“A RELATIONAL MODEL OF DATA FOR LARGE SHARED DATA BANKS”

Through the internet, I find more information about Edgar F. Codd. He is a mathematician and computer scientist who laid the theoretical foundation for relational databases--the standard method by which information is organized in and retrieved from computers. In 1981, he received the A. M. Turing Award, the highest honor in the computer science field for his fundamental and continuing contributions to the theory and practice of database management systems.

This paper is concerned with the application of elementary relation theory to systems which provide shared access to large banks of formatted data. It is divided into two sections. In section 1, a relational model of data is proposed as a basis for protecting users of formatted data systems from the potentially disruptive changes in data representation caused by growth in the data bank and changes in traffic. A normal form for the time-varying collection of relationships is introduced. In Section 2, certain operations

on relations are discussed and applied to the problems of redundancy and consistency in the user's model.

Relational model provides a means of describing data with its natural structure only--that is, without superimposing any additional structure for machine representation purposes. Accordingly, it provides a basis for a high level data language which will yield maximal independence between programs on the one hand and machine representation and organization of data on the other.

A further advantage of the relational view is that it forms a sound basis for treating derivability, redundancy, and consistency of relations. Finally, the relational view permits a clearer evaluation of the scope and logical limitations of present formatted data systems, and also the relative merits.

If the user's relational model is set up in normal form, names of items of data in the data bank can take a simpler form than would otherwise be the case. A general name would take a form such as

R (g).r.d

where R is a relational name; g is a generation identifier (optional); r is a role name (optional); d is a domain name. Since g is needed only when several generations of a given relation exist, or are anticipated to exist, and r is needed only when the relation R has two or more domains named d, the simple form R.d will often be adequate.

The operation of the relational data can use relational algebra or relational calculus to express. They are the basis of the actual relationship language and are equivalent to each other.

The adoption of a relational model of data permits the development of a universal data sublanguage based on an applied predicate calculus. There are two forms: tuple relational calculus and domain relational calculus, which are tuples with relations as predicates and the domain variables.

In the second section , it mainly talks about the redundancy and consistency.

Firstly, Edgar F. Codd introduced a few additional operations about set. As we all know, traditional set operations include union, intersection, difference, product and so on. The operations discussed in this section are specially for relations. It play a key role in deriving relations from other relations.

Permutation, projection, join, composition, restriction form a complete set of operations. And the other relational algebra operations can be a combination of these five operations to achieve. Thus we are now in a position to consider various applications of these operations on relations.

Secondly, we can know something about redundancy. A set of relations is strongly redundant if it contains at least one relation that possesses a projection which is derivable from other projections of relations in the set.

A collection of relations is weakly redundant if it contains a relation that has a projection which is not derivable from other members but is at all times a projection of some join of other projections of relations in the collection.

Thirdly, Relational database can maintain the data consistency. Enabling data sharing, Relational database reduces an amount of duplicate data and data redundancy. Whenever the named set of relations is redundant in either sense, we shall associate with that set a collection of statements which define all of the redundancies which hold independent of time between the member relations.

After writing the article" A Relational Model of Data for Large Shared Data Banks", he continued to carefully develop series of papers. Dr. Codd built upon this space and in doing so has provided the impetus for widespread research into numerous related areas, including database languages, query subsystems, database semantics, locking and recovery, and inferential subsystems.

Since the 1980s, almost all database management systems which computer makers launched have supported the relational model. Banks rely on relational database to track liquidity; retailers use them to monitor inventory levels; HR staff uses them to manage the accounts; libraries, hospitals and government agencies rely on them to store millions of records. Over 40 years passed, relational database has already been the basis for today's businesses.

When I am writing this review, there are still many puzzles in my mind. Although it is not difficult to know the meaning of each word, it is still not easy to well understand the meaning of the whole article after reading it several times. I know there are many things I have to learn. Just like one of my classmates said, "the adventure in the database begins." We may meet many challenges latter on. Just facing and beating them, we will get a further understand about the joy of learning and life.

Source: http://toyhouse.cc/profiles/blogs/personal-review-of-a-relational-model-of-data-for-large-shared