First, second, third Normal Form

COURSE : 23CAT- Database Management System
UNIT III : Database Design
CLASS : I Semester / I MCA

Evils of Redundancy

Repetition of data may leads to

- Making relations very large
- Not easy to maintain and update data
- Errors and inconsistencies increases
- Searching is too slow
- Poor utilization of disk space and resources

Normalization

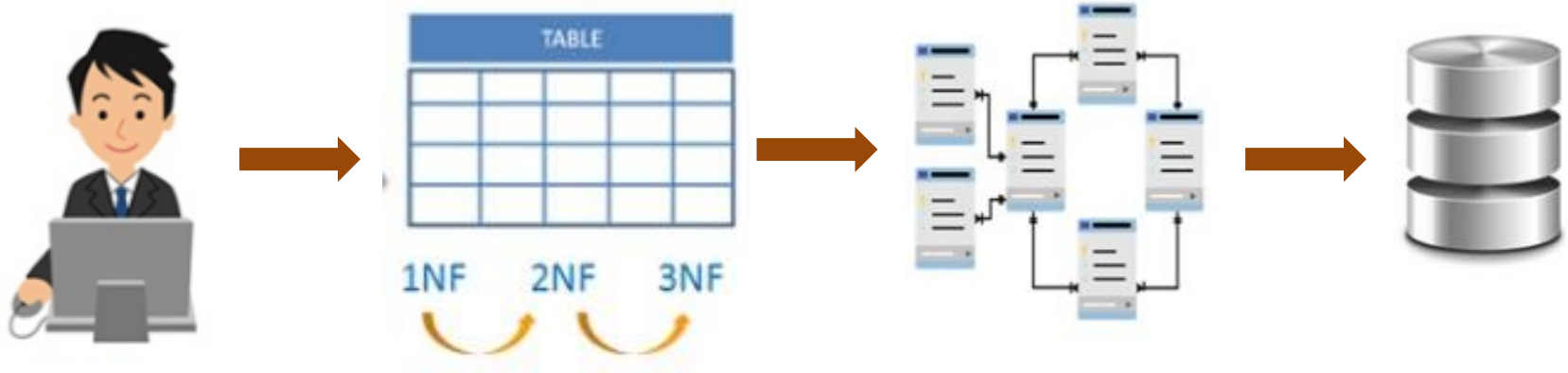
What is Normalization?

- Process of organizing the data in the database
- Divides the larger table into smaller and links them using relationships

Why Normalization?

- Used to minimize the redundancy
- Eliminate undesirable characteristics like Insertion, Update, and Deletion Anomalies

Normalization



Data Modification Anomalies

Insertion Anomaly

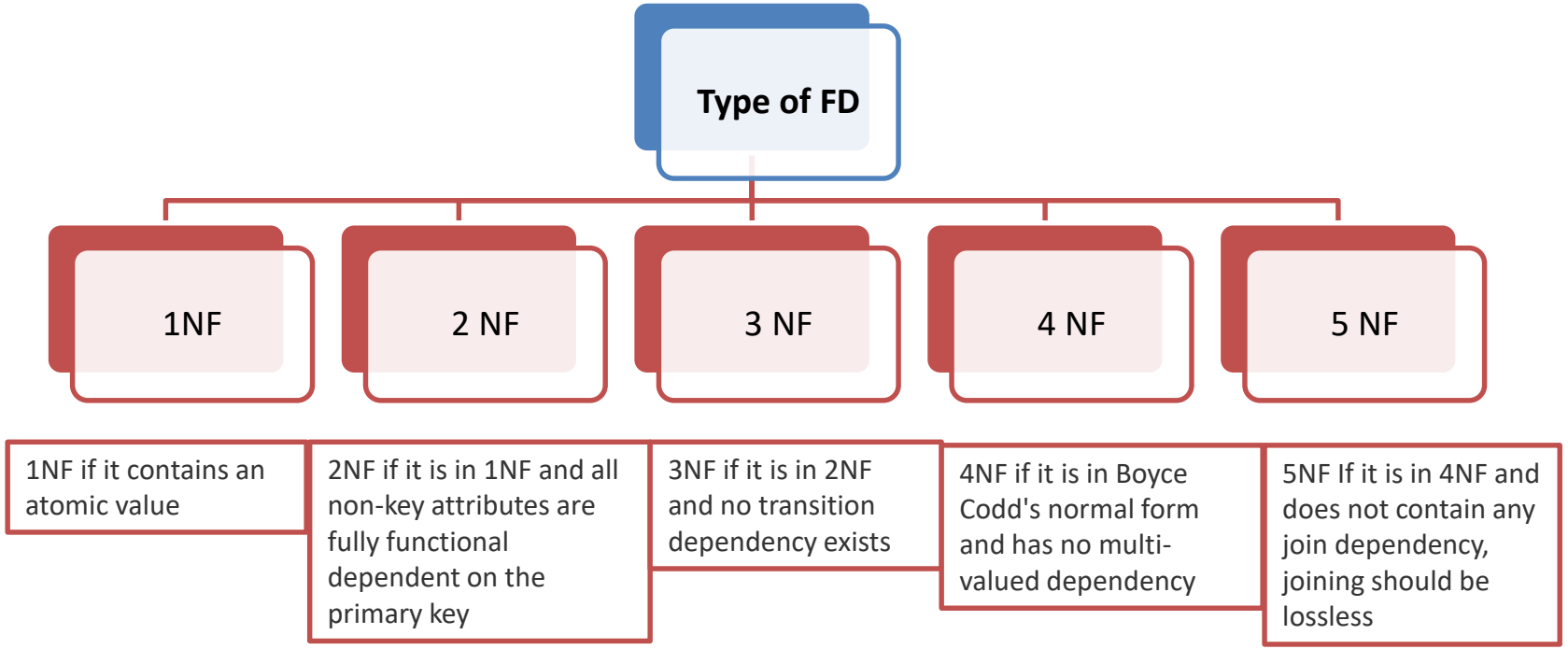
One cannot insert a new tuple into a relationship due to lack of data

Deletion Anomaly

Deletion of data results in the unintended loss of some other important data

Updation Anomaly

When an update of a single data value requires multiple rows of data to be updated.


Type of FD
1NF
2 NF
3 NF
4 NF
5 NF
1NF if it contains an atomic value
2NF if it is in 1NF and all non-key attributes are fully functional dependent on the primary key
3NF if it is in 2NF and no transition dependency exists
4NF if it is in Boyce Codd's normal form and has no multi-valued dependency
5NF If it is in 4NF and does not contain any join dependency, joining should be lossless

Normalization

- Greater database organization
- Minimize data redundancy
- Ensures Data consistency
- More flexible database design

Enforces the relational integrity

- Greater database organization
- Minimize data redundancy
- Ensures Data consistency
- More flexible database design

Enforces the relational integrity

1 Normal Form (1 NF)

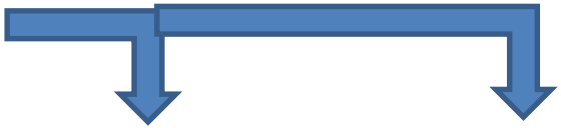
- A relation will be 1NF if it contains an atomic value
- An attribute of a table cannot hold multiple values. It must hold only single-valued attribute.
- Disallows the multi-valued attribute, composite attribute, and their combinations.

Emp_ID	Emp_name	Phone_no	State
A101	John	123456789 157856511	TN
A102	Hari	2548796 7598456	KA

Emp_ID	Emp_name	Phone_no	State
A101	John	123456789	TN
A101	John	157856511	TN
A102	Hari	7598456	KA
A102	Hari	2548796	KA

Teacher_ID	Course	Age
A101	DBMS	30
A101	JAVA	30
A102	C	45
A102	Cloud	45

- relational must be in 1NF
- all non-key attributes are fully functional dependent on the primary key



COURSE

Teacher_ID	Course
A101	DBMS
A101	JAVA
A102	C
A102	Cloud

TEACHER

Teacher_ID	Age
A101	30
A102	45

- if it is in 2NF and not contain any transitive partial dependency
- Reduces data duplication and achieve the data integrity
- A relation is in third normal form if it holds atleast one of the following conditions for every non-trivial function dependency X → Y.
  - X is a super key.
  - Y is a prime attribute, i.e., each element of Y is part of some candidate key

3 Normal Form (3 NF) Example

Order ID	Customer ID	Customer Name	Customer City	Order Date	Order Total
1	100	John Smith	New York	2022-01-01	100
2	101	Jane Doe	Los Angeles	2022-01-02	200
3	102	Bob Johnson	San Francisco	2022-01-03	300

"Customer City" is transitively dependent on the primary key. That is, it depends on "Customer ID", which is not part of the primary key, instead of depending directly on the primary key "Order ID"

Table 1: Customers

Customer ID	Customer Name	Customer City
100	John Smith	New York
101	Jane Doe	Los Angeles
102	Bob Johnson	San Francisco

Table 2: Orders

Order ID	Customer ID	Order Date	Order Total
1	100	2022-01-01	100
2	101	2022-01-02	200
3	102	2022-01-03	300