

# Module 8

## Normal Forms

© Louis D. Nel 1996    95.305 Introduction to Databases    Normal Forms 8 - 1

### 95.305

#### Objectives

- Learn the Normal Forms that are based on keys and functional dependencies
- Learn how to apply the normal forms for decompose relational schemas

© Louis D. Nel 1996    95.305 Introduction to Databases    Normal Forms 8 - 2

## 95.305

### Topics

- **Normal Forms**
  - 1st Normal Form
  - 2nd and 3rd Normal Forms
  - Boyce-Codd Normal Form

### Case Study [Rob & Coronel]

- Construction company manages several building projects
- Each project has project number, name, employees assigned to it, ...
- Employees have an employee number, name, job classification (engineer, computer tech., ...)
- Clients are billed based on hours spent on a project
- Employee's billing rate depends on their job classification
- A project report showing project, hours billed, total charges is produced periodically

## Sample Report

Proj Number	Project Name	Employee Number	Employee Name	Job Class	Chg/Hour	Hours Billed	Total Charges
1	Hurricane	101	John News	Elect. Eng.	65	13	845
		102	David Senior	Comm. Tech.	60	16	960
		104	Anne Ramoras	Comm. Tech.	60	19	1,140
Subtotals							2,245
2	Coast	101	John News	Elect. Eng.	65	15	975
		103	June Arbough	Biol. Eng.	55	17	935
Subtotals							1,910
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18	1,080
		102	David Senior	Comm. Tech.	60	14	840
Subtotals							1,920
Total							6,775

## Initial Solution -Organize Relation Like the Report

P_NO	P_NAME	E_NO	E_NAME	JOB_CLASS	CHG_HOUR	HOURS
1	Hurricane	101	John News	Elect. Eng.	65	13
		102	David Senior	Comm. Tech.	60	16
		104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
		103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
		102	David Senior	Comm. Tech.	60	14

- Is this a good relation?

## Problems

P_NO	P_NAME	E_NO	E_NAME	JOB_CLASS	CHG_HOUR	HOURS
1	Hurricane	101	John News	Elect. Eng.	65	13
		102	David Senior	Comm. Tech.	60	16
		104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
		103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
		102	David Senior	Comm. Tech.	60	14

- Project number, P\_NO, is apparently intended to be a key, or part of a key, but has NULL values
- Table has redundancies
- Table invites inconsistencies (Elect. Eng. may be entered as El. Eng. or EE ...)

## Problems: Update Anomalies

P_NO	P_NAME	E_NO	E_NAME	JOB_CLASS	CHG_HOUR	HOURS
1	Hurricane	101	John News	Elect. Eng.	65	13
		102	David Senior	Comm. Tech.	60	16
		104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
		103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
		102	David Senior	Comm. Tech.	60	14

- Modifying the JOB\_CLASS of an employee probably requires many alterations (update anomaly)
- Employee can't exist if not assigned to a project (insertion anomaly)
- If employee 101 quits, many deletions must be made, which is a problem if emp. 101 is the only one on a project (deletion anomaly)

## Problems: Redundancy

P_NO	P_NAME	E_NO	E_NAME	JOB_CLASS	CHG_HOUR	HOURS
1	Hurricane	101	John News	Elect. Eng.	65	13
		102	David Senior	Comm. Tech.	60	16
		104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
		103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
		102	David Senior	Comm. Tech.	60	14

- Every time a new employee is added to a project the hourly rate for their job classification is repeated
- Info. about employees appears many times, what if they become inconsistent
- However, the report is easy to generate from this table!

## Fixing the Problems

P_NO	P_NAME	E_NO	E_NAME	JOB_CLASS	CHG_HOUR	HOURS
1	Hurricane	101	John News	Elect. Eng.	65	13
		102	David Senior	Comm. Tech.	60	16
		104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
		103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
		102	David Senior	Comm. Tech.	60	14

- Fix the table so it is a legitimate relation
- Probably involves breaking the table up into smaller relations -but how
- What are the guide-lines or theory for decomposing a table like this so it makes sense

## Normalization

- **Process of decomposing relational schemas by breaking up their attributes into smaller relations so that the smaller relations will have desirable properties**
- **Among the desirable objectives is having relations which are free of update anomalies**

## Normalization Theory

- **Normalization Theory provides:**
- **A formal framework for analysing relations based on their keys and on their functional dependencies**
- **A series of tests that can be carried out on individual relation schemas so the database can be normalized to the desired degree**
- **When a relations fails a test it is decomposed into smaller relations that pass the test.**

## Normalization Process

- The Normalization process must also ensure:
- That no spurious tuples will be created by ensuring lossless joins
- The functional dependencies must be preserved (each dependency must be reflected in one to the decomposed relations)
- These issues will be deferred for now

## Normal Forms

- Based on Keys and Functional Dependencies

1st Normal Form  
2nd Normal Form  
3rd Normal Form  
Boyce-Codd Normal Form

- Based on Multi-valued Dependencies etc.

4th Normal Form  
5th Normal Form

## Keys

- Relation is a set of tuples, so tuples must be distinct
- $R(A_1, A_2, \dots, A_n) \quad S = \{A_1, A_2, \dots, A_n\}$

$u[S] \neq v[S]$  for any tuples  $u, v$  in  $r(R)$  ... (1)

- Any subset of  $S$  for which (1) holds is a superkey
- A key is a minimal superkey
- Relation may have more than one key; each is called a candidate key
- A primary key is a designated key (usually underlined)

## Prime Attributes

- Prime Attribute:

If  $R = (A_1, A_2, \dots, A_n)$  is a table, attribute  $A_i$  is prime if there exists a key  $K$  of  $R$  such that  $A_i$  is an element of  $K$ .

- If an attribute is not prime it is called non-prime.



## 95.305

### Topics

- Normal Forms
- 1st Normal Form
- 2nd and 3rd Normal Forms
- Boyce-Codd Normal Form

### Fixing the Problems: Repeating Groups

P_NO	P_NAME	E_NO	E_NAME	JOB_CLASS	CHG_HOUR	HOURS
1	Hurricane	101	John News	Elect. Eng.	65	13
		102	David Senior	Comm. Tech.	60	16
		104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
		103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
		102	David Senior	Comm. Tech.	60	14

- Table cannot be a relation yet because it does not appear to have a key
- It has repeating groups: three groups of information grouped around P\_NO, P\_NAME
- Solution: define a key, and replace NULLs to avoid repeating groups (at least the NULL values of prime attributes must be removed)

## Fixing the Problems: Repeating Groups

<u>P_NO</u>	<u>P_NAME</u>	<u>E_NO</u>	<u>E_NAME</u>	<u>JOB_CLASS</u>	<u>CHG_HOUR</u>	<u>HOURS</u>
1	Hurricane	101	John News	Elect. Eng.	65	13
1	Hurricane	102	David Senior	Comm. Tech.	60	16
1	Hurricane	104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
2	Coast	103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
3	Satellite	102	David Senior	Comm. Tech.	60	14

- Table is now a relation
- Key: P\_NO, E\_NO
- Table still has some problems but its a good start
- Notice the problems with redundancy still exists
- The functional dependencies will provide the clues for what to do next

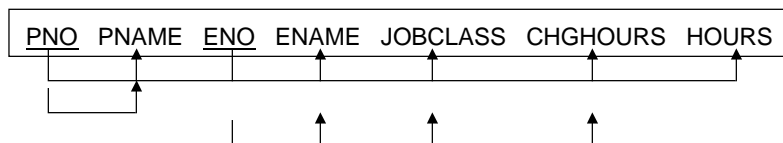
© Louis D. Nel 1996

95.305 Introduction to Databases

Normal Forms 8 - 19

## Dependency Diagram

<u>P_NO</u>	<u>P_NAME</u>	<u>E_NO</u>	<u>E_NAME</u>	<u>JOB_CLASS</u>	<u>CHG_HOUR</u>	<u>HOURS</u>
1	Hurricane	101	John News	Elect. Eng.	65	13
1	Hurricane	102	David Senior	Comm. Tech.	60	16
1	Hurricane	104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
2	Coast	103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
3	Satellite	102	David Senior	Comm. Tech.	60	14



PNO, ENO -> PNAME, ENAME, JOBCLASS, CHG\_HOUR, HOURS

PNO -> PNAME

ENO -> ENAME, JOBCLASS, CHG\_HOUR

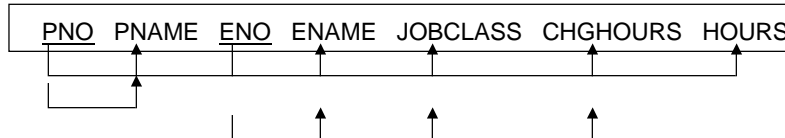
© Louis D. Nel 1996

95.305 Introduction to Databases

Normal Forms 8 - 20

## Dependency Diagram

<u>P_NO</u>	<u>P_NAME</u>	<u>E_NO</u>	<u>E_NAME</u>	<u>JOB_CLASS</u>	<u>CHG_HOUR</u>	<u>HOURS</u>
1	Hurricane	101	John News	Elect. Eng.	65	13
1	Hurricane	102	David Senior	Comm. Tech.	60	16
1	Hurricane	104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
2	Coast	103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
3	Satellite	102	David Senior	Comm. Tech.	60	14



- Every non-prime attribute is functionally determined by the key -good
- Some non-prime attribute are functionally determined by only part of the key -not so good (leads to redundancy)

## First Normal Form

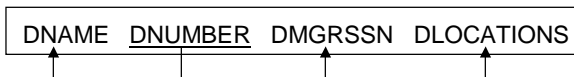
- Defn: [Rob & Coronel]
- A relation is in First Normal Form (1NF) if:
  - All the key attributes are defined
  - There are no repeating groups in the table. That is, each row-column intersection can contain one and only one value, rather than a set of values
  - All attributes are dependent on the primary key

## First Normal Form

- Defn: [Elmasri & Navathe]
- First Normal Form is now considered part of the basic definition of a relation.
- The relation must have a key
- 1NF requires that all attribute domains be atomic, and attribute values not be multi-valued
- In other words, does not allow “relations within relations”

## Example -not in 1NF

DEPARTMENT			
DNAME	DNUMBER	DMGRSSN	DLOCATIONS
Research	5	333445555	{Bellaire, Sugarland, Houston}
Administration	4	987654321	{Stafford}
Headquarters	1	888665555	{Houston}



- Department locations are not determined by the key DNUMBER (they are multi-valued)
- Table is not in 1NF

## Solutions which are in 1NF

DEPARTMENT			
DNAME	<u>DNUMBER</u>	DMGRSSN	<u>DLOCATIONS</u>
Research	5	333445555	Bellaire
Research	5	333445555	Sugarland
Research	5	333445555	Houston
Administration	4	987654321	Stafford
Headauarters	1	888665555	Houston

DEPARTMENT		
DNAME	<u>DNUMBER</u>	DMGRSSN
Research	5	333445555
Administration	4	987654321
Headauarters	1	888665555

DEPT_LOCATIONS	
<u>DNUMBER</u>	<u>DLOCATIONS</u>
5	Bellaire
5	Sugarland
5	Houston
4	Stafford
1	Houston

- Which solution is better
- Does it matter?

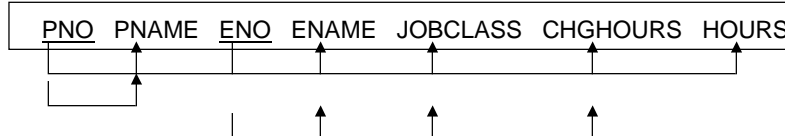
## 95.305

### Topics

- Normal Forms
- 1st Normal Form
- 2nd and 3rd Normal Forms
- Boyce-Codd Normal Form

### 1st Normal Form table still has problems

<u>P_NO</u>	<u>P_NAME</u>	<u>E_NO</u>	<u>E_NAME</u>	<u>JOB_CLASS</u>	<u>CHG_HOUR</u>	<u>HOURS</u>
1	Hurricane	101	John News	Elect. Eng.	65	13
1	Hurricane	102	David Senior	Comm. Tech.	60	16
1	Hurricane	104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
2	Coast	103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
3	Satellite	102	David Senior	Comm. Tech.	60	14



- There is still a problem with redundancy
- Clue: some non-prime attributes depend on a prime attribute, but not the whole key
- To eliminate redundancy, and achieve 2NF, we must decompose based on these partial dependencies

### Trivial Dependencies and Closure

- Trivial Dependency: A dependency  $X \rightarrow Y$  is trivial if  $Y$  is a subset of  $X$ .
- Closure of  $F$ : if  $F$  is a set of functional dependencies, the closure of  $F$ , denoted  $F^+$ , is  $\{X \rightarrow Y \mid F \text{ logically implies } X \rightarrow Y\}$ .

## Partial Dependencies

- **Partial Dependency:**  
Suppose  $X$  is a key of table  $R$  and  $Y$  is a proper subset of  $X$ , and  $A$  is an attribute not in  $Y$ .  
Then  $Y \rightarrow A$  is a partial dependency.
- e.g.  
 $\{SSN, PNUMBER\} \rightarrow HOURS$  is a full dependency because  $HOURS$  is neither dependent on  $SSN$  or  $PNUMBER$  alone  
  
In table  $\{SSN, PNUMBER, ENAME, HOURS\}$   
 $SSN \rightarrow ENAME$  is a partial dependency

## 2nd Normal Form

- Defn: [Rob & Coronel]
- A table is in 2NF if
- It is in 1NF *and*
- It includes no partial dependencies; that is, no attribute is dependent on only a portion of the primary key

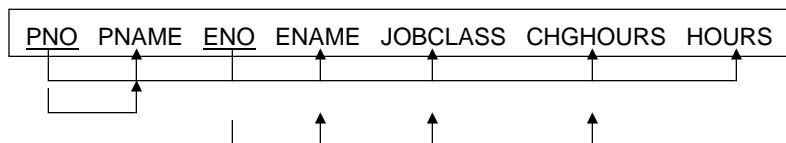
## 2nd Normal Form

- **2<sup>nd</sup> Normal Form:**

A table R with associated functional dependencies F is in 2<sup>nd</sup> normal form if F+ contains no partial dependencies  $Y \rightarrow A$  where A is non-prime.

## 1st Normal Form table

<u>P_NO</u>	<u>P_NAME</u>	<u>E_NO</u>	<u>E_NAME</u>	<u>JOB_CLASS</u>	<u>CHG_HOUR</u>	<u>HOURS</u>
1	Hurricane	101	John News	Elect. Eng.	65	13
1	Hurricane	102	David Senior	Comm. Tech.	60	16
1	Hurricane	104	Anne Ramoras	Comm. Tech.	60	19
2	Coast	101	John News	Elect. Eng.	65	15
2	Coast	103	June Arbough	Biol. Eng.	55	17
3	Satellite	104	Anne Romoras	Comm. Tech.	60	18
3	Satellite	102	David Senior	Comm. Tech.	60	14



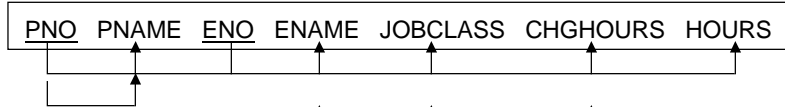
$PNO, ENO \rightarrow PNAME, ENAME, JOBCLASS, CHG_HOUR, HOURS$

$PNO \rightarrow PNAME$

$ENO \rightarrow ENAME, JOBCLASS, CHG_HOUR$



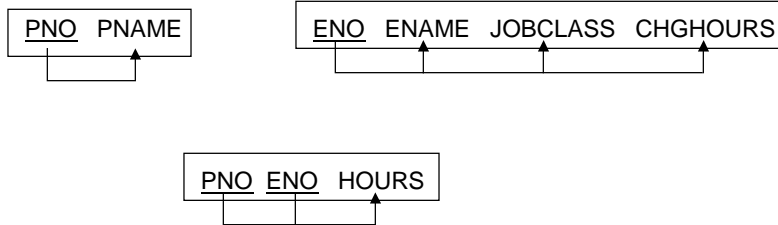
## Decomposition to 2NF



$PNO, ENO \rightarrow PNAME, ENAME, JOBCLASS, CHGHOURLS, HOURS$

$PNO \rightarrow PNAME$

$ENO \rightarrow ENAME, JOBCLASS, CHGHOURLS$



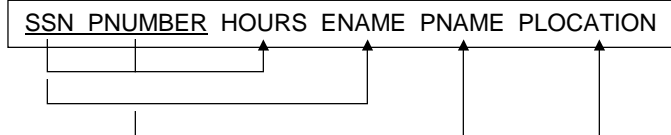
© Louis D. Nel 1996

95.305 Introduction to Databases

Normal Forms 8 - 33

## example [Elmasri & Navathe]

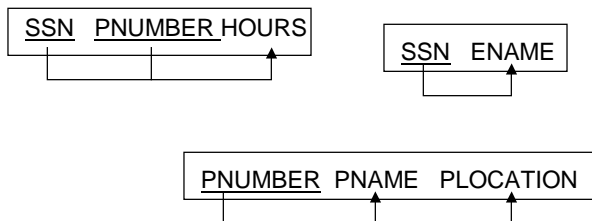
EMP PROJ



$SSN \rightarrow ENAME$

$PNUMBER \rightarrow \{PNAME, PLOCATION\}$

$\{SSN, PNUMBER\} \rightarrow HOURS$



© Louis D. Nel 1996

95.305 Introduction to Databases

Normal Forms 8 - 34

## Tables in 2NF

PROJECT	
P_NO	P_NAME
1	Hurricane
2	Coast
3	Satellite

EMPLOYEE			
E_NO	E_NAME	JOB_CLASS	CHG_HOUR
101	John News	Elect. Eng.	65
102	David Senior	Comm. Tech.	60
104	Anne Ramoras	Comm. Tech.	60
103	June Arbough	Biol. Eng.	55

ASSIGN		
P_NO	E_NO	HOURS
1	101	13
1	102	16
1	104	19
2	101	15
2	103	17
3	104	18
3	102	14

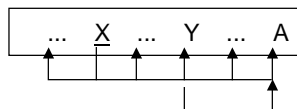
- Unfortunately there is still some repetition in the **EMPLOYEE** table
- Clue: **JOB\_CLASS** -> **CHG\_HOURS** but **JOB\_CLASS** is not a prime attribute (Transitive Dependency)

© Louis D. Nel 1996    95.305 Introduction to Databases    Normal Forms 8 - 35

## Transitive Dependencies

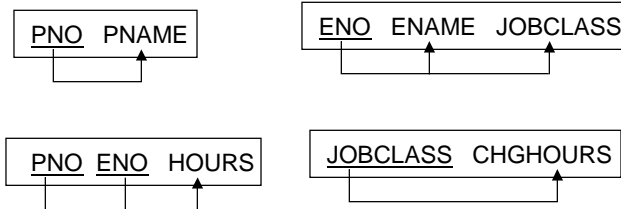
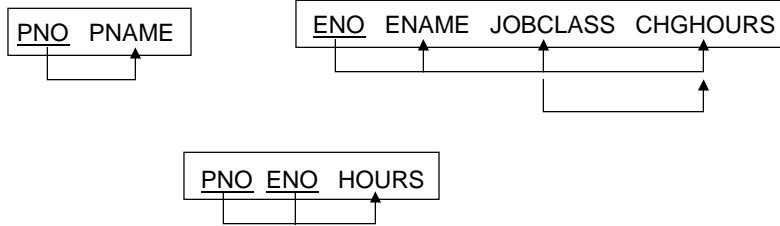
- Transitive Dependency:

Let Y be a set of attributes from table R and A be an attribute not contained in Y. The functional dependency  $Y \rightarrow A$  is a transitive dependency if Y is neither a superkey of R nor a proper subset of a key of R.



© Louis D. Nel 1996    95.305 Introduction to Databases    Normal Forms 8 - 36

## Conversion to 3NF



© Louis D. Nel 1996

95.305 Introduction to Databases

Normal Forms 8 - 37

## Tables in 3NF

PROJECT	
P_NO	P_NAME
1	Hurricane
2	Coast
3	Satellite

EMPLOYEE		
E_NO	E_NAME	JOB_CLASS
101	John News	Elect. Eng.
102	David Senior	Comm. Tech.
104	Anne Ramoras	Comm. Tech.
103	June Arbough	Biol. Eng.

ASSIGN		
P_NO	E_NO	HOURS
1	101	13
1	102	16
1	104	19
2	101	15
2	103	17
3	104	18
3	102	14

JOB	
JOB_CLASS	CHG_HOUR
Elect. Eng.	65
Comm. Tech.	60
Biol. Eng.	55

© Louis D. Nel 1996

95.305 Introduction to Databases

Normal Forms 8 - 38

### 3rd Normal Form

- Defn: [Rob & Coronel]
- A table is in 3NF if
- It is in 2NF *and*
- It contains no transitive dependencies

### ...Tables in 3NF

EMPLOYEE		
<u>E_NO</u>	E_NAME	JOB_CLASS
101	John News	Elect. Eng.
102	David Senior	Comm. Tech.
104	Anne Ramoras	Comm. Tech.
103	June Arbough	Biol. Eng.

- **Potential Problem:** as the number of employees gets large it is likely a Job class may be entered incorrectly (e.g. El. Eng. instead of Elect. Eng.)
- **Suggestion:** create a JOB\_CODE to act as a primary key in the JOB table and a foreign key in the EMPLOYEE table

## ...Tables in 3NF (Better)

PROJECT	
P_NO	P_NAME
1	Hurricane
2	Coast
3	Satellite

EMPLOYEE		
E_NO	E_NAME	JOB_CODE
101	John News	502
102	David Senior	501
104	Anne Ramoras	500
103	June Arbough	501

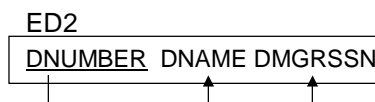
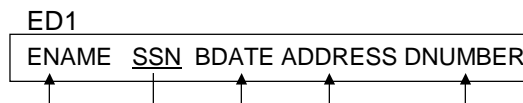
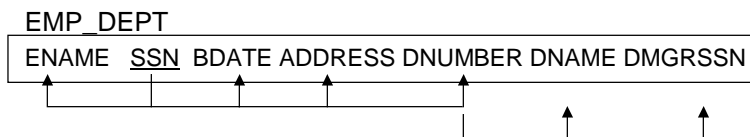
ASSIGN		
P_NO	E_NO	HOURS
1	101	13
1	102	16
1	104	19
2	101	15
2	103	17
3	104	18
3	102	14

JOB		
JOB_CODE	JOB_CLASS	CHG_HOUR
500	Biol. Eng.	55
501	Comm. Tech.	60
502	Elect. Eng.	65

- Notice: ASSIGN is probably the most active table, yet it requires only the P\_NO, E\_NO and HOURS be entered
- This will help data entry

© Louis D. Nel 1996    95.305 Introduction to Databases    Normal Forms 8 - 41

## example [Elmasri & Navathe]



© Louis D. Nel 1996    95.305 Introduction to Databases    Normal Forms 8 - 42

## Definition of 3NF

- **3<sup>rd</sup> Normal Form:**

A table, with dependencies  $F$ , is in 3<sup>rd</sup> normal form if it is in 2<sup>nd</sup> normal form and if  $F^+$  contains no transitive dependencies  $Y \rightarrow A$  where  $A$  is non-prime.

(Equivalently, a table is in 3<sup>rd</sup> normal form if, for each non-trivial dependency  $Y \rightarrow A$ ,  $Y$  is a superkey or  $A$  is prime)

## 95.305

### Topics

- Normal Forms
- 1st Normal Form
- 2nd and 3rd Normal Forms
- Boyce-Codd Normal Form

### Example Scenario

- Bicycle retailer maintains two stores: Hull, QE and Ottawa ON.
- Each store assembles bicycles from different components and bicycle frames classified by manufacturer

#### Component

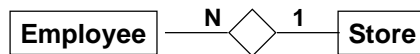
Shimano  
Campagnolo  
Ritchie  
...

#### Frame

Marinoni  
Specialized  
Eclipse  
...

### ...Example Scenario

- Each store has many employees, but an employee works for only one store

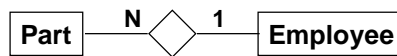


- A part (component group or frame) can be in either store's inventory



### ...Example Scenario

- At each store employees specialize in parts by manufacturer. One employee may manage many parts, but a part is managed by a particular employee.
- e.g. Employee John manages the Shimano components in the Ottawa store, whereas Sue manages all the Marinoni and Specialized frames in the Ottawa store



### Sample Data

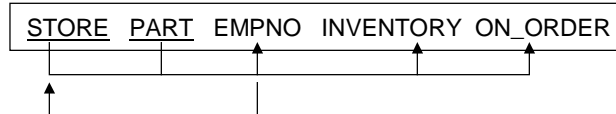
<u>STORE</u>	<u>PART</u>	<u>EMPNO</u>	<u>INVENTORY ON_ORDER</u>	
Ottawa	Mar100	101	10	0
Ottawa	Spe100	103	15	0
Ottawa	Spe102	103	10	5
Ottawa	Ecl300	101	12	7
Hull	Mar100	102	3	6
Hull	Ecl300	102	14	0
Ottawa	Shi105	104	20	0
Ottawa	Shi 600	104	12	10
Hull	Shi 105	105	10	0
Hull	Shi 600	106	11	0
Hull	Cmp100	106	3	2

- What would the functional dependencies look like for this table?



## Sample Data

<u>STORE</u>	<u>PART</u>	EMPNO	INVENTORY	ON_ORDER
Ottawa	Mar100	101	10	0
Ottawa	Spe100	103	15	0
Ottawa	Spe102	103	10	5
Ottawa	Ecl300	101	12	7
Hull	Mar100	102	3	6
Hull	Ecl300	102	14	0
Ottawa	Shi105	104	20	0
Ottawa	Shi 600	104	12	10
Hull	Shi 105	105	10	0
Hull	Shi 600	106	11	0
Hull	Cmp100	106	3	2



<u>STORE</u>	<u>PART</u>	EMPNO	INVENTORY	ON_ORDER
Ottawa	Mar100	101	10	0
Ottawa	Spe100	103	15	0
Ottawa	Spe102	103	10	5
Ottawa	Ecl300	101	12	7
Hull	Mar100	102	3	6
Hull	Ecl300	102	14	0
Ottawa	Shi105	104	20	0
Ottawa	Shi 600	104	12	10
Hull	Shi 105	105	10	0
Hull	Shi 600	106	11	0
Hull	Cmp100	106	3	2

- Which Normal Forms does this table adhere to?

### Sample Data -Normal Forms

<u>STORE</u>	<u>PART</u>	<u>EMPNO</u>	<u>INVENTORY ON_ORDER</u>	
Ottawa	Mar100	101	10	0
Ottawa	Spe100	103	15	0
Ottawa	Spe102	103	10	5
Ottawa	Ecl300	101	12	7
Hull	Mar100	102	3	6
Hull	Ecl300	102	14	0
Ottawa	Shi105	104	20	0
Ottawa	Shi 600	104	12	10
Hull	Shi 105	105	10	0
Hull	Shi 600	106	11	0
Hull	Cmp100	106	3	2

- **1NF -because it is a relation with a key**
- **2NF -because it has no partial dependencies (attribute dependent on part of a key)**
- **3NF -because it has no transitive dependencies (non-prime attribute dependent on another non-prime attribute)**

© Louis D. Nel 1996    95.305 Introduction to Databases    Normal Forms 8 - 51

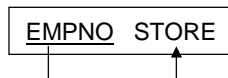
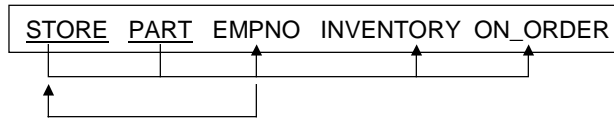
### So what's the problem?

<u>STORE</u>	<u>PART</u>	<u>EMPNO</u>	<u>INVENTORY ON_ORDER</u>	
Ottawa	Mar100	101	10	0
Ottawa	Spe100	103	15	0
Ottawa	Spe102	103	10	5
Ottawa	Ecl300	101	12	7
Hull	Mar100	102	3	6
Hull	Ecl300	102	14	0
Ottawa	Shi105	104	20	0
Ottawa	Shi 600	104	12	10
Hull	Shi 105	105	10	0
Hull	Shi 600	106	11	0
Hull	Cmp100	106	3	2

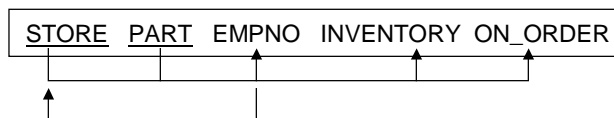
- **Because EMPNO -> STORE, the fact that an employee works at a particular store is repeated many times (redundancy)**
- **But, EMPNO -> STORE is neither a full, partial, nor transitive dependency**
- **We can eliminate this with more decomposition, but we must go beyond 3NF**

© Louis D. Nel 1996    95.305 Introduction to Databases    Normal Forms 8 - 52

## Decompose further to BCNF

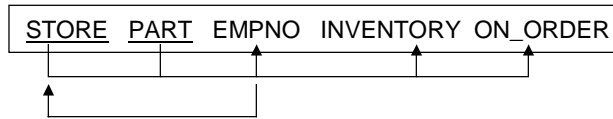


## Boyce-Codd Normal Form



- So what kind of dependency is EMPNO → STORE
- It's not a full, partial, nor transitive dependency
- It's a non-prime attribute functionally determining a prime attribute

### ...Boyce-Codd Normal Form



- Notice that we probably could have used PART, EMPNO as a key instead of STORE, PART
- The table has more than one candidate key
- This looks like a partial dependency -if we had used the key PART,EMPNO instead
- CLUE: the determinant EMPNO is not a candidate key in the relation

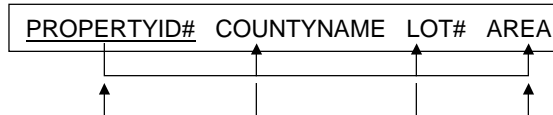
(A determinant is any attribute whose value determines others)

### Boyce-Codd Normal Form Definition

[Rob & Coronel]

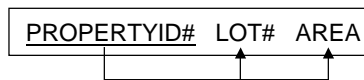
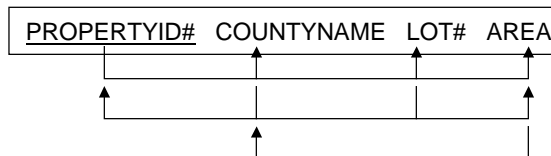
- A table is in BCNF if every determinant in that table is a candidate key.
- If a table contains only one candidate key, 3NF and BCNF are equivalent

### Example from [Elmasri & Navathe 12.5]



- Suppose there are only two counties involved: “Marion County” and “Liberty County”
- Lots in Marion County only have areas of 0.5, 0.6, 0.7, 0.8, 0.9 or 1.0 acres
- Lots in Liberty County only have areas of 1.1, 1.2, ..., 1.9, 2.0 acres
- So AREA → COUNTYNAM
- Suppose also there are many Lots (much data, so redundancy is important)

### Example from [Elmasri & Navathe 12.5]



### Definition of Boyce-Codd Normal Form

- **Boyce-Codd Normal Form:**

A table, with dependencies  $F$ , is in BCNF if  $F^+$  contains no partial or transitive dependencies.

(Equivalently, a table is in BCNF if the left side of each non-trivial dependency in  $F^+$  is a superkey.)

### Normalization during E-R modelling

- A good database designer won't first create a "bad" set of tables and then normalize
- Normalization is taken into account at the E-R data modelling stage

## Employee-Project example [Rob & Coronel]

### Requirements Scenario

- The company manages many projects
- Each project requires the services of many employees
- An employee may be assigned to several different projects
- To be considered an employee, a person must be assigned to at least one project
- Each employee has a (single) primary job classification; this job classification determines their hourly billing rate
- Many employees may have the same job classification

© Louis D. Nel 1996

95.305 Introduction to Databases

Normal Forms 8 - 61

### First Iteration



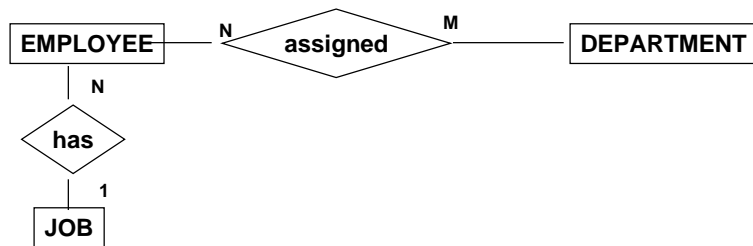
- **Two entities initially defined**
- PROJECT(PNO, PNAME)  
EMPLOYEE (ENO, ENAME, JOBCLASS, CHG HOUR)
- PROJECT is in 3NF, no need to change it.
- EMPLOYEE is not in 3NF, the following apply  
JOBCLASS -> CHG HOUR (transitive dependency)

© Louis D. Nel 1996

95.305 Introduction to Databases

Normal Forms 8 - 62

## Second Iteration -Normalization yields a new entity



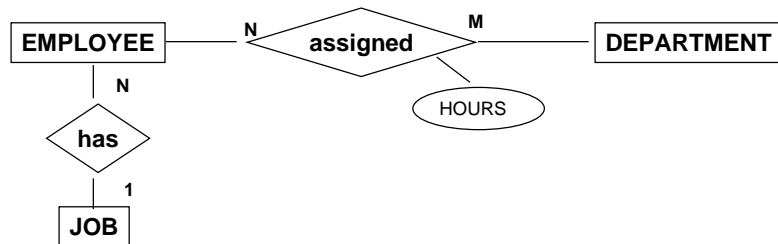
- PROJECT(PNO, PNAME)  
EMPLOYEE (ENO, ENAME, JOBCLASS)  
JOB (JOBCLASS, CHG HOUR)
- The assigned relationship will have to be implemented in a table (ASSIGN)
- The hours worked by an employee on a project is not yet represented

© Louis D. Nel 1996

95.305 Introduction to Databases

Normal Forms 8 - 63

## Second Iteration -Normalization yields a new entity



- PROJECT(PNO, PNAME)  
EMPLOYEE (ENO, ENAME, JOBCLASS)  
JOB (JOBCLASS, CHG HOUR)  
ASSIGN (PNO, ENO, HOURS)

- All of the relations in this design are now in 3NF

© Louis D. Nel 1996

95.305 Introduction to Databases

Normal Forms 8 - 64



## Normalization and E-R modeling

### Summary

- Don't necessarily separate E-R modeling and Normalization
- They are thought about together when developing the data model

## Definitions