Normalization - I I

Gagan Deep

Founder & Director

Rozy Computech Services

Kurukshetra , rozygag@yahoo.com

www.rozyph.com

In our previous lecture we discussed

- **Basics of Relational databases**
- **Objectives of Normalization**
- **Modification anomalies**
  - **Deletion Anomaly**
  - **Insertion Anomaly**
  - **Updation Anolmaly**
- **Functional Dependencies**
- **Armstrong's axioms of FD etc.**

Now we have to talk about Normalization & Different forms of normalization.

NORMALIZATION

Normalization is

- **Refinement of the database specifically ER model.**
- **Process of removing the anomalies of database/s**
- **Segregation of data over many entities or relations.**

Need for Normalization

- Improves database design.
- Ensures minimum redundancy of data.
- To make database structure flexible
- Reduces need to reorganize data when design is modified or enhanced.
- To achieve consistency - Removes modification anomalies. Normalized model converted to physical database tables.

Classes of Relation

- **Relation can be classified by the types of modification anomalies to which they are vulnerable.**
- **The classes of relations and techniques for preventing anomalies are called *normal forms*.**
- **Depending on its structure, a relation might be in *first normal form*, *second normal form*, *third normal form* or some other normal form.**

- **E.F. Codd defined first, second, and third normal forms (1NF, 2NF, 3NF). Later, *Boyce-Codd normal* form (BCNF) was postulated, and then *fourth and fifth normal forms* were designed.**
- **Thus *a relation that is in second form is also in first normal form*. A relation in 5NF is also in 4NF, BCNF, 3NF, 2NF and 1NF.**

First Normal Forms (1NF)

- **Every relation is, by definition, in first normal form. That is because a relation, has no repeating groups in it - that is the characteristic of the first normal form.**
- **In easy words, only one value is associated with each attribute and the value is not a set of values or a list of values.**
- **Let us study with the help of example about this non repeating group characteristic**

Repeating Group

Department	Teacher	Paper	Class
Computer	A Kush	101	BCAI
		102	BCAI
		202	BSCII
	P Malik	102	BCAI
		105	BCAI
Math	S B Malik	123	BSCII
		234	BSCII
	N Bhatti	124	BCAI
		125	BCAI
		233	BSCII

Non Repeating Group

Table shows the 1st normal form

Department	Teacher	Paper	Class
Computer	A Kush	101	BCAI
Computer	A Kush	102	BCAI
Computer	A Kush	202	BSCII
Computer	Gagan Deep	102	BCAI
Computer	Gagan Deep	105	BCAI
Math	S B Malik	123	BSCII
Math	S B Malik	234	BSCII
Math	N Bhatti	124	BCAI
Math	N Bhatti	125	BCAI
Math	N Bhatti	233	BSCII

Anomalies in 1NF

The first normal form was related to the structure of the representation but the second normal form is related to eliminate data redundancy in these relations.

Before going to 2NF let us discuss the problems of 1NF. The problems can be listed as.

1. **Deletion anomaly**
2. **Insertion anomaly**
3. **Updation anomaly**

Let us consider another example Library order with attribute order_no, B_name, qty and cost. Here PK is composed of two attribute (order_no & B_name) non key attributes are qty and cost. Let us consider the following table.

order_no	B_name	qty	cost
2000	Let Us C	12	190
2000	C++	14	200
2002	DBMS	23	245
2003	C++	10	150
2008	DBMS	20	300

Although the Library Order is in 1NF but it suffers from all anomalies like Deletion, Insertion and Updation because of some attributes are functionally dependant on others.

To overcome such problems we have to convert this relation into 2NF.

Second Normal Form (2NF)

- **The database must meet all the requirements of the first normal form.**
- **A relation is in Second Normal Form if all non-key attributes are dependent on all of the key.**
- **If the key is a single attribute, then the relation is automatically in the second normal form.**
- **GAME Table(as in modification anomaly) can be decomposed to form two relations is in second normal form, namely S-GAME and G-FEE.**
- **These relations are in 2NF because they have single-attribute keys.**

2NF

ROLLNO	GAME	FEE
101	CRICKET	500
110	FOOTBALL	150
135	SWIMMING	100
180	TENNIS	300

ROLLNO	GAME
101	CRICKET
110	FOOTBALL
135	SWIMMING
180	TENNIS

GAME	FEES
CRICKET	500
FOOTBALL	150
SWIMMING	100
TENNIS	300

Steps to convert 1NF into 2NF with the Library Order example : as you seen Game example

To convert a 1NF table into 2NF we may split the table into two or more relations where any resulting has no partial key dependency. The procedure may be as follows.

1. Create a new table which contains the attribute FFD on the concatenated PK and PK itself.

order_no	B_name	qty	cost
2000	Let Us C	12	190
2000	C++	14	200
2002	DBMS	23	245
2003	C++	10	150
2008	DBMS	20	300

Library order

order_no	B_name	qty
2000	Let Us C	12
2000	C++	14
2002	DBMS	23
2003	C++	10
2008	DBMS	20

order relation

B_name	Cost
Let Us C	190
C++	200
DBMS	300

Information relation

2. Create another table which contains the attributes that are not fully FD on PK and members of PK on which they are dependent.

The relation is broken into two relations as follows.

order relation(_order_no, B_name ,qty)

Information (B_name , cost)

Although many shortcomings of 1NF has been removed in 2NF still in certain cases all these three anomalies encounter which are discussed and removed in 3NF.

Third Normal Form (3NF)

- **Third normal form (3NF) requires that data stored in a table be dependent only on the primary key, and not on any other field in the table.**
- **Any field which is dependent not only on the primary key but also on another field is moved out to a separate table.**
- **A relation is in the third normal form if it is in second normal form and has no transitive dependencies.**

Example

- **Previous examples of 2NF is also for 3NF.**
- **Consider another example, the Teacher relation in Table on next slide. Suppose the requirements underlying this relation are that a student (ROLLNO) can have one or more subjects (SUBJECT), a subject can have several faculty members (Tname) as Teachers, and a faculty member (Tname) Teaches in only one subject area.**

Teacher(ROLLNO, SUBJECT, Tname)

Key (primary) : (ROLLNO, SUBJECT)

Key(candidate) : (ROLLNO, Tname)

Functional dependencies : Tname ---- SUBJECT

Relation in 3NF

In this relation there is no transitive dependency

$A \rightarrow B$ and $A \rightarrow C$ but not $B \rightarrow C$

ROLLNO	SUBJECT	TNAME
100	Math	Ramesh Lal
120	Hindi	Hari Chand
150	Math	Rajneesh Kumar
160	Hindi	Hari Chand

Data Anomalies in 3NF

- **3NF helped us to remove some shortcomings that occurred in 2NF however still there certain anomalies in 3NF particularly when the relation has when the relation has two overlapping candidate keys , composite candidate keys or multiple candidate key.**
- **All these shortcomings are removed in BCNF and other higher normalization processes like 4NF & 5NF.**

Boyce Codd Normal Form (BCNF)

- ***A relation is in BCNF if and only if the determinants are candidate key.***
- **The difference between 3NF and BCNF is that for a Functional dependency A → B, 3NF allows this dependency in a relation if B is a primary key attribute and A is not a candidate key, whereas BCNF insists that for this dependency to remain in a relation, A must be a candidate key.**

Example of BCNF

- Relation in 3NF but not in BCNF.

ROLLNO	SUBJECT	TNAME
100	Math	Ramesh Lal
120	Hindi	Hari Chand
150	Math	Rajneesh Kumar
160	Hindi	Hari Chand

S-Teacher (ROLLNO,Tname) TEACHER-SUBJECT (Tname, Subject)

Key: ROLLNO, Tname Key : Tname

ROLLNO	Tname	Tname	Subject
100	Ramesh Lal	Ramesh Lal	Math
120	Hari Chand	Hari Chand	Hindi
150	Rajneesh Kumar	Rajneesh Kumar	Math
160	Hari Chand		

Comparison of BCNF & 3NF

Since a BCNF is modification of 3NF so a relation in BCNF is also in 3NF but the converse is not true.

1. **In BCNF the relation has multiple composite key and two or more candidate keys share a common attribute .**
2. **The FD existing in source relation may not be preserved in broken relations.**

3. If a relation has only one candidate key than it can be normalized with 3NF or BCNF

4. Basically BCNF is a stronger definition of 3NF.

Thanks !

If you have any query please mail me at

rozygag@yahoo.com

www.rozyph.com