Relational Model

Hrudaya Kumar Tripathy
School of Computer Engineering,
KIIT University

Relational Model

- Relational data model is the primary data model for commercial data- processing applications
- A relational database consists of a collection of tables, each of which is assigned a unique name
- A row in a table represents a relationship among a set of values. Thus, a table is an entity set and a row is an entity
- The columns or properties are called attributes
- For each attribute, there is a set of permitted values, called the domain of that attribute. Same domain can be shared by more than one attribute
- Degree is the number of attributes in the relation/ table, where as Cardinality is the number of tuples or rows in the relation/table
- The attribute values are required to be atomic, i.e. indivisible

Relational Model...

- Let D_1, D_2, and D_3 are the domains. Any row of the table consists of a 3-tuple (v_1, v_2, v_3) where $v_1 \in D_1$, $v_2 \in D_2$ and $v_3 \in D_3$. Thus, the table will contain only a subset of the set of all possible rows. Therefore, the table is a subset of D_1 x D_2 x D_3
- Each attribute of a relation has a unique name
- NULL Value is a domain value which is a member of any possible domain
- Database Schema is the logical design of the database. If (a_1, a_2 ...a_n) be the attributes, then the relation schema will be R=(a_1,a_2 ...a_n)
- Database Instance is the snapshot of the data in the database at a given instant of time
- *Relation is denoted by lower case names and Relation Schema is the name beginning with an uppercase letter*

Relational Database

Relational Database

Relational database is a database consisting of multiple relations or tables. The information about an enterprise is broken up into parts, with each relation storing one part of the information

The normalization process deals with how to design relational schemas

Relational Data Integrity

Relational Data Integrity

Candidate key is an attribute or set of attributes that can uniquely identify a row or tuple in a table. Let R be the relation with attributes a_1, a_2 ... a_n. The set of attributes of R is said to be a candidate key of R iff the following two properties holds:

- **Uniqueness**: At any given time, no two distinct tuples or rows of R have the same value for a_i, the same value for a_j ...a_n
- **Minimality**: No proper subset of the set (a_i , a_j ... a_n) has the uniqueness property

The major types of integrity constraints are:

1. Domain Constraints

- All the values that appear in a column of a relation must be taken from the same domain
- This constraint can be applied by specifying a particular data type to a column

Relational Data Integrity...

2. Entity Integrity

- The entity integrity rule is designed to assure that every relation has a primary key, and that the data values for that primary key are all valid
- Usually, the primary key of each relation is the first column
- Entity integrity guarantees that every primary key attribute is NOT NULL
- Primary key performs the unique identification function in a relational model

3. Referential Integrity

- In relational data model, associations between tables are defined by using foreign keys
- A referential integrity constraint is a rule that maintains consistency among the rows of two relations

Relational Data Integrity...

3. Referential Integrity...

- The rule states that if there is a foreign key in one relation, either each foreign key value must match a primary key value in the other table or else the foreign key value must be NULL
- A foreign key that references its own relation is known as recursive foreign key
- The linking between the foreign key and primary key allows a set of relations to form an integrated database

4. Operational Constraints

- These are the constraints enforced in the database by the business rules or real world limitations

Database Languages

DDL (Data Definition Language)

- DDL is used to define the conceptual schema. The definition includes the information of all the entity sets and their associated attributes as well as the relationships between the entity sets
- The data values stored in the database must specify certain consistency constraints. The database systems check these constraints every time the database is updated
- The output of the DDL is placed in the Data Dictionary which contains the metadata (data about data)
- The data dictionary is considered to be a special type of table, which can only be accessed and updated by the database system itself
- The database system consults the data dictionary, before querying or modifying the actual data, for the validation purpose
- *CREATE, ALTER, DROP, RENAME & TRUNCATE*

Database Languages...

DML (Data Manipulation Language)

- DML is used to manipulate data in the database
- A query is a statement in the DML that requests the retrieval of data from the database
- *SELECT, INSERT, UPDATE & DELETE*

DCL (Data Control Languages)

- DCL allows in changing the permissions on database structures
- *GRANT & REVOKE*

TCL (Transaction Control Language)

- TCL allows permanently recording the changes made to the rows stored in a table or undoing such changes
- *COMMIT, ROLLBACK & SAVEPOINT*

CODD's Rules

Codd's rules are a set of 12 rules proposed by E. F. Codd designed to define what is required from a database management system in order for it to be considered relational,
i.e. RDBMS. Any database that satisfies even six rules may be categorized as RDBMS

Rule0

A relational system should be able to manage databases, entirely through its relational capabilities

Rule1: Information representation

The entire information is explicitly and logically represented by the data values of the tables in the relational data model

Rule2: Guaranteed access

In relational model, at each cell, i.e. the interaction of each row and column, it will have one and only one value of data (or NULL value). Each value of data must be addressable via the combination of a table name, primary key value and the column name

CODD's Rules...

Rule3: Systematic treatment of NULL values

NULL values are supported in fully relational DBMS for representing missing information and inapplicable information in a systematic way independent of data type

Rule4: Database description rule

The database description is represented at the logical level in the same way as ordinary data, so that authorized users can apply the same relational language to its interrogation as they apply to the regular data. This means, the RDBMS should have a data dictionary

Rule5: Comprehensive data sub-language

The RDBMS should have its own extension of SQL. The SQL should support Data Definition, View Definition, Data Manipulation, Integrity Constraint, and Authorization

CODD's Rules...

Rule6: Views updation

All views that are theoretically updatable are also updatable by the system. Similarly, the views which are theoretically non-updatable are also non-updatable by the database system

Rule7: High-level update, insert, deletes

A RDBMS should not only support retrieval of data as relational sets, but should also support insertion, updation and deletion of data as a relational set

Rule8: Physical data independence

Application programs and terminal activities are not disturbed if any changes are made either in storage representations or access methods

Rule9: Logical data independence

User programs and the user should not be aware of any changes to the structure of the tables such as the addition of extra columns

CODD’s Rules...

Rule10: Distribution independence

A relational DBMS has distribution independence. The RDBMS may spread across more than one system and across several networks. However to the end-user, the tables should appear no different to those that are local

Rule11: Integrity rule

Integrity rules must be supported by the relational data sub-language; they can be stored in the catalogue and not in the application program. **Entity integrity**: no component of a primary key may have a NULL value. **Referential integrity**: for every unique non-null ’foreign key’ values in the database, there should be a matching primary key value from the same domain

Rule12: Data integrity cannot be subverted

If a relational system has a low-level language, that low level cannot be used to subvert or bypass the integrity rules and constraints expressed in the higher level relational language