Acta Cybernetica, Vol. 10, Nr. 3, Szeged, 1992

On the interaction between closure operations and choice functions with applications to relational databases*

János Demetrovics † Gusztáv Hencsey† Leonid Libkin‡
Ilya Muchnik§

Abstract

A correspondence between closure operations and special choice functions on a finite set is established. This correspondence is applied to study functional dependencies in relational databases.

1 Introduction

Having been introduced in connection with some topological problems, *closure operations* were applied in various branches of mathematics. In the last years they were successfully applied to study so-called *functional dependencies* (FDs for short) in relational databases. Now we recall some definitions and facts; they can be found in [DK],[DLM1].

Let $U = \{a_1, \ldots, a_n\}$ be a finite set of *attributes* (e.g. name, age etc.) and $W(a_i)$ the domain of a_i. Then a subset $R \subseteq W(a_i) \times \ldots \times W(a_n)$ is called a *relation over* U.

A *functional dependency* (FD) is an expression of form $X \longrightarrow Y$, $X, Y \subseteq U$. We say that FD $X \longrightarrow Y$ holds for a relation R if for every two elements of R with identical projections onto X, the projections of t hese elements onto Y also coincide. According to [Ar], the family $\mathcal{F}$ of all FD's that hold for R satisfies the properties (F1)-F4):

(F1) $(X \longrightarrow X) \in \mathcal{F}$;
(F2) $(X \longrightarrow Y) \in \mathcal{F}$ and $(Y \longrightarrow V) \in \mathcal{F}$ imply $(X \longrightarrow V) \in \mathcal{F}$;

*Research partially supported by HungarianNational Foundation for Scientific Research, Grant 2575.

†Computer and Automation Institute, P.O.Box 63, Budapest H-1518, Hungary; E-mail: h935dem@ella.hu and h103hen@ella.hu.

‡Department of Computer and Information Science, University of Pennsylvania, Philadelphia PA 19104, USA; E-mail: libkin@saul.cis.upenn.edu.

§24 Chestnut St., Waltham MA 02145, USA; E-mail: ilya@darwin.bu.edu.

(F3) $(X \longrightarrow Y) \in \mathcal{F}$ and $X \subseteq V, W \subseteq Y$ imply $(V \longrightarrow W) \in \mathcal{F}$;
(F4) $(X \longrightarrow Y) \in \mathcal{F}$ and $(V \longrightarrow W) \in \mathcal{F}$ imply $(X \cup V \longrightarrow Y \cup W) \in \mathcal{F}$.

Conversely, given a family of FD's satisfying (F1)-(F4) (so-called *full family*), there is a relation R over U generating exactly this family of FD's, see [Ar] and also [BDFS] for a constructive proof.

We shall write a_i instead of $\{a_i\}$ throughout the paper. Let R be a relation over $U, X \subseteq U$ and put $L_R(X) = \{a \in U | X \longrightarrow a$ holds for $R\}$. Then L_R satisfies

(C1) $X \subseteq L_R(X)$;
(C2) $X \subseteq Y \Longrightarrow L_R(X) \subseteq L_R(Y)$;
(C3) $L_R(L_R(X)) = L_R(X)$,

i.e. L_R is a *closure operation*. Note that the properties (C1)-(C3) may be concisely expressed as $X \subseteq L_R(Y)$ iff $L_R(X) \subseteq L_R(Y)$. Given a closure L (sometimes we shall omit the word "operation"), there is a relation R over U with $L = L_R$, see [De1].

A set $X \subseteq U$ is called *closed* (w.r.t. a closure L) if $L(X) = X$. Let $Z(L)$ stand for the family of all closed sets w.r.t. L. Then

(S1) $U \in Z(L)$,
(S2) $X, Y \in Z(L)$ implies $X \cap Y \in Z(L)$,

i.e. $Z(L)$ is a *semilattice*. Given a semilattice $Z \subseteq 2^U$ define $L(X) = \cap\{Y | X \subseteq Y, Y \in Z\}$. Then L is a closure with $Z(L) = Z$. Therefore, we can think of semilattices providing an equivalent description of closures and full families of FD's.

A closure is an extensive operation $(X \subseteq L(X))$. The operations satisfying the reverse inclusion (called *choice functions*) were also widely studied in connection with the theory of rational behaviour of individuals and groups, see [AM],[Ai],[Mo]. We give some necessary definitions.

A mapping $C : 2^U \longrightarrow 2^U$ satisfying $C(X) \subseteq X$ for every $X \subseteq U$, is called a choice function. U is interpreted as a set of alternatives, X as a set of alternatives given to the decision-maker to choose the best and $C(X)$ a s a choice of the best alternatives among X.

There were introduced some conditions (or properties) to characterize the rational behaviour of a decision-maker. The most important conditions are the following (see [AM],[Ai],[Mo]):

Heredity (<u>H</u> for short):

$$\forall X, Y \subseteq U : X \subseteq Y \Longrightarrow C(Y) \cap X \subseteq C(X);$$

Concordance (<u>C</u> for short):

$$\forall X, Y \subseteq U : C(X) \cap C(Y) \subseteq C(X \cup Y);$$

Out casting (<u>O</u> for short):

$$\forall X, Y \subseteq U : C(X) \subseteq Y \subseteq X \Longrightarrow C(X) = C(Y);$$

Monotonicity ($\underline{M}$ for short):

$$\forall X, Y \subseteq U : X \subseteq Y \Longrightarrow C(X) \subseteq C(Y).$$

Let P be a *binary relation* on U, i.e. $P \subseteq U \times U$. Let $C_P(X) = \{a \in X | (\nexists b \in X : (b, a) \in P)\}$.

One of the central results of the theory of choice functions states that a choice function can be represented as C_P for some P iff it satisfies $\underline{H}$ and $\underline{C}$.

Given a closure operation L, we can define choice functions $C(X) = L(U - X) \cap X$ and $C(X) = U - L(U - X)$. In Section 2 we characterize the choice functions of the second type as satisfying $\underline{M}$ and $\underline{O}$. In the other sections we use this correspondence to transfer the properties of choice functions to closures and to apply them to the study of FD's. In Section 3 we use the logical representation of choice functions (see [VR],[Li1]) to construct a similar representation and characterization of closure operations.

In Section 4 new properties of closure operations are obtained and studied by new properties being added to $\underline{M}$ and $\underline{O}$.

Finally, in the Section 5, we use choice functions to construct a structural representation for so-called *functional independencies* (cf. [Ja]) in the same way as closures were used to represent FD's.

2 The main correspondence

Let L be a closure operation. Define two choice functions associated with L as follows:

$$C_L(X) = L(U - X) \cap X,$$

$$C^L(X) = U - L(U - X), X \subseteq U.$$

Note that both C_L and C^L uniquely determine the closure L, in fact, $L(X) = X \cup C_L(U - X)$ and $L(X) = U - C^L(U - X)$. For every $X \subseteq U$ the sets $C_L(X)$ and $C^L(X)$ form a partition of X, i.e. $C_L(X) \cap C^L(X) = \emptyset$ and $C_L(X) \cup C^L(X) = X$.

Theorem 1 *The mapping $L \longrightarrow C^L$ establishes a one-to-one correspondence between the closure operations and the choice functions satisfying $\underline{O}$ and $\underline{M}$.*

Proof. Let L be a closure operation. We prove that C^L satisfies $\underline{M}$ and $\underline{O}$.

Let $x \in C^L(X)$ and $X \subseteq Y$. Then $x \notin L(U - X)$ and since $U - Y \subseteq U - X$, we have $x \notin L(U - Y)$, i.e. $x \in C^L(Y)$. Hence, C^L satisfies $\underline{M}$.

Let $X \subseteq U$. Then $L(L(U-X)) = L(U-X)$. Using $L(U-X) = U - C^L(X)$, we obtain that $U - C^L(U-(U-C^L(X)) = U - C^L(X)$, i.e. $C^L(C^L(X)) = C^L(X)$. Now let $C^L(X) \subseteq Y \subseteq X$; Since C^L satisfies $M, C^L(C^L(X)) \subseteq C^L(Y)) \subseteq C^L(X)$ and $C^L(X) = C^L(Y)$. Therefore, C^L satisfies $\underline{O}$.

Let C be a choice function satisfying $\underline{O}$ and $\underline{M}$. Consider $L(X) = U - C(U-X)$. We prove that L is a closure. Clearly, $X \subseteq L(X)$. If $X \subseteq Y$ and $x \in L(X)$, then $x \notin C(U-X)$ and $x \notin C(U-Y)$, i.e. $x \in L(Y)$. Since C satisfies $O, C(C(U-X)) = C(U-X)$. Applying $C(U-X) = U - L(X)$ we obtain $L(L(X)) = L(X)$. Hence, L is a closure and $C^L = C$.

To finish the proof, note that the mapping $L \longrightarrow C^L$ is injective, because for two distinct closures L_1 and L_2 with $L_1(X) \neq L_2(X)$ one has $C^{L_1}(U-X) \neq C^{L_2}(U-X)$. The theorem is proved.

□

Let $\mathcal{K}$ be a property of choice functions. We say that a choice function C satisfies $\overline{\mathcal{K}}$ if its complement $\overline{C}$ satisfies $\mathcal{K}$. (The complementary function $\overline{C}$ of C is defined as follows: $\overline{C}(X) = X - C(X)$ for $X \subseteq U$.)

Corollary 1 *The mapping $L \longrightarrow C_L$ establishes one-to-one correspondence between the closure operations and the choice functions satisfying $\underline{H}$ and $\underline{O}$.*

Proof. It follows from the facts that C_L and C^L are complementary choice functions and that $\underline{H} = \overline{\underline{M}}, \underline{M} = \overline{\underline{H}}$, see [Ai].

□

3 On logical representation of closure operations and choice functions

The family of all choice functions on U equipped with the operations $\cup, \cap$ and $^-$, is a Boolean algebra. Logical representation of the choice functions was introduced in [VR] to show that this Boolean algebra is isomorphic to one consisti ng of tuples of n Boolean functions, each depending on at most $n-1$ variables.

Let $U = \{a_1, \ldots, a_n\}$, $X \subseteq U$. Define

$$\beta_i(X) = \begin{cases} 1, & a_i \in X, \\ 0, & a_i \notin X, \end{cases}$$

$$\begin{aligned} \beta^i(X) &= (\beta_1(X), \ldots, \beta_{i-1}(X), \beta_{i+1}(X), \ldots, \beta_n(X)) \text{ and} \\ \beta^Z(X) &= (\beta_{i_1}(X), \ldots, \beta_{i_k}(X)) \end{aligned}$$

where $\{a_{i_1}, \ldots, a_{i_k}\} = U - Z$ and $i_1 < \ldots < i_k$.

Definition [VR]. *A family $\langle f_1^C, \ldots f_n^C \rangle$ of Boolean functions, each depending on $n-1$ variables, is called a first logical form of a choice function C if for every $a_i \in U$ and $X \subseteq U$:*

$$a_i \in C(X) \text{ iff } a_i \in X \text{ and } f_i^C(\beta^i(X)) = 1.$$

Definition [Lil]. *A family $\langle f_\emptyset^C, \ldots, f_U^C \rangle$ of Boolean functions indexed by subsets of U, is called a second logical form of a choice function C if for every $Z, X \subseteq U$:*

$$Z = C(X) \text{ iff } Z \subseteq X \text{ and } f_Z^C(\beta^Z(X)) = 1.$$

Note that f_Z^C depends on $n - |Z|$ variables.

Each logical form uniquely determines a choice function. By [VR], every tuple of Boolean functions, each depending on $n-1$ variables, is a first logical form of some choice function, moreover, $C \longrightarrow \langle f_1^C, \ldots, f_n^C \rangle$ is an isomorphism of Boolean algebras.

A family $\langle f_\emptyset, \ldots, f_U \rangle$, f_Z depends on $n - |Z|$ variables, is a second logical form of some choice function iff for each $Z \subseteq U$ the set $\{f_Z(\beta^Z(X)) : Z \subseteq X\}$ contains a unique one.

Let L be an operation satisfying (C1), i.e. $X \subseteq L(X)$ for all $X \subseteq U$. We can introduce two logical forms as before.

Definition. *A family $\langle f_1^L, \ldots, f_n^L \rangle$ of Boolean functions, each depending on at most $n-1$ variables, is called a first logical form of L if for every $a_i \in U$ and $X \subseteq U$:*

$$a_i \in L(X) \text{ iff } a_i \in X \text{ or } f_i^L(\beta^i(X)) = 1.$$

Let $Z = \{a_{i_1}, \ldots, a_{i_k}\}, i_1 < \ldots < i_k$, and $\beta_Z(X) = (\beta_{i_1}(X), \ldots, \beta_{i_k}(X))$.

Definition. *A family $\langle f_\emptyset^L, \ldots f_U^L \rangle$ of Boolean functions indexed by subsets of U, f_Z^L depends on $|Z|$ variables, is called second logical form of L if for every $Z, X \subseteq U$:*

$$Z = L(X) \text{ iff } X \subseteq Z \text{ and } f_Z^L(\beta_Z(X)) = 1.$$

We use these logical forms to characterize the closure operations among all the operations satisfying (C1).

Theorem 2 *Let L satisfy (C1). Then L is a closure operation iff all the functions $f_i^L, i = 1, \ldots, n; f_Z^L, Z \subseteq U$, are monotonic.*

Proof. Since $a_i \in \underline{L}(X)$ iff $a_i \in X$ or $a_i \in C_L(U - X)$, we have $f_i^L(\beta^i(X)) = f_i^{C_L}(\beta^i(U - X))$, i.e. $\overline{f_i^L} = (f_i^{C_L})^*$, where f^* stands for the dual function. Analogously, we obtain that $\overline{f_Z^L} = (f_{U-Z}^{C^L})^*$ (note that f_Z^L and $f_{U-Z}^{C^L}$ depend on the same variables). Since $f_i^{C_L} = \overline{f_i^{C^L}}$ theorem 1 and the following facts imply the theorem: (1) C^L satisfies $\underline{M}$ iff all the functions $f_i^{C^L}$, $i = 1, \ldots, n$, a re monotonic (cf. [VR]);

(2) C^L satisfies $\underline{O}$ iff all the functions $\overline{f}_Z^{C^L}$, $Z \subseteq U$, are monotonic (cf. [Li1]). The theorem is proved.

□

Remark. A set of attributes $X \subseteq U$ is called a *candidate key* (w.r.t. a relation R) if $L_R(X) = U$ and for every $Y \subset X : L_R(Y) \neq U$. The problem of finding the candidate keys (or a candidate key) is one of the most important problems in the theory of relational databases, see e.g. [BDFS],[De2]. According to the previous theorem, the candidate keys are exactly the lower units of monotonic function $f_U^{L_R}$. Hence, we can apply a recognition algorithm for monotonic Boolean functions to construct an algorithm of finding the candidate keys. (Note that if we are given a set of FD's, we can calculate a value $f_U^{L_R}$ in polynomial time in the size of the set of FDs. However, the problem of finding all the candidate keys is NP-hard, see [BDFS]).

Some other aspects of the applications of recognition of monotonic Boolean functions to the study of choice functions satisfying $\underline{M}$ and $\underline{O}$ (and, hence, closure operations) can be found in [Li2].

4 On the properties of closures induced by the properties of choice functions

In this section we consider the closures for which choice functions C_L and C^L defined in Section 2 satisfy some additional properties. Note that in the theory of choice functions such properties are ussually studied in some fixed combinations. These combinations explain the use of C_L and C^L. E.g., the property $\underline{C}$ (Concordance) is usually studied together with $\underline{H}$ (see [Ai], [AM],[Mo],[Li1]). Thus, studying this property we consider C_L (moreover, the property $\underline{C}$ implies monotonicity and there is no reason to consider C^L).

Property $\underline{C}$. As it was mentioned, we consider the functions C_L.

Let L be a closure and $\mathcal{F}_L$ a corresponding full family of FD's. Recall that an FD $X \longrightarrow Z$ is called *nontrivial* [De2],[DLM1] if $X \cap Z = \emptyset$. Let P_6 stand for the (Post) class consisting of conjunctions and constants, cf. [Po].

Proposition 1 *Let L be a closure operation on U. The following are equivalent:*

1) C_L satisfies the property $\underline{C}$;

2) $L(X) \cap L(Y) - (X \cup Y) \subseteq L(X \cap Y)$ for all $X, Y \subseteq U$;

3) If $X \longrightarrow Z$ and $Y \longrightarrow Z$ are nontrivial FD's from $\mathcal{F}_L$, then $X \cap Y \longrightarrow Z \in \mathcal{F}_L$;

4) $(X \longrightarrow a) \in \mathcal{F}_L$ iff $U - \{a, b\} \longrightarrow a$ for all $b \notin X$, where $a \notin X$;

5) For all $i = 1, \ldots, n : f_i^L \in P_6$.

Proof. $1 \longrightarrow 2$. Let C_L satisfy $\underline{C}$. Then for all $X, Y \subseteq U : C_L(U - X) \cap C_L(U - Y) \subseteq C_L(U - X \cap Y)$. Using $C_L(Z) = L(U - Z) \cap Z$ we obtain $L(X) \cap L(Y) - (X \cup Y) \subseteq L(X \cap Y) - (X \cap Y)$. Hence, 2 hold s.

$2 \longrightarrow 3$. Let $X \longrightarrow Z$ and $Y \longrightarrow Z$ be nontrivial FD's from $\mathcal{F}_L$. Then so are $X \longrightarrow a$ and $Y \longrightarrow a$ for all $a \in Z$. Since $a \in L(X) \cap L(Y) - (X \cup Y)$, we have that $a \in L(X \cap Y)$, i.e. $X \cap Y \longrightarrow a \in \mathcal{F}_L$. Then by (F4) $X \cap Y \longrightarrow Z \in \mathcal{F}_L$.

$3 \longrightarrow 1$. Let 3) hold and $a \in C_L(X) \cap C_L(Y)$, $X, Y \subseteq U$. Then $U - X \longrightarrow a \in \mathcal{F}_L$ and $U - Y \longrightarrow a \in \mathcal{F}_L$ and both FD's are nontrivial. Hence, $U - (X \cup Y) \longrightarrow a \in \mathcal{F}_L$ and $a \in L(U - (X \cup Y))$. Since $a \in (X \cup Y)$, we have $a \in C_L(X \cup Y)$. Therefore, C_L satisfies $\underline{C}$.

$1 \longleftrightarrow 4$. Let $a \notin X$. Then $X \longrightarrow a \in \mathcal{F}_L$ iff $a \in C_L(U - X)$, and $U - \{a, b\} \longrightarrow a \in \mathcal{F}_L$ iff $a \in C_L(\{a, b\})$. Hence, 4) is equivalent to: $a \in C_L(Z)$ iff $a \in C_L(\{a, b\})$ for all $b \in Z$. According to [AM],[Mo] the last property holds iff C_L satisfies $\underline{C}$.

$1 \longleftrightarrow 5$. Since C_L satisfies $\underline{H}$, it satisfies $\underline{C}$ iff all the functions $f_i^{C_L} i = 1, \ldots, n$ can be represented as $\overline{f}^*$, where $f \in P_6$, see [VR],[Li1]. Since $f_i^L = \overline{f_i^{C_L}}^*$, we have that C_L satisfies $\underline{C}$ iff $f_i^L \in P_6$ for all i. The proposition is proved.

□

Property of submission. This property was introduced in [Li1] as a dual form of $\underline{C}$. We say that a choice function satisfies the submission property ($\underline{S}$ for short) if

$$\forall X, Y \subseteq U : C(X \cap Y) \subseteq C(X) \cup C(Y).$$

Recall that a closure is called *topological* if $L(X \cup Y) = L(X) \cup L(Y)$ for all $X, Y \subseteq U$.

Let S_6 stand for the class of Boolean functions consisting of disjunctions and constants, cf. [Po].

Proposition 2 *Let L be a closure operation. Then the following are equivalent:*

1) C_L *satisfies* $\underline{S}$;

2) L *is a topological closure;*

3) $X \longrightarrow a \in \mathcal{F}_L$ *iff* $b \longrightarrow a \in \mathcal{F}_L$ *for some* $b \in X$;

4) *For all* $i = 1, \ldots, n : f_i^L \in S_6$.

Proof. $1 \longrightarrow 2$. Let C_L satisfy $\underline{S}$. Then for all $X, Y \subseteq U : L(X \cup Y) = X \cup Y \cup C_L(U - X \cup Y) = X \cup Y \cup C_L((U - X) \cap (U - Y)) \subseteq (X \cup C_L(U - X)) \cup (Y \cup C_L(U - Y)) = L(X) \cup L(Y)$. Since (C2) holds, $L(X) \cup L(Y) \subseteq L(X \cup Y)$, i.e. L is topological.

$2 \longrightarrow 1$. Let L be topological. Then for all $X, Y \subseteq U : C_L(X \cap Y) = L(U - X \cap Y) \cap X \cap Y = L((U - X) \cup (U - Y)) \cap X \cap Y \subseteq (L(U - X) \cup L(U - Y)) \cap X \cap Y \subseteq (L(U - X) \cap X) \cup (L(U - Y) \cap Y)) = C_L(X) \cup C_L(Y)$, i.e. C_L satisfies $\underline{S}$.

$2 \longleftrightarrow 3$. It was proved in [DLM2].

$1 \longleftrightarrow 4$. According to [Li1], C_L satisfies $\underline{S}$ iff for all $i = 1, \ldots, n : (f_i^{C_L})^* \in S_6$, i.e. iff $f_i^L \in S_6$. The proposition is proved.

□

The topological closures are known to have simple matrix representations. Consider two binary relations P_L and T_L on U as follows:

$(a_i, a_j) \in P_L$ iff every closed subset X (w.r.t. L) either contains a_j or does not contain a_i.

$(a_i, a_j) \in T_L$ iff $a_j \in L(a_i)$.

For a closure L, P_L is a reflexive relation. Given a reflexive relation P suppose that $L(X)$ is the intersection of all $Y \supseteq X$ such that for all $(a_i, a_j) \in P$ either $a_i \notin Y$ or $a_j \in Y$. Then L thus constructed is a topological closure with $P_L = P$, see [DLM2].

For a topological closure L, T_L is a transitive binary relation. Conversely, given a transitive binary relation T, define $L(X) = X \cup \{a \in U | \exists b \in X : (b, a) \in T\}$. Then L is a topological closure with $T_L = T$. Moreover, T_L is t he minimal transitive binary relation containing P_L, see [DLM2]

It is known that the choice functions satisfying $\underline{H}$ and $\underline{S}$ can be represented by binary relation as follows [Li1]:

$$C^P(X) = \{a \in X | \exists b \in X : (b, a) \in P \Longrightarrow \exists c \notin X : (c, a) \in P\}.$$

Hence, P thus constructed can be considered as a representation of a topological closure with $C_L = C^P$.

Proposition 3 *$C_L = C^{T_L}$ holds for any topological closure L.*

Proof. Let $a \in X$. Since T_L is reflexive, $a \in C^{T_L}(X)$ iff for some $c \notin X$: $(c, a) \in T_L$, i.e. iff $a \in L(c)$. Since L is topological, the last is equivalent to $a \in L(U - X) \cap X$, i.e. $a \in C_L(X)$.

□

Property of multi-valued concordance. This property also has been introduced in [Li1] in order to be studied together with the property $\underline{O}$.

A subset of $U \times 2^U$ was called in [AM] a *hyper-relation.* We will call a hyper-relation *correct* [Li1] if for every $X \subseteq U$ there is a unique $Y \subseteq X$ such that for all $a \in X - Y$ the pairs (a, Y) belong to the hype r-relation.

Proposition 4 *Let L be a closure operation. Then the following are equivalent:*

1. *C^L satisfies the property of multivalued concordance, i.e. if $Z = C^L(X) = C^L(Y)$ then $Z = C^L(X \cup Y)$;*
2. *For all $X, Y \subseteq U : L(X) = L(Y)$ implies $L(X) = L(X \cap Y)$;*
3. *For all $Z \subseteq U : f_Z^L \in P_6$;*
4. *For all $X \subseteq U : C^L(X) = Y$, where $(a, Y) \in D$ for all $a \in X - Y$ and D is a correct hyper-relation.*

Proof. The equivalence of 1 and 2 is evident. The equivalences $1 \longleftrightarrow 3$ and $1 \longleftrightarrow 4$ follow from [Li1].

□

5 Structural representation of functional independencies

Let R be a relation over U. We say that a *functional independency* (FID for short) $X \longrightarrow Y$ holds for R if there are two elements of R with coinciding projections onto X and distinct projections onto Y (i.e. FD $X \longrightarrow Y$ does not hold), see [Ja]. A review of properties of FID's can be found in [Ja]. In this section we construct the representations of FID's via operations on a power set and semilattices.

Let R be a relation and $\mathcal{FI}_R$ the family of all FID's that hold for R. A family $\mathcal{FI}$ of FID's is called *full* if for some relation R one has $\mathcal{FI} = \mathcal{FI}_R$.

Given a full family $\mathcal{FI}$, define for $X \subseteq U$ $C_{\mathcal{FI}}(X) = \{a \in X | (U - X) \longrightarrow a \in \mathcal{FI}\}$. Conversely, given a choice function C, define a family of FID's $\mathcal{FI}_C$ as follows:

$$X \longrightarrow Y \in \mathcal{FI}_C \text{ iff } Y \subseteq C(U - X).$$

Let C be a choice function. Define $\mathcal{L}(C) = \{X \subseteq U | C(X) = X\}$. For a join-semilattice $\mathcal{L}$, $(\mathcal{L} \subseteq 2^U, \emptyset \in \mathcal{L}, X, Y \in L \Longrightarrow X \cup Y \in L)$ define $C_{\mathcal{L}}$ as follows:

$$C_{\mathcal{L}}(X) = \cup(Y | Y \subseteq X, Y \in \mathcal{L}).$$

Theorem 3 *a) The mappings $\mathcal{FI} \longrightarrow C_{\mathcal{FI}}$ and $C \longrightarrow \mathcal{FI}_C$ establish mutually inverse one-to-one correspondences between full families of FID's and choice functions satisfying $\underline{M}$ and $\underline{O}$.*

b) The mappings $C \longrightarrow \mathcal{L}(C)$ and $\mathcal{L} \longrightarrow C_{\mathcal{L}}$ establish mutually inverse one-to-one correspondences between choice functions satisfying $\underline{M}$ and $\underline{O}$ and join-semilattices.

Proof. a) Let $\mathcal{FI} = \mathcal{FI}_R$ be a full family of FID's. Then $a \in C_{\mathcal{FI}}(X)$ iff $a \notin L_R(U - X)$, i.e. $C_{\mathcal{FI}}(X) = U - L_R(U - X)$ and C satisfies $\underline{O}$ and $\underline{M}$ by theorem 1.

Let C satisfy O and M. Then $C = C^L$ for some closure L, and $X \longrightarrow Y \in \mathcal{FI}_C$ iff $Y \cap L(X) = \emptyset$, i.e. $(X \longrightarrow Y) \notin \mathcal{F}_L$. Hence $\mathcal{FI}_C$ is a full family. Moreover, $a \in C_{\mathcal{FI}_C}(X)$ iff $(U - X) \longrightarrow a \in \mathcal{FI}_C$ iff $a \in C(X)$. Part *a* is proved.

b) Let L be a closure. Then $\mathcal{L}(C^L) = \{X \subseteq U | C^L(X) = X\} = \{X \subseteq U | L(U - X) = U - X\} = \{X \subseteq U | U - X \in Z(L)\}$. Hence, part *b* follows from theorem 1 and the well-known correspondence bet ween (meet)-semilattices and closure operations, see [DK],[DLM1]. The theorem is proved.

□

The last question to be considered is as follows: when is a full family of FID's also a full family of FD's? In other words, when is a closure operation $L(X) = X \cup \{a \notin X | X \longrightarrow a \in \mathcal{FI}_R\}$?

Proposition 5 *Let R be a relation over U. Then the following are equivalent:*

1. *$L(X) = X \cup \{a \notin X | X \longrightarrow a \in \mathcal{FI}_R\}$ is a closure operation;*

2. There is $Z \subseteq U$ such that $L_R(X) = X \cup Z$ for all $X \subseteq U$.

Proof. Let $L(X) = X \cup \{a \notin X | X \longrightarrow a \in \mathcal{F}I_R\}$ be a closure. Then C^L satisfies $\underline{H}$ (see theorem 1) and since C^L satisfies $\underline{M}$ we have that for some $V \subseteq U : C^L(X) = X \cap V$ for all $X \subseteq U$, see [AM]. Therefore, for $Z = U - V$ one has $L_R(X) = X \cup Z$.

Conversely, if L_R is as in 2, then $L(X) = X \cup \{a \notin X | X \longrightarrow a \in \mathcal{F}I_R\} = X \cup \{a \notin X | X \longrightarrow a \notin \mathcal{F}_R\}$ is obviously a closure operation. The proposition is proved.

□

References

[Ai] M.A. Aizermann, New problems in the general choice theory (Review of a research trend), J. Social Choice and Welfare 2 (1985), 235-282.

[AM] M.A. Aizerman and A.V. Malishevski, General theory of best variants choice: Some aspects, IEEE Trans. Automat. Control 26 (1981), 1030-1041.

[Ar] W.W. Armstrong, Dependency structure of data-base relationship, Information Processing 74, North Holland, Amsterdam, (1974), 580-583.

[BDFS] C. Beeri, M. Dowd, R. Fagin and R. Statman, On the structure of Armstrong relations for functional dependencies, J. ACM 31 (1984), 30-46.

[De1] J. Demetrovics, Candidate keys and antichains, SIAM J. Alg. Disc. Meth. 1 (1980), 92.

[De2] J. Demetrovics, On the equivalence of candidate keys with Sperner systems, Acta Cybernetica 4 (1979), 247-252.

[DK] J. Demetrovics and G.O.H. Katona, Extremal combinatorial problems of database models, MFDBS 87, Springer LNCS 305 (1988), 99-127.

[DLM1] J. Demetrovics, L.O. Libkin and I.B. Muchnik, Functional dependencies and the semilattice of closed classes, MFDBS 89, Springer LNCS 364 (1989), 136-147.

[DLM2] J. Demetrovics, L.O. Libkin and I.B. Muchnik, Closure operations and database models (in Russian) to appear in Kibernetika, Kiew.

[Ja] J. M. Janas, Covers for functional independencies, MFDBS 89, Springer LNCS 364(1989), 254-268.

[Li1] L.O. Libkin, Algebraic methods for construction and analysis of the choice function classes (in Russian), Individual Choice and Fuzzy Sets, Trudy VNIISI, Moscow 14(1987), 46-53.

[Li2] L.O. Libkin, Recognition of choice functions (in Russian), Automatika i Telemehanika, 1988, No. 10, p. 128-132. English Translation in: J. Automation and Remote Control.

[Mo] H. Moulin, Choice functions over a finite set: A summary. IMA Preprint series, University of Minnesota, Minneapolis, 1984.

[Po] E. Post, Two-valued Iterative Systems, 1941.

[VR] T.M. Vinogradskaja and A.A. Rubchinski, Logical forms of choice functions (in Russian), Dokl. Akad. Nauk. SSR 254(1980), 1362-1366. English Translation in: Soviet Math. Dokl.

Received February 1, 1990.