomplete. A set of FDs is monodependent if it satisfies both the intersection property and the split-freeness property. (The two defining properties of monodependent sets of FDs correspond to the two defining properties of conflict-free sets of multivalued data dependencies.)

We investigate the properties of the lattice $\mathcal{L}(F)$ of closed sets of a monodependent set of FDs F over a relation schema R . We show an interesting connection between monodependent sets of FDs and exchange and antiexchange lattices. In addition, we give a characterisation of the intersection property in terms of the existence of certain distributive sublattices of $\mathcal{L}(F)$. Assume that a set of FDs F satisfies the intersection property. We show that the cardinality of the family $\mathcal{M}(F)$ of meet-irreducible closed sets in $\mathcal{L}(F)$ is polynomial in the number of attributes associated with R ; in general, this number is exponential. Thus an Armstrong relation for F having a polynomial number of tuples in the number of attributes associated with R can be generated. As a corollary we show that the prime attribute problem can be solved in polynomial time in the size of F ; in general, the prime attribute problem is NP-complete. We also show that F satisfies the intersection property if and only if the cardinality of each element in $\mathcal{M}(F)$ is greater than or equal to the cardinality of the attribute set of R minus two. Using this result we are able to show that the superkey of cardinality k problem is still NP-complete when F is restricted to satisfy the intersection property. Finally, we show that separatory sets of FDs are monodependent.

1 Introduction

In order to handle incomplete information, Codd [CODD79] suggested the addition to the database domains of an unmarked null value, denoted by *unk*, whose meaning is “value at

* Appeared in *Acta Cybernetica* **12**, (1995), 181-207

present exists but is unknown”. We call relations, whose tuples may contain the null value *unk*, *incomplete relations*. The semantics of an incomplete relation r are defined in terms of the *possible worlds* relative to r . Each possible world relative to r is a *complete relation*, i.e. a relation without any occurrence of *unk*, emanating from a possible substitution of all the occurrences of *unk* in r by nonnull values in the underlying database domains.

Functional Dependencies (or simply FDs) are by far the most common integrity constraint in the real world [ULLM88, ATZE93, MANN92] and the notion of a key (derived from a given set of FDs) [CODD79] is fundamental to the relational model. Given a set of FDs F over a relation schema R and an incomplete relation r over R , it is therefore natural to say that r satisfies F if there is a complete relation s , in the set of possible worlds relative to r , such that s satisfies each of the FDs in F . This gives rise to the *additivity problem*, which is the problem of whether the statement that r satisfies F is equivalent to the statement that r satisfies each FD in a reduced cover G of F [LEVE94, LEVE95a] (cf. [ATZE93]); if these two statements are equivalent for a class of incomplete relations and a class of sets of FDs then we say that *satisfaction is additive* with respect to these classes. It is well known that, in general, satisfaction of FDs is not additive [ATZE86, LEVE94, LEVE95a]. If satisfaction is not additive, then a set of FDs F in this nonadditive class may be viewed as contradictory. Thus we consider the solution of the additivity problem to be an important prerequisite for any relational database system supporting FDs in the context of incomplete information, since otherwise semantic anomalies may arise.

In [LEVE94] we introduced the class of *monodependent* sets of FDs. A set of FDs F over a relation schema R is monodependent if the following two properties are satisfied. The first property, called the *intersection property*, informally states that for each attribute A in the attribute set associated with R , there is a unique nontrivial and reduced FD in the closure of F that functionally determines A . The second property, called the *split-freeness property*, informally states that there are no two nontrivial FDs in the closure of F such that the right-hand side of each of the two FDs splits the left-hand side of the other FD. The main result in [LEVE94] shows that satisfaction is additive with respect to the class of all incomplete relations and a class of sets of FDs \mathbf{FC} , if and only if all the sets of FDs in \mathbf{FC} are monodependent sets of FDs. Therefore, monodependence provides a solution to the additivity problem.

In [LEVE95b] we studied the impact on normalisation theory in relational databases of assuming that sets of FDs are monodependent, and in [LEVE95a] we extended the results in [LEVE94] to the wider class of sets of FDs and *unary inclusion dependencies* [COSM90].

It is well known that the family of all closed sets, with respect to a set of FDs F , is a lattice partially ordered by set inclusion; we denote this lattice by $\mathcal{L}(F)$ [DEME92, DEME93]. An in-depth investigation concerning the connection between the structure of a set of FDs and the type of lattice of closed sets it induces was carried out in [DEME92]. Herein we investigate the properties of $\mathcal{L}(F)$ when F is monodependent.

We next briefly outline the main results of this paper. The set of *equivalence classes* of a set of FDs F over R is a partition of F such that two FDs are in the same equivalence class if and only if the closures of their left-hand sides are the same [MAIE80, MANN83]. Assume that F satisfies the intersection property. We then show that $\mathcal{L}(F)$ is *exchange* [GRAT78] if and only if F satisfies the split-freeness property and the cardinality of all the nonempty equivalence classes of F is maximal. Correspondingly, we show that $\mathcal{L}(F)$ is *antiexchange* [JAMI85] if and only if F satisfies the split-freeness property and the cardinality of all the

nonempty equivalence classes of F is minimal, i.e. one. We conclude that the lattice of closed sets of a monodependent set of FDs is something in between an exchange and antiexchange lattice according to the cardinalities of its equivalence classes.

We also investigate some of the characteristics of the lattice $\mathcal{L}(F)$ when the set of FDs F satisfies the intersection property but not necessarily the split-freeness property. We give a characterisation of the intersection property in terms of the existence of certain distributive sublattices of $\mathcal{L}(F)$. We then present a polynomial time algorithm in the size of F to compute the set of *meet-irreducible* closed sets in $\mathcal{L}(F)$, which we denote by $\mathcal{M}(F)$ (see Definition 6.1). Let n be the cardinality of the attribute set of R . As a corollary of this algorithm we show that the cardinality of $\mathcal{M}(F)$ is at most $\binom{n}{n-2}$; in general, this number is exponential in n . Thus an Armstrong relation having $\binom{n}{n-2} + 1$ tuples can be generated [MANN86]. As an additional corollary of this algorithm we show that testing whether an attribute is *prime* (see Definition 4.2) when F satisfies the intersection property can be done in polynomial time in the size of F ; in general, testing whether an attribute is prime is NP-complete [LUCC78]. We also show that F satisfies the intersection property if and only if the cardinality of each element in $\mathcal{M}(F)$ is greater than or equal to $n - 2$. Another well known problem, which is NP-complete in the general case, is the problem of deciding whether there exists a superkey for R of cardinality k or less [LUCC78, DEME88]. Utilising this result we are able to show that this problem is still NP-complete when F satisfies the intersection property. Finally, we show that separatory sets of FDs are monodependent.

The layout of the rest of the paper is as follows. In Section 2 we formalise the notion of incomplete relations. In Section 3 we define the notion of a functional dependency being satisfied in an incomplete relation. In Section 4 we present the relevant properties of FDs which are utilised in the paper. In Section 5 we introduce monodependent sets of FDs and give some technical results, which are utilised in the following sections. In Section 6 we introduce the lattice-theoretic concepts that are used in the remaining sections. In Section 7 we give some negative results concerning the structure of $\mathcal{L}(F)$ when F is monodependent. In Section 8 we investigate the connection between exchange and antiexchange lattices of closed sets and monodependent sets of FDs. In Section 9 we investigate some of the characteristics of lattices of closed sets of FDs that satisfy the intersection property. In Section 10 we show that separatory sets of FDs are monodependent. Finally, in Section 11 we give our concluding remarks.

2 Relations that model incomplete information

Herein we formalise the notion of an incomplete relation, which allows us to model incomplete information of the form “value at present exists but is unknown”.

We use the notation $|S|$ to denote the cardinality of a set S . If S is a subset of T we write $S \subseteq T$ and if S is a proper subset of T we write $S \subset T$. Furthermore, S and T are *incomparable* if $S \not\subseteq T$ and $T \not\subseteq S$. At times we denote the singleton $\{A\}$ simply by A , and the union of two sets S and T , i.e. $S \cup T$, simply by ST . The power set of a set S is denoted by $\mathcal{P}(S)$.

Definition 2.1 (Relation schema and relation) A *relation schema* R is a finite set of attributes which we denote by $\text{schema}(R)$; we denote the cardinality of $\text{schema}(R)$ by $\text{type}(R)$. From now on we abbreviate $\text{schema}(R)$ to $\text{sch}(R)$.

We assume a countably infinite domain of constants, Dom , containing a distinguished constant unk , denoting the null value “unknown”.

A tuple over R is a total mapping t from $sch(R)$ into Dom such that $\forall A_i \in sch(R), t(A_i) \in Dom$. The *projection* of a tuple t over R onto a set of attributes $Y \subseteq sch(R)$, denoted by $t[Y]$, is the restriction of t to Y .

An *incomplete relation* (or simply a relation) over R is a finite set of tuples over R . A relation over R having no occurrences of unk is called a *complete relation*.

From now on we let R be a relation schema and r is a relation over R . As usual uppercase letters (which may be subscripted) from the end of the alphabet such as X, Y, Z will be used to denote sets of attributes, while those from the beginning of the alphabet such as A, B, D will be used to denote single attributes.

In [LEVE94, LEVE95a] we defined the semantics of incomplete relations in terms of *possible worlds* by defining a partial order, \sqsubseteq , in Dom , such that $u \sqsubseteq v$ if and only if either $u = unk$ or $u = v$, where $u, v \in Dom$. The partial order \sqsubseteq is extended to tuples over R in a natural way. The set of all possible worlds relative to r , denoted by $POSS(r)$, is the set of all complete relations that emanate from all possible substitutions of occurrences of unk in r by nonnull values in $Dom - \{unk\}$.

3 Functional dependencies in incomplete relations

Herein we formalise the notion a functional dependency being satisfied in an incomplete relation.

Definition 3.1 (Functional dependency) A *functional dependency* over R (or simply an FD) is a statement of the form $X \rightarrow Y$, where $X, Y \subseteq sch(R)$.

We call an FD of the form $X \rightarrow Y$, where $Y \subseteq X$, a *trivial* FD. Two nontrivial FDs of the forms $X \rightarrow A$ and $Y \rightarrow A$ are said to be *incomparable* if X and Y are incomparable. Two nontrivial FDs of the forms $XB \rightarrow A$ and $YA \rightarrow B$ are said to be *cyclic*.

We stress the fact that we allow FDs whose left-hand side is the empty set. From now on we let F be a set of FDs over R . We define the *size* of an FD $X \rightarrow Y$ to be $|X| + |Y|$, and the *size* of F , denoted by $||F||$, to be the sum of the sizes of all the FDs in F .

Definition 3.2 (Satisfaction of an FD) An FD $X \rightarrow Y$ is satisfied in a relation r , denoted by $r \models X \rightarrow Y$, whenever $\forall t_1, t_2 \in r$, if $\forall A \in X, t_1[A] \neq unk$ and $t_1[X] = t_2[X]$ then $\forall B \in Y$, either $t_1[B] = unk, t_2[B] = unk$ or $t_1[B] = t_2[B]$.

The reader can verify that when the relation, r , in Definition 3.2 is a complete relation then the definition of satisfaction of an FD in r reduces to the standard definition of satisfaction of an FD [ULLM88, MANN92, ATZE93]. It was shown in [LEVE94, LEVE95a] that $X \rightarrow Y$ is satisfied in r if and only if there exists a complete relation $s \in POSS(r)$ that satisfies the FD in the standard way.

Definition 3.3 (Logical implication) A set of FDs F over R *logically implies* an FD $X \rightarrow Y$, written $F \models X \rightarrow Y$, if whenever r is a relation over R then the following condition is true:

if $\forall W \rightarrow Z \in F, r \models W \rightarrow Z$ holds then $r \models X \rightarrow Y$ also holds.

4 Some properties of sets of functional dependencies

We assume that the reader is familiar with Armstrong's axiom system for FDs [ARMS74, ULLM88, MANN92, ATZE93], consisting of the inference rules: reflexivity, augmentation and transitivity. A fundamental result in relational database theory is that Armstrong's axiom system is sound and complete for FDs holding in complete relations. We denote the closure of a set of FDs F over R with respect to Armstrong's axiom system by F^+ . Lien [LIEN82], and Atzeni and Morfuni [ATZE86] have shown that the inference rules: reflexivity, augmentation, decomposition and union, are sound and complete for FDs holding in incomplete relations; we call this axiom system, Lien and Atzeni's axiom system. That is, by dropping the transitivity rule from Armstrong's axiom system and adding the decomposition and union rules, we obtain Lien and Atzeni's axiom system. We denote the closure of a set of FDs F over R with respect to Lien and Atzeni's axiom system by F^* . The soundness and completeness of Lien and Atzeni's axiom system for FDs holding in incomplete relations can be written symbolically as the statement: $F \models X \rightarrow Y$ if and only if $X \rightarrow Y \in F^*$.

The following useful property of derivations of FDs, using Armstrong's axiom system, which appears as Lemma 2 in [BEER79], will be used in subsequent proofs.

Proposition 4.1 Let F be a set of FDs and assume that $W \rightarrow Z \in F$ is used nonredundantly in a derivation of an FD $X \rightarrow Y \in F^+$ from F by using Armstrong's axiom system. Then $X \rightarrow W \in (F - \{W \rightarrow Z\})^+$. \square

Definition 4.1 (Closure of a set of attribute) The closure of a set of attributes $X \subseteq \text{sch}(R)$, with respect to a set of FDs F , denoted by $C_F(X)$ (or simply $C(X)$ whenever F is understood from context), is given by

$$C(X) = \bigcup \{Y \mid X \rightarrow Y \in F^+\}.$$

A set of attributes $X \subseteq \text{sch}(R)$ is *closed* with respect to F (or simply closed whenever F is understood from context) if $C_F(X) = X$.

We note that $C(X)$ can be computed in linear time in the size of F [BEER79]. In the sequel we will use the equivalent statements $Y \subseteq C_F(X)$ and $X \rightarrow Y \in F^+$, interchangeably.

Definition 4.2 (Superkey, key and antikey) A set of attributes $X \subseteq \text{sch}(R)$ is a *superkey* for R with respect to F (or simply a superkey for R whenever F is understood from context), if $C_F(X) = \text{sch}(R)$. A set of attributes $X \subseteq \text{sch}(R)$ is a *key* for R with respect to F (or simply a key for R whenever F is understood from context), if X is a superkey for R with respect to F and, in addition, for no proper subset $Y \subset X$, is it the case that Y is a superkey for R with respect to F . We denote the set of all keys for R with respect to F by $\mathcal{K}(F)$.

An attribute $A \in \text{sch}(R)$ is *prime* with respect to F (or simply prime whenever F is understood from context) if $A \in X$ for some $X \in \mathcal{K}(F)$; otherwise A is *nonprime* with respect to F .

An *antikey* for R with respect to F (or simply an antikey for R whenever F is understood from context) is a maximal subset X of $\text{sch}(R)$ such that X is not a superkey for R . We denote the set of all antikeys for R with respect to F by $\mathcal{A}(F)$.

Definition 4.3 (A cover of a set of FDs) A set of FDs G over R is a *cover* of F if $F^+ = G^+$.

By Definition 4.1 if G is a cover of a set of FDs F then $C_F(X) = C_G(X)$.

Definition 4.4 (Reduced and canonical sets of FDs) An FD $X \rightarrow Y \in F^+$ is *reduced* [BEER79] if there does not exist a set of attributes $W \subset X$ such that $W \rightarrow Y \in F^+$. A set of FDs F is *reduced* if all the FDs in F are reduced; F is *canonical* if it is reduced and the right-hand sides of all the FDs in F are singletons.

A reduced cover G of F can be obtained in polynomial time in the size of F [BEER79].

Definition 4.5 (A minimum set of FDs) A set of FDs F is a *minimum* [MAIE80] set of FDs if there is no cover G of F such that G has fewer FDs than F , all the FDs in F are reduced and for every FD $X \rightarrow Y \in F$ and for every $Z \subset Y$, $((F - \{X \rightarrow Y\}) \cup \{X \rightarrow Z\})^+ \neq F^+$.

In [MAIE80] a minimum set of FDs is called an *LR-minimum* set of FDs. Furthermore, a minimum cover G of a set of FDs F can be obtained in polynomial time in the size of F [MAIE80].

Definition 4.6 (An optimum set of FDs) A set of FDs F is an *optimum* [MAIE80, MANN83] set of FDs if there does not exist a cover G of F such that $\|G\| < \|F\|$. We denote an optimum cover of a set of FDs F by $\text{opt}(F)$.

In [MAIE80] it was shown that, in general, finding an optimum cover is NP-complete [MAIE80].

Definition 4.7 (Equivalent sets of attributes) Given a set of FDs F , the sets of attributes $X, Y \subseteq \text{sch}(R)$, are *equivalent* under F , if $X \rightarrow Y, Y \rightarrow X \in F^+$. We denote the subset of FDs in F whose left-hand sides are equivalent to a set of attributes $X \subseteq \text{sch}(R)$ by $E_F(X)$; we call the sets $E_F(X)$ the *equivalence classes* of F .

5 Monodependent Sets of Functional Dependencies

Given a set of FDs F and an incomplete relation r it is natural to say that r satisfies F if there is some complete relation, $s \in \text{POSS}(r)$, such that s satisfies each of the FDs in F . This gives rise to the *additivity problem*, which is the problem of whether the statement that r satisfies F is equivalent to the statement that r satisfies each FD in a reduced cover G of F [LEVE94, LEVE95a] (cf. [ATZE93]); if these two statements are equivalent for a class of incomplete relations and a class of sets of FDs then we say that *satisfaction is additive* with respect to these classes. If satisfaction is not additive, then F may be viewed as contradictory. Thus we consider the solution of the additivity problem to be an important prerequisite for any relational database system supporting FDs in the context of incomplete information, since otherwise semantic anomalies may arise.

Obviously satisfaction is additive with respect to the class of complete relations and the class of all sets of FDs. On the other hand, it is well known that satisfaction is *not* additive with respect to the class of incomplete relations and the class of all sets of FDs [ATZE86, LEVE94]. In [LEVE94] we introduced the class *monodependent* sets of FDs. Informally, a set of FDs F over R is monodependent if for each attribute $A \in \text{sch}(R)$, there is a unique nontrivial and reduced FD in F^+ that functionally determines A , and in addition there are no two nontrivial FDs in F^+ such that the right-hand side of each of the two FDs splits the left-hand side of the other FD. The main result in [LEVE94] shows that satisfaction is additive with respect to the class of all incomplete relations and a class of sets of FDs, **FC**, if and only if all the sets of FDs in **FC** are monodependent sets of FDs.

In [LEVE95b] we studied the impact on normalisation theory in relation databases of assuming that sets of FDs are monodependent, and in [LEVE95a] we extended the results in [LEVE94] to the wider class of sets of FDs and *unary inclusion dependencies* [COSM90].

Definition 5.1 (A monodependent set of FDs) A set of FDs F is a *monodependent set* of FDs over R (or simply monodependent whenever R is understood from context) if $\forall A \in \text{sch}(R)$, the following two conditions are true:

1. Whenever there exist incomparable FDs, $X \rightarrow A, Y \rightarrow A \in F^+$, then $X \cap Y \rightarrow A \in F^+$; we call this property the *intersection property*.
2. Whenever there exist cyclic FDs, $XB \rightarrow A, YA \rightarrow B \in F^+$, then either $Y \rightarrow B \in F^+$ or $(X \cap Y)A \rightarrow B \in F^+$; we call this property the *split-freeness property*.

An immediate consequence of the above definition is that if G is a cover of F then F is monodependent if and only if G is monodependent. In addition, we have shown in [LEVE94] that monodependence of a set of FDs F can be checked in polynomial time in the size of F .

We observe that the two defining properties of monodependent sets of FDs correspond to the two defining properties of *conflict-free* sets of *multivalued dependencies* (MVDs) [SCIO81, LIEN82, BEER86]. We further observe that the set of MVDs that is logically implied by a monodependent set of FDs may not be conflict-free and thus monodependence is a weaker notion than conflict-freeness. For example, let $F = \{A \rightarrow B, B \rightarrow A\}$, with $\text{sch}(R) = \{A, B, D\}$. It can easily be verified that R is monodependent but that the set of MVDs logically implied by R is not conflict-free.

The next theorem from [LEVE94] shows that if F satisfies the intersection property, then the closure of F with respect to Armstrong's axiom system (i.e. F^+) is equal to the closure of F with respect to Lien and Atzeni's axiom system (i.e. F^*). This result is fundamental to the theory of FDs in incomplete relations, since it justifies the use of Armstrong's axiom system in the context of incomplete relations when F is monodependent.

Theorem 5.1 If F is reduced and satisfies the intersection property then $F^+ = F^*$. \square

The converse of Theorem 5.1 is, in general, false. For example, let $F = \{A \rightarrow D, B \rightarrow D\}$, with $\text{sch}(R) = \{A, B, D\}$. It can be easily verified that $F^+ = F^*$, however, F does not satisfy the intersection property, since $\emptyset \rightarrow D \notin F^+$.

The following technical results, which are utilised in the sequel, are proved in [LEVE95b].

Lemma 5.2 Let F be a set of FDs that is minimum and satisfies the intersection property. Then $\forall A \in \text{sch}(R)$, there is at most one FD, $X \rightarrow Y \in F$, such that $A \in Y$. \square

Lemma 5.3 Let F be a set of FDs which is monodependent and minimum, and let $E_F(X)$ be an equivalence classes of F such that $|E_F(X)| > 1$. Then any two FDs in $E_F(X)$ are reduced and of the form, $WA \rightarrow B$ and $WB \rightarrow A$. \square

Lemma 5.4 Let F be a set of FDs which is monodependent and minimum, and let $E_F(X)$ and $E_F(V)$ be distinct and nonempty equivalence classes of F , with $W \rightarrow Z \in E_F(X)$. Then $\forall A \in WZ$, A does not appear in the right-hand side of any FD in $E_F(V)$. \square

Lemma 5.5 Let F be a set of FDs which is monodependent and minimum, and let $E = \{E_F(X_1), E_F(X_2), \dots, E_F(X_k)\}$ be the set of all nonempty equivalence classes of F . Then the number of keys for R is given by

$$|\mathcal{K}(F)| = \prod_{i=1}^k |E_F(X_i)|. \quad \square$$

Theorem 5.6 If a set of FDs F is minimum and satisfies the intersection property then it is also an optimum set of FDs. \square

An immediate result of Theorem 5.6 is that finding an optimum cover of a set of FDs which satisfies the intersection property can be computed in polynomial time. This is due to the fact that finding a minimum cover of a set of FDs can be computed in polynomial time [MAIE80, WILD95]. In general, when a set of FDs does not satisfy the intersection property, then finding an optimum cover is NP-complete [MAIE80]. We note that in [LEVE95b] we have also shown that the optimal cover of a set of FDs F that satisfies the intersection property is unique, implying that the minimum cover of F is also unique.

The next corollary follows from Theorem 5.6 Lemma 5.2, and the fact that if $X \rightarrow Y$ is an FD, with $X \cap Y = \emptyset$, then the size of $X \rightarrow Y$ is less than or equal to $\text{type}(R)$.

Corollary 5.7 If a set of FDs F is optimum and satisfies the intersection property, then $|F| \leq \text{type}(R)$ and $\|F\| \leq (\text{type}(R))^2$. \square

We close this section with an interesting result showing that monodependent sets of FDs which are also optimum are closed under the proper subset operation.

Proposition 5.8 Let F be a monodependent set of FDs and let $G = \text{opt}(F)$. Then $\forall H \subset G$, H is a monodependent and optimum set of FDs over R .

Proof. Let $H \subset G$. By Lemmas 5.3 and 5.4, and Proposition 4.1 we can deduce that $X \rightarrow A \in G^+$ and $A \in Z$ for some FD $W \rightarrow Z \in H$ if and only if $X \rightarrow A \in H^+$. We call this statement Observation 1. Therefore, H must satisfy the intersection property, since otherwise there must exist incomparable FDs $X \rightarrow A, Y \rightarrow A \in H^+$, but $X \cap Y \rightarrow A \in G^+ - H^+$, which contradicts Observation 1. Similarly, H must satisfy the split-freeness property, since otherwise there must exist cyclic FDs, $XB \rightarrow A, YA \rightarrow B \in H^+$, but either $Y \rightarrow B \in G^+ - H^+$ or $(X \cap Y)A \rightarrow B \in G^+ - H^+$, which again contradicts Observation 1.

Next, suppose that H is not optimum and that $J = \text{opt}(H)$, with $\|J\| < \|H\|$. Therefore, $(G - H) \cup J)^+ = G^+$. This leads to a contradiction that G is optimum, since $\|(G - H) \cup J\| < \|G\|$. The result that H is optimum follows. \square

6 The Lattice of Closed Sets

Herein we give the definitions of the lattice-theoretic concepts used in the rest of the paper. The reader is referred to [DAVE90] for an introduction to lattice theory and to [GRAT78] for more advanced material.

The operator C_F (see Definition 4.1) which closes sets of attributes in $\text{sch}(R)$ is a *closure operator* in the lattice-theoretic sense [DAVE90]. It follows by [DAVE90, Theorem 2.21] that the family of all the closed sets in the power set of $\text{sch}(R)$ is a lattice partially ordered by set inclusion, which we denote by $\mathcal{L}(F)$ (see also [DEME92, DEME93]). The lattice $\mathcal{L}(F)$ is, by definition, cover insensitive and thus G is a cover of F if and only if $\mathcal{L}(F) = \mathcal{L}(G)$. It is easy to see that $\mathcal{L}(F)$ is closed under intersection and thus the greatest lower bound of two closed sets in $\mathcal{L}(F)$ is just their intersection. On the other hand, it can be verified that the least upper bound, denoted by \sqcup , of two closed sets $X, Y \in \mathcal{L}(F)$ is given by $X \sqcup Y = C(X \cup Y)$. We refer the reader to [DEME92] for an in-depth investigation concerning the connection between the structure of a set of FDs and the type of lattice of closed sets it induces.

The following result shown in [DEME92, DEME93] shows the basic connection between a set of FDs F over R and its induced lattice of closed sets $\mathcal{L}(F)$.

Proposition 6.1 There is a one-to-one correspondence between F^+ and $\mathcal{L}(F)$. \square

Definition 6.1 (Meet-irreducible elements) A closed set $X \in \mathcal{L}(F)$ is *meet-irreducible* [DAVE90] if $\forall Y, Z \in \mathcal{L}(F), X = Y \cap Z$ implies that either $X = Y$ or $X = Z$. The family of all meet-irreducible closed sets in $\mathcal{L}(F)$ is denoted by $\mathcal{M}(F)$.

The following result shows the basic connection between $\mathcal{L}(F)$ and $\mathcal{M}(F)$ [BEER84, MANN86, WILD95].

Proposition 6.2 $\mathcal{M}(F)$ is the unique minimal subset of $\mathcal{L}(F)$ such that $X \in \mathcal{L}(F)$ if and only if X is the intersection of all the closed sets in $\mathcal{M}(F)$ that are supersets of X . \square

The following result, which was shown in [MANN86], gives an alternative characterisation of $\mathcal{M}(F)$.

Lemma 6.3 Let $\text{MAX}(\mathcal{F}, A)$ be the family of all the maximal closed sets $\mathcal{L}(\mathcal{F})$ such that $\forall X \in \text{MAX}(\mathcal{F}, A), A \notin X$. Then the following equality holds:

$$\mathcal{M}(\mathcal{F}) = \bigcup_{A \in \text{sch}(\mathcal{R})} \text{MAX}(\mathcal{F}, A). \quad \square$$

For completeness of the paper we include the definitions of the various types of lattices referred to hereafter. In particular, we define distributive, modular [GRAT78, DAVE90], semimodular [GRAT78], exchange [GRAT78] and antiexchange [JAMI85] lattices.

Definition 6.2 (Distributive lattice) $\mathcal{L}(\mathcal{F})$ is *distributive* if

$$\forall X, Y, Z \in \mathcal{L}(\mathcal{F}), X \cap (Y \sqcup Z) = (X \cap Y) \sqcup (X \cap Z).$$

Definition 6.3 (Semimodular and modular lattice) We say that X is *covered by* Y , denoted by $X \prec Y$, where $X, Y \in \mathcal{L}(\mathcal{F})$, if $X \subset Y$ and $X \subseteq Z \subset Y$ implies that $Z = X$, with $Z \in \mathcal{L}(\mathcal{F})$.

$\mathcal{L}(\mathcal{F})$ satisfies the *upper covering condition* if

$$\forall X, Y, Z \in \mathcal{L}(\mathcal{F}), X \prec Y \text{ implies that } X \sqcup Z \prec Y \sqcup Z \text{ or } X \sqcup Z = Y \sqcup Z.$$

The *lower covering condition* is the dual statement of the upper covering condition.

$\mathcal{L}(\mathcal{F})$ is *semimodular* if it satisfies the upper covering condition. $\mathcal{L}(\mathcal{F})$ is *modular* if it satisfies both the upper and lower covering conditions.

Definition 6.4 (Exchange property) $\mathcal{L}(\mathcal{F})$ satisfies the *exchange property* (or simply $\mathcal{L}(\mathcal{F})$ is exchange) whenever

$$\forall A, B \in \text{sch}(\mathcal{R}), \forall X \subseteq \text{sch}(\mathcal{R}), \text{ if } A, B \notin C(X) \text{ and } A \in C(XB) \text{ then } B \in C(XA).$$

Definition 6.5 (Antiexchange property) $\mathcal{L}(\mathcal{F})$ satisfies the *antiexchange property* (or simply $\mathcal{L}(\mathcal{F})$ is antiexchange) whenever

$$\forall A, B \in \text{sch}(\mathcal{R}), \forall X \subseteq \text{sch}(\mathcal{R}), \text{ if } A, B \notin C(X) \text{ and } A \in C(XB) \text{ then } B \notin C(XA).$$

The reader can also verify that the intersection property can be redefined as follows in terms of a property of the lattice $\mathcal{L}(\mathcal{F})$ of closed sets.

Definition 6.6 (Intersection property) Let \div be the symmetric difference operator, i.e. $X \div Y = (X - Y) \cup (Y - X)$, where $X, Y \subseteq \text{sch}(\mathcal{R})$. Then $\mathcal{L}(\mathcal{F})$ satisfies the *intersection property* if

$$C(X \cap Y) - (X \div Y) = (C(X) \cap C(Y)) - (X \div Y).$$

The reader can also verify that the split-freeness property can be redefined as follows in terms of a property of the lattice $\mathcal{L}(\mathcal{F})$ of closed sets.

Definition 6.7 (Split-freeness property) $\mathcal{L}(F)$ satisfies the *split-freeness property* if

$\forall A, B \in \text{sch}(R), \forall X, Y \subseteq \text{sch}(R)$, if $B \in C(YA), B \notin C(Y)$ and $B \notin C((X \cap Y)A)$ then $A \notin C(XB)$.

We say that a lattice $\mathcal{L}(F)$ embeds the figure \mathcal{N} , if $\exists W, X, Y, Z \in \mathcal{L}(F)$ such that $W \subset X, Y \subset X$ and $Y \subset Z$. It can be verified that if $\mathcal{L}(F)$ does not embed the figure \mathcal{N} , then $\mathcal{L}(F)$ satisfies the split-freeness property.

When F is a monodependent set of FDs, we say that $\mathcal{L}(F)$ is *monodependent*. We close this section by defining the concept of a sublattice.

Definition 6.8 (Sublattice) A subset $S \subseteq \mathcal{L}(F)$ is a *sublattice* [DAVE90] of $\mathcal{L}(F)$ if $X, Y \in S$ implies that both $X \sqcup Y \in S$ and $X \cap Y \in S$.

7 Counterexamples for monodependent set of FDs

Herein we present some negative results concerning the structure of $\mathcal{L}(F)$ when F is monodependent. We first show that $\mathcal{L}(F)$ may not even be semimodular and that a distributive lattice of closed sets may not be monodependent. We then show that, in general, the concepts of monodependence and exchange, and also of monodependence and antiexchange are incomparable.

Proposition 7.1 The lattice of closed sets of a monodependent set of FDs F is *not*, in general, semimodular.

Proof. Let R be a relation schema with $\text{sch}(R) = \{A, B, D\}$ and let $F = \{AB \rightarrow D, BD \rightarrow A\}$. It can easily be verified that F is monodependent. Furthermore, $\emptyset \prec B$ but it is *not* true that $(\emptyset \sqcup A = A) \prec (A \sqcup B = ABD)$. Therefore, $\mathcal{L}(F)$ is not semimodular. \square

In [WILD89] it was shown that when $\mathcal{L}(F)$ is modular then an optimum cover of F can be obtained in polynomial time in the size of F . By Proposition 7.1, Theorem 5.6 is incomparable with Wild's result, since $\mathcal{L}(F)$ may satisfy the intersection property but not be modular. Furthermore as the next example shows $\mathcal{L}(F)$ being modular does not imply that a minimum cover of F is also optimum.

Example 7.1 Let $F = \{D \rightarrow AB, E \rightarrow AB, AB \rightarrow DE\}$, with $\text{sch}(R) = \{A, B, D, E\}$. It can easily be verified that $\mathcal{L}(F)$ is modular and that F is minimum. On the other hand, F is not optimum, since it can be verified that $G = \{D \rightarrow AB, E \rightarrow D, AB \rightarrow DE\}$ is an optimum cover of F .

Proposition 7.2 Distributive lattices of closed sets are *not*, in general, monodependent.

Proof. We give two counterexamples of a relation schema R and a set of FDs F such that $\mathcal{L}(F)$ is distributive but F is not monodependent. In the first example F violates the intersection property and in the second example F violates the split-freeness property.

Counterexample 1. Let R be a relation schema with $\text{sch}(R) = \{A, B, D\}$ and let $F = \{B \rightarrow A, D \rightarrow A\}$. It can easily be verified that $\mathcal{L}(F)$ is distributive. Furthermore, the set of

FDs F is not monodependent, since it violates the intersection property due to the fact that $\emptyset \rightarrow A \notin F^+$.

Counterexample 2. Let R be a relation schema with $\text{sch}(R) = \{A, B, D\}$ and let $F = \{B \rightarrow AD, AD \rightarrow B\}$. It can easily be verified that $\mathcal{L}(F)$ is distributive. Furthermore, the set of FDs F is not monodependent, since it violates the split-freeness property due to the fact that both $A \rightarrow B \notin F^+$ and $D \rightarrow B \notin F^+$. \square

Proposition 7.3 The lattice of closed sets of a monodependent set of FDs F is *not*, in general, exchange.

Proof. Let R be a relation schema with $\text{sch}(R) = \{A, B\}$ and let $F = \{A \rightarrow B\}$. It can easily be verified that F is monodependent. Furthermore, $A, B \notin C(\emptyset)$ and $A \in C(B)$ but $B \notin C(A)$. \square

Proposition 7.4 Exchange lattices of closed sets are *not*, in general, monodependent.

Proof. Let R be a relation schema with $\text{sch}(R) = \{A, B, D\}$ and let $F = \{A \rightarrow B, B \rightarrow A, A \rightarrow D, D \rightarrow A\}$. It can easily be verified that $\mathcal{L}(F)$ is exchange. Furthermore, the set of FDs F is not monodependent, since it violates the intersection property. \square

Proposition 7.5 The lattice of closed sets of a monodependent set of FDs F is *not*, in general, antiexchange.

Proof. Let R be a relation schema with $\text{sch}(R) = \{A, B\}$ and let $F = \{A \rightarrow B, B \rightarrow A\}$. It can easily be verified that F is monodependent. Furthermore, $A, B \notin C(\emptyset)$, $A \in C(B)$ and also $B \in C(A)$. \square

Proposition 7.6 Antiexchange lattices of closed sets are *not*, in general, monodependent.

Proof. Let R be a relation schema with $\text{sch}(R) = \{A, B, D\}$ and let $F = \{B \rightarrow A, D \rightarrow A\}$. It can easily be verified that $\mathcal{L}(F)$ is antiexchange. Furthermore, the set of FDs F is not monodependent, since it violates the intersection property. \square

8 The connection between exchange and antiexchange lattices and monodependence

Herein we investigate the connection between exchange and antiexchange lattices of closed sets, and sets of FDs that satisfy the split-freeness property. We first show that if F satisfies the intersection property and $\mathcal{L}(F)$ is either exchange or antiexchange then F is monodependent. We then show that when F satisfies the intersection property then $\mathcal{L}(F)$ is exchange if and only if F satisfies the split-freeness property and the cardinality of all the nonempty equivalence classes of F is maximal, i.e. for each such equivalence class the said cardinality is the size of any FD in the class (see Lemma 5.3). Finally we show that when F satisfies the intersection property then $\mathcal{L}(F)$ is antiexchange if and only if F satisfies the split-freeness property and the cardinality of all the nonempty equivalence classes of F is minimal, i.e. the said cardinality is

one. We conclude that the structure of the lattice of closed sets of a monodependent set of FDs is something in between an exchange and antiexchange lattice according to the cardinalities of its equivalence classes.

Several properties of exchange and antiexchange lattices of closed sets have been investigated in [DEME92]. When $\mathcal{L}(F)$ is exchange then *Boyce-Codd normal form* [ULLM88, MANN92, ATZE93] can be characterised in terms of a *uniform closure*. In addition, if $\mathcal{L}(F)$ is exchange and $C(\emptyset) = \emptyset$, then second normal form and third normal form are equivalent. (See [ULLM88, MANN92, ATZE93] for the definitions of the various normal forms.) When $\mathcal{L}(F)$ is antiexchange then for every subset $X \subseteq \text{sch}(R)$, there is a unique reduced FD such that $Y \rightarrow X \in F^+$. In particular, when $\mathcal{L}(F)$ is antiexchange, then $|\mathcal{K}(F)| = 1$ [BISK91].

Lemma 8.1 Let F be a set of FDs that satisfies the intersection property. Then if $\mathcal{L}(F)$ is either exchange or antiexchange, then F satisfies the split-freeness property, i.e. F is monodependent.

Proof. Assume to the contrary that F does not satisfy the split-freeness property. Therefore, by Definition 5.1 there exist cyclic FDs, $XB \rightarrow A, YA \rightarrow B \in F^+$, but both $Y \rightarrow B \notin F^+$ and $(X \cap Y)A \rightarrow B \notin F^+$. We can assume without loss of generality that $XB \rightarrow A$ and $YA \rightarrow B$ are reduced FDs. Thus it is also the case that $X \rightarrow A \notin F^+$. Now, $Y \not\subseteq X$ holds, otherwise $(X \cap Y)A \rightarrow B$ is simply $YA \rightarrow B$, which is assumed to be in F^+ . There are two case to consider.

Firstly, assume that $X \subset Y$ and thus $YB \rightarrow A \in F^+$ but $XA \rightarrow B \notin F^+$, since the FDs are reduced. Thus $\mathcal{L}(F)$ is not exchange, since $A, B \notin C(X)$ and $A \in C(XB)$ but $B \notin C(XA)$. Furthermore, $\mathcal{L}(F)$ is not antiexchange, since $A, B \notin C(Y)$ and $B \in C(YA)$ but also $A \in C(YB)$.

Secondly, assume that X and Y are incomparable. Now, we have that $A, B \notin C(X)$ and $A \in C(XB)$. Assume that $\mathcal{L}(F)$ is exchange and thus $B \in C(XA)$. Thus, $XA \rightarrow B \in F^+$ and $YA \rightarrow B \in F^+$ are incomparable FDs. It follows that $(X \cap Y)A \rightarrow B \in F^+$ by the intersection property, which contradicts the fact that F does not satisfy the split-freeness property. Thus $B \notin C(XA)$ and $\mathcal{L}(F)$ is not exchange. Now, if $A \in C(XY)$, then $X \rightarrow A \in F^+$ by the intersection property, and similarly if $B \in C(XY)$, then $Y \rightarrow B \in F^+$ also by the intersection property. If $X \rightarrow A \in F^+$ then fact that $XB \rightarrow A$ is reduced is contradicted, and correspondingly, if $Y \rightarrow B \in F^+$ then the fact that $YA \rightarrow B$ is reduced is contradicted. So, we conclude that $A, B \notin C(XY)$. It follows that $\mathcal{L}(F)$ is not antiexchange, since $A \in C(XYB)$ but also $B \in C(XYA)$. The result that F satisfies the split-freeness property and is thus monodependent follows as required. \square

Theorem 8.2 Let F be a set of FDs that satisfies the intersection property, and let $E = \{E_G(X_1), E_G(X_2), \dots, E_G(X_k)\}$ be the set of all nonempty equivalence classes of G , where

$G = \text{opt}(F)$. Then the following statements are equivalent:

1. $\mathcal{L}(F)$ is exchange.
2. F satisfies the split-freeness property, i.e. F is monodependent, and $\forall E_G(X_i) \in E, |E_G(X_i)| = |XY|$, for some FD $X \rightarrow Y \in E_G(X_i)$.

Proof. (1 \Rightarrow 2.) By Lemma 8.1 \mathcal{F} satisfies the split-freeness property. Now, assume that for some $E_G(X_i) \in \mathcal{E}$ and $X \rightarrow Y \in E_G(X_i)$, $|E_G(X_i)| < |XY|$. Thus, by Lemma 5.3, $\exists A, B \in \text{sch}(\mathcal{R})$ such that $WB \rightarrow A \in E_G(X_i)$ but $WA \rightarrow B \notin E_G(X_i)$. Furthermore, also by Lemma 5.3, $W \rightarrow A, W \rightarrow B \notin F^+$, leading to a contradiction of the fact that $\mathcal{L}(\mathcal{F})$ is exchange.

(2 \Rightarrow 1.) Suppose to the contrary that $\mathcal{L}(\mathcal{F})$ is not exchange. Then for some set of attributes, $V \subset \text{sch}(\mathcal{R})$, $\exists A, B \in \text{sch}(\mathcal{R})$ such that $A, B \notin C(V)$, $A \in C(VB)$ but $B \notin C(VA)$. Thus $VB \rightarrow A \in F^+$ and there is some equivalence class $E_G(X) \in \mathcal{E}$ such that $X \rightarrow Y \in E_G(X)$, with $A \in Y$. If $|XY| = 1$, then $X = \emptyset$ and $V \rightarrow A \in F^+$, leading to a contradiction. So, it must be the case that $|XY| > 1$ and thus by Lemma 5.3 $Y = \{A\}$. Now, if $B \notin X$, then by the intersection property, it follows that $(X \cap VB) \rightarrow A \in F^+$ and thus $V \rightarrow A \in F^+$, since $B \notin (X \cap VB)$. So, it must be the case that $B \in X$ and thus $(X-B)A \rightarrow B \in E_G(X)$, since $|E_G(X)| = |XA|$. Let $W = (X-B) \cap V$. Then by the intersection property $WB \rightarrow A \in F^+$, with $W \subseteq V$. Furthermore, by Lemma 5.3, $X = WB$, since $X \rightarrow A$ is reduced. Thus $WA \rightarrow B \in E_G(X)$ and $VA \rightarrow B \in F^+$ leading to a contradiction. It follows that $B \in C(VA)$ and thus $\mathcal{L}(\mathcal{F})$ is exchange as required. \square

Theorem 8.3 Let \mathcal{F} be a set of FDs that satisfies the intersection property, and let $\mathcal{E} = \{E_G(X_1), E_G(X_2), \dots, E_G(X_k)\}$ be the set of all nonempty equivalence classes of \mathcal{G} , where

$\mathcal{G} = \text{opt}(\mathcal{F})$. Then the following statements are equivalent:

1. $\mathcal{L}(\mathcal{F})$ is antiexchange.
2. \mathcal{F} satisfies the split-freeness property, i.e. \mathcal{F} is monodependent, and $\forall E_G(X_i) \in \mathcal{E}$, $|E_G(X_i)| = 1$.

Proof. (1 \Rightarrow 2.) By Lemma 8.1 \mathcal{F} satisfies the split-freeness property. Furthermore, by [JAMI85, DEME92] $|\mathcal{K}(\mathcal{F})| = 1$, since $\mathcal{L}(\mathcal{F})$ is antiexchange. Now, assume that for some $E_G(X_i) \in \mathcal{E}$, $|E_G(X_i)| > 1$. Thus, by Lemma 5.5 $|\mathcal{K}(\mathcal{F})| > 1$, leading to a contradiction of the fact that $\mathcal{L}(\mathcal{F})$ is antiexchange.

(2 \Rightarrow 1.) Suppose to the contrary that $\mathcal{L}(\mathcal{F})$ is not antiexchange. Then for some set of attributes, $V \subset \text{sch}(\mathcal{R})$, $\exists A, B \in \text{sch}(\mathcal{R})$ such that $A, B \notin C(V)$, $A \in C(VB)$ but also $B \in C(VA)$. Thus $VA \rightarrow B, VB \rightarrow A \in F^+$. Now, there exist equivalence classes $E_G(X_1), E_G(X_2) \in \mathcal{E}$ such that $X \rightarrow Y \in E_G(X_1)$, with $A \in Y$, and $W \rightarrow Z \in E_G(X_2)$, with $B \in Z$.

Assume that $E_G(X_1) = E_G(X_2)$. There are two cases to consider. Firstly, if $X \rightarrow Y$ and $W \rightarrow Z$ are distinct FDs then $|E_G(X_1)| > 1$, leading to a contradiction. Secondly, if $X \rightarrow Y$ and $W \rightarrow Z$ are in fact the same FD, then $X \rightarrow AB \in F^+$, with $A, B \notin X$. Now, $VA \not\subseteq X$, since $A \notin X$, and also $X \not\subseteq VA$, since otherwise $X \subseteq V$ and $V \rightarrow A \in F^+$ leading to a contradiction. Therefore, $X \rightarrow B \in F^+$ and $VA \rightarrow B \in F^+$ are incomparable FDs. It follows by the intersection property that $(X \cap VA) \rightarrow B \in F^+$, with $(X \cap VA) \subseteq V$. Therefore, $V \rightarrow B \in F^+$ again leading to a contradiction.

So, assume that $E_G(X_1)$ and $E_G(X_2)$ are distinct equivalence classes of \mathcal{G} . Thus, by Lemma 5.4, $A \notin WZ$ and $B \notin XY$. Now, $VA \not\subseteq W$, since $A \notin W$, and also $W \not\subseteq VA$, since otherwise $W \subseteq V$ and $V \rightarrow B \in F^+$ leading to a contradiction. Therefore, $W \rightarrow B \in F^+$ and $VA \rightarrow B \in F^+$ are incomparable FDs. It follows by the intersection property that $(W \cap VA) \rightarrow B \in F^+$, with $(W \cap VA) \subseteq V$. Therefore, $V \rightarrow B \in F^+$ again leading to a contradiction. It thus follows that $B \notin C(VA)$ and thus $\mathcal{L}(\mathcal{F})$ is antiexchange as required. \square

9 Characteristics of lattices satisfying the intersection property

Herein we investigate some of the characteristics of lattices of closed sets of FDs that satisfy the intersection property. We first utilise the concept of an interval, which is defined below, to investigate how the lattice of closed sets changes from one that does not necessarily satisfy the intersection property to one that does (cf. [BURO87]). We then give a characterisation of the intersection property in terms of the existence of certain distributive sublattices of $\mathcal{L}(F)$. We also present a polynomial time algorithm in the size of F in order to compute the set of meet-irreducible closed sets, $\mathcal{M}(F)$, when F satisfies the intersection property. As a corollary of this algorithm we show that when F satisfies the intersection property then the cardinality of $\mathcal{M}(F)$ is at most $\binom{\text{type}(R)}{\text{type}(R)-2}$. As an additional corollary of this algorithm we show that testing whether an attribute is prime can be done in polynomial time in the size of F , when F satisfies the intersection property; in general, the problem of testing whether an attribute is prime is known to be NP-complete [LUCC78].

Definition 9.1 (Intersection property descriptor) The *intersection property descriptor* of a set of FDs F over a relation schema R , denoted by $\mathcal{I}(F)$, is defined by

$$\mathcal{I}(F) = \{X \cap Y \rightarrow A \mid \text{there exist incomparable FDs, } X \rightarrow A, Y \rightarrow A \in F^+, \text{ but } X \cap Y \rightarrow A \notin F^+\}.$$

The next lemma, which characterises the lattice of closed sets of a set of FDs that satisfies the intersection property, follows from Definition 6.6 and [DEME92, Theorem 3.1]. We begin by defining the concept of an interval.

Definition 9.2 (Interval) The *interval* between X and Y , where $X \subseteq Y \subseteq \text{sch}(R)$, denoted by $[X, Y]$, is given by $[X, Y] = \{Z \mid X \subseteq Z \subseteq Y\}$.

Lemma 9.1 Given a set of FDs F , let $G = F \cup \mathcal{I}(F)$. Then the lattice $\mathcal{L}(G)$ of closed sets of G is given by

$$\begin{aligned} \mathcal{L}(G) &= \mathcal{L}(F) - \bigcup_{X \cap Y \rightarrow A \in \mathcal{I}(F)} [X \cap Y, \text{sch}(R) - A] \\ &= \mathcal{L}(F) - \bigcup_{X \cap Y \rightarrow A \in \mathcal{I}(F)} \bigcup_{\substack{B \in (X-Y) \\ D \in (Y-X)}} [X \cap Y, \text{sch}(R) - ABD]. \quad \square \end{aligned}$$

A closed set $X \in S$, where $S \subseteq \mathcal{L}(F)$, is *maximum* if $\forall Y \in S, Y \subseteq X$.

Lemma 9.2 Let F be a set of FDs and \mathcal{H} be the family of closed sets defined by

$$\mathcal{H} = \mathcal{L}(F) \cap \bigcup_{X \cap Y \rightarrow A \in \mathcal{I}(F)} \bigcup_{\substack{B \in (X-Y) \\ D \in (Y-X)}} [X \cap Y, \text{sch}(R) - ABD].$$

Then all the maximum elements in \mathcal{H} are in $\mathcal{M}(F)$, i.e. all the maximum elements in \mathcal{H} are meet-irreducible closed sets in $\mathcal{L}(F)$.

Proof. Let X be a maximum element of \mathcal{H} and $Y, Z \in \mathcal{L}(F)$ be two closed sets such that $X = Y \cap Z$. We need to show that either $X = Y$ or $X = Z$. Suppose to the contrary that $X \neq Y$ and $X \neq Z$ and thus both $X \subset Y$ and $X \subset Z$ hold. Now, by Lemma 9.1, $Y, Z \in \mathcal{L}(G)$, where $G = F \cup \mathcal{I}(F)$. A contradiction has arisen, since it must be the case that $X \in \mathcal{L}(G)$, due to the fact that $\mathcal{L}(G)$ is closed under intersection. \square

Definition 9.3 (The family of left-hand sides of a set of FDs) The *family of left-hand sides* of a set of FDs F with respect to $A \in \text{sch}(\mathbf{R})$, denoted by $F(A)$, is defined by

$$F(A) = \{X \mid X \rightarrow A \in F^+ \text{ is a nontrivial FD}\}.$$

The schema of $F(A)$, denoted by $\text{sch}(F(A))$, is defined by

$$\text{sch}(F(A)) = \bigcup \{X \mid X \in F(A)\}.$$

We observe that $F(A) \subseteq \mathcal{P}(\text{sch}(\mathbf{R})) - \mathcal{L}(F)$. In other words, the family of left-hand sides of F with respect to A is a subset of the complement of the lattice of closed sets of F .

Definition 9.4 (Lattice of sets) A *lattice of sets* over a finite set S is a subset of the power set, $\mathcal{P}(S)$, which is closed under union and intersection [DAVE90].

The *if part* of the next theorem follows from the definition of the intersection property and the *only if part* of the theorem follows from Definition 9.3.

Theorem 9.3 A set of FDs F satisfies the intersection property if and only if $\forall A \in \text{sch}(\mathbf{R})$, $F(A)$ is a lattice of sets over $\text{sch}(F(A))$. \square

An immediate consequence of Theorem 9.3 is that $F(A)$ is distributive, since it is well known that a lattice of sets over S is distributive [DAVE90].

We now give the pseudo-code of an algorithm, designated MEET_IRR(F), which will be shown to return the family $\mathcal{M}(F)$ of meet-irreducible closed sets of $\mathcal{L}(F)$, where F satisfies the intersection property.

Algorithm 1 (MEET_IRR(F))

```

1. begin
2.   Meet_irr :=  $\emptyset$ ;
3.    $G := \text{opt}(F)$ ;
4.   for each  $A \in \text{sch}(\mathbf{R})$  do
5.     if  $\nexists X \rightarrow Y \in G$ , with  $A \in Y$  then
6.       Meet_irr := Meet_irr  $\cup$   $\{\text{sch}(\mathbf{R}) - A\}$ ;
7.     else
8.       let  $X \rightarrow Y$  be the FD in  $G$  with  $A \in Y$ ;
9.       for each  $B \in X$  do
10.        Meet_irr := Meet_irr  $\cup$   $(\text{sch}(\mathbf{R}) - AB)$ ;
11.      end for
12.    end if
13.  end for
14.  return Meet_irr;
15. end.

```

On using Corollary 5.7 the reader can verify that Algorithm 1 executes in polynomial time in $\text{type}(\mathbf{R})$. The next theorem establishes the correctness of Algorithm 1.

Theorem 9.4 If F is a set of FDs that satisfies the intersection property, then Algorithm 1 returns $\mathcal{M}(F)$.

Proof. We need to show that $M = \mathcal{M}(F)$, where $M = \text{MEET_IRR}(F)$ is the set returned by Algorithm 1.

$M \subseteq \mathcal{M}(F)$. Let $W \in M$. By Lemma 6.3 it remains to show that $W \in \text{MAX}(F, A)$ for some $A \in \text{sch}(\mathbf{R})$. Consider the *for loop* beginning at line 4 and ending at line 13, with $A \in \text{sch}(\mathbf{R})$. If W was added to M at line 6, then the condition of the *if statement* beginning at line 5 is true, and obviously $W = \text{sch}(\mathbf{R}) - A \in \text{MAX}(F, A)$. Otherwise, let $W = \text{sch}(\mathbf{R}) - AB$ be the set added to M at line 10. Now, $W \rightarrow A \notin F^+$, otherwise by the intersection property $W \cap X \rightarrow A \in F^+$, with $|W \cap X| < |X|$, contradicting the fact that G is optimum. Furthermore, W is a maximal set of attributes such that $W \rightarrow A \notin F^+$, since $X \subseteq WB$.

$\mathcal{M}(F) \subseteq M$. Let $W \in \mathcal{M}(F)$. Then by Lemma 6.3 $W \in \text{MAX}(F, A)$ for some $A \in \text{sch}(\mathbf{R})$. It remains to show that $W \in M$. Suppose to the contrary that $W \notin M$.

Consider the *if statement* beginning at line 5 and ending at line 12. There are two cases to consider. Firstly the condition of line 5 is true and thus $\exists X \rightarrow Y \in G$, with $A \in Y$. It follows that $W \subset \text{sch}(\mathbf{R}) - A$ and thus it is not a maximal subset of $\text{sch}(\mathbf{R})$ such that $W \rightarrow A \notin F^+$, contradicting the fact that $W \in \text{MAX}(F, A)$. Secondly the condition of line 5 is false and thus by Lemma 5.2 there is a unique $X \rightarrow Y \in F$, with $A \in Y$. Let $X \rightarrow Y$ be the FD, with $A \in Y$, that is chosen in line 8. Therefore, by the *for loop* beginning at line 9 and ending at line 11, it follows that either $X \subset W$ or $W \subset Z$, for some $Z \in M$, due to the fact that $|W| \leq \text{type}(\mathbf{R}) - 2$ and $A \notin W$. Both cases lead to a contradiction of the fact that $W \in \text{MAX}(F, A)$. The result that $M = \mathcal{M}(F)$ follows. \square

The next corollary, which gives a polynomial upper bound in $\text{type}(\mathbf{R})$ for the cardinality of $\mathcal{M}(F)$, is an immediate consequence of Theorem 9.4 on inspecting Algorithm 1. In general, when a set of FDs does not satisfy the intersection property, the cardinality of $\mathcal{M}(F)$ may be exponential in $\text{type}(\mathbf{R})$ [BEER84, MANN86].

Corollary 9.5 If a set of FDs F satisfies the intersection property, then $|\mathcal{M}(F)| \leq \binom{\text{type}(\mathbf{R})}{\text{type}(\mathbf{R})-2}$. \square

An immediate consequence of Corollary 9.5 is that an Armstrong relation having $\binom{\text{type}(\mathbf{R})}{\text{type}(\mathbf{R})-2} + 1$ tuples can be generated [MANN86]. The following result, which is immediate from Theorem 9.4 and Proposition 5.8, shows that when removing an FD $X \rightarrow Y$ from an optimum set of FDs that satisfies the intersection property, then the sets $\text{MAX}(F, A)$ attain their simplest structure.

Corollary 9.6 Let F be set of FDs that is optimum and satisfies the intersection property. Then $\forall X \rightarrow Y \in F, \forall A \in Y, \text{MAX}(F - \{X \rightarrow Y\}, A) = \{\text{sch}(\mathbf{R}) - A\}$. \square

In general, the problem of deciding whether an attribute $A \in \text{sch}(\mathbf{R})$ is prime with respect to F is known to be NP-complete [LUCC78]. Our next result shows that when F satisfies the intersection property this problem, known as the prime attribute problem, can be decided in polynomial time in the size of F .

Corollary 9.7 If a set of FDs F satisfies the intersection property, then deciding whether an attribute $A \in \text{sch}(R)$ is prime can be solved in polynomial time in the size of F .

Proof. By [MANN89, Theorem 2] an attribute $A \in \text{sch}(R)$ is prime with respect to F if and only if for some $W \in \text{MAX}(F, A)$, $C(WA) = \text{sch}(R)$; recall that $C(WA)$ can be computed in linear time in the size of F [BEER79]. Furthermore, by [MANN89, Lemma 1], given a set of attributes $X \subseteq \text{sch}(R)$, testing whether $X \in \text{MAX}(F, A)$ can be done in polynomial time in the size of F . The result now follows by Corollary 9.5 on using Algorithm 1 to compute $\mathcal{M}(F)$. \square

The next theorem gives a characterisation of the intersection property in terms of the cardinality of the elements in $\mathcal{M}(F)$.

Theorem 9.8 A set of FDs F satisfies the intersection property if and only if $\forall X \in \mathcal{M}(F)$, $|X| \geq \text{type}(R) - 2$.

Proof. The *only if* part of the theorem is an immediate consequence of Theorem 9.4 on inspecting Algorithm 1.

We prove the *if* part of the theorem by contraposition. Suppose that F violates the intersection property. Therefore, for some $A \in \text{sch}(R)$, there exist incomparable FDs, $X \rightarrow A$, $Y \rightarrow A \in F^+$, but $X \cap Y \rightarrow A \notin F^+$. We can assume without loss of generality that $X \rightarrow A$ and $Y \rightarrow A$ are reduced FDs.

Now, since X and Y are incomparable there is an attribute $B \in X - Y$ and an attribute $D \in Y - X$. Let $W = \text{sch}(R) - ABD$, and thus $|W| = \text{type}(R) - 3$. There are two cases to consider.

Firstly, $W \rightarrow A \in F^+$ and thus there exists some $Z \subset W$ such that $Z \in \text{MAX}(F, A)$; note that $\emptyset \rightarrow A \notin F^+$, since $X \cap Y \rightarrow A \notin F^+$. The result now follows by Lemma 6.3.

Secondly, $W \rightarrow A \notin F^+$ and thus by the construction of W we have that $W \in \text{MAX}(F, A)$. The result now follows by Lemma 6.3. \square

The next proposition establishes which meet-irreducible elements of $\mathcal{L}(F)$ are antikeys (in [DEME92] antikeys are called *coatoms*).

Proposition 9.9 A set of attributes $X \subset \text{sch}(R)$ is an antikey for R if and only if X is a maximal set in $\mathcal{M}(F)$. \square

It follows from Theorem 9.4 on using Proposition 9.9 that the set of antikeys for R , i.e. $\mathcal{A}(F)$, can be computed in polynomial time in $\text{type}(R)$; in general, computing $\mathcal{A}(F)$ can only be done in exponential time in $\text{type}(R)$ [THI86]. We observe that an alternative proof to Corollary 9.7 can utilise a result in [DEME87] which states that an attribute $A \in \text{sch}(R)$ is nonprime if and only if A is a member of the intersection of all the antikeys in $\mathcal{A}(F)$.

The next result establishes the connection between superkeys and antikeys [DEME88].

Proposition 9.10 A set of attributes $X \subset \text{sch}(R)$ is a superkey for R if and only if $\forall Y \in \mathcal{A}(F)$, $X \not\subseteq Y$. \square

In [DEME88] Proposition 9.10 is used to derive an algorithm which computes the set of all keys $\mathcal{K}(F)$ given the set of all antikeys $\mathcal{A}(F)$.

In general, the problem of finding a superkey for R with respect to F , whose cardinality is less than or equal to a natural number k , is known to be NP-complete [LUCC78, DEME88]. Our next result shows that this problem, known as the superkey of cardinality k problem, is still NP-complete when F satisfies the intersection property

Theorem 9.11 The superkey of cardinality k problem is NP-complete, when F is a set of FDs that satisfies the intersection property.

Proof. The problem is known to be in NP [LUCC78]. It remains to show that the problem is NP-hard.

By [DEME88, Lemma 2.4] the vertex cover problem, which is known to be NP-complete [GARE79], can be reduced to the following problem. Given a set of antikeys for R , say \mathbf{S} , such that $\forall X \in \mathbf{S}, |X| = \text{type}(R) - 2$, solve the superkey of cardinality k problem.

By the remark made after Proposition 9.10 it follows that \mathbf{S} can be used to derive the set of keys $\mathcal{K}(F)$ for some set of FDs F over R . Furthermore, we can assume, without loss of generality, that the set $\{X \rightarrow \text{sch}(R) \mid X \in \mathcal{K}(F)\}$ is a cover of F . It follows by Proposition 9.9 that $\mathbf{S} = \mathcal{M}(F)$. The result that F satisfies the intersection property now follows by Theorem 9.8. \square

It is interesting to note that when F is monodependent then the superkey of cardinality k problem can be solved in polynomial time in the size of F [LEVE95b]. This is a corollary of the fact that when F satisfies the split-freeness property then all the keys for R have the same cardinality [LEVE95b].

10 Separatory sets of FDs are monodependent

Several properties of separatory lattices of closed sets are investigated in [DEME92]. In particular when $\mathcal{L}(F)$ is separatory, then $|\mathcal{K}(F)| = 1$ and F has a cover whose cardinality is at most $(\text{type}(R))^2$ [BISK91, DEME92]. Herein we show that separatory sets of FDs are monodependent. We also give an example of a set of FDs which is monodependent but not separatory.

Definition 10.1 (Separatory set of FDs) A set of FDs F is *separatory* [DEME92] if it has a cover of the form $\{X_1 \rightarrow A_1, X_2 \rightarrow A_2, \dots, X_m \rightarrow A_m\}$, where $X_1 \subseteq X_2 \subseteq \dots \subseteq X_m$. We let $\text{RHS}(F)$ denote the set $\{A_1, A_2, \dots, A_m\}$.

The next lemma is useful in proving the ensuing theorem.

Lemma 10.1 A set of FDs is separatory if and only if it has a canonical cover F of the form $\{X_1 \rightarrow A_1, X_2 \rightarrow A_2, \dots, X_m \rightarrow A_m\}$, where $X_1 \subseteq X_2 \subseteq \dots \subseteq X_m$ and $\forall i \in \{1, 2, \dots, m\}, X_i \cap \text{RHS}(F) = \emptyset$.

Proof. We can assume without loss of generality that $\forall A \in \text{sch}(R)$, F does not contain distinct FDs of the form $X \rightarrow A$ and $Y \rightarrow A$. If this were the case then one of $X \rightarrow A$ or $Y \rightarrow A$ is redundant, since either $X \subset Y$ or $Y \subset X$. Next, let $X \rightarrow A$ be an FD in F .

Claim 1. The FD $X \rightarrow A$ is *not* reduced but $Y \rightarrow A \in F^+$ is reduced, with $Y \subset X$, if and only if $X = YZ$, where $Z \neq \emptyset$, $Z \subseteq \text{RHS}(F)$ and $Y \cap \text{RHS}(F) = \emptyset$. (We observe that $|F| > |Z|$.)

For the *if part* of the claim let $Z = \{B_1, B_2, \dots, B_k\}$, with $k > 0$. We use an induction on k to prove the result. For the basis step assume that $k = 1$ and thus $X = YB_1$. It follows that $W \subseteq Y$, where $W \rightarrow B_1 \in F$, since F is separatory, and also that $YB_1 \not\subseteq W$. Therefore, on using Armstrong's axiom system, $B_1 \in C(Y)$ and thus $A \in C(Y)$ also. Furthermore, $Y \rightarrow A$ is reduced, since $Y \cap \text{RHS}(F) = \emptyset$.

For the induction step assume that the result holds when $|Z| = k$, with $k \geq 1$; we then need to prove that the result holds when $|Z| = k+1$. Let $V = Y(Z-B_k)$. It follows that $W \subseteq V$, where $W \rightarrow B_k \in F$, since F is separatory, and also that $V \not\subseteq W$. Therefore, on using Armstrong's axiom system, $B_k \in C(V)$ and thus $A \in C(V)$ also. The result follows by inductive hypothesis.

For the *only if part* of the claim consider a nonredundant derivation of $Y \rightarrow A$ from F that uses n FDs, with $n > 0$, in the following order: $Y_1 \rightarrow B_1, Y_2 \rightarrow B_2, \dots, Y_n \rightarrow B_n$ and $X \rightarrow A$, all of which are in F . It follows that $Y_1 \subseteq Y$, $Y_2 \subseteq YB_1$, \dots , $Y_n \subseteq YB_1B_2 \dots B_{n-1}$ and finally $X \subseteq YB_1B_2 \dots B_n$. Therefore, since $Y \subset X$, we have that $X = YZ$, where $Z \subseteq \text{RHS}(F)$. It remains to show that $Y \cap \text{RHS}(F) = \emptyset$. Suppose that this is not the case and hence there is an attribute $B \in Y \cap \text{RHS}(F)$. Thus there is an FD $W \rightarrow B \in F$, with $W \subseteq Y-B$, since F is separatory, and $Y \not\subseteq W$. Therefore, on using Armstrong's axiom system, $B \in C(Y-B)$, contradicting the fact that $Y \rightarrow A$ is reduced. The claim now follows.

From Claim 1 it follows that we can rewrite $X_1 \subseteq X_2 \subseteq \dots \subseteq X_m$ as $Y_1Z_1 \subseteq Y_2Z_2 \subseteq \dots \subseteq Y_mZ_m$, where $\forall i \in \{1, 2, \dots, m\}$, $Y_i \cap \text{RHS}(F) = \emptyset$ and $Z_i \subseteq \text{RHS}(F)$. The result now follows, since by Claim 1 $\{Y_1 \rightarrow A_1, Y_2 \rightarrow A_2, \dots, Y_m \rightarrow A_m\}$ is a reduced cover of F , with $Y_1 \subseteq Y_2 \subseteq \dots \subseteq Y_m$. \square

The lattice $\mathcal{L}(F)$ of closed sets is said to be *separatory* if $\mathcal{P}(\text{sch}(R)) - \mathcal{L}(F)$ is a *semilattice*, i.e. if it is closed under intersection [GOTT90, LIBK92]. It was shown in [DEME92, Proposition 6.10] that a set of FDs F is separatory if and only if the lattice of closed sets $\mathcal{L}(F)$ is separatory. The next result shows that separatory sets of FDs are also monodependent.

Theorem 10.2 If a set of FDs F is separatory, then it is monodependent.

Proof. Assume by Lemma 10.1 that F is canonical and has the form $\{X_1 \rightarrow A_1, X_2 \rightarrow A_2, \dots, X_m \rightarrow A_m\}$, where $X_1 \subseteq X_2 \subseteq \dots \subseteq X_m$ and $\forall i \in \{1, 2, \dots, m\}$, $X_i \cap \text{RHS}(F) = \emptyset$.

By [DEME92, Proposition 6.10] $\mathcal{P}(\text{sch}(R)) - \mathcal{L}(F)$ is a semilattice. Let $A \in \text{sch}(R)$. It remains to show that $F(A)$ is a sublattice of $\mathcal{P}(\text{sch}(R)) - \mathcal{L}(F)$, whereupon by Theorem 9.3 F satisfies the intersection property. If there is no FD in F of the form $X \rightarrow A$ then the result follows, since $F(A) = \emptyset$. So, let $X \rightarrow A$ in F be the reduced FD whose right-hand side is A .

We claim that if $Y \rightarrow A \in F^+$ is a nontrivial FD, then $X \subseteq Y$. Suppose that $X \not\subseteq Y$ and that $B \in X-Y$. By Proposition 4.1 it must be the case that $Y \rightarrow X \in F^+$ and thus $Y \rightarrow B \in F^+$ is a nontrivial FD. However, by Claim 1 in the proof of Lemma 10.1 it is also true that $B \notin \text{RHS}(F)$ and therefore there cannot be a nontrivial FD in F^+ whose right-hand side is B , which leads to a contradiction. The result that F satisfies the intersection property now follows, since we have that $F(A) = [X, \text{sch}(R)-A]$.

It remains to show that F satisfies the split-freeness property. Suppose to the contrary that there exist cyclic FDs $XB \rightarrow A$, $YA \rightarrow B \in F^+$, but $Y \rightarrow B \notin F^+$ and $(X \cap Y)A \rightarrow B \notin F^+$. We assume without any loss of generality that $YA \rightarrow B$ is reduced. Now, by Claim 1 in the proof of Lemma 10.1 it follows that there is an FD $W \rightarrow A \in F$ such that $B \notin W$, due to the fact that $XB \rightarrow A \in F^+$ is a nontrivial FD. Therefore, on using Armstrong's axiom system, $WY \rightarrow B \in F^+$. It follows by the intersection property that $Y \rightarrow B \in F^+$, since $A \notin W$. Hence $YA \rightarrow B$ is not reduced leading to a contradiction of our assumption. The result that F satisfies the split-freeness property follows as required. \square

As the following example shows a set of FDs may be monodependent but not separatory.

Example 10.1 Let $F = \{A \rightarrow B, D \rightarrow E\}$, with $\text{sch}(R) = \{A, B, D, E\}$. It can easily be verified that F is monodependent but not separatory.

11 Concluding Remarks

Monodependence is a desirable property of sets of FDs when assuming that relations may be incomplete. We have investigated the structure of the lattice of closed sets $\mathcal{L}(F)$ when F is monodependent. As a consequence of this investigation we have shown that monodependent sets of FDs give rise to several desirable properties. Moreover, several difficult problems in relational database theory become tractable when F is monodependent. The connection between lattice theory and relational database theory is important, since it provides us with additional insight into the semantics of data dependencies such as FDs. A lattice-theoretic investigation of MVDs was carried out in [DAY93]. We conclude by giving a brief summary of the main results.

Assume that F satisfies the intersection property. In Theorem 8.2 we show that $\mathcal{L}(F)$ is exchange if and only if the cardinality of all the nonempty equivalence classes of F is maximal. On the other hand, in Theorem 8.3 we show that $\mathcal{L}(F)$ is antiexchange if and only if the cardinality of all the nonempty equivalence classes of F is minimal, i.e. it is one.

In Theorem 9.3 we give a characterisation of the intersection property in terms of the existence of certain distributive sublattices of $\mathcal{L}(F)$. In Corollary 9.5 we show that the cardinality of $\mathcal{M}(F)$ is at most m , where $m = \binom{\text{type}(R)}{\text{type}(R)-2}$. Thus an Armstrong relation having $m + 1$ tuples can be generated. In Corollary 9.7 we show that the prime attribute problem can be solved in polynomial time in the size of F . In Theorem 9.8 we show that F satisfies the intersection property if and only if the cardinality of each element in $\mathcal{M}(F)$ is greater than or equal to $\text{type}(R) - 2$. Using this result we are able to show in Theorem 9.11 that the superkey of cardinality k problem is still NP-complete, when F is restricted to be a set of FDs that satisfies the intersection property. Finally, in Theorem 10.2 we show that separatory sets of FDs are monodependent.

References

- [ARMS74] W.W. Armstrong, Dependency structures of data base relationships. In: *Proceedings of the IFIP Congress*, Stockholm, pp. 580-583, 1974.

- [ATZE86] P. Atzeni and N.M. Morfuni, Functional dependencies and constraints on null values in database relations. *Information and Control* **70**, (1986), 1-31.
- [ATZE93] P. Atzeni and V. De Antonellis, *Relational Database Theory*. Benjamin/Cummings, Redwood City, CA., 1993.
- [BEER79] C. Beeri and P.A. Bernstein, Computational problems related to the design of normal form relational schemas. *ACM Transactions on Database Systems* **4**, (1979), 30-59.
- [BEER84] C. Beeri, M. Dowd, R. Fagin and R. Statman, On the structure of Armstrong relations for functional dependencies. *Journal of the ACM* **31**, (1984), 30-46.
- [BEER86] C. Beeri and M. Kifer, Elimination of intersection anomalies for database schemes. *Journal of the ACM* **33**, (1986), 423-450.
- [BISK91] J. Biskup, J. Demetrovics, L.O. Libkin and I.B. Muchnik, On relational database schemes having unique minimal key. *Journal of Information Processing Cybernetics* **27**, (1991), 217-225.
- [BURO87] G. Burosch, J. Demetrovics and G.O.H. Katona, The poset of closures as a model of changing databases. *Order* **4**, (1987), 127-142.
- [CODD79] E.F. Codd, Extending the database relational model to capture more meaning. *ACM Transactions on Database Systems* **4**, (1979), 397-434.
- [COSM90] S.S. Cosmadakis, P.C. Kanellakis and M.Y. Vardi, Polynomial-time implication problems for unary inclusion dependencies. *Journal of the ACM* **37**, (1990), 15-46.
- [DAVE90] B.A. Davey and H.A. Priestly, *Introduction to Lattices and Order*. Cambridge University Press, Cambridge, U.K., 1990.
- [DAY93] A. Day, A lattice interpretation of database dependencies. In: *Semantics of Programming Languages and Model Theory*, M. Droste and Y. Gurevich (Eds.), Gordon and Breach, pp. 305-325, 1993.
- [DEME87] J. Demetrovics and V.D. Thi, Keys, antikeys and prime attributes. *Annales Univ. Sci. Budapest, Sect. Comp.* **8**, (1987), 35-52.
- [DEME88] J. Demetrovics and V.D. Thi, Relations and minimal keys. *Acta Cybernetica* **8**, (1988), 279-285.
- [DEME92] J. Demetrovics, L. Libkin and I.B. Muchnik, Functional dependencies in relational databases: A lattice point of view. *Discrete Applied Mathematics* **40**, (1992), 155-185.
- [DEME93] J. Demetrovics and G.O.H. Katona, A survey of some combinatorial results concerning functional dependencies in database relations. *Annals of Mathematics and Artificial Intelligence* **7**, (1993), 63-82.
- [GARE79] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness*. W.H. Freeman, New York, 1979.

- [GOTT90] G. Gottlob and L. Libkin, Investigations on Armstrong relations, dependency inference and excluded functional dependencies. *Acta Cybernetica* **9**, (1990), 385-402.
- [GRAT78] G. Grätzer, *General Lattice Theory*. Academic Press, New York, 1978.
- [JAMI85] R.E. Jamison, The theory of convex geometries. *Geometriae Dedicata* **19**, (1985), 247-270.
- [LEVE94] M. Levene and G. Loizou, The additivity problem for functional dependencies in incomplete relations. Research Note RN/94/41, Department of Computer Science, University College London, 1994. To appear in *Acta Informatica*.
- [LEVE95a] M. Levene and G. Loizou, The additivity problem for data dependencies in incomplete relational databases. In: *Proceedings of the Workshop on Semantics in Databases*, Prague, pp. 49-62, 1995.
- [LEVE95b] M. Levene and G. Loizou, Database design for incomplete relations. Research Note RN/95/18, Department of Computer Science, University College London, 1995. Submitted for publication.
- [LIBK92] L. Libkin and I. Muchnik, Separatory sublattices and subsemilattices. *Studia Scientiarum Mathematicarum Hungarica* **27**, (1992), 471-477.
- [LIEN82] Y.E. Lien, On the equivalence of database models. *Journal of the ACM* **29**, (1982), 333-362.
- [LUCC78] C.L. Lucchesi and S.L. Osborn, Candidate keys for relations. *Journal of Computer and System Sciences* **17**, (1978), 270-279.
- [MANN83] H. Mannila and K.-J. Rähä, On the relationship of minimum and optimum covers for a set of functional dependencies. *Acta Informatica* **20**, (1983), 143-158.
- [MANN86] H. Mannila and K.-J. Rähä, Design by example: an application of Armstrong relations. *Journal of Computer and System Sciences* **33**, (1986), 126-141.
- [MANN89] H. Mannila and K.-J. Rähä, Practical algorithms for finding prime attributes and testing normal forms. In: *Proceedings of the ACM Symposium on Principles of Database Systems*, Philadelphia, PA., pp. 128-133, 1989.
- [MANN92] H. Mannila and K.-J. Rähä, *The Design of Relational Databases*. Addison-Wesley, Reading, MA., 1992.
- [MAIE80] D. Maier, Minimum covers in the relational database model. *Journal of the ACM* **27**, (1980), 664-674.
- [SCIO81] E. Sciore, Real world mvd's. In: *Proceedings of ACM SIGMOD Conference on Management of Data*, Ann Arbor, pp. 121-132, 1981.
- [THI86] V.D. Thi, Minimal keys and antikeys. *Acta Cybernetica* **7**, (1986), 361-371.
- [ULLM88] J.D. Ullman, *Principles of Database and Knowledge-Base Systems, Vol. I*. Computer Science Press, Rockville, MD., 1988.

- [WILD89] M. Wild, Optimal implicational bases for finite modular lattices. Preprint 1211, Technische Hochschule Darmstadt, Germany, 1989.
- [WILD95] M. Wild, Computations with finite closures systems and implications. In: *Proceedings of the International Conference on Computing and Combinatorics*, Xian, China, 1995, pp. 111-120.