Order **4** (1987), 127–142.

The Poset of Closures as a Model of Changing Databases

G. BUROSCH
Goethestr. 21, Rostock, G.D.R.

J. DEMETROVICS*
Folyondàr köz 10, H-1037, Hungary

and

G. O. H. KATONA*
Mathematical Institute, P.F. 127, H-1364 Budapest, Hungary

Communicated by R. P. Dilworth

(Received: 2 May, 1985; accepted: 30 April 1987)

Abstract. The combinatorial properties of the poset of closures are studied, especially the degrees in the Hasse diagram.

AMS subject classifications (1980). 06A10, 68-XX.

Key words. Poset, closures, Hasse diagram.

1. Introduction

Armstrong [1] and Codd [2] introduced closures as models of databases. Let X be the (finite) set of types of data. That is, the elements of X are words like 'name', 'date of birth', 'age', etc. Some of the data determine some other data uniquely. For instance, the date of birth determines the age (in a given moment). Let $A \subset X$, $a \in X$, $a \notin A$. We say that A *determines* a and write $A \to a$ iff the set of data in A determines the data in a, more precisely, there are no two individuals having the same data in A and different in a. The function $\mathscr{L}: 2^X \to 2^X$ is defined by

$$\mathscr{L}(A) = \{a : A \to a\}.$$

This function obviously possesses the following properties:

$$A \subset \mathscr{L}(A), \tag{1}$$

$$A \subset B \text{ implies } \mathscr{L}(A) \subset \mathscr{L}(B), \tag{2}$$

* The research of these authors was supported by the Hungarian National Foundation for Scientific Research, grant numbers 1812 and 700021.

$$\mathscr{L}(\mathscr{L}(A)) = \mathscr{L}(A). \tag{3}$$

Such a function is known as a *closure*. Therefore a closure is a possible model of a database bringing some of its important features into relief.

$$\mathscr{L}(\phi) = \phi \tag{4}$$

is a rather natural assumption for closures formed from databases. In the present paper we will use the name closure for the closures satisfying (4).

A database is constantly changing during its life. It also changes the corresponding closure. A typical change is to delete the data of some individuals. If $A \to a$ is true then it remains true after the change. This implies

$$\mathscr{L}_1(A) \subset \mathscr{L}_2(A) \quad (\text{for all } A \subset X) \tag{5}$$

if $\mathscr{L}_1$ and $\mathscr{L}_2$ denote the closures before and after the change. We say that $\mathscr{L}_1$ is *richer* than or equal to $\mathscr{L}_2$ ($\mathscr{L}_1 \geqslant \mathscr{L}_2$) if they satisfy (5). It is easy to see that this property is transitive, consequently the closures of a fixed n-element set X form a *partially ordered set (poset)* for the ordering given in (5).

The aim of the present paper is to study this poset P. We emphasize the term poset, because the properties investigated here arise typically at posets. The investigation of these properties is motivated below. We do not know, however, any database motivation for the usual questions interesting from a lattice theoretical point of view (distributive elements, semimodularity, etc.).

Section 2 contains some preliminary results. Most of them (the Lemmas and the Propositions) belong to the folklore. We collected them here to make the paper selfcontained. We found no proper reference. As the referee kindly informed us, a lecture note was prepared and circulated by R. P. Dilworth in 1942, which was the most complete treatment of general closure operators. Unfortunately, it was never published. These results give a more clear picture of P by terms of closed sets ($\mathscr{L}(A) = A$). It also follows that P has a rank-function.

It should be mentioned that P is defined and studied in a paper of Dilworth and McLaughlin [4]. In that paper, however, $\mathscr{L}_1 \leqslant \mathscr{L}_2$ iff (5) holds. We chose, after some hesitation, the present 'unnatural' opposite way because it ensured that richer closures had 'more' closed sets (Proposition 1).

The main results of the paper are in Section 4. Consider the closures of a fixed rank in P and try to find the minimum (maximum) number of neighbouring closures from above (from below). That is, lower and upper estimates of the degrees (from above and from below) in a fixed level of the Hasse diagram of P are sought.

The database motivation for this question is the following. A temporary state of the database corresponds to an element of P. A change in the database corresponds to a move along an edge of this Hasse diagram. The life of the database therefore corresponds to a random walk along the Hasse diagram of P. As a first model, we might suppose that all the edges at a given element are chosen with equal probabilities. To obtain any (probabilistic) statement in this model we obviously need information about the degrees.

The other motivation for the questions of Section 4 comes from the theory of posets. What is the maximum number of incomparable elements in P? (Maximum size of an antichain.) A well-known method giving good estimates for this number uses the number of chains (total number of chains and the number of chains passing through a given element). To obtain estimates on these numbers, we again need information about the degrees.

The results of Section 3 are not along the main line of the paper. $K \subset X$ is said to be a *minimal key* of $\mathscr{L}$ iff $\mathscr{L}(K) = X$ but $\mathscr{L}(A) \neq X$ for any proper subset of K. The minimal keys play an important role in the theory of databases. The data of a minimal key determine all other data. It is natural to ask what is the relationship between the closure and its system of minimal keys. We give a necessary and sufficient condition, under which the system of minimal keys determine the closure $\mathscr{L}$ uniquely.

2. The Poset of Closures

The set A is called closed in the closure $\mathscr{L}$ iff $\mathscr{L}(A) = A$. The family of closed sets is denoted by $\mathscr{F} = \mathscr{F}(\mathscr{L})$. We use the following well-known facts (Lemmas 1–3) about $\mathscr{F}$ without proof.

LEMMA 1.

$$A, B \in \mathscr{F} \quad \textit{implies } A \cap B \in \mathscr{F};\ \phi \in \mathscr{F};\ X \in \mathscr{F}.$$

If a family satisfies (6), we say that $\mathscr{F}$ is *closed under intersection*.

LEMMA 2. *$\mathscr{L}(A)$ is equal to the smallest closed set containing A.*

LEMMA 3. *For any family $\mathscr{F} \subset 2^X$ satisfying* (6) *there is a unique closure $\mathscr{L}$ such that $\mathscr{F} = \mathscr{F}(\mathscr{L})$.*

The main statement of this section is the following

PROPOSITION 1. $\mathscr{L}_1 \leqslant \mathscr{L}_2$ *iff* $\mathscr{F}(\mathscr{L}_1) \subset \mathscr{F}(\mathscr{L}_2)$.

Proof. Suppose that $\mathscr{L}_1 \leqslant \mathscr{L}_2$ and A is closed in $\mathscr{L}_1$. By definition, $\mathscr{L}_1(A) = A$ holds. Formula (5) implies $A \supset \mathscr{L}_2(A)$ and, by (1), we obtain $A = \mathscr{L}_2(A)$. That is, $A \in \mathscr{F}(\mathscr{L}_2)$ and the first part is proved.

Conversely, suppose now $\mathscr{F}(\mathscr{L}_1) \subset \mathscr{F}(\mathscr{L}_2)$. By Lemma 2, $\mathscr{L}_1(A)$ and $\mathscr{L}_2(A)$ are the smallest closed sets with respect to $\mathscr{L}_1$ and $\mathscr{L}_2$, containing A, resp. $\mathscr{F}(\mathscr{L}_1) \subset \mathscr{F}(\mathscr{L}_2)$, implies $\mathscr{L}_2(A) \subset \mathscr{L}_1(A)$ and this the definition of $\mathscr{L}_1 \leqslant \mathscr{L}_2$. The proof is complete. □

We say that $\mathscr{L}_2$ *covers* $\mathscr{L}_1$ and write $\mathscr{L}_2 \succ \mathscr{L}_1$ iff $\mathscr{L}_2 > \mathscr{L}_1$ and there is no $\mathscr{L}_3$

satisfying $\mathscr{L}_2 > \mathscr{L}_3 > \mathscr{L}_1$. The function r associating nonnegative integers with the elements of a given poset satisfying the following two conditions is called a rank function:

$$r \text{ is zero for some element.} \tag{7}$$

$$\text{if } \mathscr{L}_2 \text{ covers } \mathscr{L}_1 \text{ then } r(\mathscr{L}_2) = r(\mathscr{L}_1) + 1. \tag{8}$$

PROPOSITION 2.

$$\mathscr{L}_1 \prec \mathscr{L}_2 \quad \textit{iff}\, \mathscr{F}(\mathscr{L}_1) \subset \mathscr{F}(\mathscr{L}_2) \quad \textit{and} \quad |\mathscr{F}(\mathscr{L}_2) - \mathscr{F}(\mathscr{L}_1)| = 1.$$

Proof. If $\mathscr{L}_1 \prec \mathscr{L}_2$ then $\mathscr{F}(\mathscr{L}_1) \subset \mathscr{F}(\mathscr{L}_2)$ follows by Proposition 1. We have to prove $|\mathscr{F}(\mathscr{L}_2) - \mathscr{F}(\mathscr{L}_1)| = 1$, only.

Suppose indirectly that $|\mathscr{F}(\mathscr{L}_2) - \mathscr{F}(\mathscr{L}_1)| \geqslant 2$. Choose any member $A \in \mathscr{F}(\mathscr{L}_2) - \mathscr{F}(\mathscr{L}_1)$. Consider a minimal $C = A \cap B \notin \mathscr{F}(\mathscr{L}_1)$, where $B \in \mathscr{F}(\mathscr{L}_1)$. Then $\mathscr{F}(\mathscr{L}_1) \cup \{C\}$ is closed under intersection.

The rest of the proof is a trivial consequence of Proposition 1 and Lemma 3. □

PROPOSITION 3. $r(\mathscr{L}) = |\mathscr{F}(\mathscr{L})| - 2$ *is a rank-function on* P.

Proof. $\mathscr{L}(X) = X$ and $\mathscr{L}(\phi) = \phi$ follow by (1) and (4), resp. This implies $|\mathscr{F}(\mathscr{L})| \geqslant 2$ for any element of P. Condition (7) is fulfilled. Proposition 2 implies (8). The proof is complete. □

The rank $r(P)$ of P is the maximum value of the rank function. It is easy to see that it is

$$r(P) = 2^n - 2$$

where $|X| = n$.

Let $\mathscr{F}$ be a family closed under intersection. Then let $\mathscr{M}(\mathscr{F})$ denote the family of such members $A \in \mathscr{F}$ which are not intersections of two other members of $\mathscr{F}$, different from A. $\mathscr{M}(\mathscr{F}(\mathscr{L}))$ is shortly denoted by $\mathscr{M}(\mathscr{L})$.

LEMMA 4. *If* $\mathscr{F}$ *is closed under intersection then any member of* $\mathscr{F}$ *is an intersection of some members of* $\mathscr{M}(\mathscr{F}) - \{X\}$.

Proof. We use induction on the size of the member $A \in \mathscr{F}$. X is the void intersection. If $|A|$ is maximum in $\mathscr{F} - \{X\}$, then $A \in \mathscr{M}(\mathscr{F})$ and A is a one-term intersection. Suppose that $|A|$ is smaller than the maximum and the statement is proved for larger sizes. If $A \in \mathscr{M}(\mathscr{F})$ were are done. If $A \notin \mathscr{M}(\mathscr{F})$, A is an intersection $B_1 \cap B_2$ of members different from A. Their size is larger than that of A, consequently we may use the inductional hypothesis for them and this gives an intersection of the members of $\mathscr{M}(\mathscr{F})$. The proof is complete. □

PROPOSITION 4. *Let* $\mathscr{F}$ *be a family closed under intersection and* $A \in \mathscr{F}$. $\mathscr{F} - \{A\}$ *is closed under intersection iff* $A \in \mathscr{M}(\mathscr{F}) - \{\phi, X\}$.

Proof. Suppose first that $A \in \mathscr{M}(\mathscr{F}) - \{\phi, X\}$. $B, C \in \mathscr{F} - \{A\}$ implies $B \cap C \in$

$\mathscr{F}$, but $B \cap C$ can be equal to A only when one of them is equal to A. This contradiction shows that $\mathscr{F} - \{A\}$ is closed under intersection.

The other implication will be proved in an indirect way. Suppose that $A \in \mathscr{F} - \mathscr{M}(\mathscr{F})$. Then there are B and C satisfying $B \neq A \neq C$, $A = B \cap C$ and $B, C \in \mathscr{F}$. Hence, we obtain $B, C \in \mathscr{F} - \{A\}$ and $B \cap C \notin \mathscr{F} - \{A\}$. $\mathscr{F} - \{A\}$ is not closed under intersection. If $A = \phi$ or X, this holds by (6). This contradiction completes the proof. □

Lemma 4 and Proposition 4 can be summarized in the following way. If $\mathscr{F}$ is a family closed under intersection, then exactly the members of $\mathscr{M}(\mathscr{F}) - \{\phi, X\}$ can be omitted from $\mathscr{F}$ without violating this property. On the other hand, $\mathscr{F}$ and $\mathscr{M}(\mathscr{F}) - \{\phi, X\}$ determine each other uniquely. This is almost true with the family of sets which can be added to $\mathscr{F}$ preserving the property that it is closed under intersection. We will explain it more in detail.

Let $H \subset X$, $H \notin \mathscr{F}$ and suppose that both $\mathscr{F}$ and $\mathscr{F} \cup \{H\}$ are closed under intersection. Consider the sets A satisfying $A \in \mathscr{F}$, $H \subset A$. The intersection of all these sets is in $\mathscr{F}$, therefore it is different from H. Denote it by $L(H)$. $H \subsetneq L(H)$ is obvious. Let $\mathscr{H}(\mathscr{F})$ denote the set of all pairs $(H, L(H))$ where $H \subset X$, $H \notin \mathscr{F}$ but $\mathscr{F} \cup \{H\}$ is closed under intersection. We will prove that $\mathscr{H}(\mathscr{F})$ uniquely determines $\mathscr{F}$ and then we characterize the set of $\mathscr{H}(\mathscr{F})$'s for all $\mathscr{F}$ closed under intersection.

THEOREM 1.

$$\mathscr{F} = \{A : A \subset X, H \subset A \Rightarrow L(H) \subset A \textit{ for all } (H, L(H)) \in \mathscr{H}(\mathscr{F})\}.$$

Proof. By definition, any member A of $\mathscr{F}$ has to satisfy the condition

$$H \subset A \Rightarrow L(H) \subset A \quad \text{for all } (H, L(H)) \in \mathscr{H}(\mathscr{F}). \tag{9}$$

We have to prove only that all sets satisfying (9) belong to $\mathscr{F}$. We prove this by induction on $k = k_A = |\{H : H \subset A, (H, L(H)) \in \mathscr{H}(\mathscr{F})\}|$.

Consider first the sets A with $k_A = 0$, that is, the sets A containing no $H((H, L(H)) \in \mathscr{H}(\mathscr{F}))$ as a subset. Choose the one not being in $\mathscr{F}$ and minimizing $|A|$. Then the intersection of A with any member of $\mathscr{F}$ is either A or a member of $\mathscr{F}$, therefore $\mathscr{F} \cup \{A\}$ is closed under intersection. $A = H$ must hold for some $(H, L(H)) \in \mathscr{H}(\mathscr{F})$, contradicting $k_A = 0$. This contradiction proves $A \in \mathscr{F}$ if $k_A = 0$.

Suppose now that $k > 1$ and that the statement is proved for all sets A satisfying (9) with $k_A < k$. Further suppose that $k_A = k$ and A satisfies (9),

$$A \supset H_1, \ldots, H_k, \quad A \not\supset H_{k+1}, \ldots \tag{10}$$

and

$$A \notin \mathscr{F}. \tag{11}$$

If there are more such sets A, take the one minimizing $|A|$.

We will verify that $\mathscr{F} \cup \{A\}$ is closed under intersection. Let $B \in \mathscr{F}$. We have to prove that $M = A \cap B$ is either A or a member of $\mathscr{F}$. If $M = A$ then we are done, $M \subsetneq A$ can be supposed. We distinguish two cases:

(i) $B \supset H_1, \ldots, H_k$. This condition combined with $B \in \mathscr{F}$ imply $B \supset L(H_1), \ldots, L(H_k)$. $A \supset L(H_1), \ldots, L(H_k)$ follows by (9) and (10). Hence, we have $M \supset L(H_1), \ldots, L(H_k)$. $M \not\supset H_{k+1}, \ldots$ is obvious. Hence, we see that M satisfies (9), (10) and $|M| < (A)$, therefore $M \in \mathscr{F}$ must hold.

(ii) $B \not\supset H_j$ for some $1 \leqslant j \leqslant k$. As before, it can be seen that $M = A \cap B$ satisfies (9), and $M \not\supset H_j$, $H_{k+1}, \ldots$. Hence, $k_M < k$ and by the inductional hypothesis we have $M \in \mathscr{F}$.

$\mathscr{F} \cup \{A\}$ is really closed under intersection. However, A cannot be equal to any H because of (10), (9) and $H_i \subsetneq L(H_i)$. This contradiction proves $A \in \mathscr{F}$ and the theorem. (□

THEOREM 2. *The set* $\{(A_i, B_i)\}_{i=1}^m$ *is equal to* $\mathscr{H}(\mathscr{F})$ *for some* $\mathscr{F}$ *closed under intersection on* X *if and only if the following conditions are satisfied*

$$\phi \neq A_i \subsetneq B_i \subset X, \tag{12}$$

$$A_i \subset A_j \textit{ implies either } B_i \subset A_j \textit{ or } B_i \supset A_j, \tag{13}$$

$$A_i \subset B_j \textit{ implies } B_i \subset B_j, \tag{14}$$

for any i *and* $C \subset X$ *satisfying* $A_i \subsetneq C \subsetneq B_i$ *there is a* j *such that either* $C = A_j$ *or* $A_j \subset C$, $B_j \not\subset C$, $B_j \not\supset C$ *all hold.* (15)

Proof. First we prove that the conditions are necessary. Suppose that $\mathscr{H}(\mathscr{F}) = \{(A_i, B_i)\}_{i=1}^m$ for some $\mathscr{F}$ closed under intersection. First prove (12). $A_i \in \mathscr{F}$ and $\phi \in \mathscr{F}$ prove $\phi \neq A_i$. As $B_i = L(A_i)$, it is the intersection of all members of $\mathscr{F}$ containing A_i. Therefore, $A_i \subset B_i$ holds, but we cannot have equality since $B_i = L(A_i) \in \mathscr{F}$ ($\mathscr{F}$ is closed under intersection) and $A_i \notin \mathscr{F}$.

Suppose $A_i \subset A_j$. As $\mathscr{F} \cup \{A_j\}$ is closed under intersection, $B_i = L(A_i) \in \mathscr{F}$ implies that $A_j \cap B_i$ is either A_j or is an element of $\mathscr{F}$. In the second case, $A_i \subset A_j$ and $A_i \subset B_i$ imply $A_i \subset A_j \cap B_i \in \mathscr{F}$. By the definition of $L(A_i)$ we obtain $L(A_i) \subset A_j \cap B_i$, that is, the desired relation $B_i \subset A_j$. Condition (13) is proved.

Condition (14) is a consequence of $B_j \in \mathscr{F}$ and of the definition of $(B_i =) L(A_i)$.

To prove (15) suppose $A_i \subsetneq C \subsetneq B_i$. (It can happen only when $|B_i - A_i| \geqslant 2$.) If $C = A_j$ for some j we are done. Otherwise $\mathscr{F} \cup \{C\}$ is not closed under intersection. There must exist a $D \in \mathscr{F}$ such that $D \cap C$ is neither C nor an element of $\mathscr{F}$. By Theorem 1 there is an A_j such that $A_j \subset D \cap C$ but $B_j = L(A_j) \not\subset D \cap C$. $A_j \subset D \cap C$ implies $A_j \subset D$ and $A_j \subset C$. Hence, we have $B_j = L(A_j) \subset D$, by $D \in \mathscr{F}$. $B_j \not\subset D \cap C$ yields $B_j \not\subset C$. The last relation we have to prove is $C \not\subset B_j$. This is a consequence of $D \cap C \neq C$ and $B_j \subset D$. This proves that our conditions are necessary.

Conversely, we have to prove that if (12)–(15) are fulfilled for the set $\{(A_i, B_i)\}_{i=1}^m$, then there is a family $\mathscr{F}$ closed under intersection and $\mathscr{H}(\mathscr{F}) = \{(A_i, B_i)\}_{i=1}^m$. We determine $\mathscr{F}$ by

$$\mathscr{F} = \{E : E \subset X, A_i \subset E \Rightarrow B_i \subset E \text{ for all } i (1 \leqslant i \leqslant m)\} \tag{16}$$

(see Theorem 1).

It is obvious that $\mathscr{F}$ is closed under intersection and $A_i \notin \mathscr{F}$ $(1 \leqslant i \leqslant m)$. Prove first that $\mathscr{F} \cup \{A_i\}$ is closed under intersection. Let $D \in \mathscr{F}$ and suppose that $A_j \subset D \cap A_i$. We have to prove that $D \cap A_i$ is either A_i or it contains B_j. $A_j \subset D \in \mathscr{F}$ implies $B_j \subset D$ by (16). $A_j \subset A_i$ implies either $B_j \subset A_i$ or $B_j \supset A_i$ by (13). In the first case we obtain $B_j \subset D \cap A_i$, while in the second case $A_i \subset B_j \subset D$ results in $D \cap A_i = A_i$.

We also have to prove that B_i is the intersection of the members of $\mathscr{F}$ containing A_i. Indeed, if $A_i \subset D \in \mathscr{F}$ then by (16) $B_i \subset D$ follows. On the other hand, we have to see that $B_i \in \mathscr{F}$. This is a consequence of (16) and (14).

It remained to prove only that no other sets can be added to $\mathscr{F}$, only the A_i. Suppose that $C \notin \mathscr{F}$, $\mathscr{F} \cup \{C\}$ is closed under intersection and $C \neq A_i$ $(1 \leqslant i \leqslant m)$. By (16) there is an i $(1 \leqslant i \leqslant m)$ such that $A_i \subsetneq C$ but $B_i \not\subset C$. The intersection $C \cap B_i$ is either C or is a member of $\mathscr{F}$. The latter case contradicts (16): $A_i \subset C \cap B_i \not\subset B_i$. This shows $C \subsetneq B_i$. We obtained $A_i \subsetneq C \subsetneq B_i$. By (15) there is a j such that either $C = A_j$ or $A_j \subset C$, $B_j \subset C$, $B_j \not\supset C$ all hold. The first case contradicts the assumption that C is different from the sets A_i. In the second case let us investigate $B_j \cap C$. It is neither equal to C (since $B_j \supset C$) nor is in $\mathscr{F}$ ($A_j \subset B_j \cap C$ but $B_j \subset B_j \cap C$). Therefore, $\mathscr{F} \cup \{C\}$ is not closed under intersection. This contradiction proves $\mathscr{H}(\mathscr{F}) = \{(A_i, B_i)\}_{i=1}^m$ and the theorem. □

COROLLARY 1. *If* $|\mathscr{H}(\mathscr{F})| = 1$ *then* $|L(H) - H| = 1$, *otherwise* $|L(H) - H| < |\mathscr{H}|$ *holds for all* $H \in \mathscr{H}$.

Proof. If $\mathscr{H}(\mathscr{F}) \times \{(H, L(H))\}$, then condition (15) cannot be fulfilled for a C satisfying $H \subsetneq C \subsetneq L(H)$. Therefore, $L(H) - H$ must be a one-element set.

Suppose $|\mathscr{H}(\mathscr{F})| \geqslant 2$ and prove that the set A_j in condition (15) cannot be a subset of A_i. This is obviously true if A_j satisfies $A_i \subsetneq C = A_j$. In the other case $A_j \subset A_i$ implies either $B_j \subset A_i$ or $B_j \supset A_i$. In the first subcase $B_j \subset A_i \subset C$ contradicts $B_j \not\subset C$, while in the second subcase $B_j \supset A_i$ implies $B_j \supset B_i$ by (14) and the contradiction is between $B_j \supset B_i \supset C$ and $B_j \not\supset C$.

Fix an element $(H, L(H)) \in \mathscr{H}(\mathscr{F})$. Take the sets $H(a) = H \cup \{a\}$ for all $a \in L(H) - H$. Apply (15) for them. There must exist an $H(a)$, $(H(a), L(H(a))) \in \mathscr{H}(\mathscr{F})$, playing the role of A_j. By the above section $H(a) \not\subset H$, therefore $a \in H(a)$. Hence, the sets $H(a)$ are all different. This implies the desired inequality $|L(H) - H| < |\mathscr{H}|$. □

In the present section we described the poset qualitatively. In the last section we come back to the quantitative part of this problem. Before this, we investigate a question of somewhat another nature.

3. When do the Keys Determine the Closure?

Let $\mathscr{K}(\mathscr{L})$ denote the family of minimum keys in the closure $\mathscr{L}$. By the minimality, it follows that $A, B \in \mathscr{K}(\mathscr{L})$, $A \neq B$ imply $A \not\subset B$. The families having this property are called *Sperner families*. A Sperner family $\mathscr{S}$ is *saturated* if $\{A\} \cup \mathscr{S}$ is not a Sperner family for any $A \notin \mathscr{S}$.

In this section we ask the converse question. Given a Sperner family $\mathscr{S}$, does there exist a closure satisfying $\mathscr{S} = \mathscr{K}(\mathscr{L})$. If yes, is it unique? It seems to be evident that there is always such a closure. We will prove it later. If $\mathscr{S}$ contains only a few members then $\mathscr{S}$ is obviously nonunique. One can, however, think that any saturated $\mathscr{S}$ determines the closure uniquely. This is true (Corollary 2) however, not only the saturated ones have this property. We show later a counterexample. The situation is more complex.

Some necessary definitions. If $\mathscr{A}$ is an arbitrary family of subsets of X, then $\underline{\mathscr{A}}$ and $\overline{\mathscr{A}}$ are used in the following ways:

$$\underline{\mathscr{A}} = \{A : \exists B \text{ such that } A \subset B \in \mathscr{A}\},$$
$$\overline{\mathscr{A}} = \{A : \exists B \text{ such that } A \supset B \in \mathscr{A}\}.$$

Moreover, if $\mathscr{S}$ is a Sperner family then let $\mathscr{S}^{-1}$ denote the family of maximal subsets not in $\overline{\mathscr{S}}$.

THEOREM 3. *Let $\mathscr{S}$ be a nonempty Sperner family. A closure $\mathscr{L}$ satisfies $\mathscr{S} = \mathscr{K}(\mathscr{L})$ iff*

$$\{X\} \cup \mathscr{S}^{-1} \subset \mathscr{F}(\mathscr{L}) \subset \{X\} \cup \underline{\mathscr{S}^{-1}} \tag{17}$$

Proof. Suppose that $\mathscr{S} = \mathscr{K}(\mathscr{L})$ and prove (17). X is always closed, we need not consider it. Let us first verify that $\mathscr{S}^{-1} \subset \mathscr{F}(\mathscr{L})$. Take an arbitrary member $A \in \mathscr{S}^{-1}$. Property (1) implies $A \subset \mathscr{L}(A)$. Suppose that $A \neq \mathscr{L}(A)$. The definition of $\mathscr{S}^{-1}$ yields $\mathscr{L}(A) \notin \mathscr{S}^{-1}$, therefore $\mathscr{L}(A) \in \overline{\mathscr{S}}$ and there is a $K \in \mathscr{S}$ such that $K \subset \mathscr{L}(A)$. (3), (2) and $\mathscr{S} = \mathscr{K}(\mathscr{L})$ imply $\mathscr{L}(A) = \mathscr{L}(\mathscr{L}(A)) \supset \mathscr{L}(K) = X$ contradicting $A \notin \mathscr{S} = \mathscr{K}(\mathscr{L})$. This proves the left-hand side of (17). The right-hand side is trivial: If $A \in \mathscr{F}(\mathscr{L}) - \underline{\mathscr{S}}^{-1}$ then there is a $K \in \mathscr{K}(\mathscr{L})$ such that $K \subset A$. (2) and $A \in \mathscr{F}(\mathscr{L})$ imply $A = \mathscr{L}(A) \supset \mathscr{L}(K) = X$, consequently $A = X$.

Conversely, suppose that (17) holds and prove $\mathscr{S} = \mathscr{K}(\mathscr{L})$. Let $A \in \mathscr{S}$ and show $A \in \mathscr{K}(\mathscr{L})$. $\mathscr{L}(A)$ is a closed set by (3). Condition (17) implies that either $\mathscr{L}(A) = X$ or $\mathscr{L}(A) \in \mathscr{S}^{-1}$ holds. The latter one contradicts $A \in \mathscr{S}$, we may conclude $\mathscr{L}(A) = X$. However, a proper subset B of A is in $\underline{\mathscr{S}}^{-1}$ and there is a C such that $B \subset C \subset A$ and $X \neq C \in \mathscr{S}^{-1}$. Property (2) implies $\mathscr{L}(B) \subset \mathscr{L}(C) \subset \mathscr{L}(A)$. C is a closed set by (17), therefore $\mathscr{L}(B) \subset \mathscr{L}(C) = C \neq X$ implies that A is a minimal key: $A \in \mathscr{K}(\mathscr{L})$.

Suppose now that $K \in \mathscr{K}(\mathscr{L})$. $\mathscr{L}(K) = X$ follows. By Lemma 1, K cannot be a subset of a closed set other than X. (17) implies that it cannot be a subset of a member of $\mathscr{S}^{-1}$. Hence, $K \in \overline{\mathscr{S}}$ follows. There is an $A \in \mathscr{S}$ with $A \subset K$.

$\mathscr{L}(A)$ is a closed set $\in \overline{\mathscr{S}}$, therefore the right-hand side of (17) leads to $\mathscr{L}(A)=X$. If $A \neq K$, this contradicts $K \in \mathscr{K}(\mathscr{L})$. We may conclude $K \in \mathscr{S}$. This proves $\mathscr{S}=\mathscr{K}(\mathscr{L})$ and the theorem. □

In an application of the theorem we have to take into account that $\mathscr{F}(\mathscr{L})$ is closed under intersection. So the smallest and largest $\mathscr{F}(\mathscr{L})$ are the ones formed from all the possible intersections of $\{X\} \cup \mathscr{S}^{-1}$ and $\{X\} \cup \underline{\mathscr{S}}^{-1}$, resp. The corresponding closures are denoted by $\mathscr{L}_{\min}(\mathscr{S})$ and $\mathscr{L}_{\max}(\mathscr{S})$, resp.

A consequence of the above theorem is the next one.

THEOREM 4. *Let $\mathscr{S}$ be a nonempty Sperner family. $\mathscr{S}=\mathscr{K}(\mathscr{L})$ determines $\mathscr{L}$ uniquely iff any member of $\underline{\mathscr{S}}^{-1}$ is an intersection of members of $\mathscr{S}^{-1}$.*

This theorem was independently discovered by Z. Füredi [5], too.

The next condition is uglier than Theorem 3 but sometimes it is algorithmically better. We do not bother the reader with its proof.

THEOREM 5. *Let $\mathscr{S}$ be a nonempty Sperner family. $\mathscr{S}=\mathscr{K}(\mathscr{L})$ determines $\mathscr{L}$ uniquely iff there are $A \subset X$ and $a \in X$ such that*

$a \notin A$,
$\exists B \in \mathscr{S}$ such that $B \subset X-\{a\}$,
$\nexists B \in \mathscr{S}$ such that $B \subset A$,
for any C satisfying $a \in C \in \mathscr{S}$ and $A \cup C \neq X$ there is a $D \in \mathscr{S}$ such that $D \subset A \cup B-\{a\}$.

COROLLARY 2. *If $\mathscr{S}$ is a saturated Sperner family then $\mathscr{S}=\mathscr{K}(\mathscr{L})$ uniquely determines $\mathscr{L}$.*

Proof. It can be deduced from both Theorems 4 and 5 but it has a simpler proof.

If $A \in \overline{\mathscr{S}}$ then the condition $\mathscr{S}=\mathscr{K}(\mathscr{S})$ ensures $\mathscr{L}(A)=X$. We will prove that for the sets $A \in \underline{\mathscr{S}}^{-1}$ the closure is also uniquely determined, namely $\mathscr{L}(A)=A$.

$\mathscr{S}$ is saturated, either $B \supset \mathscr{L}(A)$, $B \neq \mathscr{L}(A)$ or $B \subset \mathscr{L}(A)$ holds for some $B \in \mathscr{S}$. In the latter case (3) and (2) imply $\mathscr{L}(A)=\mathscr{L}(\mathscr{L}(A)) \supset \mathscr{L}(B)=X$ contradicting $A \in \underline{\mathscr{S}}^{-1}$. If $B \supsetneq \mathscr{L}(A) \supset A$ then (1), (2) and (3) imply

$$B=(B-\mathscr{L}(A)) \cup \mathscr{L}(A) \subset \mathscr{L}(B-\mathscr{L}(A)) \cup \mathscr{L}(A) \subset$$
$$\subset \mathscr{L}((B-\mathscr{L}(A)) \cup A)=\mathscr{L}(B-(\mathscr{L}(A)-A)).$$

Hence $X=\mathscr{L}(B)=\mathscr{L}(B-(\mathscr{L}(A)-A))$ follows. $B \in \mathscr{S}=\mathscr{K}(\mathscr{L})$ is a minimal key, therefore $\mathscr{L}(A)=A$ as we wanted to show. The proof is complete. □

Finally, we show an example of a nonsaturated Sperner family uniquely determining its $\mathscr{L}$. Put $|X|=n=5$ and $\mathscr{S}=\{\{1,2\},\{2,3\},\{3,4\}\},\{4,5\},\{1,5\}\}$. It is easy to see that all three-element sets are in $\overline{\mathscr{S}}$. So $\mathscr{S}^{-1}$ consists of the remaining two-element sets. Their intersections give us all the one-element and zero-element sets. Theorem 4 can be applied to show that $\mathscr{L}$ is unique. One could also use Theorem 5.

4. The Number of Neighbours of an Element in P

Let $\deg_a(\mathscr{L})$ and $\deg_b(\mathscr{L})$ denote the number of elements of P covering $\mathscr{L}$ and covered by $\mathscr{L}$, respectively. We define the following functions:

$$f_1(n, k) = \max\{\deg_a(\mathscr{L}) : r(\mathscr{L}) = k\},$$
$$f_2(n, k) = \min\{\deg_a(\mathscr{L}) : r(\mathscr{L}) = k\},$$
$$f_3(n, k) = \max\{\deg_b(\mathscr{L}) : r(\mathscr{L}) = k\},$$
$$f_4(n, k) = \max\{\deg_b(\mathscr{L}) : r(\mathscr{L}) = k\},$$
$$(1 \leqslant n, 0 \leqslant k \leqslant 2^n - 2).$$

THEOREM 6. $f_1(n, k) = 2^n - k - 2$.

Proof. If $r(\mathscr{L}) = k$, then $|\mathscr{F}(\mathscr{L})| = k + 2$ by Proposition 3. Proposition 2 implies that the closures $\mathscr{L}'$ covering $\mathscr{L}$ satisfy $|\mathscr{F}(\mathscr{L}')| = k + 3$, $\mathscr{F}(\mathscr{L}') \supset \mathscr{F}(\mathscr{L})$. The number of possible choices of the member $\mathscr{F}(\mathscr{L}') - \mathscr{F}(\mathscr{L})$ is at most $2^n - k - 2$. We construct now a $\mathscr{F}(\mathscr{L})$ allowing all these choices. Suppose that

$$k + 2 = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{r} + a, \quad \text{where } 0 \leqslant a < \binom{n}{r+1}.$$

Then let $\mathscr{F}(\mathscr{L})$ consist of the empty set, all one-element, two-element, ..., r-element subsets any a pieces of the $r + 1$-element subsets and X. It is easy to see that this family is closed under intersection. Moreover $\mathscr{F}(\mathscr{L}) \cup \{A\}$ $(A \in \mathscr{F}(\mathscr{L}))$ also has this property. We can choose A in $2^n - k - 2$ different ways and all these choices lead to closures covering $\mathscr{L}$. The proof is complete. □

We are not able to determine $f_2(n, k)$ in general, but we show that it can take small values for many k's. On the other hand, we give good upper estimates for the upper half of k's. Before, we state a useful lemma.

LEMMA 5. *Let* $r = 2^{\alpha_1} + \cdots + 2^{\alpha_l} > 1$, $l \geqslant 1$, $\alpha_1 > \cdots > \alpha_l \geqslant 0$ *be integers. One can find nonempty subsets* $A_1, \ldots, A_l$ *in the* $(\alpha_1 + 1)$*-element set* A *so that* $A_i \not\subset A_j$ $(i \neq j)$ *and*

$$|\{B : \exists i\, (1 \leqslant i \leqslant l), B \subset A_i\}| = r.$$

Proof. We use induction on l. If $l = 1$, the statement is obvious, $|A_1|$ should be α_1.

Suppose that the lemma is proved for $l - 1$, that is, there are subsets $B_2, \ldots, B_l$ in an $(\alpha_2 + 1)$-element set B so, that the total number of subsets is $2^{\alpha_2} + \cdots + 2^{\alpha_l}$. The $(\alpha_1 + 1)$-element set A is chosen satisfying $B \subset A$. Since $\alpha_1 > \alpha_2$, there is an element $a \in A - B$. Then let $A_1 = A - \{a\}$, $A_2 = B_2 \cup \{a\}, \ldots, A_l = B_l \cup \{a\}$. The number of subsets of A_1 is 2^{α_1}. We have to count only such subsets of $A_2, \ldots, A_l$ which are not in A_1, that is, the ones containing a. The number of such subsets of $B_2, \ldots, B_l$: $2^{\alpha_2} + \cdots + 2^{\alpha_l}$. The total number of subsets is $2^{\alpha_1} + \cdots + 2^{\alpha_l}$, as desired. The proof is complete. □

THEOREM 7.

$$f_2(n, k) = 0 \text{ iff } k = 2^n - 2, \tag{18}$$

$$f_2(n, k) = 1 \text{ iff } k = 2^n - 2^{n-a-1} - 2 \text{ for some } 0 < a < n, \tag{19}$$

$f_2(n, k) = 2$ iff k is equal to one of the following expressions but it is not of the form (19):

$$2^n - 2^{n-a-1} - 2^{n-b-1} - 2 \ (0 < a \leqslant b < n), \tag{20}$$

$$2^n - 2^{n-a-1} - 2^{n-b-1} + 2^{n-a-b+c-2} - 2$$
$$(0 \leqslant c < a \leqslant b, a + b - c \leqslant n - 2), \tag{21}$$

$$2^n - 2^{n-a-1} - 2^{n-b-1} + 2^{n-a-b+c-1} - 2$$
$$(0 \leqslant c < a \leqslant b, a + b - c \leqslant n - 1). \tag{22}$$

Proof. $f_2(n, 2^n - 2) = 0$ is obvious. If $k < 2^n - 2$, then let $\mathscr{L}$ be any closure of rank k and take a minimal set M not being in $\mathscr{F}(\mathscr{L})$. It is clear that $\mathscr{F}(\mathscr{L}) \cup \{M\}$ is closed under intersection, therefore it is $\mathscr{F}(\mathscr{L}')$ for some $\mathscr{L}'$ covering $\mathscr{L}$. This proves (18).

To determine the values k giving $f_2(n, k) = 1$ we have to find the closures $\mathscr{L}$ with $\mathscr{H} = \mathscr{H}(\mathscr{F}(\mathscr{L}))$ satisfying $|\mathscr{H}| = 1$. Corollary 1 states that $|L(H) - H| = 1$ holds in this case. Suppose $|H| = a$ $(0 < a < n)$. Using Theorem 1, the number of members of $\mathscr{F}$ is $2^n - 2^{n-a-1}$. Therefore, the rank is $k = 2^n - 2^{n-a-1} - 2$. H is obvious, however, that for such k's there is a $\mathscr{F}$ with $|\mathscr{H}(\mathscr{F})| = 1$.

To determine the $\mathscr{F}$'s with $|\mathscr{H}(\mathscr{F})| = 2$ we use Corollary 1 again. $\mathscr{H} = \{H_1, L(H_1)), (H_2, L(H_2))\}$ implies $|L(H_1) - H_1| = |L(H_2) - H_2| = 1$. In this case (15) automatically holds, therefore one should find the pairs satisfying (12), (13) and (14). It is easy to see that for the possible constructions, the number of members of $\mathscr{F}$ based on Theorem 1 is either (20) or (21) or (22). In the constructions $a = |H_1|$, $b = |H_2|$, $c = |H_1 \cap H_2|$. When $H_1 \not\subset H_2$, $H_2 \not\subset H_1$, $L(H_1) \subset H_2$ (20) can be obtained. If

$$L(H_1) - H_1 \neq L(H_2) - H_2, \ L(H_1) - H_1, \ L(H_2) - H_2 \not\subset H_1 \cup H_2, \ H_1 \not\subset H_2,$$
$$H_1 \not\supset H_2$$

then the rank is given by (21). Finally, if

$$L(H_1) - H_1 = L(H_2) - H_2 \not\subset H_1 \cup H_2, H_1 \not\subset H_2, H_1 \not\supset H_2$$

then we obtain (22). The proof is complete. □

THEOREM 8. *Suppose $k > 2^{n-1} + 2$. Then $f_2(n, k) \leqslant$ number of bits 1 in the binary expansion of $2^n - k - 2$. What is at most $n - 1$.*

Proof. Suppose that k satisfies the conditions of the theorem. Then

$$2^n - k - 2 = 2^{\alpha_1} + 2^{\alpha_2} + \cdots + 2^{\alpha_l} \quad (n-1 > \alpha_1 > \alpha_2 > \cdots > \alpha_l \geqslant 0).$$

Let $x \in X$ be a fixed element. According to Lemma 5, we can find sets $A_1, \ldots, A_l$ in the $(n-1)$-element set $X - \{x\}$ so that

$$|\{B : B \subset A_i \text{ for some } i(1 \leqslant i \leqslant l)\}| = 2^n - k - 2. \tag{23}$$

Obviously, none of A_i is equal to $X - \{x\}$. Choose $H_i = X - A_i - \{x\}$ $(1 \leqslant i \leqslant l)$ and $L(H_i) = H_i \cup \{x\}$. It is easy to see that the set $\{(H_i, L(H_i))\}$ satisfies the conditions of Theorem 2. Therefore, there is a $\mathscr{F}$ closed under intersection with $\mathscr{H}(\mathscr{F}) = \{(H_i, L(H_i))\}$. By Theorem 1

$$|\mathscr{F}| = |\{A : H_i \subset A \subset X \text{ implies } L(H_i) \subset A \ (1 \leqslant i \leqslant l)\}|.$$

This is equal to

$$|\{A : H_i \subset A \subset X \text{ implies } x \in A\}| = 2^n - |\{A : H_i \subset A \subset X - \{x\}$$

for some $i(1 \leqslant i \leqslant l)\}$. Using the complements with respect to $X - \{x\}$ and (23) we obtain

$$2^n - |\{\overline{A} : \overline{A} \subset \overline{H}_i \text{ for some } i(1 \leqslant i \leqslant l)\}|$$
$$= 2^n - |\{B : B \subset A_i \text{ for some } i(1 \leqslant i \leqslant l)\}| = 2^n - (2^n - k - 2) = k + 2.$$

This proves that the rank of $\mathscr{F}$ is k. As $|\mathscr{H}(\mathscr{F})| = l$ is the number of sets which can be added to $\mathscr{F}$, this proves $f_2(n, k) \leqslant l$ and the theorem. □

We know practically nothing about $f_3(n, k)$. Let

$$\mathscr{F} = \{\phi, \{1\}, \{2\}, \ldots, \{n\}, \{1, 2\}, \{1, 2, 3\}, \ldots, \{1, 2, \ldots, n\}\}.$$

It is obvious that $\mathscr{F}$ is closed under intersection and omitting any member $(\neq \phi, X)$ of it, this property is preserved. This proves $f(n, 2n-2) = 2n-2$. The subfamilies of $\mathscr{F}$ prove $f(n, k) = k$ for $0 \leqslant k \leqslant 2n-2$. On the other hand, it is easy to see that all members (except ϕ and X) can be omitted only then when any two members of $\mathscr{F}$ are either in inclusion or are disjoint. Lynch [6] proved that there are, at most, $2n-2$ sets in an n-element X satisfying this property. This proves $f(n, k) < k$ for $k > 2n-2$.

Finally, we study $f_4(n, k)$. $f_4(n, k)$ is the minimum number of members A of a family $\mathscr{F} = \mathscr{F}(\mathscr{L})$ closed under intersection, satisfying (i) $|\mathscr{F}| = k + 2$ and (ii) $\mathscr{F} - \{A\}$ is closed under intersection. Proposition 4 answers the question about which members can be omitted from $\mathscr{F}$. Hence, we have

$$f_4(n, k) = \min |\mathscr{M}(\mathscr{F}) - \{\phi\}| - 1. \tag{24}$$

Since $X \in \mathscr{M}(\mathscr{F})$ always while $\phi \in \mathscr{M}(\mathscr{F})$ not necessarily holds. Lemma 4 says that any member of $\mathscr{F}$ is an intersection of some members of $\mathscr{M}(\mathscr{F}) - \{X\}$. This implies

$$k+2=|\mathscr{F}|\leqslant 2^{|\mathscr{M}(\mathscr{F})|-1} \tag{25}$$

If $\phi\notin\mathscr{M}(\mathscr{F})$, then (24) and (25) result in

$$\lceil\log_2(k+2)\rceil\leqslant f_4(n,k). \tag{26}$$

If $\phi\in\mathscr{M}(\mathscr{F})$, then all the intersections containing ϕ are empty, therefore we have

$$k+2=|\mathscr{F}|\leqslant 1+2^{|\mathscr{M}(\mathscr{F})|-2}$$

in place of (25). This leads to

$$\log(k+1)\leqslant f_4(n,k). \tag{27}$$

First we show that (27) is sharp if n is large enough relative to k. This is easy in the case $k+1$ is a power of 2, say $k+1=2^\nu$. Suppose that $n\geqslant\nu+1$ and fix an element x of the n-element X, and choose a $(\nu+1)$-element subset N of X containing x. Let $\mathscr{F}$ consist of ϕ, X and of all proper subsets of N containing x. Then $|\mathscr{F}|$ is obviously equal to $1+2^\nu=k+2$ while

$$|\mathscr{M}(\mathscr{F})|=|\{\phi,X\}\cup\{\nu\text{-element subsets of } N \text{ containing } x\}|=2+\nu.$$

We have $f_4(n,k)=\nu$ by (24). Formula (27) is sharp in this case.

To prove the sharpness of (27) in the rest of the cases we need a lemma.

LEMMA 6. *If* $2^{r-1}+1\leqslant m\leqslant 2^r$ $(r\geqslant 1)$ *then there is a family* $\mathfrak{F}(r,m)$ *of* r *subsets of a* 2^{r-1}*-element set* $X(r,m)$ *such that* (i) *the number of all different intersections is exactly* m, (ii) *the intersection of all members of the family is empty,* (iii) $X(r,m)$ *itself is not in the family.*

Proof. We construct the family recursively. If $r=1$, then $m=2$. $\mathfrak{F}(1,2)=\{\phi\}$ in a 1-element $X(1,2)$. It is easy to see that it satisfies the conditions. Suppose now that $r>1$ and that the statement is proved for smaller r and any m.

First assume that m is even. By the induction hypothesis, there is a family $\mathfrak{F}(r-1,m/2)$. Let $x\in X(r-1,m/2)$ a completely new element and define

$$X(r,m)=X(r-1,m/2)\cup\{x\}$$

and

$$\mathfrak{F}(r,m)=\{F\cup\{x\}:F\in\mathfrak{F}(r-1,m/2)\}\cup\{X(r-1,m/2)\}.$$

The number $|\mathfrak{F}(r,m)|$ is obviously r. The intersections are either equal to the intersections generated by $\mathfrak{F}(r-1,m/2)$ (except the empty intersection, which is changed) or such sets completed by x. Therefore the number of intersection is exactly m. It is easy to see that the other conditions are also fulfilled.

If m is odd, take another copy $X'(r-1)$, $(m+1)/2)$ of $X(r-1,(m+1)/2)$. If $F\subset X(r-1,(m+1)/2)$ then let F' denote the set of copies of the elements of F. Define

$$X(r, m) = X(r-1, (m+1)/2) \cup X'(r-1, (m+1)/2)$$

$$\mathfrak{F}(r, m) = \{F \cup F' : F \in \mathfrak{F}(r-1, (m+1)/2)\} \\ \cup \{X(r-1, (m+1)/2)\}$$

where $\mathfrak{F}'$ is a copy of $\mathfrak{F}$ in $X'(r-1, (m+1))$. $|\mathfrak{F}(r, m)|$ is obviously r. The empty set is an intersection in two different ways, therefore the number of intersections is m. The other properties trivially hold. The lemma is proved.

Let us formulate our results on $f_4(n, k)$ as Theorem 9.

THEOREM 9.

$$f_4(n, k) \geq \lceil \log_2(k+1) \rceil$$

and

$$f_4(n, k) = \lceil \log_2(k+1) \rceil \text{ if } n \geq k+2.$$

Proof. The inequality is proved in (27). The equality is constructed when $k+1$ is a power of 2. The condition for the construction was $n \geq \log_2(k+1)+1$. We have to show a construction for the rest of the cases. $\lceil \log_2(k+1) \rceil$ can be replaced by $\lceil \log_2(k+2) \rceil$. Let $m = k+2$ and $r = \lceil \log_2(k+2) \rceil$. Apply Lemma 6 for these parameters. This can be done if $|X| \geq 2^{r-1}$, however the latter inequality certainly holds if $n \geq k+2$. The proof is complete. □

We know some upper estimates on $f_4(n, k)$ for smaller values of n, as well:

THEOREM 10.

$$f_4(n, k) \leq \lfloor \log_2(k+2) \rfloor - 1 + (\textit{number of bits } 1 \textit{ in the binary expansion of } (k+2)) \leq 2\lfloor \log_2(k+2) \rfloor.$$

We will use two lemmas, one of them (Lemma 5) has been stated earlier.

LEMMA 7. *Using the notations of Lemma 5, one can find nonempty subsets $C_2, \ldots, C_{l+\alpha_1}$ in the (α_1+1)-element subset C so, that the total number of intersections formed from $C_2, \ldots, C_{l+\alpha_1}$ is exactly r.*

Proof. Using Lemma 5, choose an (α_2+1)-element set A with the subsets $A_2, \ldots, A_l$ so that the number of all of their subsets is $2^{\alpha_2} + \cdots + 2^{\alpha_l}$. Let $|C| = \alpha_1 + 1$, $C \supset A$ and choose an element $a \in C - A$. This can be done since $\alpha_1 > \alpha_2$. Let $A_{l+1}, \ldots, A_{l+\alpha_1}$ be all $(\alpha_1 - 1)$-element subsets of $C - \{a\}$ and define

$$C_i = A_i \quad (2 \leq i \leq l),$$
$$C_i = A_i \cup \{a\} \quad (l+1 \leq i \leq l+\alpha_1).$$

It is easy to see that any set containing a is an intersection of some of $C_{l+1}, \ldots, C_{l+\alpha_1}$. The number of these subsets is 2^{α_1}. From the sets not containing a exactly those sets are intersections which are subsets of one of $A_2, \ldots, A_l$. Their number is $2^{\alpha_2} + \cdots + 2^{\alpha_l}$ by the construction. The total number of intersections is $r = 2^{\alpha_1} + 2^{\alpha_2} + \cdots + 2^{\alpha_l}$, as desired. The proof is complete. □

Proof of Theorem 10. Consider the binary expansion of $k+2 = 2^{\alpha_1} + \cdots + 2^{\alpha_l}$. Lemma 7 constructs $\alpha_1 + l - 1$ nonempty sets in a set X of size $\geqslant \alpha_1 + 1$ with $k+2$ possible intersections. Let $\mathscr{M}(\mathscr{F})$ consist of these sets and X, while $\mathscr{F}$ is the family of all intersections. Here $|\mathscr{F}| = k+2$ and $|\mathscr{M}(\mathscr{F})| - 1 = \alpha_1 + l - 1$. Since $\phi \notin \mathscr{M}(\mathscr{F})$ this proves

$$f_4(n, k) \leqslant \alpha_1 + l - 1$$

by (24). Here $\alpha_1 = \lfloor \log_2(k+2) \rfloor$ and the first inequality of the theorem is proved.

It is easy to prove that the number of bits 1 in the binary expansion of $k+2$ is at most $\lfloor \log_2(k+2) \rfloor + 1$. This proves the second inequality and the theorem.

We saw earlier that

$$f_4(n, 2^\nu - 1) = \log_2 2^\nu = \nu \tag{28}$$

independently of n. We need, of course, the condition $2^n \geqslant 2^\nu + 1$, that is, $n \geqslant \lceil \log_2(k+2) \rceil = \nu + 1$. However, (28) holds for any such n. Theorems 9 and 10 give

$$f_4(n, 2^\nu + 2^\mu - 2) = \lceil \log_2(2^\nu + 2^\mu) \rceil \tag{29}$$

if $\nu > \mu > 0$. (29) holds, supposing $n \geqslant \log_2(2^\nu + 2^\mu) = \nu + 1$, independently of n.

One might think that (28) and (29) can be generalized. However, it is not true when the number of bits 1 in the binary expansion of $k+2$ is more than 2. More exactly, choosing the smallest possible n, namely $n = \lceil \log_2(k+2) \rceil$, the value of $f_4(n, k)$ is larger than $\lceil \log_2(k+2) \rceil$ if $l > 2$. The inequality

$$f_4(\lceil \log_2(k+2) \rceil, k) > \lceil \log_2(k+2) \rceil \quad (l > 2)$$

can be formulated in terms of sets:

If $A_1, \ldots, A_n$ are subsets of an n-element set and the number of all intersections formed from $A_1, \ldots, A_n$ is more than 2^{n-1} then it has the form $2^{n-1} + 2^u$, where $n - 1 \geqslant u \geqslant 0$.

This statement is proved in [3].

Acknowledgement

We are indebted to the anonymous referee for his valuable remarks, especially for calling our attention to reference [4].

References

1. W. W. Armstrong (1974) Dependency structures of data base relationship, *Information Processing* **74**, North-Holland, Amsterdam, pp. 580–583.
2. E. F. Codd (1970) A relational model of data for large shared data banks, *Comm. ACM* **13**, 377–387.
3. J. Demetrovics, G. O. H. Katona, and P. P. Pálfy, On the number of unions in a family of sets, *Ann. NY Acad. Sci.* (submitted).
4. R. P. Dilworth and J. M. McLaughlin (1952) *Duke Math. J.* **19**, 683–694.
5. Z. Füredi, personal communication.
6. J. F. Lynch, Taxonomic classes of sets (preprint).