Relational Model of Data Large Shared Data Banks

--- E.F. Codd

Summary

The main problems Codd addressed in this paper are

- The application programs running by the user at end terminal are getting effected when there is a change in Internal Representation and some cases of external representation of data in data banks which have shared access.
- The changes in data representation depend on factors like changes in query, Types of stored information and query traffic.
- Inadequacies of existing tree-structured files or general network models.

It is Important to make application programs and user activities independent on growth in data type and data representation because the applications may only work for a particular representation of data and may fail if there is a change in data representation. It is also important that the data should be consistent when there is a change in data representation.

The Main Ideas which Codd discussed are

- The problems with existing general network models or tree-structured files like Data Dependency and Data Inconsistency.
- Data dependencies in existing systems like ordering dependency, indexing dependency and Access Path dependency.
- He proposed a relational view of data with an n-ary relation R which is set of n-tuples and explained the significance of relationship than domain ordering by introducing concept of primary key and foreign key.
- He proposed a Normal form which can be represented with a simple two dimensional homogeneous array by a simple elimination procedure called normalization.
- He also introduced a concept of universal sub data language based on predicate calculus.

- Certain operations are applied on relations (sets) to the problems of data redundancy and consistency :
  - Permutation
  - Projection
  - Join (Natural Join , Linear Join & Cyclic Join)
  - Composition
  - Restriction
- He applied the above operations on the set of relations and classify the redundancy into Strong redundancy and Weak Redundancy.
- He applied the same above operations to classify a relation as consistent or inconsistent. He also gave information on when the system will check for possible inconsistencies and how it responds.

The final results are

- Though there is no specific implemented solution for his proposals, He explained how relational models protect application programs running by user at end terminal affect by changes in data representation.
- He introduced a Normalization Technique to eliminate non simple designs which need complex data structures to represent.
- He classified the redundancy into two types: Strong and Weak. And also determined the consistency levels by applying operations on relations.

Submitted By

Lakshman Madhav Kollipara

SID- 932655504

Submitted To

Arash Termehchy

CS540- Database Management Systems