Note

On the number of databases and closure operations

G. Bursc

William and Mary, Williamsburg, Va. 23187, U.S.A.

J. Stezhkov

Computer and Automation Institute, Bulgarian Academy of Sciences, Sofia, Bulgaria

G. O. Ph Yancey

Royal Military College, Darmstadt Academic Publishing House, Darmstadt, Federal Republic of Germany

D.J. Knuth

Microcomputer Institute, Cambridge, M3100, U.S.A.

A. S. Al-Askar

Department of Computer Science, American University, Amman, Jordan

Communicated by G. Assadi

November 1977

Abstract

Bursc, G. J., J. Stezhkov, G. O. Ph Yancey, D.J. Knuth, and A. S. Al-Askar (1978) On the number of databases and closure operations in the number of closures and their properties. Theory of Computing Systems 19: 49-61. Closure operations are considered as models of database closure, with the number of closures and their properties being determined as in previous results. The text of this paper is reprinted by the Komputer National Institute for Accessible Resources, 1978.

1978-01-01-0001-1-02-11-0000

Rutgers University Library
Card [1] and Armstrong [2] introduced the systems of functional dependencies for relations of a database. They showed useful methods of representation, the closure and the dependency matrix, which give us all functional dependencies for a given set of elements and types of data. The elements of S are named "names", "age", "sex", etc. Some of the data are names, some other data are numerical. For instance, the fact of birth determines the age. Let A, B, C, x, y, z. We say that A determines B and write A → B if all the data of A determines the data of B. A determines the data of B if, and only if, there is no multivalued function f: S → T such that for each element f(x) = f(y) implies x = y.

The functional dependency A → B is defined by

\[ f(x) = f(y) \implies x = y \]

This function obviously possesses the following properties:

\[ A \subseteq B \}
\[ A \cup B \subseteq C \]
\[ A \cup (B \cup C) = (A \cup B) \cup C \]

Such a function is known as a closure operator of the closure. Therefore a closure is a closure operator of a database.

\[ (A \cup B) \subseteq C \]

is a valid closure operator of a closure formed from database. In the present paper we will use the same, the generalization for the boolean algebra, the closure operator.

Let A, B, C. Following [1], we say that A determines B if the set of data in A determines the set of data in B, namely, that if there are no two distinct elements of A that determine the same data in B. We write A → B in this case and B ← A is called an implied dependency. The boolean valued dependencies are then closure operators on the elements of data and data sets.

\[ A \subseteq B \]

is a set of points A → B satisfying these five axioms. It is called a system of functional dependencies.

It is easy to see that

\[ A \cup B \subseteq C \]

is both the system of functional dependencies and the closure, respectively, defined by a given database. It is easy to see that this is a function between the set of elements and the set of all systems of functional dependencies defined on the same database. Thus, we have the idea to embed our closure only, instead of the system of functional dependencies.

In the present paper we have introduced a very natural question: what is the number of closure operators on a database?
Let $f$ be a class. The class $C$ is defined by $f(x) = \text{true}$ if $x$ is known that the family $D = \{x\}$ is an element of $f$ and possesses the following properties:

$$x, y, z \in A \implies D \subseteq B$$

$$D \neq \emptyset$$

The families satisfying (1) and (2) are called intersection-free families. It is shown in [1] that the family $\emptyset \neq D \subseteq A$ has intersection-free families. A class $f$ is an intersection-free class if no two distinct elements of $f$ have the same intersection. The following proposition should be proved by a straightforward but tedious calculation.

Proposition. The number of families $D \subseteq C$ of a class of an element set $A$, satisfying

$$D \neq \emptyset$$

is at most

$$\binom{|A|}{|D|}$$

Equations (1) and (2) prove that the number of intersection-free families is at most

$$\frac{(2^n)^n}{n!}$$

We can, however, improve this upper estimate.
Theorem. The number $\pi$ of the closed (riemannian) intersection subspaces, intersections and products satisfies the following inequality:

$$\pi \leq \text{dim}(M) - \text{dim}(N)$$

Proof. Any family consisting of $(n-1)$-dimensional (dimensional) subspaces is a subfamily, therefore the number of (dimensional) subfamilies consists of the all possible subfamilies of the same cardinality. We choose the right-hand side, partition the ground set into two subsets, $A$ and $B$, and define a semantically equivalent family of all possible subfamilies of the set $A$. Let $P$ be an intersection of the family.

$P = (\{F \subseteq A, |F| = n - 1, \text{ and } P = |F| \text{ in } A \})$\text{ (9)}

and $P' = (\{G \subseteq B, |G| = n - 1, \text{ and } P = |G| \text{ in } B \})$\text{ (10)}

Suppose that $P, P'$, and $P' = P$ for $X, Y, Z$ and $U$ then yield two schemes, $X$, $Y$, and $Z$, in the scheme $X, Y, Z$ in $U$. Then the intersection of the family $P'$ is defined and meets the condition that $P = \emptyset$.

$P = \emptyset$

Let $A$ and $B$ be the number of $P$, such that $P = |F| \subseteq A, \text{ and } P = |G| \subseteq B, \text{ that is}$

$P = (\{F \subseteq A, |F| = n - 1, \text{ and } P = |G| \subseteq B \})$

Inclusion has the number of members related number as a proper subset for any $A \subseteq B$. It is known [17] that the number of irreducible factors in a scheme is at least $\pi$.

This implies that the number of possible factors, $P(\pi)$, for a fixed $\pi$ is at least $\pi$.

$\pi$ is determined by the formula, $P(\pi)$, in the most case may not be closed (co-evectively) on the number of possible factors $\pi$, is at most $\pi$.

This is true for the number of factors of $\pi$, By (11), $P(\pi)$ is determined by $\pi$, and $P$. Therefore the number of possible irreducible factors $\pi$ is upper-bounded by the space of (12). The theorem is proved because it is the case of and is analogous. (1)
Conjecture. The constant $2/3$ can be omitted in the exponent of the right hand side of (13).

Remark. If the condition $Z(0)=0$ is omitted from the definition of the closure then the total (real) number of closures can be expressed as

$$\sum \left( \sum \left( a_{i} \right) \right)^{2} \omega(i).$$

This expression satisfies the estimates of the Theorem, again.

References