CS105.3
Database Management Systems

Normalization

By: Chalani Oruthotaarachchi

Normalization

The objective of normalization:

"to create relations where every dependency is on the key, the whole key, and nothing but the key".

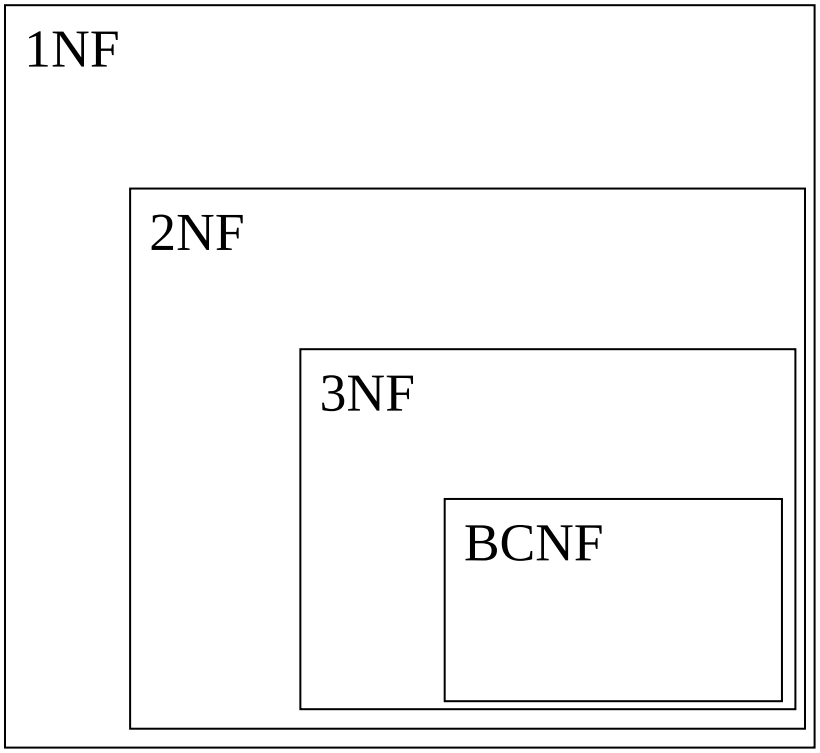
- Evaluating and correcting table structures to minimize data redundancies
- Reduces data anomalies

Normalization

There is a sequence to normal forms:
- 1NF is considered the weakest,
- 2NF is stronger than 1NF,
- 3NF is stronger than 2NF, and
- BCNF is considered the strongest

Also,
- any relation that is in BCNF, is in 3NF;
- any relation in 3NF is in 2NF; and
- any relation in 2NF is in 1NF.

Normalization



a relation in BCNF, is also in 3NF

a relation in 3NF is also in 2NF

a relation in 2NF is also in 1NF

First Normal Form

The following in **not** in 1NF

EmpNum	**EmpPhone**	**EmpDegrees**
123	233-9876	
333	233-1231	BA, BSc, PhD
679	233-1231	BSc, MSc

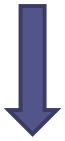
EmpDegrees is a multi-valued field:

employee 679 has two degrees: *BSc* and *MSc*

employee 333 has three degrees: *BA*, *BSc*, *PhD*

First Normal Form

EmpNum	EmpPhone	EmpDegrees
123	233-9876	
333	233-1231	BA, BSc, PhD
679	233-1231	BSc, MSc



1NF

EmpNum	EmpPhone	EmpDegree
123	233-9876	
333	233-1232	BA
333	233-1232	BSc
333	233-1232	PhD
679	233-1231	BSc
679	233-1231	MSc

Exercise

- Convert below table in to 1NF

Emp_ID	Emp_Name	Qualifications	Extension
101	Ann	BSc, MSc	1226
102	Tom	BSc	1269
103	Emmy	BSc, MSc, PhD	1358
104	James	BSc	1457

Second Normal Form

- For a table to be in the Second Normal Form,
  - It should be in the First Normal form.
  - And, it should not have **Partial Dependency**.

Dependency

- Determination
  - “A determines B” indicates that if you know the value of attribute A, you can look up (determine) the value of attribute B.
    - A→B

- Functional Dependency:
  - The attribute B is functionally dependent on the attribute A if each value in column A determines one and only one value in column B.

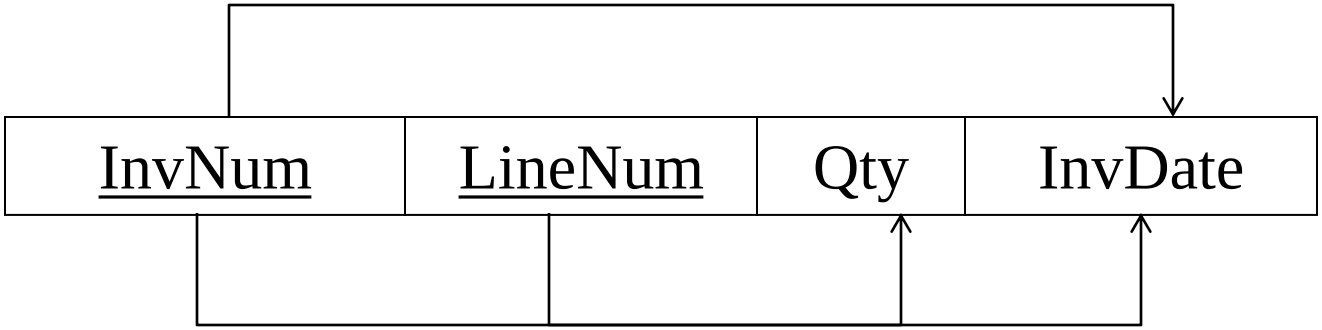
subject_id	subject_name	teacher
1	Java	Java Teacher
2	C++	C++ Teacher
3	Php	Php Teacher

Dependency contd.

- Fully functional dependency (composite key)
  - If attribute B is functionally dependent on a composite key A but not on any Subset of that composite key, the attribute B is fully functionally dependent on A.

Partial dependency

A **partial dependency** exists when an attribute B is functionally dependent on an attribute A, and A is a component of a multipart candidate key.

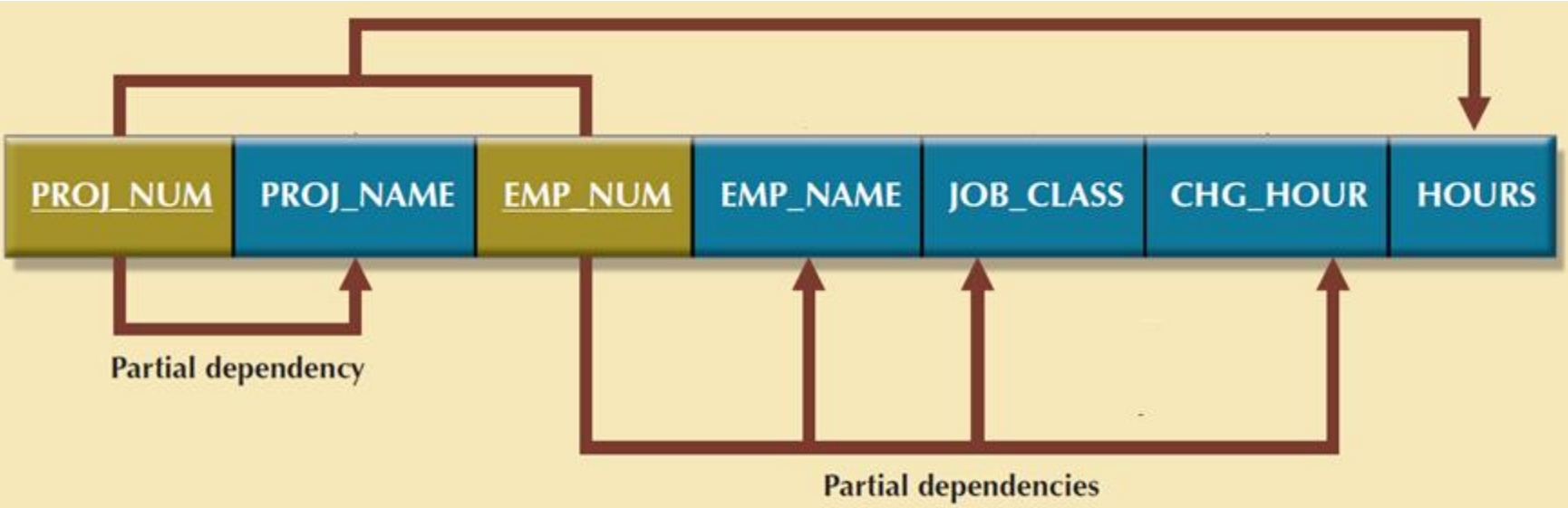


Candidate keys: {InvNum, LineNum} InvDate is *partially dependent* on {InvNum, LineNum} as InvNum is a determinant of InvDate and InvNum is part of a candidate key

Sample employee working table for identifying partial dependency

Project No	Employee No	Employee Name	Employee Project Work Hours
11	110	Michele	10
12	120	Tina	18
13	110	Michele	15
14	115	Gerard	12
15	121	Neil	14

First Normal Form (1NF) Dependency Diagram



1NF (PROJ_NUM, EMP_NUM, PROJ_NAME, EMP_NAME, JOB_CLASS, CHG_HOURS, HOURS)

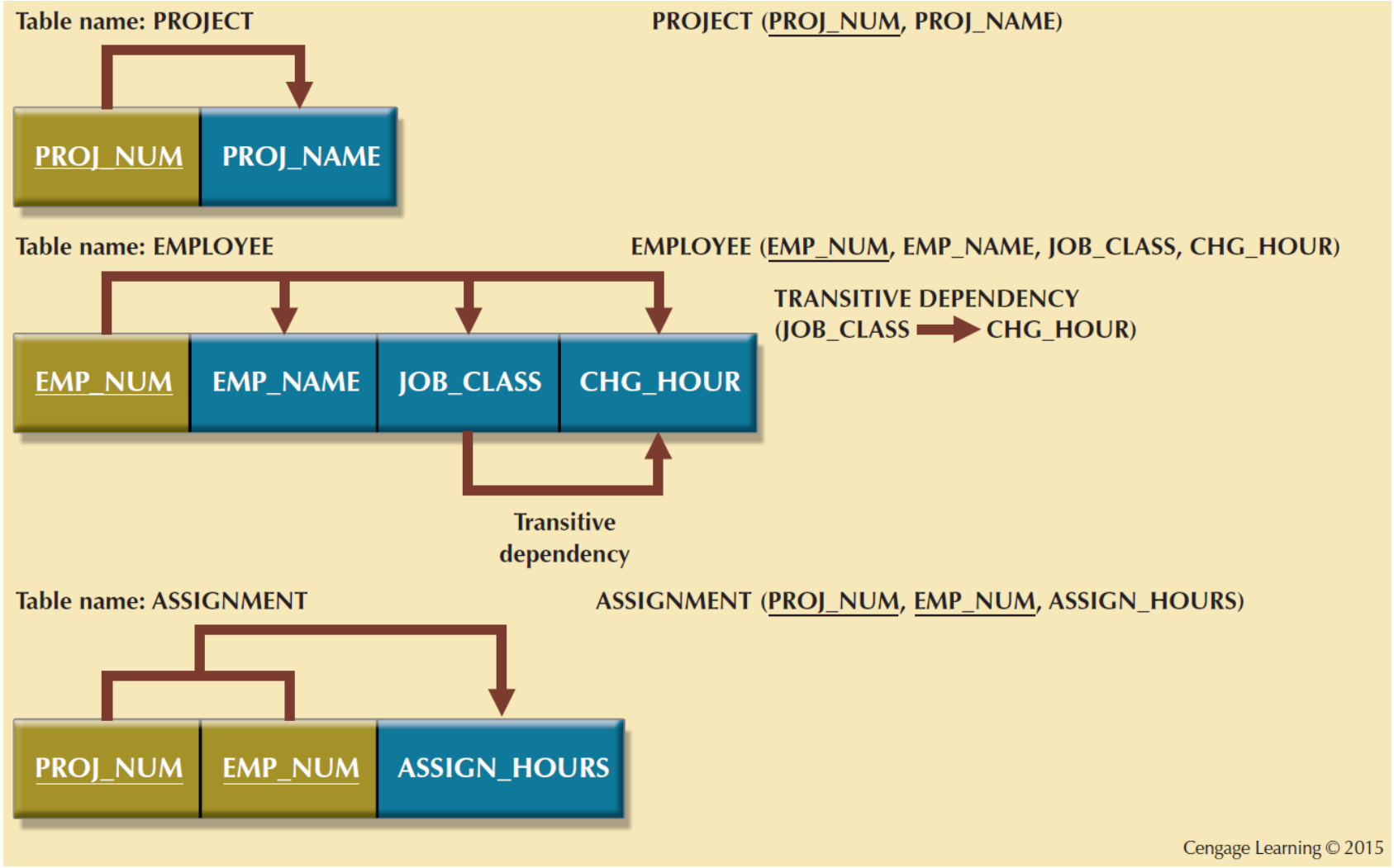
PARTIAL DEPENDENCIES:

(PROJ_NUM → PROJ_NAME)

(EMP_NUM → EMP_NAME, JOB_CLASS, CHG_HOUR)

(EMP_NUM, PROJ_NAME → HOURS)

Second Normal Form (2NF) Conversion Results



Exercise

- Determine whether there are any partial dependencies in the below table which is in 1NF and convert it into 2NF.

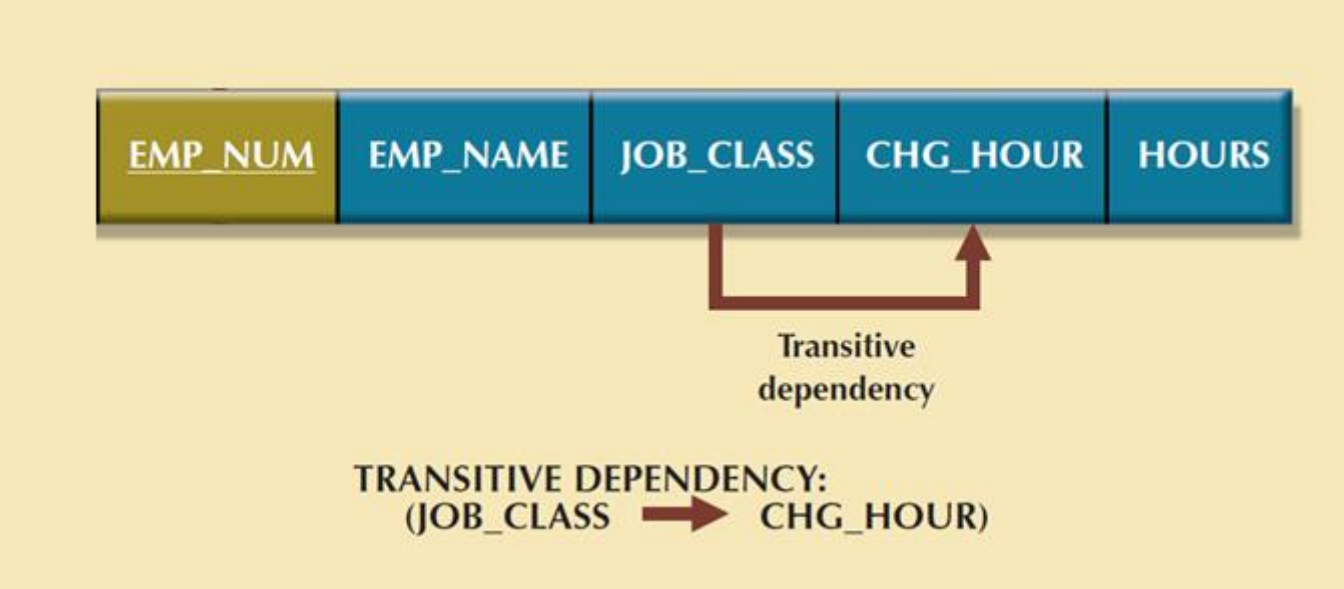
Student_ID	Subject_ID	Subject_Name	Marks	Teacher
10	1	Java	70	Java teacher
10	2	C++	75	C++ teacher
11	1	Java	80	Java teacher
12	3	PHP	85	PHP teacher

Conversion to Third Normal Form

- Steps
  - Make new tables to eliminate transitive dependencies
    - **Determinant**: Any attribute whose value determines other values within a row
  - Reassign corresponding dependent attributes
- Table is in 3NF when it:
  - Is in 2NF
  - Contains no **transitive dependencies**

Transitive dependency

Occurs when a non-prime attribute depends on other non-prime attributes rather than depending upon the prime attributes or primary key.



Third Normal Form (3NF) Conversion Results

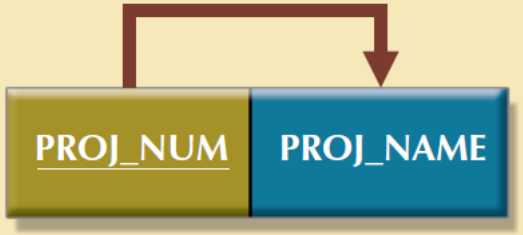


Table name: PROJECT

PROJECT (PROJ_NUM, PROJ_NAME)

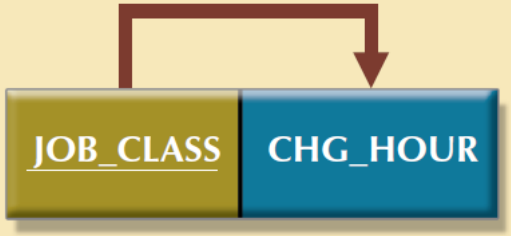


Table name: JOB

JOB (JOB_CLASS, CHG_HOUR)

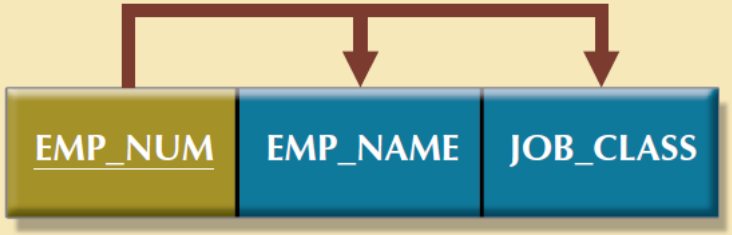


Table name: EMPLOYEE

EMPLOYEE (EMP_NUM, EMP_NAME, JOB_CLASS)

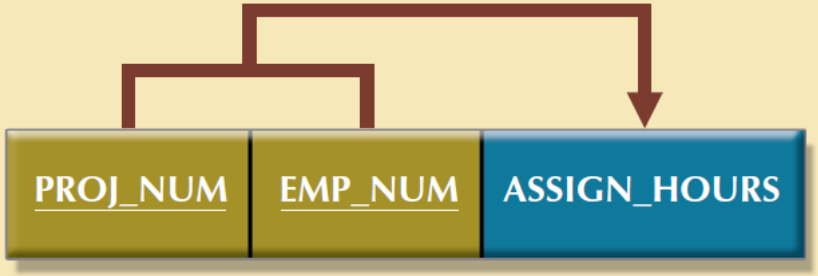


Table name: ASSIGNMENT

ASSIGNMENT (PROJ_NUM, EMP_NUM, ASSIGN_HOURS)

Exercise

- Determine whether there are any transitive dependencies in the below table which is in 2NF and convert it into 3NF.

UnitCode	UnitName	CourseCode	CourseName
101	Database	COMP2009	Computing
131	Network	COMP2009	Computing
112	OS	COMP2009	Computing
123	Project	BIS3001	Business Management

End of the session