Subject: Database management system & RDBMS
Topic: Normalization
Name of the teacher: Lisna Thomas
Academic year:2020-2021

NORMALIZATION

NORMALIZATION & TYPES OF NORMALIZATION

1) DEFINE NORMALIZATION

Normalization can be defined as :-

- A process of organizing the data in database to avoid data redundancy, insertion anomaly, update anomaly & deletion anomaly.
- A process of organizing data into tables in such a way that the results of using the database are always unambiguous and as intended. Such normalization is intrinsic to relational database theory. It may have the effect of duplicating data within the database and often results in the creation of additional tables.

Types of normalization

- First Normal Form (1NF)
- Second Normal Form (2NF)
- Third Normal Form (3NF)
- Boyce-Codd Normal Form (BCNF)
- Fourth Normal Form (4NF)
- Fifth Normal Form (5NF)

First Normal Form-

A given relation is called in First Normal Form (1NF) if each cell of the table contains only an atomic value.

OR

A given relation is called in First Normal Form (1NF) if the attribute of every tuple is either single valued or a null value.

Example-

The following relation F-

Student_id	**Name**	**Subjects**
100	Akshay	Computer Networks, Designing
101	Aman	Database Management System
102	Anjali	Automata, Compiler Design

However,

- This relation can be brought into 1NF.
- This can be done by rewriting the relation such that each cell of the table contains only one value.

Student_id	Name	Subjects
100	Akshay	Computer Networks
100	Akshay	Designing
101	Aman	Database Management System
102	Anjali	Automata
102	Anjali	Compiler Design

Second Normal Form-

A given relation is called in Second Normal Form (2NF) if and only if-

1. Relation already exists in 1NF.
2. No partial dependency exists in the relation.

Partial Dependency

A partial dependency is a dependency where few attributes of the candidate key determines non-prime attribute(s).

OR

A partial dependency is a dependency where a portion of the candidate key or incomplete candidate key determines non-prime attribute(s).

In other words,

A → B is called a partial dependency if and only if-

1. A is a subset of some candidate key
2. B is a non-prime attribute.

If any one condition fails, then it will not be a partial dependency.

Example-

Consider a relation- R (V , W , X , Y , Z) with functionaldependencies-

$$VW \rightarrow XY$$

$$Y \rightarrow V$$

$$WX \rightarrow YZ$$

The possible candidate keys for this relation are-

VW , WX , WY

From here,

- Prime attributes = { V , W , X , Y }
- Non-prime attributes = { Z }

Now, if we observe the given dependencies-

- There is no partial dependency.
- This is because there exists no dependency where incomplete candidate key determines any non-prime attribute.

Thus, we conclude that the given relation is in 2NF.

Third Normal Form-

A given relation is called in Third Normal Form (3NF) if and only if-

1. Relation already exists in 2NF.
2. No transitive dependency exists for non-prime attributes.

Transitive Dependency

A → B is called a transitive dependency if and only if-

1. A is not a super key.
2. B is a non-prime attribute.

If any one condition fails, then it is not a transitive dependency.

NOTE-

- Transitive dependency must not exist for non-prime attributes.
- However, transitive dependency can exist for prime attributes.

OR

A relation is called in Third Normal Form (3NF) if and only if-

Any one condition holds for each non-trivial functional dependency $A \rightarrow B$

1. A is a super key
2. B is a prime attribute

Example-

Consider a relation- R (A , B , C , D , E) with functional dependencies-

$$A \rightarrow BC$$

$$CD \rightarrow E$$

$$B \rightarrow D$$

$$E \rightarrow A$$

The possible candidate keys for this relation are-

$$A , E , CD , BC$$

From here,

- Prime attributes = { A , B , C , D , E }
- There are no non-prime attributes

Now,

- It is clear that there are no non-prime attributes in the relation.
- In other words, all the attributes of relation are prime attributes.
- Thus, all the attributes on RHS of each functional dependency are prime attributes.

Thus, we conclude that the given relation is in 3NF.