Functional Dependencies in Relational Databases: A Lattice Point of View *

János Demetrovics

Computer and Automation Institute Hungarian Academy of Sciences
H-1132 Budapest, Victor Hugo u. 18-22., Hungary

Leonid Libkin

Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104-6389, USA

Ilya B. Muchnik

Institute of Control Sciences, Profsojuznaya 65, Moscow 117806, USSR

Abstract

A lattice theoretic approach is developed to study the properties of functional dependencies in relational databases. The particular attention is paid to the analysis of the semilattice of closed sets, the lattice of all closure operations on a given set and to a new characterization of normal form relation schemes. Relation schemes with restrictions on functional dependencies are also studied.

1. Introduction

The relational datamodel was defined by E.F. Codd [14] in 1970, and it is still one of the most powerful database models. In this model a relation is a matrix (table) every row of which corresponds to a record and every column to an attribute. This model has been widely studied. One of the most important branches in the theory of relational databases is that dealing with the design of database schemes. This branch is based on the theory of dependencies and constraints.

In this paper we study the functional dependencies. Informally, functional dependency means that some attributes' values can be reconstructed unambiguously by the others. A pair consisting of a set of attributes and a set of functional dependencies on it is called a *relational database scheme,* or relation scheme.

The concept of functional dependency was introduced by W.W. Armstrong [2]. It was shown in [2] that the families of functional dependencies (or, equivalently, relation schemes) can be described by closure operations on the attributes' set. This representation was successfully applied to find many properties of functional dependencies.

There is another representation of relation schemes. In fact, the closed sets of a closure form a semilattice. Hence, the semilattices with greatest elements give an equivalent description of functional dependencies. Sometimes this representation is very useful, for instance, in order to construct a relation representing a given relation scheme (so-called Armstrong relation). However the representation of relation schemes by semilattices is not developed well enough in contrast to that by closure operations.

---

* Research partially supported by Hungarian National Foundation for Scientific Research, Grant 1066. The second author was also partially supported by NSF Grants IRI-86-10617 and CCR-90-57570

The main purpose of this paper is to develop a lattice point of view for the study of the relation schemes. The rest of the paper is organized in six parts.

In the second section some necessary definitions and facts about relational databases and lattice theory are given.

Section 3 deals with the semilattice of closed sets. It is shown how to construct the semilattice if a relation scheme is given. This construction is applied to find the lattice theoretic form of such concepts as cover, FD-implication, nonredundancy etc. It is also used to estimate the number of non-equivalent relation schemes.

It was proposed in [11] to study the poset of all closures on an attributes' set as a model of changing databases. In the section 4 we show that this poset is in fact a lattice (moreover, the lattice of subsemilattices of a semilattice). The properties of this lattice are used to establish the new properties of relation schemes. For instance, it is shown how to implement the lattice operations for closures and how to construct arbitrary relation schemes from the simple ones.

Section 5 deals with a lattice theoretic characterization of normal form relation schemes. In fact we characterize the semilattices of closed sets if a relation scheme is in the second, third or Boyce-Codd normal form [44]. This characterization has practical applications. It is well-known that the recognizing the third and Boyce-Codd normal forms are NP-complete problems for relation schemes [8,32]. More precisely, it is NP-complete to find out if a proper subset of attributes' set is in Boyce-Codd normal form. However, the new characterization being used, it is easy to construct algorithms recognizing these normal forms in polynomial time if we are given a relation instead of a relation scheme. Besides, we give a new characterization of relation schemes which are uniquely determined by their candidate keys.

In the section 6 we study the relation schemes with restrictions on functional dependencies. These restrictions are of two types: either the size of lefthand sides of functional dependencies is limited or a relation scheme has to provide the closure belong to a given class of closures. It is shown that relation schemes with restrictions have some nice properties. For instance, for some schemes it is easy to find a compact representation of closures or to construct an Armstrong relation with small number of tuples. Some problems which are generally NP-complete can be solved in polynomial time for special schemes. Sometimes the structure of candidate keys can be described very clearly. Moreover, the database concepts being studied for known types of closures, we obtain some new results about these closures and related mathematical objects.

In the last section we briefly recall the main results of the paper and outline some ideas of the further development.

The extended abstract of this paper was published in [19].

2. Basic definitions

In this section we present briefly the main concepts of the relational design theory which will be needed in sequel. The main concepts of a relation [14] and a functional dependency [2] are given. The other concepts and facts given in this section can be found in [8-10, 19-22, 25, 43-47, 50, 51].

Let $\mathcal{U}$ be a finite set of *attributes* (e.g. name, age etc.). The elements of $\mathcal{U}$ will be denoted by $a, b, c, \dots x, y, z$ or, if an ordering on $\mathcal{U}$ is needed, by $a_1, \dots a_n$. A map *dom* associates with each $a \in \mathcal{U}$ its *domain* dom(a). A *relation* R over $\mathcal{U}$ is a subset of Cartesian product $\prod_{a\in A} dom(a)$.

We can think of a relation R over $\mathcal{U}$ as being a set of tuples : $R = \{h_1, \dots, h_m\}$,

$$h_i : \mathcal{U} \to \cup_{a\in\mathcal{U}}\mathrm{dom}(a),\ \ h_i(a) \in \mathrm{dom}(a),\ \ i = 1, \dots, m.$$

A *functional dependency* (FD for short) is an expression of form $X \to Y$, where $X, Y \subseteq \mathcal{U}$. We say that FD $X \to Y$ holds for a relation $R = \{h_1, \dots, h_m\}$ (or R obeys $X \to Y$) if $h_i(a) = h_j(a)$ for all $a \in X$ implies $h_i(a) = h_j(a)$ for all $a \in Y$, $\forall h_i, h_j \in R, i \neq j$.

Let F_R be a family of all FD*s* that hold for R.

Then $F = F_R$ satisfies

(F1) $X \to X \in F$;

(F2) $(X \to Y \in F, Y \to Z \in F) \Longrightarrow (X \to Z \in F)$;

(F3) $(X \to Y \in F, X \subseteq V, W \subseteq Y) \Longrightarrow (V \to W \in F)$;

(F4) $(X \to Y \in F, V \to W \in F) \Longrightarrow (X \cup V \to Y \cup W \in F)$.

A family of FD*s* satisfying $(F1) - (F4)$ is called *a full family.* F_R is a full family and for every full family F there is a relation R with $F = F_R$.

Given a family F of FD*s*, there is unique minimal full family F^+ that contains F. In fact, F^+ consists of all FD*s* that can be derived from FD*s* of F by using $(F1)-(F4)$.

A family G of FD*s* is called a *cover* of F if $G^+ = F^+$.

A pair $< \mathcal{U}, F >$ consisting of an attributes' set $\mathcal{U}$ and a family F of FD*s* on $\mathcal{U}$ is called a *relation scheme.* A relation R over U is called an *instance* of $< \mathcal{U}, F >$ if R obeys F and does not obeys any FD not from F^+. Clearly, R is an instance of $< \mathcal{U}, F >$ iff it is an instance of $< \mathcal{U}, G >$ for G a cover of F.

Further we will not distinguish an element $a \in \mathcal{U}$ and one-element set $\{a\}$. We will write simply a instead of $\{a\}$.

Let F be a family of FD*s*. Define the mapping $C_F : P(\mathcal{U}) \to P(\mathcal{U})$, where $P(\mathcal{U})$ is the set of all subsets of U, as follows:

$$C_F(X) = \{a \in X | X \to a \in F^+\}, X \subseteq \mathcal{U}.$$

C_F thus constructed satisfies the properties:

(C1) $X \subseteq C_F(X)$;

(C2) $X \subseteq Y \Longrightarrow C_F(X) \subseteq C_F(Y)$,

(C3) $C_F(C_F(X)) = C_F(X)$,

i.e. C_F is a *closure operation* closure on $\mathcal{U}$ (or simply *closure* for short). Conversely, given a closure C on U, there is a family F of FD*s* with $C = C_F$. Clearly, $C_F = C_G$ iff G is a cover of F.

Define $S_F = \{X \subseteq \mathcal{U} | C_F(X) = X\}$. S_F satisfies the properties:

(S1) $\mathcal{U} \in S_F$;

(S2) $X, Y \in S_F \Longrightarrow X \cap Y \in S_F$,

i.e. S_F is a *meet-semilattice* (SL for short) with the greatest element. Conversely, if S satisfies $(S1) - (S2)$, there is F such that $S = S_F$.

An element $X \in S_F$ is called (meet)-*irreducible* if $X = Y \cap Z$, and $Y, Z \in S_F$ imply $Y = X$ or $Z = X$.

The set of all irreducible elements is denoted by $M(S_F)$. Every element of S_F is an intersection of elements from $M(S_F)$.

Thus, closures and SLs satisfying $(S1)$ are the models of full families of FDs, that is, of families of FDs holding for relations over $\mathcal{U}$.

Let C be a closure on U. A subset $X \subseteq U$ is called *closed* if $C(X) = X$. The family of closed sets is denoted S_C. If S is a SL containing $\mathcal{U}$, define $C_S(X) = \cap(Y|X \subseteq Y, Y \in S)$. Then $C \rightarrow S_C$ and $S \rightarrow C_S$ are mutually inverse one-to-one correspondences between closures on $\mathcal{U}$ and SLs containing $\mathcal{U}$.

If we are given a relation scheme $< \mathcal{U}, F >$ (or, equivalently, if we are given a closure or a SL), a set $K \subseteq \mathcal{U}$ is called a *key* if $K \rightarrow \mathcal{U} \in F^+$ $(C_F(K) = \mathcal{U})$. Minimal keys are called *candidate keys*. The candidate keys of a relation R are the candidate keys of $< \mathcal{U}, F_R >$.

The candidate keys obviously form an antichain. Conversely, given an antichain of subsets of $\mathcal{U}$, there is a relation scheme (and, of course, a relation) whose candidate keys are exactly the elements of this antichain.

Maximal non-key is called an *antikey*. Maximal element of a SL $S_F - \{\mathcal{U}\}$ is called its *coatom*. The antikeys of $< \mathcal{U}, F >$ are exactly the coatoms of S_F.

An attribute $a \in \mathcal{U}$ is *prime* if for some candidate key K of $< \mathcal{U}, F >$ one has $a \in K$, and *nonprime* otherwise. The sets of prime and nonprime attributes are denoted by $\mathcal{U}_p$ and $\mathcal{U}_n$ respectively ($\mathcal{U}_p(F)$ and $\mathcal{U}_n(F)$ if F is not understood). The following holds: $\mathcal{U}_n(F)$ is the intersection of all antikeys.

A relation scheme $< \mathcal{U}, F >$ is in

1) *second normal form* ($2NF$ for short) if for every candidate key K and $a \in \mathcal{U}_n$, $K' \rightarrow a \in F^+$ for no proper subset $K' \subset K$;
2) *third normal form* ($3NF$ for short) if $X \rightarrow a \in F^+, a \in \mathcal{U}_n, a \notin X$ imply that X is a key;
3) *Boyce-Codd normal form* ($BCNF$ for short) if $X \rightarrow a \in F^+, a \notin X$ imply that X is a key.

Now recall some basic facts about lattices and SLs. A SL is an algebra $< S, \cdot >$ with one binary idempotent commutative associative operation. A partial order on S is defined as follows: $x \geq y$ iff $x \cdot y = y$. $S' \subseteq S$ is called a *subsemilattice* (SSL for short) of S if S' is closed under $\cdot$. A finite SL is *free* if it is isomorphic to $< P(\mathcal{U}) - \{\mathcal{U}\}, \cap >$ for some $\mathcal{U}$.

An algebra $< \mathcal{L}, \vee, \wedge >$ with two semilattice operations $\vee$ and $\wedge$ satisfying $x \vee (x \wedge y) = x, x \wedge (x \vee y) = x$ is called a *lattice*. $\vee$ and $\wedge$ sometimes are called supremum and infimum. A partial order on L is defined as follows: $x \leq y \iff x \vee y = y \iff x \wedge y = x$. The lattices are isomorphic as algebras iff they are isomorphic as posets.

A lattice is called *distributive* iff $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ for all $x, y, z \in \mathcal{L}$. Finite distributive lattices and only they can be represented as sublattices of $< P(\mathcal{U}), \cup, \cap >$, where $\mathcal{U}$ is finite.

An element $x \in \mathcal{L}$ is said to be *join-* (meet-) *irreducible* if $x = y \vee z$ $(x = y \wedge z)$ implies $x = y$ or $x = z$. The sets of join- (meet) irreducible elements are denoted by $J(\mathcal{L})$ and $M(\mathcal{L})$.

The last concept to be used is that of an *interval*. If $X, Y \subseteq \mathcal{U}$ and $X \subseteq Y$, then $[X, Y] = \{Z \subseteq \mathcal{U} |\ X \subseteq Z \text{ and } Z \subseteq Y\}$.

Finally, recall the main abbreviations. FD stands for functional dependency, SL for semilattice, SSL for subsemilattice, $2NF$, $3NF$, $BCNF$ for the second, third and Boyce-Codd normal forms respectively. Remind also that all the sets are finite throughout the paper.

3. The semilattice of closed sets.

In this section we find the formula that gives us an immediate representation of the SL of closed sets by FDs. As it was mentioned in the previous section, closure operators and SLs give the equivalent descriptions of the families of FDs. However, if we are given a family of FDs, the closure operator corresponding to this family can be constructed. That is, given a family of FDs and a set $X \subseteq \mathcal{U}$, we can find the closure of X (note that it can be done in polynomial time [8]). On the other hand, to find the SL of closed sets we must check up all the sets in order to find out if they are closed or not. Thus the closure is used as an intermediate step to construct SL. In order to avoid this step we find the direct representation of SL by FDs.

This representation will show us that the use of lattice theoretic concepts is not poorer than that of closure operators in order to describe FDs on a given set of attributes. For instance, we will give the structural representation of FD implication and some problems related to covers of FDs. It also will be shown how to find a relation representing given SL.

Making use of the semilattice terminology also allows us to transfer some results of lattice theory to relational databases. E.g., some different algebras has been studied on the set of SSLs of a SL, cf. [41], [49]. The results obtained in these works will be applied in the next section. The other idea is to consider some known classes of lattices and SLs in order to study special families of FDs. A part of this program of research will be carried out in the section 6.

Now we are ready to formulate the main result of this section.

Theorem 3.1. *Let F be a family of* FDs *on $\mathcal{U}$ and S_F the* SL *of closed subsets of $\mathcal{U}$. Then*

$$S_F = P(\mathcal{U}) - \bigcup_{\substack{X \to Y \in F \\ a \in Y - X}} [X, \mathcal{U} - a]. \tag{3.1}$$

Proof. First prove that

$$S_F = P(\mathcal{U}) - \bigcup_{\substack{X \to Y \in F^+ \\ a \in Y - X}} [X, \mathcal{U} - a]. \tag{3.2}$$

Denote $\bigcup([X,\mathcal{U} - a] | X \to Y \in F, a \in Y - X)$ by D_F. Let $Z \in D_{F^+}$. Then $Z \in [X,\mathcal{U} - a]$ for $X \to a \in F^+$ and $Z \to a \in F^+$, i.e. $Z \notin S_F$. If $Z \notin S_F$, there is nontrivial FD $Z \to a \in F^+$ and $Z \in [Z,\mathcal{U} - a] \subseteq D_{F^+}$. This proves (3.2).

Now we must prove $D_F = D_{F^+}$. Because of $F \subseteq F^+$, $D_F \subseteq D_{F^+}$ is obvious. In order to prove $D_{F^+} \subseteq D_F$ we must show that for F_1 obtained from F by a single FD being derived according to one of the rules $(F1)-(F4)$ it holds: $D_{F_1} \subseteq D_F$. Let $X \to Y$ be this single FD and $a \in Y - X$. If it is obtained by $(F1)$, the inclusion is evident. Suppose it is obtained by $(F2)$. Then for a set Z we have $X \to Z \in F$ and $Z \to Y \in F$. If $a \in Z$ then $[X,\mathcal{U} - a] \subseteq D_F$. If $a \notin Z$, suppose $V \in [X,\mathcal{U} - a]$. Suppose there is $b \in Z - V$. Then $V \in [X,\mathcal{U} - b] \subseteq D_F$. If $Z \subseteq V$, then $V \in [Z,\mathcal{U} - a] \subseteq D_F$. Hence in all the cases $V \in D_F$, and $[X,\mathcal{U} - a] \subseteq D_F$.

Let $X \to Y$ be obtained by $(F3)$, i.e. $V \to W \in F$ and $V \subseteq X, Y \subseteq W$. Then $a \in W - V$ and $[X,\mathcal{U} - a] \subseteq [V,\mathcal{U} - a] \subseteq D_F$.

Finally, if $X \to Y$ is obtained by $(F4)$, i.e. $X = X_1 \cup X_2$, $Y = Y_1 \cup Y_2$, $X_i \to Y_i \in F$, $i = 1,2$, suppose $a \in Y_1$. Then $[X,\mathcal{U} - a] \subseteq [X_1,\mathcal{U} - a] \subseteq D_F$.

Hence $D_{F_1} \subseteq D_F$ and since $(F1) - (F4)$ is sound and complete system for FD implication, $D_{F^+} \subseteq D_F$. This inclusion together with (3.2) proves (3.1).

Remark. Formula (3.1) is the special case of the interval representation of SSLs of distributive SLs which was established first for Boolean in [36] and afterwards for arbitrary distributive SLs in [40, 41].

Now we are going to give the structural representation of such concepts as cover, FD implication, e.t.c.

Corollary 3.2. *Let F and G be two families of* FDs *on $\mathcal{U}$. Then F is a cover of G iff*

$$\bigcup_{\substack{X \to Y \in F \\ a \in Y - X}} [X,\mathcal{U} - a] = \bigcup_{\substack{X \to Y \in G \\ a \in Y - X}} [X,\mathcal{U} - a]. \tag{3.3}$$

Remind that a family of FDs is called *open* [26] if every FD has one-element right-hand part. A family F of FDs is called *nonredundant* [26, 44] if for any FD $f \in F$ one has $f \notin (F - f)^+$.

Corollary 3.3. *An open family F of* FDs *is nonredundant iff*

$$[X,\mathcal{U} - a] \nsubseteq \bigcup_{\substack{Y \to b \in F \\ Y \to b \neq X \to a}} [Y,\mathcal{U} - b] \ \forall X \to a \in F. \tag{3.4}$$

A single FD f is implied by a family of FDs $F (F \models f)$ if $f \in F^+$ (cf. [26, 44, 47, 51]).

Corollary 3.4. *$F \models X \to Y$ holds iff*

$$\bigcup_{a \in Y - X} [X,\mathcal{U} - a] \subseteq \bigcup_{\substack{Z \to W \in F \\ b \in W - Z}} [Z,\mathcal{U} - b]. \tag{3.5}$$

Notice that (3.3) is a weak form of (3.5). The formula (3.3) gives rise to an algorithm for checking up if two SLs are identical or not. In fact, it was proved in [36, 41] that each SL of sets can be represented as

$$S = P(\mathcal{U}) - \cup[X, \mathcal{U} - a]$$

for some collection of pairs (X, a). Then, given two SLs, construct two families of FDs consisting of FDs $X \to a$ for representing collections of pairs. Then the SLs coincide iff one family of FDs is a cover of the other according to (3.3). Note also that, given two families F_1 and F_2, we can find out if $F_1^+ = F_2^+$ in polynomial time in size of $F_1 \cup F_2$ [8].

Two following corollaries may be valuable for practical purposes because they both establish concrete covers for full families of FDs.

Remind that a FD $X \to a$ is called *primitive* and *maximal* (cf. [10]) for a full family F^+ if $X \to a \in F^+$, $a \notin X$ and for each proper subset $X' \subset X$ we have $X' \to a \notin F^+$.

Corollary 3.5. *For a full family F^+ the subfamily of primitive maximal* FDs *is a cover of F^+.*

Proof follows immediately from (3.1).

In [11] it was proposed to describe a family F of FDs by a collection of sets $\{H_j | j \in J\}$ such that $S_F \cup \{H_j\}$ is again a SL for all $j \in J$ and no other set $H \in P(\mathcal{U}) - S_F$ satisfies this condition. It was proved that $\{H_j | j \in J\}$ unambiguously determines the closure and SL. Now we show how to construct the interval representation of a SL by using the family of H_j's.

Proposition 3.6. *Let F be a family of* FDs. *Suppose $\{H_j | j \in J\}$ is constructed as above. Then*

$$S_F = P(\mathcal{U}) - \cup([H_j, \mathcal{U} - a] | j \in J, a \in C_F(H_j) - H_j). \tag{3.6}$$

Proof. Let $X \in [H_j, \mathcal{U} - a]$ for $j \in J, a \in C_F(H_j) - H_j$. Then $a \in C_F(X) - X$ and $X \notin S_F$. Conversely, if X belongs to the right-hand side of (3.6), suppose $X \notin S_F$. According to [8, th.1] for some $j \in J : H_j \subseteq X$ and $C_F(H_j) \nsubseteq X$, i.e. there is $a \in C_F(H_j) - H_j$ such that $X \in [H_j, \mathcal{U} - a]$. Proposition is proved.

The sets $H_j, j \in J$, were called *quasiclosed* in [52]. A quasiclosed set H_j is called *pseudoclosed* if there is no $H_i \subset H_j$ with $C_F(H_i) = C_F(H_j)$.

It follows immediately from (3.1) and (3.6)

Corollary 3.7. *Let F be a family of* FDs *and $\{H_j | j \in J'\}$ the family of pseudo-closed sets. Then $\{H_j \to C_F(H_j) | j \in J'\}$ is a cover of F.*

In fact, the cover constructed above is a minimum one [52].

In the first part of the section we have shown that the lattice theoretic language is equivalent to that of closures in order to describe FDs, and that it is rather clear. Moreover, the use of SLs allows us to construct a relation over $\mathcal{U}$ representing a given family F of FDs (so-called Armstrong relation for F [9,27,45]).

Armstrong relations are very useful for practical purposes as they reveal concealed (to database designer) FDs. Now we remind the main idea of construction of Armstrong relations because it will be used further.

Let $M(S_F)$ be the set of irreducible elements of S_F. (It was called $GEN(F)$ in [9]. In [46] $M(S_F)$ was represented as a collection of so-called MAX-sets, which express, in fact, the concept of copoints (cf. [25]) for FD terminology).

The earliest example of Armstrong relation for F was found in [16] and it contained $2 \cdot |M(S_F)|$ tuples. Afterwards, the lower and upper bounds for the size of Armstrong relation were found in [9] as $\lceil(1 + \sqrt{1 + 8|M(S_F)|})/2\rceil + 1$ and $|M(S_F)|$, where $\lceil x\rceil$ stands for the greatest integer less than x. These bounds show that the number of tuples of an Armstrong relation representing F is polynomial in $|\mathcal{U}|$ iff so is $|M(S_F)|$.

Mannila and Raiha [45] presented an algorithm to construct an Armstrong relation. Suppose $|M(S_F)| = \{X_0, \ldots, X_k\}$, where $X_0 = \mathcal{U}$. With each X_i associate the tuple h_i

$$h_i(a) = \begin{cases} 0, & a \in X_i, \\ i, & a \notin X_i. \end{cases}$$

Then $R = \{h_0, \ldots, h_k\}$ is an Armstrong relation for F. Notice that in general the size of R is exponential because so is the size of $M(S_F)$. The example of exponential $M(S_F)$ was given in [9] and the upper bound was established in [34]:

$$|M(S_F)| \leq (\,{n \atop \lceil n/2\rceil}\,)(1 + o(1)).$$

We finish this section by the calculation of lower bound for the number of full families of FDs on an n-element set which do not contain FDs of type $\emptyset \rightarrow X$, $X \neq \emptyset$. Denote the number of such families by $\alpha(n)$. Clearly, $\alpha(n)$ is the number of SLs on $\mathcal{U}, |\mathcal{U}| = n$, containing $\{\emptyset\}$.

The following was proved in [12]: $\alpha(n) \geq 2^{\binom{n}{\lceil n/2\rceil}}$.

Consider $\binom{\mathcal{U}}{k} = \{X \subseteq \mathcal{U} | |X| = k\}$ and the SLs $S_{\mathcal{A}} = P(\mathcal{U}) - \cup\{[X, \mathcal{U} - a] | a \in \mathcal{U}, X \in \mathcal{A} \subseteq (\binom{\mathcal{U}}{k})\}$.

Clearly, $S_{\mathcal{A}}$ is a SL and $\mathcal{A}_1 \neq \mathcal{A}_2$ implies $S_{\mathcal{A}_1} \neq S_{\mathcal{A}_2}$.

Let $S^a_{\mathcal{A}} = P(\mathcal{U}) - \cup\{[X, \mathcal{U} - a] | X \in \mathcal{A} \subseteq \binom{\mathcal{U}-a}{k}\}$.

Then $S^a_{\mathcal{A}}$ is a SL and for each $a \in \mathcal{U}$ $\mathcal{A}_1 \neq \mathcal{A}_2$ implies $S^a_{\mathcal{A}_1} \neq S^a_{\mathcal{A}_2}$. The above constructions immediately lead us to

Proposition 3.8. *For each* $n \geq 3$

$$\alpha(n) \geq \sum_{k=1}^{n-1} 2^{\binom{n}{k}} + n \sum_{k=1}^{n-2} 2^{\binom{n-1}{k}}.$$

This lower bound is more precise then one given in [12]. However, it is still unknown if $log_2 \alpha(n) \sim \binom{n}{\lceil n/2\rceil}$.

4. The lattice of subsemilattices as a model of changing databases.

Usually databases are constantly changing during their lives. For instance, each update such as insertion, deletion etc. leads to a new state of a database, and, of course, to a new family of FDs. Thus, it is quite natural to describe how the families of FDs can change. First the efforts have been carried out in this direction in the paper [11], where the partially ordered set (*poset* for short) of all closure operators on a fixed set was studied in some detail. In this section we continue the study of this model of changing databases which is based on the fact the poset of closures is a lattice isomorphic to the lattice of SSLs of a free SL.

Before giving the formal results we are going to set forth some arguments why the study of this model seems to be useful for database design theory. There are many algorithms related to database design theory which cannot be solved in polynomial time, for instance, testing third and Boyce-Codd normal forms [8, 32], prime attribute and key cardinality problems [20, 23, 43], problem of G. Gottlob (that is, given a relation scheme $< \mathcal{U}, F >$ and a relation over $\mathcal{U}$, decide whether $F_R \subseteq F^+$ [27]) and others. However, as it will be shown in section 6, some of these problems can be solved in polynomial time if a scheme satisfies some additional properties (e.g. if all FDs are unary, testing normal forms can be done in polynomial time in $|\mathcal{U}|$, see [46]). Some additional conditions being added, the corresponding closures and SLs have to belong to some special classes. Thus, if polynomiality of some algorithms is needed, we can propose to a database designer to choose families of FDs corresponding to a given class. Moreover, if it is not possible, we can approximate a given scheme in some of "good" classes (for normal forms it has been done in [18]). However, in order to solve these problems we must know the structure of the set of all schemes (closures, SLs).

To give the other reason, notice that in databases theory the mathematical concepts are used in the main in order to describe some database problems. Here we propose another approach. The poset of closures being well-studied algebraic object, we can interpret its properties in the context of database problems and get some new properties and concepts related to FDs in relational databases.

To begin with, we establish the algebraic characterization of poset of closures (schemes, SLs). Let $|\mathcal{U}| = n$. Suppose C_1, C_2 are two closures on $\mathcal{U}$. According to [11], introduce the partial order on the set of all closures on $\mathcal{U}$ as follows:

$$C_1 \geq C_2 \Longleftrightarrow \forall X \subseteq \mathcal{U} : C_1(X) \subseteq C_2(X) \tag{4.1}$$

There are two equivalent descriptions of this order.

Lemma 4.1 [11]. $C_1 \geq C_2$ *iff* $S_{C_2} \subseteq S_{C_1}$

Lemma 4.2 [19]. $C_1 \geq C_2$ *iff* $C_1 \circ C_2 = C_2$.

Denote the poset of closures on $\mathcal{U}$ by Cl_n. Consider the SL $< P(\mathcal{U}) - \{\mathcal{U}\}, \cap >$. According to [19, 28] it is isomorphic to the free SL with n generators denoted by $F(n)$, and the mapping $X \to z_{i_1} \cdot \ldots \cdot z_{i_k}$, where $\mathcal{U} - X = \{a_{i_1}, \ldots, a_{i_k}\}$ and $z_1, \ldots, z_n$ are the generators of $F(n)$, is the isomorphism.

Let S be an arbitrary finite SL. Then the set *Sub* S of SSLs of S is a lattice in which the inf and sup operations can be expressed as follows (cf. [41, 49]):

$$S_1 \wedge S_2 = S_1 \cap S_2, \tag{4.2}$$

$$S_1 \vee S_2 = S_1 \cup S_2 \cup \{s_1 \cdot s_2 | s_1 \in S_1, s_2 \in S_2\}. \tag{4.3}$$

Now we are ready to formulate the characterization theorem.

Theorem 4.3. *The poset Cl_n of closures on $\mathcal{U}$ is a lattice isomorphic to $SubF(n)$.*

Proof. According to the above remark, we have to prove that Cl_n is isomorphic to $Sub < P(\mathcal{U}) - \{\mathcal{U}\}, \cap >$. Consider the mapping $\varphi : Cl_n \to Sub < P(\mathcal{U}) - \{\mathcal{U}\}, \cap >$: $\varphi(C) = S_C - \{\mathcal{U}\}$. According to lemma 4.1 $C_1 \le C_2$ iff $\varphi(C_1) \subseteq \varphi(C_2)$. Moreover, φ is one-to-one because so is $C \to S_C$ and each S_C contains $\{\mathcal{U}\}$. Hence, φ is an order isomorphism and, therefore, a lattice isomorphism. Q.E.D.

We are going now to calculate the operations $\wedge$ and $\vee$ for Cl_n and also to find join- and meet-irreducible elements.

Let $C_1, C_2 \in Cl_n$. Let $C_{12}(X) = C_1(X) \cup C_2(X)$, $C_{12}^2(X) = C_{12}(C_{12}(X))$, $\ldots$, $C_{12}^{k+1}(X) = C_{12}(C_{12}^k(X))$ for all $X \subseteq \mathcal{U}$.

Proposition 4.4. *For every $X \subseteq \mathcal{U}$ the following hold:*

$$C_1 \vee C_2(X) = C_1(X) \cap C_2(X); \tag{4.5}$$

$$C_1 \wedge C_2(X) = C_{12}^{n-|X|}(X). \tag{4.6}$$

Proof. Let $S = S_{C_1} \vee S_{C_2}$. According to theorem 4.3 $C_1 \vee C_2 = C_S$ and (4.5) follows immediately from (4.3).

Let $S = S_{C_1} \wedge S_{C_2} = S_{C_1} \cap S_{C_2}$ by (4.2). Then according to theorem 4.3, $C_1 \wedge C_2 = C_S$, i.e.

$$C_1 \wedge C_2(X) = \cap(Y | X \subseteq Y,\ Y \in S_{C_1},\ Y \in S_{C_2}). \tag{4.7}$$

Let $Z \subseteq C_1 \wedge C_2(X)$. Since $C_1 \wedge C_2(X) \in S_{C_1}$ (see (4.7)), $C_1(Z) \subseteq C_1(C_1 \wedge C_2(X)) = C_1 \wedge C_2(X)$. Analogously $C_2(Z) \subseteq C_1 \wedge C_2(X)$. Hence $C_{12} \subseteq C_1 \wedge C_2(X)$. Because of $X \subseteq C_1 \wedge C_2(X)$ we have $C_{12}^k(X) \subseteq C_1 \wedge C_2(X)$ for all k.

Because of the finiteness of $\mathcal{U}$, $C_{12}^{k+1}(X) = C_{12}^k(X)$ for some $k \le n - |X|$. Then $C_{12}^k(X) \subseteq C_1 \wedge C_2(X)$ and simultaneously $C_1(C_{12}^k(X)) \cup C_2(C_{12}^k(X)) = C_{12}^k(X)$, i.e. $C_{12}^k(X) \in S_{C_1} \cap S_{C_2}$. It means $C_{12}^k(X) = C_1 \wedge C_2(X)$. Since $C_{12}^m(X) = C_{12}^k(X)$ for $m \ge k$ and $k \le n - |X|$, (4.6) is proved.

Corollary 4.5. $C_1 \vee C_2 = C_1 \cap C_2; C_1 \wedge C_2 = C_{12}^n$.

Before finding irreducible elements, let us give the interpretation of formulas (4.5) and (4.6). Formula (4.5) states that if we are given two families F_1 and F_2 of FDs there exists unique maximal full family of FDs that is contained in both F_1^+ and F_2^+; in fact, $F_1^+ \cap F_2^+$.

Moreover, there exists unique minimal full family that contains both F_1^+ and F_2^+; in fact, $(F_1 \cup F_2)^+$ and the closure corresponding to this family can be calculated by (4.6).

However, we can find more clear formula for $C_1 \vee C_2$ if the families F_1 and F_2 of FDs are used to represent the closures. In fact, we find a cover of $(F_1 \cup F_2)^+$ if F_1 and F_2 are given.

Remind that sometimes a family F is treated as a binary relation on $P(\mathcal{U})$: $(X,Y) \in F$ iff $X \to Y \in F$, cf. [51]. Suppose without loss of generality that F is supplemented by all the pairs $(X,X), X \subseteq \mathcal{U}$, i.e. F is reflexive.

Proposition 4.6. *Let $C_1 = C_{F_1}$, $C_2 = C_{F_2}$ and the binary relations F_1 and F_2 are reflexive. Let $F = F_1 \circ F_2$ be superposition of binary relations F_1 and F_2. Then $C_1 \wedge C_2 = C_F$, i.e. F is a cover of $(F_1 \cup F_2)^+$.*

Proof. According to theorem 3.1,

$$S_i = S_{F_i} = S_{C_i} = P(\mathcal{U}) - \bigcup_{\substack{X \to Y \in F_i \\ a \in Y - X}} [X, \mathcal{U} - a],\ i = 1, 2.$$

Family F of FDs contains such FDs $X \to Y$ that for some $Z \subseteq \mathcal{U}$ we have $X \to Z \in F_1$ and $Z \to Y \in F_2$. Because of reflexivity, $F_1 \cup F_2 \subseteq F$. It shows $S_F \subseteq S_1 \cap S_2$.

Suppose $V \notin S_F$. Then for some $X \to Y \in F$ and $a \in Y - X$ we have $V \in [X, \mathcal{U} - a]$. If $X \to Y \in F_i$ then $V \notin S_i, i = 1, 2$. If $X \to Y \notin F_i, i = 1, 2$, then for some Z we have $X \to Z \in F_1$ and $Z \to Y \in F_2$. If $a \in Z$ then $V \in [X, \mathcal{U} - a] \subseteq P(\mathcal{U}) - S_1$. If $a \notin Z$, there are two cases. Either $V \supseteq Z$ and $V \in [Z, \mathcal{U} - a] \subseteq P(\mathcal{U}) - S_2$, or there is $b \in Z - V$, and then $b \in Z - X$, i.e. $V \in [X, \mathcal{U} - b] \subseteq P(\mathcal{U}) - S_1$. Hence, $V \in (P(\mathcal{U}) - S_1) \cup (P(\mathcal{U}) - S_2)$ and $S_1 \cap S_2 \subseteq S_F$. $S_1 \cap S_2 = S_F$ having been proved, $C_1 \wedge C_2 = C_F$ is valid by (4.3) and (4.7).

Thus, in order to find $C_1 \wedge C_2$ (or $(F_1 \cup F_2)^+$) we must find the superposition $F_1 \circ F_2$. Notice that though the superposition is not commutative, both $F_1 \circ F_2$ and $F_2 \circ F_1$ are the covers of $(F_1 \cup F_2)^+$.

Propositions 4.4 and 4.6 show the importance of finding irreducible elements. Really, if $F_1, ..., F_p$ represent the join-irreducible elements of Cl_n (i.e. $\{C_{F_1}, \ldots, C_{F_p}\} = J(Cl_n)$) then each full family of FDs can be represented as $\bigcap_{i \in I \subseteq \{1,\ldots,p\}} F_i^+$.

If $F_1, \ldots, F_r$ represent the meet-irreducible elements of Cl_n (i.e. $\{C_{F_1}, \ldots, C_{F_r}\} = M(Cl_n)$) and each F_i contains all FDs $X \to X$, then each full family has a cover which is the superposition of some $F_i s$, i.e. for every full family F there is F_0 such that

$$F_0 = \ \circ_{i \in I \subseteq \{1,\ldots,r\}} F_i \text{ and } F_0^+ = F.$$

Let X be a subset of $\mathcal{U}$. Define C_X as follows:

$$C_X(Y) = \begin{cases} X, & Y \subseteq X, \\ \mathcal{U}, & Y \not\subseteq X. \end{cases} \tag{4.8}$$

Let $X \subseteq \mathcal{U}$ and $a \in \mathcal{U}$. Define C_X^a as follows:

$$C_X^a(Y) = \begin{cases} Y \cup a, & \text{if} Y \in [X, \mathcal{U} - a], \\ Y, & \text{otherwise}. \end{cases} \tag{4.9}$$

(If $a \in X$, $[X,\mathcal{U}-a]$ is empty).

Proposition 4.7. *Join-irreducible elements of Cl_n are exactly the closures C_X (4.8) and meet-irreducible elements are exactly the closures C_X^a (4.9).*

Proof. Clearly, $S_{C_X} = \{X,\mathcal{U}\}$ and $S_{C_X^a} = P(\mathcal{U}) - [X,\mathcal{U}-a]$. Hence, for isomorphism φ from the proof of theorem 4.1 we have $\varphi(C_X) = \{X\}$ and $\varphi(C_X^a) = P(\mathcal{U}) - ([X,\mathcal{U}-a] \cup \{\mathcal{U}\})$. Since the single elements are exactly the join-irreducible elements of $Sub < P(\mathcal{U}) - \{\mathcal{U}\}, \cap >$ and the SL*s* of type $P(\mathcal{U}) - (\{\mathcal{U}\} \cup [X,\mathcal{U}-a])$ are exactly the meet-irreducible elements of $Sub < P(\mathcal{U}) - \{\mathcal{U}\}, \cap >$ (see [40, 41]), the corresponding closures are join- and meet-irreducible elements of Cl_n.

We do not study in detail the lattice theoretic properties of Cl_n here. Some of them are common properties of the lattices of *S*SL*s* and can be found in [41]. Other properties are the properties of so-called meet-distributive lattices and can be found in [24]. Here we only recall some properties given in [19].

The dual lattice Cl_n^* is semimodular. Hence, Cl_n has a rank function and $r(C) = |S_C| - 1$ (cf. [11]). This rank function satisfies the inequalities $r(C_1) + r(C_2) \leq r(C_1 \vee C_2) + r(C_1 \wedge C_2) \leq r(C_1) + r(C_2) + (r(C_1) - r(C_1 \wedge C_2))\ (r(C_2) - r(C_1 \wedge C_2))$. Also, $r(C)$ is the number of join-irreducible elements under C. Every semimodular sublattice of Cl_n is distributive. If an ideal of Cl_n is a distributive sublattice of Cl_n, then it is Boolean.

Finishing this section, we characterize the subsets of Cl_n corresponding to the following restriction

$$F \subseteq F', \tag{4.10}$$

where F is a fixed family of FD*s*. Let $Cl_n(F) = \{C_{F'} | F \subseteq F'\}$. Suppose $C = C_F$. Then $C_{F'} \leq C$. Conversely, if $C' \leq C$, then $\{X \to Y | Y \subseteq C'(X)\}$ contains F and $C' \in Cl_n(F)$. Hence, $Cl_n(F) = \{C' | C' \leq C\}$, i.e.$Cl_n(F)$ is the principal ideal $(C]$ in Cl_n.

Proposition 4.8. *Let F be a family of* FD*s. Then $Cl_n(F)$ is a lattice isomorphic to $Sub < S_F - \{\mathcal{U}\}, \cap >$. Moreover, for arbitrary finite* SL *S there is a number n, an n-element set $\mathcal{U}$ and a family F of* FD*s on $\mathcal{U}$ such that $Cl_n(F)$ is isomorphic to $Sub\ S$.*

Proof. The first part follows immediately from the proof of theorem 4.1. To prove the second part, consider an arbitrary finite SL S. Suppose without loss of generality that S is meet-SL (because $SubS \simeq SubS^*$). Then S can be embedded in $< P(\mathcal{U}) - \{\mathcal{U}\}, \cap >$ [1, 28]. Suppose $F = \{X \to Y | Y \subseteq C_{S \cup \{\mathcal{U}\}}(X)\}$. Then $S_F = S \cup \{\mathcal{U}\}$ and $Cl_n(F) \simeq SubS$. Proposition is proved.

Remark. The direct product decompositions of the lattices $Cl_n(F)$ were completely characterized in [19]. In particular, the lattices Cl_n are directly indecomposable.

5. A lattice theoretic characterization of normal form relation schemes.

In this section we study the properties of a SL S_F if a relation scheme $< \mathcal{U}, F >$ is in second, third or Boyce-Codd normal form. The subsets of Cl_n corresponding to these normal forms will be investigated in the next section.

The main idea of normalization was proposed by E.F. Codd. That is, to construct the relations with "good" families of FDs if we are given a relation R over $\mathcal{U}$ such that R is a join of "good" relations which are in fact projections onto some attributes' sets, see [44]. Here we study only the most widely used second, third and Boyce-Codd normal forms ($2NF$, $3NF$ and $BCNF$ in sequel) which were introduced in the early 70th. Afterwards these normal forms have been studied both theoretically and practically (cf. [8,9,13,32,44,45,46,47]).

There are some stimuli to study the lattice properties of normal forms. First, presenting a lattice theoretic characterization of S_F if F is in $2NF, 3NF$ or $BCNF$, we continue the line of research that has been proposed in the section 3. That is, to formulate the main results about FDs using semilattice terminology. In section 3 we generally described S_F and studied covers, FD implication and some related problems. Here we use the semilattice representation in order to give a new characterization of normal form relation schemes.

The other result is more close to practical purposes and has to do with the problem of complexity. It was proved that the problems $3NFTEST$ of testing third normal form and $BCNFTEST$ of testing if a proper subset $X \subset \mathcal{U}$ is in Boyce-Codd normal form are NP-complete [8,32]. One related problem is NP-complete too. It is the prime attribute problem, i.e. given an attribute $a \in \mathcal{U}$, decide whether a is prime or not (recall that a is prime if it belongs to a candidate key) [44]. Using the representation of the SL of closed sets by means of an equality set of a relation [22, 23], Demetrovics and Thi proved that the prime attribute problem can be solved in polynomial time if we are given a relation instead of a relation scheme [20,23]. Here we prove the analogous result for normal forms. That is, $3NFTEST$ and $BCNFTEST$ can be done in polynomial time if we are given a relation over $\mathcal{U}$. The lattice characterization of normal forms plays an essential role in the construction of these polynomial algorithms.

We use the lattice characterization in order to give a new solution of the problem if the candidate keys determine closure uniquely [11]. Also, we show that the problem of G. Gottlob [27] (to find out if a relation R is an Armstrong relation of F, i.e. $F_R = F^+$) can be solved in polynomial time if a scheme is known to be in $BCNF$ and the number of candidate keys or tuples of a relation is bounded by a constant. Note that for arbitrary schemes it is unknown if this problem has polynomial complexity.

Now we are going to give the characterization of normal forms. To do this, we need one definition. A set $X \in S_F$ is called *prime* if $X = C_F(Y)$ where Y is a subset of a candidate key.

Remind that the sets $\mathcal{U}_p(F)$ of prime attributes and $\mathcal{U}_n(F)$ of nonprime ones can be obtained as follows: $\mathcal{U}_n(F)$ is the intersection of all coatoms of S_F, $\mathcal{U}_p(F) = \mathcal{U} - \mathcal{U}_n(F)$ [20,23].

Theorem 5.1. *Let $<\mathcal{U}, F>$ be a relation scheme. Then $<\mathcal{U}, F>$ is in 1) $2NF$ 2) $3NF$ 3) $BCNF$ iff*

1) For every prime $X \in S_F, X \neq \mathcal{U} : [X \cap \mathcal{U}_p(F), X] \subseteq S_F$;

2) For every $X \in S_F, X \neq \mathcal{U} : [X \cap \mathcal{U}_p(F), X] \subseteq S_F$;

3) For every $X \in S_F, X \neq \mathcal{U} : [\emptyset, X] \subseteq S_F$.

Proof. 1) Let $[X \cap \mathcal{U}_p, X] \subseteq S_F$ for all prime $X \in S_F, X \neq \mathcal{U}$ (we will write $\mathcal{U}_p$ instead of $\mathcal{U}_p(F)$ if F is understood). Suppose $<\mathcal{U}, F>$ is not in $2NF$, i.e. $K' \to a \in$

F^+, where $K' \subset K$, K is a candidate key and $a \in \mathcal{U}_n$. Let $X = C_F(K')$. Clearly, X is prime and $X \neq \mathcal{U}$. Since $a \in X$ and $a \notin K'$, $X - a \to a \in F^+$ and $X - a \notin S_F$. Simultaneously $X - a \in [X \cap \mathcal{U}_p, X] \subseteq S_F$, a contradiction. Hence, $< \mathcal{U}, F >$ is in $2NF$.

Suppose $< \mathcal{U}, F >$ is in $2NF$ and $X = C_F(Y)$ is prime, $Y \subset K$, K is a candidate key, $X \neq \mathcal{U}$. Let $a \in \mathcal{U}_n$. Then $a \notin Y$. If $X - a \to a \in F^+$ then $Y \to a \in F^+$, a contradiction. Hence $X - a \in S_F$. Since S_F is a SL, $[X - \mathcal{U}_n, X] = [X \cap \mathcal{U}_p, X] \subseteq S_F$. Case 1 is proved.

2) Let $< \mathcal{U}, F >$ be in $3NF$, $X \in S_F$, $X \neq \mathcal{U}$, $a \in \mathcal{U}_n$. Suppose $X - a \notin S_F$. Since $X \in S_F$, $C_F(X - a) = X$ and $X - a \to a \in F^+$. Hence, $X \to \mathcal{U} \in F^+$ and $X = C_F(X) = \mathcal{U}$, a contradiction. It shows $X - a \in S_F$ and therefore $[X - \mathcal{U}_n, X] = [X \cap \mathcal{U}_p, X] \subseteq S_F$.

Conversely, let $[X \cap \mathcal{U}_p, X] \subseteq S_F$ for all $X \in S_F, X \neq \mathcal{U}$. Let $X \to a \in F^+$, $a \in \mathcal{U}_n$, $a \notin X$. We must prove $C_F(X) = \mathcal{U}$. Suppose $C_F(X) = Y \neq \mathcal{U}$. We have $X \subseteq Y - a \subseteq Y$ and $C_F(Y - a) = Y$. But $Y \cap \mathcal{U}_p \subseteq Y - a \subseteq Y$ and $Y - a \in S_F$, a contradiction. Hence $Y = \mathcal{U}$ and $< \mathcal{U}, F >$ is in $3NF$.

3) Let $< \mathcal{U}, F >$ be in $BCNF$. Suppose $X \in S_F$,$X \neq \mathcal{U}$, $a \in X$. If $X - a \notin S_F$ then $X - a \to a \in F^+$ and $X - a \to \mathcal{U} \in F^+$, i.e. $X = C_F(X) = \mathcal{U}$. Therefore, $X - a \in S_F$ for all $a \in X$ and $[\emptyset, X] \subseteq S_F$ because S_F is a SL.

Conversely, let $[\emptyset, X] \subseteq S_F$ for all $X \in S_F, X \neq \mathcal{U}$. Suppose $X \to a \in F^+$, $a \notin X$. If $C_F(X) = Y \neq \mathcal{U}$, then $X \in [\emptyset, Y] \subseteq S_F$ and $a \in X$, a contradiction. Hence $Y = \mathcal{U}$ and $X \to \mathcal{U} \in F^+$. Thus, $< \mathcal{U}, F >$ is in $BCNF$. Theorem is completely proved.

The result about $BCNF$ can be expressed in a more clear form.

Corollary 5.2. *Let $< \mathcal{U}, F >$ be a relation scheme and $X_1, \ldots, X_t$ its antikeys. Then $< \mathcal{U}, F >$ is in $BCNF$ iff $S_F = \bigcup_{i=1}^{t} [\emptyset, X_i] \cup \{\mathcal{U}\}$.*

Of course, this corollary is equivalent to the following one.

Corollary 5.3. *[5]. Let $< \mathcal{U}, F >$ be a relation scheme and $K_1, \ldots, K_r$ its candidate keys. Then $< \mathcal{U}, F >$ is in $BCNF$ iff $P(\mathcal{U}) - S_F = \bigcup_{i=1}^{r} [K_i, \mathcal{U}] - \{\mathcal{U}\}$.*

It is well-known that the problems $3NFTEST$ and $BCNFTEST$ are NP-complete if we are given a relation scheme [8,32]. Now we are going to prove that these problems can be solved in polynomial time if we are given a relation instead of a relation scheme.

Algorithm 5.4.
Input: a relation $R = \{h_1, \ldots, h_m\}$ over $\mathcal{U}$;
Output: $3NF(R) \in \{0, 1\}$.

Step 1. Construct the equality set $E_R = \{h_{ij} | 1 \leq i < j \leq m\}$, $h_{ij} = \{a \in \mathcal{U} | h_i(a) = h_j(a)\}$.
Step 2. Find E_R^+ as the family of maximal elements of $E_R - \{\mathcal{U}\}$.
Step 3. Find $\mathcal{U}_p = \mathcal{U} - \cap(X | X \in E_R^+)$.
Step 4. Put

$$3NF(R) = \begin{cases} 1, & \text{if for all } X \in E_R, a \notin \mathcal{U}_p, a \notin X \text{ we have } C_R(X - a) = X - a; \\ 0, & \text{otherwise.} \end{cases}$$

Step 5. Stop.

It is almost obvious that this algorithm requires time $0(m^4n^2)$. According to [21,22], $M(S_R) \subseteq E_R \cup \{\mathcal{U}\} \subseteq S_R$ and E_R^+ is the set of antikeys of R, and $\mathcal{U}_p = \mathcal{U}_p(R)$ by [20,23]. Hence, $3NF(R) = 1$ iff R is in $3NF$ according to theorem 5.1 (2). Thus, we have

Theorem 5.5. *There exists an algorithm that given a relation R over $\mathcal{U}$, decides if R is in $3NF$ or not in polynomial time in the number of attributes and tuples of R.*

Consider the following algorithm.

Algorithm 5.6.
Input: a relation $R = \{h_1, \ldots, h_m\}$ over $\mathcal{U}$; a proper subset $X \subset \mathcal{U}$;
Output: $BCNF(R, X) \in \{0, 1\}$.

Step 1. Find the projection R' of R onto X.
Step 2. Construct the equality set $E_{R'} = \{h'_{ij} | 1 \leq i < j \leq n\}$, $h'_{ij} = \{a \in X | h_i(a) = h_j(a)\}$.
Step 3. Find $E^+_{R'}$ as the family of maximal elements of $E_{R'} - \{X\}$.
Step 4.

$$BCNF(R, X) = \begin{cases} 1, \text{ if for every } Y \in E^+_{R'}, \text{and } a \in Y \text{ we have } C_{R'}(Y - a) = Y - a; \\ 0, \text{ otherwise.} \end{cases}$$

Step 4. Stop.

Again, this algorithm requires time $0(m^4n^2)$. It follows immediately from corollary 5.2 that $BCNF(R) = 1$ iff R is in $BCNF$. Therefore, we have

Theorem 5.7. *There is an algorithm that given a relation R over $\mathcal{U}$ and $X \subseteq \mathcal{U}$, decides if the projection $R|_X$ is in $BCNF$ in polynomial time in the number of attributes and tuples of R.*

In the rest of this section we give two applications of the characterization of $BCNF$.

It was proved in [15] that the family of candidate keys of a relation (scheme) is an *antichain* (sometimes it is called a *Sperner family*), and for every antichain there exists a scheme the candidate keys of which are exactly the elements of this antichain. The following problem was formulated in [11]: find a condition which guarantees that the antichain of candidate keys uniquely determines the scheme. In other words, when does a family of candidate keys determine a closure (or SL) uniquely? In this case we say that a family of candidate keys satisfies the *uniqueness* condition. We also say that a scheme satisfies the uniqueness condition if its candidate keys satisfy one. That means, a scheme can be unambiguously reconstructed by its keys.

Theorem 5.8. *A scheme $< \mathcal{U}, F >$ satisfies the uniqueness condition iff it is in $BCNF$ and for every $X \in S_F, X \neq \mathcal{U}$ and $a \in X$ there is $b \notin X$ such that $(X - a) \cup b \in S_F$.*

Proof. According to [11], $< \mathcal{U}, F >$ satisfies the uniqueness condition iff for every X lying under an antikey, X is an intersection of antikeys.

In this case $< \mathcal{U}, F >$ is in $BCNF$ by corollary 5.2. Let $X \in S_F, X \notin \mathcal{U}$. Then $X - a$ is an intersection of antikeys and since $< \mathcal{U}, F >$ is in $BCNF$ we have $(X - a) \cup b \in S_F$ for some $b \notin X$.

Conversely, let $< \mathcal{U}, F >$ satisfy the conditions of theorem. Suppose X is an antikey and $a \in X$. Then $(X - a) \cup b \in S_F$ for some $b \notin X$. Since $(X - a) \cup b \subseteq Y$, Y is an antikey, we have $X - a = X \cap Y$. Therefore, all the sets of form $X - a$ can be represented as $X \cap Y, X, Y$ antikeys. It shows that all the elements of S_F except $\{\mathcal{U}\}$ are the intersections of antikeys, i.e. $< \mathcal{U}, F >$ satisfies the uniqueness condition. Theorem is proved.

Notice that the corollary 5.2. states exactly that $< \mathcal{U}, F >$ is in $BCNF$ iff $S_F - \{\mathcal{U}\}$ is an *independence system* [1]. One of the most important examples of an independence system is a *matroid* [1].

Corollary 5.9. *Let $S_F - \{\mathcal{U}\}$ be a family of independent sets of a matroid on $\mathcal{U}$ containing more than one base. Then $< \mathcal{U}, F >$ satisfies the uniqueness condition.*

Proof. If $S_F - \{\mathcal{U}\}$ is a family of independent sets of a matroid, then antikeys $X_1, \ldots, X_t$ of F are exactly the bases of this matroid, $t > 1$.

Consider $a \in X_i$. According to [1], $(X_i - a) \cup b$ is a base for some $b \in X_j, j \neq i$. Since $X_i - a = X_i \cap [(X_i - a) \cup b]$, each subset of an antikey is an intersection of antikeys and $< \mathcal{U}, F >$ satisfies the uniqueness condition.

Finishing this section, we prove that the problem to decide if $F^+ = F_R$ for given scheme $< \mathcal{U}, F >$ and a relation over $\mathcal{U}$ can be solved in a polynomial time if we know that $< \mathcal{U}, F >$ is in $BCNF$ and the number of keys is bounded by a constant. Really, minimum cover of F can be found in a polynomial time in $|F|$, see [44]. According to [52] we may construct such a minimum cover which consists of FD*s* $K_1 \to \mathcal{U}, \ldots, K_r \to \mathcal{U}$, where $K_1, \ldots, K_r$ are the candidate keys. If we are given a relation, we can decide whether R is in $BCNF$ in polynomial time in $|R| + n$ and also find its antikeys $X_1, \ldots, X_t$, see algorithm 5.6. Hence, $F^+ = F_R$ iff $\{K_1, \ldots, K_r\}^{-1} = \{X_1, \ldots, X_t\}$. Here $\{K_1, \ldots, K_r\}^{-1}$ is the family of all antikeys corresponding to the family of keys $\{K_1, \ldots, K_r\}$, i.e. the family of all maximal nonkeys. According to [50], the last equality can be checked up in a polynomial time in $r \cdot t \cdot n$ if r is bounded by a constant. This proves the polynomiality of checking $F^+ = F_R$. If the number of tuples of a relation is bounded by a constant and a relation is in $BCNF$, we can find out if $F_R = F^+$ in polynomial time too. See [27] for details.

6. Relation schemes with restrictions on functional dependencies

In this section we study the problem which was mentioned in [19] and, to our knowledge, has not been studied in detail formerly. That is, to study the schemes $< \mathcal{U}, F >$ such that $\mathcal{P}(F)$ is true, where $\mathcal{P}$ is a predicate. For instance, (4.10) represents a predicate $\mathcal{P}(F) =$ *'true'* iff $F \supseteq F'$, F' being a fixed family of FDs. Also, Mannila and Räihä [46] established some properties of the schemes in which the left-hand sides of all the FDs consist of one or two attributes. But we can use another idea as well. Many types of closures have been widely studied. Thus, each class of closures induces a predicate $\mathcal{P}$ such that $\mathcal{P}(F)$ is *'true'* iff C_F belongs to this class. In this section we are going to study some types of predicates that appear either by means of the restriction on the left-hand sides of FDs or by letting C_F belong to a given class of closures. The classes of closures to be studied in this section are the following: topological, exchange [1], antiexchange [25, 35], and separatory [19, 40]. Of course, these classes do not cover all the possibilities to introduce a predicate $\mathcal{P}$, but they demonstrate some typical results that can be obtained in this way.

For instance, it will be shown that some problems which are, generally speaking, NP-complete, become polynomial for the special classes of relation schemes. Also, some new results about keys, antikeys, prime attributes, normal forms e.t.c. can be obtained. Besides, for some classes we can guarantee the existence of an Armstrong relation whose number of tuples is polynomial in the number of attributes.

It is important to know what the complexity of the problem of recognizing these properties is. This is one of the topics of paper [27], and here we pay attention mostly to the structural properties of the classes of closures to be introduced, and to the complexity of known problems in the arising particular cases.

Since some classes of FDs have nice properties, one can either choose schemes of these classes or approximate a given scheme in one of this classes. In order to solve the approximation problem, we need to know the structure of the set of all closures C_F from a given class. In this paper we discuss only the problem if a given class is closed with respect to one of the operations described in proposition 4.4. That is the most important information to find approximation, cf. [18]. Notice that the approximation problem is completely solved for normal forms [18].

Now we are going to give the analysis of database concepts for some special classes of closures. We begin with topological closures.

6.1. Topological closures and unary dependencies

A FD is called *unary* if its left-hand side consists of unique element [46]. A closure C on $\mathcal{U}$ is called *topological* if

$$C(X \cup Y) = C(X) \cup C(Y) \text{ for all } X, Y \subseteq \mathcal{U}.$$

It is almost evident that C_F is topological iff there is a cover G of F consisting of unary dependencies. That is,

$$X \to a \in F^+ \Longrightarrow \exists b \in X : b \to a \in F^+. \tag{6.1}$$

It has been shown in [46] that if F consists only of unary FDs then to find a relation R with $F_R = F^+$ (i.e. an Armstrong relation for F) requires polynomial time in $|\mathcal{U}|$. Hence, prime attribute problem [43], $3NFTEST$ and $BCNFTEST$ [8,32,46] can be solved in polynomial time for unary FDs while they are NP-complete in general. Note also that Gottlob's problem mentioned above can be solved in polynomial time.

Given a family F of FDs, C_F is topological iff a minimum cover of F consists of unary FDs. Thus, if we are given a family F of FDs, we can check up if C_F is topological or not in time polynomial in $|F| + |\mathcal{U}|$.

A closure C_F is topological iff

$$X, Y \in S_F \Longrightarrow X \cup Y \in S_F, \tag{6.2}$$

i.e. S_F is a distributive lattice. Since S_F can be embedded in $P(\mathcal{U})$, it means that $|M(S_F)|$ is less than $|\mathcal{U}|$, i.e. the number of tuples of a minimal Armstrong relation is at most $|\mathcal{U}|$.

Theorem 3.1. and (6.1) immediately imply the following

Proposition 6.1. *Let F be a family of FDs. Then C_F is topological iff*

$$S_F = P(\mathcal{U}) - \bigcup_{a \to b \in F^+} [a, \mathcal{U} - b]. \tag{6.3}$$

According to [48] and (6.3), C_F is topological iff S_F is a distributive lattice and for every distributive lattice $\mathcal{L}$ we can find a scheme $< \mathcal{U}, F >$ such that C_F is topological and $\mathcal{L} \simeq S_F$.

The formula (6.3) gives rise to two matrix representations of topological closures.

Let F consist of unary FDs only. Suppose without loss of generality that the right-hand sides of FDs of F also consist of single elements. Define two $n \times n - (0,1)$-matrices $P^F = \|p^F_{ij}\|$ and $T^F = \|t^F_{ij}\|$, $i, j = 1, \ldots, n$, $\mathcal{U} = \{a_1, \ldots, a_n\}$ as follows:

$$p^F_{ij} = \begin{cases} 1, & a_i \to a_j \in F; \\ 0, & a_i \to a_j \notin F; \end{cases}$$

$$t^F_{ij} = \begin{cases} 1, & a_j \in C_F(a_i), \\ 0, & a_j \notin C_F(a_i). \end{cases}$$

Assume that $p^F_{ii} = 1$ for all i. Thus, every reflexive * $(0,1)$-matrix represents some topological closure as a matrix P^F. Note that some different matrices may represent the same closure.

Matrix T^F is transitive and reflexive. It is easy to see that each transitive and relexive matrix induces a topological closure with $C(\emptyset) = \emptyset$, and that different matrices induce different closures.

Now we are going to find the relationship between P^F and T^F.

Proposition 6.2. *If F consists of unary FDs only, then T^F is the transitive closure of P^F.*

Proof. Let $G^F = \|g^F_{ij}\|$ be the transitive closure of P^F. Suppose $g^F_{ij} = 1$. It means that $p^F_{ii_1} = 1$, $p^F_{i_1 i_2}, = 1, \ldots, p^F_{i_k j} = 1$ for some $a_{i_1}, \ldots, a_{i_k} \in \mathcal{U}$.

Then $a_i \to a_{i_1} \in F$ and $a_{i_1} \in C_F(a_i)$. Further, $a_{i_1} \to a_{i_2}$ and $a_{i_2} \in C_F(a_{i_1}) \subseteq C_F(a_1)$, e.t.c. Finally, $a_j \in C_F(a_i)$ and $t^F_{ij} = 1$.

Conversely, let $t^F_{ij} = 1$. Then $a_i \to a_j \in F^+$, i.e. it can be derived by using (F1)-(F4) from F. Clearly, (F1,F3,F4) do not lead us to new unary FDs. Hence, $a_i \to a_j$ can be derived only by (F2), i.e. there are such $a_{i_1}, \ldots, a_{i_k}$ that $a_i \to a_{i_1} \in F$, $a_{i_1} \to a_{i_2} \in F, \ldots, a_{i_k} \to a_j \in F$. That is, $p^F_{ii_1} = 1, \ldots, p^F_{i_k j} = 1$ and $g^F_{ij} = 1$. Therefore, $G^F = T^F$. Proposition is proved.

---

* We say that $(0,1) - (n \times n)$-matrix is reflexive (transitive) if so is binary relation whose adjancency matrix is the given matrix.

In the rest of the subsection we discuss the problems related to antikeys and $BCNF$.

Let F consist of unary FDs. Then the antikeys $\{X_1, \ldots, X_t\}$ can be characterized by the property that $X_i \cup X_j = \mathcal{U}$, $i \neq j$, due to (6.2) and [50]. Conversely, if $\{X_1, \ldots, X_t\}$ satisfies the above property, consider a SL generated by $\{X_1, \ldots, X_t, \emptyset, \mathcal{U}\}$. Clearly, it is a distributive lattice (i.e. it satisfies (6.2)) and its antikeys are exactly $\{X_1, \ldots, X_t\}$.

This fact immediately implies that if all the FDs in F are unary, $< \mathcal{U}, F >$ is in BCNF iff it has unique antikey X and $S_F = [\emptyset, X] \cup \{\mathcal{U}\}$.

Really, 'if' is obvious. To prove 'only if' suppose there are two antikeys X_1 and X_2. Since $X_1 \cup X_2 = \mathcal{U}$ and $< \mathcal{U}, F >$ is in BCNF, for every $a \in \mathcal{U}$ we have $a \in S_F$ and $S_F = P(\mathcal{U})$, i.e. it has unique antikey $\mathcal{U}$, a contradiction. Hence, $< \mathcal{U}, F >$ has unique antikey X and by corollary 5.2 $S_F = [\emptyset, X] \cup \{\mathcal{U}\}$.

6.2. Binary dependencies

A dependency is called *binary* if its left-hand side is a two-element set. A family of FDs is called binary if it has a cover consisting only of binary FDs.

It was proved in [9] that there exists a binary family F of FDs on $|\mathcal{U}|$ such that every Armstrong relation for F has at least exponential number of tuples in $|\mathcal{U}|$. Also, it was proved in [46] that the prime attribute problem remains NP-complete for binary FDs.

However, in order to check up if a family F of FDs is binary we only have to find a minimum cover G of F, because F is binary iff G consists only of binary FDs. Hence, this checking can be done in polynomial time.

In order to characterize the closures C_F for binary families F, remind the construction that appeared in [41]. Let C be a closure on $\mathcal{U}$. Define $C_2(X) = \cup(C(\{x, y\})|x, y \in X)$. Then C is said to have a *binary representation* iff for every $X \subseteq \mathcal{U}$ there is k such that $C(X) = C_2^k(X)$, and $C(x) = x$ for all $x \in \mathcal{U}$.

Proposition 6.3. *A family F of FDs is binary iff C_F has a binary representation.*

Proof. Let F be binary. Suppose without loss of generality that F itself consists of binary FDs. Suppose $a \in C_F(X)$. Then a can be derived as follows. At the first step, all the FDs from F being applied to X, we obtain X_1. Then, all the FDs being applied to X_1, we obtain X_2 etc. Finally, $a \in X_k$. Clearly, $X_1 \subseteq C_2(X), X_2 \subseteq C_2(X_1), \ldots, X_k \subseteq C_2(X_{k-1})$, i.e. $a \in X_k \subseteq C_2^k(X)$. Therefore, C has a binary representation.

Conversely, let C have a binary representation. Define F as the family of all FDs $\{x, y\} \to a$ such that $a \in C(\{x, y\})$. Clearly, F is binary and $C = C_F$.

There are two interesting classes of closures which are the subclasses of closures having binary representation.

Remind that S_F is in fact a lattice. If F is binary, $x \in S_F$ for all $x \in \mathcal{U}$ and S_F is atomistic [28], that is, every element of S_F is the join of atoms. An atomistic S_F (i.e. $x \in S_F\ \forall x \in \mathcal{U}$) is called *biatomic* [3] if $a \in C_F(X \cup Y)$ implies that there are $x \in C_F(X)$, $y \in C_F(Y)$ such that $a \in C_F(\{x, y\})$. (We modified the definition from [3] for our purposes). Clearly, if S_F is biatomic, C_F has a binary representation (cf. [41]). It also can be easily seen that to check up if S_F is biatomic requires polynomial time.

The other example is the following. Suppose C_F satisfies the properties: $C_F(X) = \cup(C_F(\{x,y\})|x,y \in X)$ and $C_F(x) = x$ for all $x \in \mathcal{U}$. Clearly, C_F has a binary representation, and S_F is biatomic. Moreover, the full characterization of S_F can be given. In fact, it follows from [38] that C_F satisfies the above property iff S_F is atomistic and 2-distributive (remind, that a lattice $< \mathcal{L}, \vee, \wedge >$ is called n-distributive [28,30] iff $\forall x, y_o, \ldots, y_n \in \mathcal{L} : x \wedge \bigvee_{i=0}^{n} y_i = \bigvee_{j=0}^{n}(x \wedge \bigvee_{i \neq j} y_i))$. It also can be shown that the recognizing of 2-distributivity requires polynomial time.

6.3. Exchange closures.

The closures satisfying *exchange property* were widely studied because they give one of the equivalent descriptions of matroids [1]. Remind, that C satisfies the exchange property (or it is an *exchange* closure for short) if

$$(x, y \notin C(A), x \in C(A \cup y)) \Longrightarrow y \in C(A \cup x) \ \forall A \subseteq \mathcal{U}, \forall x, y \in \mathcal{U}. \tag{6.4}$$

A pair $(\mathcal{U}, C)$, where C is an exchange closure on $\mathcal{U}$, is called a *matroid*. Note that there are many equivalent definitions of matroids [1].

In this subsection S_F is regarded to as a lattice. The lattices S_F for closures C_F satisfying (6.4) are exactly finite atomistic semimodular lattices [1]. These lattices are known to have complements.

Before presenting the properties of exchange closures, we prove one useful lemma about complemented lattices S_F.

Lemma 6.4. *Let F be a family of FDs such that S_F is a complemented lattice. Then the set $\mathcal{U}_p$ of prime attributes is $\mathcal{U} - C_F(\emptyset)$.*

Proof. If C_F is complemented, then the intersection of coatoms of S_F is the intersection of all the elements of S_F [33], i.e. $C_F(\emptyset)$. Since coatoms of S_F are antikeys [50], and the intersection of antikeys is the set of nonprime attributes [20, 23], the set of prime attributes is $\mathcal{U} - C_F(\emptyset)$.

Proposition 6.5. *Let F be a family of FDs such that S_F is a complemented lattice. Then the following are equivalent:*

1) $< \mathcal{U}, F >$ is in 2NF;
2) $< \mathcal{U}, F >$ is in 3NF;
3) $C_F(\emptyset) = \emptyset$.

Proof. If $C_F(\emptyset) = \emptyset$ then $\mathcal{U}_p = \mathcal{U}$ and $< \mathcal{U}, F >$ is in 3NF (and in 2NF) by theorem 5.1. Let $C_F(\emptyset) = X \neq \emptyset$. Suppose $< \mathcal{U}, F >$ is in 2NF. Then X is a prime set and by theorem 5.1 $\emptyset = (\mathcal{U} - C_F(\emptyset)) \cap C_F(\emptyset) = \mathcal{U}_p \cap X \in S_F$, a contradiction. Hence, $< \mathcal{U}, F >$ is neither in 2NF nor in 3NF. Proposition is proved.

Corollary 6.6. *Let F be a family of FDs such that C_F is an exchange closure. Then the following are equivalent: 1) $< \mathcal{U}, F >$ is in 2NF; 2) $< \mathcal{U}, F >$ is in 3NF; 3) $C_F(\emptyset) = \emptyset$.*

In order to characterize BCNF for exchange closures we need some new concepts.

If C_F is exchange, then candidate keys are called the bases of matroid (cf. [1]). Since the characterization of bases is well-known, it gives another characterization of families of FDs generating exchange closures. Antikeys are called copoints for exchange closures [1].

Let $|\mathcal{U}| = n$. Consider the following closure C_n^k [17], where $k \leq n$:

$$C_n^k(X) = \begin{cases} X, & |X| \leq k, \\ \mathcal{U}, & |X| > k. \end{cases}$$

C_n^k is called *uniform* (or $k - uniform$ closure). C_n^k is exchange closure whose antikeys are the sets of cardinality k and candidate keys are the sets of cardinality $k+1$ [17]. If $k = \lceil n/2 \rceil$ then there are $\binom{n}{\lceil n/2 \rceil}$ meet-irreducible elements in the SL of closed sets. Hence, if we are given a family F of FDs generating an exchange closure C_F, the minimal size of Armstrong relation for F may be exponential in $|\mathcal{U}|$.

Now we can characterize BCNF for exchange closures.

Proposition 6.7 *Let F be a family of FDs such that C_F is an exchange closure. Then $< \mathcal{U}, F >$ is in BCNF iff C_F is uniform closure.*

Proof. Clearly, if $C_F = C_n^k$ then $< \mathcal{U}, F >$ is in BCNF. Conversely, let $|\mathcal{U}| = n$, C_F be an exchange closure and $< \mathcal{U}, F >$ be in BCNF. Let X be an antikey, $X = \{a_1, \ldots, a_r\}$. Since $\emptyset$ is an independent set and $C_F(\emptyset) = \emptyset$ by theorem 5.1 (3), $\{a_1\}$ is independent by [1,6.3]. If $\{a_1, \ldots, a_{s-1}\}$ is independent, then $C_F(\{a_1, \ldots, a_{s-1}\}) = \{a_1, \ldots, a_{s-1}\}$ and again by [1,6.3] $\{a_1, \ldots, a_s\}$ is independent. Hence, X is independent, and so are the sets $X \cup a, a \notin X$. $C_F(X \cup a) = \mathcal{U}$ because X is an antikey. Since every independent set can be extended to a base, for some $a \notin X$ $X \cup a$ is a base. If for $b \notin X$ $X \cup b$ is not a base, there is a base $Y \subset X \cup b$, and $|Y| \neq |X \cup a|$, a contradiction. Hence, all the sets $X \cup a$, $a \notin X$, are the bases, i.e. candidate keys of $< \mathcal{U}, F >$. Now let there be two antikeys X_1 and X_2 with $|X_1| \neq |X_2|$. Then for some $a \notin X_1$ and $b \notin X_2$ $X_1 \cup a$ and $X_2 \cup b$ are two bases of a matroid having different cardinalities, a contradiction. Hence all the antikeys have the same cardinality k.

Let X be an antikey, $|X| = k$. Since antikeys and only they are meet-irreducible elements of S_F because they are copoints [1], for every $a \in X$ there is an antikey X' such that $X - a = X \cap X'$. Clearly, $X' = (X - a) \cup b$ for some $b \notin X$. If $x \notin X$, $x \neq b$ then $(X - a) \cup \{b, x\}$ is a candidate key.

Consider $X" = (X - a) \cup x$. If $C_F(X") \neq X"$ then $C_F(X") = Y \neq \mathcal{U}$ since $X"$ is a proper subset of a candidate key and $|Y| > k$, a contradiction. Hence $X" \in S_F$ and $X"$ is an antikey. Therefore, given an antikey X, $a \in X$ and $x \notin X$, $(X - a) \cup x$ is again an antikey. It shows that all the sets of cardinality k are the antikeys. Hence, $C_F = C_n^k$ is a uniform closure. Proposition is proved.

Mannila and Räihä [46] introduced the concept of *nonredundant set*. A set $X \subseteq \mathcal{U}$ is called nonredundant if $Y \to X \in F^+$ for no proper subset $Y \subset X$. Clearly, X is nonredundant iff $X - a \to X$ fails in F^+ for all $a \in X$, that is, $a \notin C_F(X - a)$. If C_F is exchange closure, this is the definition of independent set of a matroid. Hence, C_F is exchange iff for two nonredundant sets X and Y, $|X| > |Y|$, there is $a \in X - Y$ such that $Y \cup a$ is nonredundant [1].

It was proved in [46] that F is in BCNF iff every meet-irreducible element of S_F is nonredundant. A matroid is called uniform iff it is induced by a *uniform* closure. Combining the above result and proposition 6.6, we obtain

Corollary 6.8. *A matroid is uniform iff every copoint is independent.*

6.4. Antiexchange closures

A closure is said to satisfy the *antiexchange* property (or to be antiexchange) [24,25,31,35] if

$$(x, y \notin C(A), x \in C(A \cup y)) \Longrightarrow (y \notin C(A \cup x)) \forall A \subseteq \mathcal{U}\ \forall x, y \in \mathcal{U}$$

Let $X \subseteq \mathcal{U}$. A subset $Y \subseteq X$ is called a *minimal key* of X (w.r.t. a family F of FDs) if $C_F(Y) = C_F(X)$ and Y is a minimal set with this property.

Proposition 6.9 [24,25]. *Let F be a family of FDs. Then C_F is antiexchange iff every $X \subseteq \mathcal{U}$ has unique minimal key.*

Notice that a minimal key of $\mathcal{U}$ is a candidate key. Hence, if C_F is antiexchange, it has unique candidate key. According to [24,25], this candidate key K can be found as follows: $a \in K$ iff $a \notin C_F(\mathcal{U} - a)$. Hence, key (and prime attributes) can be found in polynomial time if we are given a relation scheme.

Consider the following example. Let $a \in \mathcal{U}$. Suppose $S = \{X \subseteq \mathcal{U} | a \in X\} \cup \{X \subseteq \mathcal{U} | a \notin X, |X| \leq k\} \cup \{\mathcal{U}\}$. Clearly S is a SL and according to [24,25] C_S is antiexchange. Therefore, the minimal size of Armstrong relation for an antiexchange closure may be exponential because $|M(S)| \geq \binom{n-1}{k}$, where $n = |\mathcal{U}|$.

Finally, notice that $\{ex_F(X) \to X | X \in S_F\}$ is a cover of F if C_F is antiexchange, where $ex_F(X) = \{a \in X | a \notin C_F(X - a)\}$.

6.5. Separatory closures

The concept of separatory SSL appeared in [40] in order to study the separation properties of SLs*. For our purposes, we will call a SL $S \subseteq P(\mathcal{U})$ *separatory* if $P(\mathcal{U}) - S$ is also a SL, i.e. if it is closed under intersection.

A closure C on $\mathcal{U}$ is called separatory if S_C is a separatory SL.

Proposition 6.10. *Let F be a family of FDs. Then C_F is separatory iff F has a cover of type $\{X_i \to a_i | i = 1, \ldots, p\}$, where $X_1 \subseteq X_2 \subseteq \ldots \subseteq X_p$.*

Proof. According to [40], a SL $S \subseteq P(\mathcal{U}), \{\mathcal{U}\} \in S$ is separatory iff it can be represented as

$$S = P(\mathcal{U}) - \bigcup_{i=1}^{p} [X_i, \mathcal{U} - a_i], \tag{6.5}$$

where $X_1 \subseteq \ldots \subseteq X_p$. Now the proposition follows from (3.1) and (6.5).

Corollary 6.11. *Let F be a family of FDs such that C_F is separatory. Then every nonredundant cover of F contains at most $(n-1)n^2$ FDs.*

Proof. According to proposition 6.10, F has a cover containing at most n^2 FDs. Hence, by [26], every nonredundant cover contains at most $(n-1)n^2$ FDs.

---

* Note that these properties had been studied formely by R.E. Jamison [31].

Similarly to the topological closures, separatory closures have a matrix representation. Let $\mathcal{U} = \{a_1, \ldots, a_n\}$. Given a closure C, define $(0,1) - n \times n-$matrix $P^C = \|p_{ij}^C\|$ as follows:

$$p_{ij}^C = \begin{cases} 1, & \mathcal{U} - \{a_i, a_j\} \in S_C; \\ 0, & \mathcal{U} - \{a_i, a_j\} \notin S_C. \end{cases} \tag{6.6}$$

Given a $(0,1) - n \times n-$matrix $P = \|p_{ij}\|$, define $C^P : P(\mathcal{U}) \to P(\mathcal{U})$ as follows:

$$C^P(X) = \begin{cases} \bigcap(\mathcal{U} - \{a_i, a_j\}, a_i, a_j \notin X,\ p_{ij} = 1) \text{ if such } a_i, a_j \text{ exist}, \\ \mathcal{U}, \text{ otherwise.} \end{cases} \tag{6.7}$$

A $n \times n-$ matrix $A = \|a_{ij}\|$ is called *absolutely determined* if every its submatrix has a saddle-point, i.e. $min\ max\ A' = max\ min\ A'$ for any submatrix A' [29].

Proposition 6.12 [29]. *The mappings (6.6) and (6.7) establish one-to-one mutually inverse correspondences between the families of separatory closures on $\mathcal{U}$ and $(0,1) - n \times n$ absolutely determined symmetrical matrices, where $n = |\mathcal{U}|$.*

Using this matrix representation, we obtain two results.

First notice that according to [40] every closure is a meet (in the sense of operation (4.6)) of separatory closures. Hence, it is interesting to know how many separatory closures exist.

Remind that $\alpha(n)$ is the number of all closures on $\mathcal{U}, |\mathcal{U}| = n$, satisfying $C(\emptyset) = \emptyset$. Let $\beta(n)$ stand for the number of all separatory closures on $\mathcal{U}$. Clearly, $\beta(n)$ is the number of $(0,1) - n \times n$ absolutely determined symmetrical matrices. According to [29], $\beta(n)$ is the number of $(0,1) - n \times n$ symmetrical matrices which can be reduced to Joung's form by some permutations of rows and columns. And this fact implies (we omit the calculations)

Proposition 6.13. $\binom{2n}{n} \leq \beta(n) \leq 2^n \cdot n! - 2n \cdot n! + 2^{n+1} - 2$.

Using $\alpha(n) \geq 2^{\binom{n}{\lceil n/2 \rceil}}$ and Stirling's formula, we obtain

Corollary 6.14. $\lim_{n \to \infty} \frac{\beta(n)}{\alpha(n)} = 0$.

The other corollary of proposition 6.12 is that the problem of recognizing $F^+ = F_R$ can be solved in polynomial time if we know that C_F is separatory. Really, given a relation $R, M(S_R)$ can be found in polynomial time in the number of tuples of R and attributes [21,22]. If C_F is separatory, all the elements of $M(S_F)$ have cardinality $n-2$, $n-1$ or n see (6.7). Hence, they also can be found in polynomial time in $|F|$. Since $F^+ = F_R$ iff $M(S_F) = M(S_R)$, the first equality can be checked up in polynomial time.

Notice also that since all the irreducible elements have cardinality $n, n-1$ or $n-2$, one can always find an Armstrong relation for a separatory closure containing at most $1 + (n+1)^2/4$ tuples.

Finishing this subsection we show that for separatory closures 3NF implies BCNF, and that every separatory closure has unique minimal key.

Proposition 6.15. *Let F be a family of FDs such that C_F is separatory. Then $< \mathcal{U}, F >$ is in 3NF iff it is in BCNF.*

Proof. Let $< \mathcal{U}, F >$ be in 3NF. We have to prove that it is in BCNF. Suppose without loss of generality $\mathcal{U} = \{a_1, \ldots, a_n\}$, $W - a_i \in S_F$ iff $i \leq k$.

According to (6.7), $W - a_i$, $i \leq k$ are exactly coatoms of S_F, i.e. antikeys. Hence, $\{a_{k+1}, \ldots, a_n\}$ is the set of nonprime attributes. According to theorem 5.1, $W - \{a_i, a_j\} \in S_F$ for all $i \leq k$ and all j, i.e. $S_F = \bigcup_{i=1}^{k} [\emptyset, W - a_i] \cup \{\mathcal{U}\}$. Therefore, $< \mathcal{U}, F >$ is in BCNF.

Corollary 6.16. *If C is a separatory closure, it has unique candidate key K which can be found as follows: $a \in K$ iff $a \notin C(\mathcal{U} - a)$.*

Proof. Use the designations of the previous corollary. Let $\mathcal{U}_p = \{a_1, \ldots, a_k\}$ be the set of prime attributes. Since $\mathcal{U}_p - a_i \subseteq W - a_i \in S_C, C(\mathcal{U}_p - a_i) \neq \mathcal{U}$. Hence, $\mathcal{U}_p$ is unique candidate key. Clearly, $a_i \notin C(\mathcal{U} - a_i)$ iff $i \leq k$.

Remarks. 1) The concept of a separatory sublattice had been introduced as well. It can be used if we study topological closures represented as distributive lattices. See [37] for details.

2) We have shown that the closures of two types have unique candidate key which can be found as follows: $a \in K$ iff $a \notin C(\mathcal{U} - a)$. In fact, a closure has unique candidate key iff K thus constructed is a key. See [7].

We finish the section by the propositions summing up all the results about subsets of Cl_n generated by closures considered above. Let TC_n Bi_n, Ex_n, AEx_n, Sep_n be the family of topological (having binary representation, exchange, antiexchange, separatory) closures in Cl_n.

Propositions 6.17.

1) TC_n and Bi_n are closed under $\wedge$ but not $\vee$.
2) Ex_n and AEx_n are closed under $\vee$ but not $\wedge$.
3) Sep_n is closed under neither $\vee$ nor $\wedge$.

Proof. 1) Let $C_1, C_2 \in TC_n$. Then $C_1 = C_{F_1}, C_2 = C_{F_2}$ where F_1 and F_2 consist only of unary FDs. According to (3.1), $C_1 \wedge C_2 = C_{F_1 \cup F_2} \in TC_n$. Analogously $C_1, C_2 \in Bi_n$ implies $C_1 \wedge C_2 \in Bi_n$. The contraexamples related to the operation $\vee$ can be easily constructed for the both cases.
2) See [1] for Ex_n and [24,25] for AEx_n.
3) See [40].

Let $2NF_n$, $3NF_n, BCNF_n \subseteq Cl_n$ be the families of closures induced by schemes in $2NF, 3NF$ and $BCNF$ respectively.

Proposition 6.18 [18]. *1) Neither $2NF_n$ nor $3NF_n$ is closed under $\vee$ or $\wedge$ in Cl_n. 2) $BCNF_n$ is a distributive sublattice of Cl_n.*

7. Conclusion

In this paper the lattice theoretic approach to the analysis of functional dependencies in relational databases has been developed. Formerly in many papers having studied formally the functional dependencies, closure operations were mostly used to represent them. Here we have proposed to make use of semilattices instead of closure operations. The use of semilattice description is formally equivalent to that of closures but sometimes it is more convenient because of the simplicity of the representation of semilattice of closed sets by functional dependencies.

Partially ordered set of closures on a set of attributes was studied in [11] as a model of changing databases. The semilattice representation having been used, we proved that this partially ordered set is a lattice and characterized it. This characterization gives rise to some application which might be useful for practical purposes. For instance, some ways to construct arbitrary families of functional dependencies from given families are proposed.

We have given a new lattice theoretic characterization of normal form relation schemes. Using this characterization we got some applications. First, we proved that recognizing relation schemes in third and Boyce-Codd normal forms can be done in polynomial time if we are given a relation instead of a relation scheme. We also have given a new characterization of schemes which are unambiguously determined by their keys as BCNF schemes satisfying an additional condition.

In the last section of the paper we have been studying the relation schemes satisfying some special conditions providing the closures to belong to a given class of closures. On this way relationships between functional dependencies and various objects having lattice representation (such as distributive lattice, matroids [1], antimatroids or convex geometries [24,25,35], separatory subsemilattices [40]) have been found.

In the rest of the paper we are going to outline some problems to be solved. First, notice that all the results related either to the representation of S_F (3.1) or to the lattice Cl_n can be interpreted for functional dependencies.

Second, third and Boyce-Codd normal forms are the main and the oldest examples of normal forms. It seems to be quite interesting to obtain a lattice theoretic characterization of other normal forms, because it may be useful, for instance in order to receive the results about complexity. Now the characterization of object normal form introduced by J. Biskup [6] is also known and this characterization gives rise to a polynomial algorithm for recognizing this normal form, see [7].

We noticed in the paper that S_F is, generally speaking, a lattice whose operations $\wedge$ and $\vee$ can be expressed as $X \wedge Y = X \cap Y, X \vee Y = C_F(X \cup Y)$. However, S_F was not investigated as being a lattice formerly. On this way we can make use of well-developed lattice theory more profoundly. The closures corresponding to distributive, 2-distributive [30,38], geometric [1,28], biatomic [3] lattices were studied in section 6.

The last mentioned class of lattices is the generalization of so-called *convexity lattices* [4]. In turn, convexity lattices were introduces to generalize the lattices of convex sets. Consider one of the most important examples of finite convexity lattice. Let $\mathcal{U} = \{a_1, \ldots, a_n\}$ and $F = \{\{a_i, a_j\} \to a_k | 1 \leq i < k < j \leq n\}$. Then S_F is so-called lattice $Co(n)$ [3,4]. It is in fact the lattice of points and segments of n collinear points in a vector space.

It was mentioned in [4] that many geometric concepts have interpretation in convexity lattices. This idea was particularly developed in [39, 42]. Thus the use of finite convexity lattices allows us to interpret some geometric concepts for functional dependencies. Notice that the idea to attract geometry to database theory was also proposed in [51].

One of the most important constructions in lattice theory is the direct product. In [17] the concepts of direct product and decomposition of closures were introduced. However, there is no one-to-one correspondence between direct product decompositions of S_F regarded to as a lattice and those of C_F. Knowing the structure of direct product decompositions of S_F seems to be useful because it might simplify the algorithm of derivation of FDs if we know a decomposition of S_F.

We plan to dedicate further research to the problems mentioned above.

References

1. M. Aigner: Combinatorial Theory. Springer Verlag, New York, 1979.

2. W.W. Armstrong: Dependency structures of data base relationships. Inf. Process. 74 North-Holland Pub. Co. (1974) 580-583.

3. M. K. Bennett: Biatomic lattices. Alg. Univ. 24(1987),60-73.

4. M. K. Bennett, G.Birkhoff: Convexity lattices. Alg. Univ. 20(1985)1-26.

5. J. Biskup: private communication. (August 1989).

6. J. Biskup: Boyce-Codd normal form and object normal forms, Inf. Process. Letters 32(1989) 29-33.

7. J.Biskup, J. Demetrovivcs, L.O. Libkin, I.B. Muchnik: On the relational database schemes having unique minimal key. Manuscript, 1990.

8. C. Beeri, P.A. Bernstein: Computational problems related to the design of normal form relational schemes. ACM Trans. Database Syst. 4, 1 (1979)30-59.

9. C. Beeri, M. Dowd, R. Fagin, R. Statman: On the structure of Armstrong relations for functional dependencies. J. ACM 31, 1(1984) 30-46.

10. A. Békéssy, J. Demetrovics: Contribution to the theory of data base relations. Discrete Mathematics, North-Holland 27(1979) 1-10.

11. G. Burosch, J. Demetrovics G.O.H. Katona: The Poset of Closures as a Model of Changing Databases. Order, D. Reidel Publ. Co. 4(1987)127-142.

12. G. Burosch, J. Demetrovics, G.O.H. Katona, D.J. Kleitman, A.A. Sapozhenko: On the number of databases and closure operations. Preprint of Math. Inst. of Hung. Acad of Sci., Budapest, 4(1988)1-33.

13. S. Ceri, G. Gottlob: Normalization of relations and Prolog. Comm. ACM 29(1986)-524-544.

14. E.F. Codd: A relational model for large shared data banks. Comm. ACM 13(1970) 377-387.

15. J. Demetrovics: On the equivalence of candidate keys with Sperner-systems. Acta Cybernetica 4, 3(1979) 247-252.

16. J. Demetrovics: Candidate keys and antichains. SIAM J. Alg.Disc.Math. 1, 1(1980)92.

17. J. Demetrovics, Z. Füredi, G.O.H. Katona: Minimum matrix representation of closure operations. Discrete Applied Math. 11(1985) 115-128.

18. J. Demetrovics, G. Hencsey, L.O. Libkin, I.B. Muchnik: Normal form relation schemes: a new characterization, paper in preparation.

19. J. Demetrovics, L.O. Libkin, I.B. Muchnik: Functional dependencies and the semilattice of closed classes. Proceedings of MFDBS 89, Springer LNCS 364(1989),136-147.

20. J. Demetrovics, V. D.Thi: Keys, antikeys and prime attributes. Annales Univ. Sci., Sect. Comp., Budapest 8(1987)37-54.

21. J. Demetrovics, V. D.Thi: Some results about functional dependencies. Acta Cybernetica 8, 3(1988)273-278.

22. J. Demetrovics, V.D. Thi: Relations and minimal keys. Acta Cybernetica 8, 3(1988) 279-285.

23. J. Demetrovics, V. D.Thi: Describing keys and prime attributes by antikeys, to appear.

24. P.H. Edelman: Abstract convexities and meet-distributive lattices. Contemp. Math. 57(1986)127-150.

25. P.H. Edelman, R.E. Jamison: The theory of convex geometries. Geom. Dedicata 19(1985)247-270.

26. G. Gottlob: On the size of nonredundant FD-covers. Information Processing Letters 24 (1987) 355-360.

27. G. Gottlob, L.O. Libkin: Investigations on Armstrong relations, dependency inference, and excluded functional dependencies. To appear in Acta Cybernetica.

28. G. Grätzer: General Lattice Theory. Springer Verlag, New York, 1978.

29. V.A. Gurvich, L.O. Libkin: Absolutely determined matrices. Math. Soc. Sci. 19 (1990) ??-??.

30. A.P. Huhn: Schwach distributive verbände. I, Acta Sci. Math. Szeged 33(1972)297-305.

31. R.E. Jamison: A perspective on abstract convexity: classifying alignments by varieties, in: Convexity and Related Combinatorial Geometry (eds. D.C. Kay and M. Breen) Dekker, New York (1982)113-150.

32. J.H. Jou and P.C. Fischer: The complexity of recognizing 3NF relation schemes. Information Processing Letters 14(1982)187-190.

33. J. Kalman: Property of algebraic lattice whose compact elements have complements. Alg. Univ. 22(1986)100-101.

34. D.J. Kleitman: Extremal properties of collections of subsets containing no two sets and their union. J. Combin. Th. (A) 20(1976)390-392.

35. B. Korte, L. Lovász: Shelling structures, convexity and a happy end, Graph Theory and Combinatorics. Proc. Cambridge Combinatorial Conference (B. Bollobás, ed.) Academic Press, London (1984)217-232.

36. L.O. Libkin: Minimal sets of choice functions generating the basic classes. Automation and Remote Control 49(12) (1988), 1662-1665.

37. L.O. Libkin: Separation theorems for lattices. MTA SZTAKI Közlemények (1988) 93-100.

38. L.O. Libkin: On the characterization of non-modular n-distributive lattices. Preprint No. 18. Mathematical Institute of the Hungarian Academy of Sciences, 1989.

39. L.O. Libkin: Parallel axiom in convexity lattices. Preprint No 79. Mathematical Institute of the Hungarian Academy of Sciences, 1989.

40. L.O. Libkin, I.B. Muchnik: Separatory subsemilattices and their properties. MTA SZTAKI Közlemények 39(1988)83-92.

41. L.O. Libkin, I.B. Muchnik: On a subsemilattice-lattice of a semilattice. MTA SZTAKI Közlemények 39(1988)101-110.

42. L.O. Libkin, I.B. Muchnik: Halfspaces and hyperplanes in convexity lattices. Preprint No. 51. Mathematical Institute of the Hungarian Academy of Sciences, 1989.

43. C.L. Lucchesi, S.L. Osborn: Candidate keys for relations. J. Comput. Syst. Sciences 17, 2(1978) 270-279.

44. D. Maier: The Theory of Relational Databases. Comp. Sci. Press, Rockville, MD, 1983.

45. H. Mannila, K.J. Räihä: Design by example: an application of Armstrong relations. J. Comp. Syst. Sci. 33(1986)126-141.

46. H. Mannila, K.J. Räihä: Practical algorithms for finding prime attributes and testing normal forms. Proc. of the eighth ACM SIGACT-SIGMOD-SIGART Symp. on Principles of Database Systems, Philadelphia, Pennsylvania, ACM Press (1989)128-133.

47. J. Paradaens, P.De Bra, M. Gyssens, D. Van Gucht: The Structure of the Relational Database Model. Springer Verlag, Berlin, 1989.

48. I. Rival: Maximal sublattices of finite distributive lattices. Proc AMS 44(1974)263-268.

49. A.B. Romanowska, J.D.H. Smith: Bisemilattices of subsemilattices. J. Algebra 70(1981)78-88.

50. V. D. Thi: Minimal keys and antikeys. Acta Cybernetica 7, 4(1986) 361-371.

51. M. S. Tsalenko: Modeling of Semantics in Databases. (Russian), Nauka, Moscow, 1989.

52. M. Wild: Implication bases for finite closure systems. Preprint No.1210. Darmstadt 1989.