

# MIS 301 RELATIONAL DATABASE MANAGEMENT SYSTEM

DATABASE MANAGEMENT SYSTEM

Database Design: Functional Dependencies, Normalization & denormalization

**Lecture 21, 22 & 23**

# FUNCTIONAL DEPENDENCIES

- ❖ *An attribute  $Y$  is said to be functionally dependent on attribute  $X$  if say  $Y$  takes up a value  $B$  whenever  $X$  takes up a value  $A$  but if  $Y$  takes up a value  $B$ , it is not essential for  $X$  to take up  $A$ .*
- ❖ *Symbolic representation:  $X \rightarrow Y$*
- ❖ The left side of FD is known as a determinant, the right side is known as a dependent.

## **Types of Functional Dependencies**

- Multivalued dependency
- Trivial functional dependency
- Non-trivial functional dependency
- Transitive dependency

# MULTIVALUED DEPENDENCY

- ❖ Multivalued dependency occurs when two attributes in a table are independent of each other, but both depend on a third attribute.
- ❖ In the given example Stream and Extracurricular are independent of each other but both depend on student.
- ❖ Symbolic representation: Student\_name  $\twoheadrightarrow$  Stream, Student\_name  $\twoheadrightarrow$  Extracurricular

Student_Name	Stream	Extracurricular
Ravi	Marketing	Cricket
Rajesh	Finance	Cricket
Dinesh	Marketing	Swimming
Dinesh	Marketing	Painting
Rajesh	Finance	Painting
Ritika	MIS	Dancing
Surabhi	Human_resource	Dancing

# TRIVIAL FUNCTIONAL DEPENDENCY

- **Trivial** – If a functional dependency  $X \rightarrow Y$  holds, where  $Y$  is a subset of  $X$ , then it is called a trivial FD.

Example             $\{\text{roll\_no, name}\} \rightarrow \text{name}$   
                          $\{\text{roll\_no, name}\} \rightarrow \text{roll\_no}$

- **Non-trivial** – If an FD  $X \rightarrow Y$  holds, where  $Y$  is not a subset of  $X$ , then it is called a non-trivial FD.

Example             $\text{name} \rightarrow \text{father\_name}$

- **Completely non-trivial** – If an FD  $X \rightarrow Y$  holds, where  $x \cap Y = \Phi$ , it is said to be a completely non-trivial FD.

Example             $\text{college} \rightarrow \text{principal}$

# TRANSITIVE DEPENDENCY

❖ When an indirect relationship causes functional **dependency** it is called **Transitive Dependency**.  
If  $P \rightarrow Q$  and  $Q \rightarrow R$  is true, then  $P \rightarrow R$  is a **transitive dependency**.

❖ Example :

college  $\rightarrow$  principal

principal  $\rightarrow$  qualification

❖ Transitive dependency says  
college  $\rightarrow$  qualification

# ADVANTAGES OF FUNCTIONAL DEPENDENCIES

- Functional Dependency removes data redundancy where the same values should not be repeated at multiple locations in the same database table ensuring elimination of data inconsistency.
- It ensures good quality of data in the database.
- It allows clearly defined meanings and constraints of databases.
- It helps in identifying bad designs of the database.

# NORMALIZATION

- ❖ Redundancy in relation may cause insertion, deletion and updation anomalies.
- ❖ **Normalization** is the process of minimizing redundancy from a relation or set of relations.
- ❖ Normalization is usually obtained by breaking tables into multiple tables.
- ❖ The core idea of database normalization is to divide the tables into smaller sub-tables and store pointers to data rather than replicating it
- ❖ Normalization helps in
  - Reducing the amount of storage needed to store the data.
  - Avoiding unnecessary data conflicts that may creep in because of multiple copies of the same data getting stored.

# 1<sup>ST</sup> DEGREE NORMALIZATION - 1NF

- ❖ A relation is said to be in 1NF i.e. normalized to the first degree if and only if all attributes in the relation contain atomic values.
- ❖ Atomic values refer to absence of multiple values in an attribute of a table. All attributes are single valued.
- ❖ The following table is not in 1NF because ***Subject taught*** is not atomic as it contains multiple values.

Teacher_code	Teacher_name	Subject_taught
T1	Radhika Sen	Physics , Mathematics
T2	Geeta Kanwal	Chemistry
T3	Rajiv Gupta	English, Biology



# 1<sup>ST</sup> DEGREE NORMALIZATION - 1NF

- ❖ A relation is said to be in 1NF i.e. normalized to the first degree if and only if all attributes in the relation contain atomic values.
- ❖ Atomic values refer to absence of multiple values in an attribute of a table. All attributes are single valued.
- ❖ The following table is in 1NF because all attributes contain indivisible, atomic values.

Teacher_code	Teacher_name	Subject_taught
T1	Radhika Sen	Physics
T2	Geeta Kanwal	Chemistry
T3	Rajiv Gupta	English
T1	Radhika Sen	Mathematics
T3	Rajiv Gupta	Biology

# 2<sup>ND</sup> DEGREE NORMALIZATION - 2 NF

- ❖ A relation is said to be normalized to the second degree if it is normalized to the first degree and every non-key attribute is functionally dependent on the key attribute.
- ❖ Here Sname, Stream, City\_cd, Ct\_name, Ct\_status are functionally dependent on Roll\_no. Tname and Tdesig are functionally dependent on T\_code. Subject and marks are functionally dependent on roll\_no+T\_code.
- ❖ Hence this relation is not in 2NF.

Roll_no	Sname	Stream	City_cd	Ct_name	Ct_status	T_code	Tname	Tdesig	Subject	Marks
101	Ranjan	Science	12	Durgapur	2	T1	NGupta	AsstProf	Physics	64
102	Geeta	Science	14	Kolkata	1	T1	NGupta	AsstProf	Physics	72
101	Ranjan	Science	12	Durgapur	2	T2	SSimgh	AsocProf	English	58
103	Nikhil	Science	13	Ondal	3	T4	PKaur	Prof	Maths	80
104	Aditya	SScience	12	Durgapur	2	T5	GSen	AsstProf	History	68
104	Aditya	SScience	12	Durgapur	2	T5	GSen	AsstProf	Civics	78
105	Biman	SScience	14	Kolkata	1	T4	PKaur	Prof	Logic	80

# 2<sup>ND</sup> DEGREE NORMALIZATION - 2 NF

- ❖ A relation is said to be normalized to the second degree if it is normalized to the first degree and every non-key attribute is functionally dependent on the key attribute.
- ❖ Here Sname, Stream, City\_cd, Ct\_name, Ct\_status are functionally dependent on Roll\_no. Tname and Tdesig are functionally dependent on T\_code. Subject and marks are functionally dependent on roll\_no+T\_code.
- ❖ Hence the relations student, teacher and result are in 2NF.

Roll_no	Sname	Stream	City_cd	Ct_name	Ct_status
101	Ranjan	Science	12	Durgapur	2
102	Geeta	Science	14	Kolkata	1
103	Nikhil	Science	13	Ondal	3
104	Aditya	SScience	12	Durgapur	2
105	Biman	SScience	14	Kolkata	1

**Student**

T_code	Tname	Tdesig
T1	NGupta	AsstProf
T2	SSimgh	AssoProf
T4	PKaur	Prof
T5	GSen	AsstProf

**Teacher**

Roll_no	T_code	Subject	Marks
101	T1	Physics	64
102	T1	Physics	72
101	T2	English	58
103	T4	Maths	80
104	T5	History	68
104	T5	Civics	78
105	T4	Logic	80

**Result**

# 3<sup>RD</sup> DEGREE NORMALIZATION - 3 NF

- ❖ A relation is said to be normalized to the third degree if it is normalized to the second degree and every non-key attribute is non-transitively dependent on the key attribute.
- ❖ Here Sname, Stream, City\_cd, Ct\_name, Ct\_status are functionally dependent on Roll\_no. Tname and Tdesig are functionally dependent on T\_code. Subject and marks are functionally dependent on roll\_no+T\_code. But there is a transitive dependence in student.
- ❖ Ct\_name and Ct\_status are functionally dependent on city\_cd also besides being dependent on roll\_no. Therefore transitive dependence exists.
- ❖ Hence these relations are not in 3NF.

Roll_no	Sname	Stream	City_cd	Ct_name	Ct_status	T_code	Tname	Tdesig	Roll_no	T_code	Subject	Marks
101	Ranjan	Science	12	Durgapur	2	T1	NGupta	AsstProf	101	T1	Physics	64
102	Geeta	Science	14	Kolkata	1	T2	SSimgh	AssoProf	102	T1	Physics	72
103	Nikhil	Science	13	Ondal	3	T4	PKaur	Prof	101	T2	English	58
104	Aditya	SScience	12	Durgapur	2	T5	GSen	AsstProf	103	T4	Maths	80
105	Biman	SScience	14	Kolkata	1	<b>Teacher</b>			104	T5	History	68
<b>Student</b>						<b>Result</b>			104	T5	Civics	78
									105	T4	Logic	80

# 3<sup>RD</sup> DEGREE NORMALIZATION - 3 NF

- ❖ A relation is said to be normalized to the third degree if it is normalized to the second degree and every non-key attribute is non-transitively dependent on the key attribute.
- ❖ Here Sname, Stream, City\_cd are functionally dependent on the key field in relation Roll\_no.
- ❖ Tname and Tdesig are functionally dependent on the key field in relation T\_code.
- ❖ Subject and marks are functionally dependent on the key field in relation roll\_no+T\_code.
- ❖ Ct\_name and Ct\_status are functionally dependent on city\_cd, the key field in relation city.
- ❖ Hence these relations are in 3NF.

Roll_no	Sname	Stream	City_cd	City_cd	Ct_name	Ct_status	T_code	Tname	Tdesig	Roll_no	T_code	Subject	Marks
101	Ranjan	Science	12	12	Durgapur	2	T1	NGupta	AsstProf	101	T1	Physics	64
102	Geeta	Science	14	14	Kolkata	1	T2	SSimgh	AssoProf	102	T1	Physics	72
103	Nikhil	Science	13	13	Ondal	3	T4	PKaur	Prof	101	T2	English	58
104	Aditya	SScience	12				T5	GSen	AsstProf	103	T4	Maths	80
105	Biman	SScience	14							104	T5	History	68
										104	T5	Civics	78
										105	T4	Logic	80

**City**

**Teacher**

**Student**

**Result**

# BOYCE CODD NORMAL FORM - BCNF

- ❖ BCNF is a stronger form of 3NF considering candidate keys and not just primary keys.
- ❖ BCNF says that every candidate key must be a determinant
- ❖ A table is in Boyce-Codd Normal form if and only if at least one of the following conditions are met for each functional dependency  $A \rightarrow B$ :
  - A is a superkey(a set of attributes uniquely identifying tuples in a relation).
  - It is a trivial functional dependency.
- ❖ All relations in BCNF are in 3NF but the reverse may not be true

# BOYCE CODD NORMAL FORM - BCNF

❖ Considering a relation Std\_tchr\_sub containing attributes student\_name, Teacher\_name and subject

where

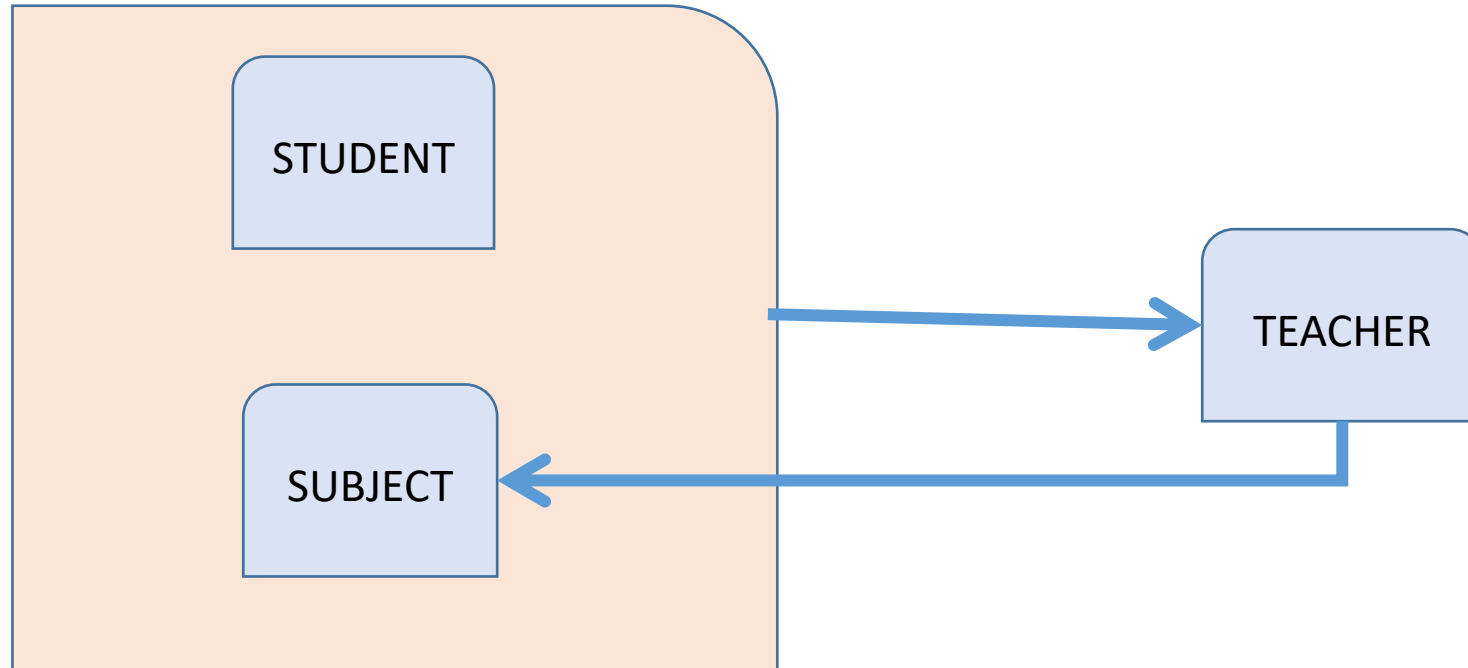
- 1. Each subject is taught to a student by a single teacher*
- 2. Each teacher teaches only one subject*
- 3. A subject is taught by several teachers*

❖ FDs in this relation are:

{Student\_name, Subject} → Teacher\_name

Teacher\_name → Subject

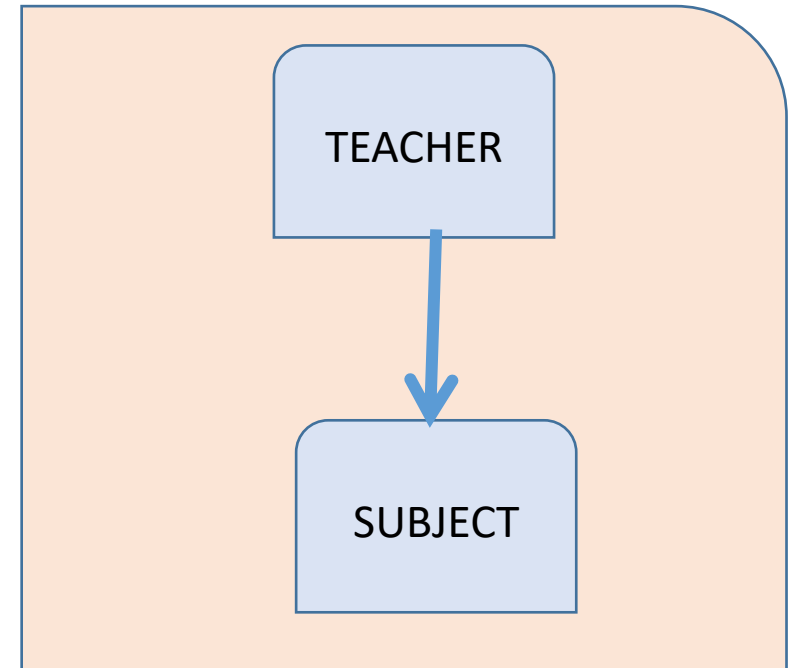
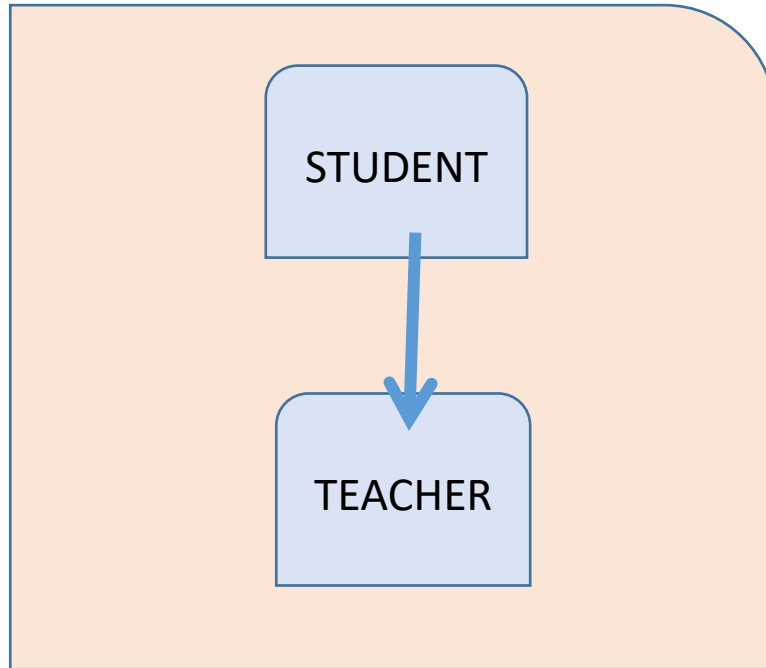
# BOYCE CODD NORMAL FORM - BCNF



- The two overlapping candidate keys are student+subject and student+teacher
- The relation is in 3NF but not in BCNF



# BOYCE CODD NORMAL FORM - BCNF



- The relations are decomposed into two relations Stud\_tchr containing student and teacher and tchr\_sub containing teacher and subject, the keys in the two tables being student and teacher respectively
- Now the relations are in 3NF as well as in BCNF

# DENORMALIZATION

- Denormalization is a database optimization technique where redundant data are added in the database to get rid of complex join operations.
- Denormalization is done after normalization for improving the performance of the database.
- Advantages of Denormalization:-Query execution is fast since fewer tables have to be joined.
- Disadvantages of Denormalization:-As data redundancy is there, update and insert operations are more expensive and take more time. Data Integrity is not maintained in denormalization.

- TILL WE MEET AGAIN IN THE NEXT CLASS.....

