

**“HỌC ĐỂ LÀM VIỆC”**

# ORACLE SERVER

**GV: Bùi Quang Đăng**

# Contents

1

Mô hình cơ sở dữ liệu quan hệ

2

Tại sao phải chuẩn hóa dữ liệu

3

Chuẩn hóa dữ liệu (Normal Form)

- Cơ sở dữ liệu

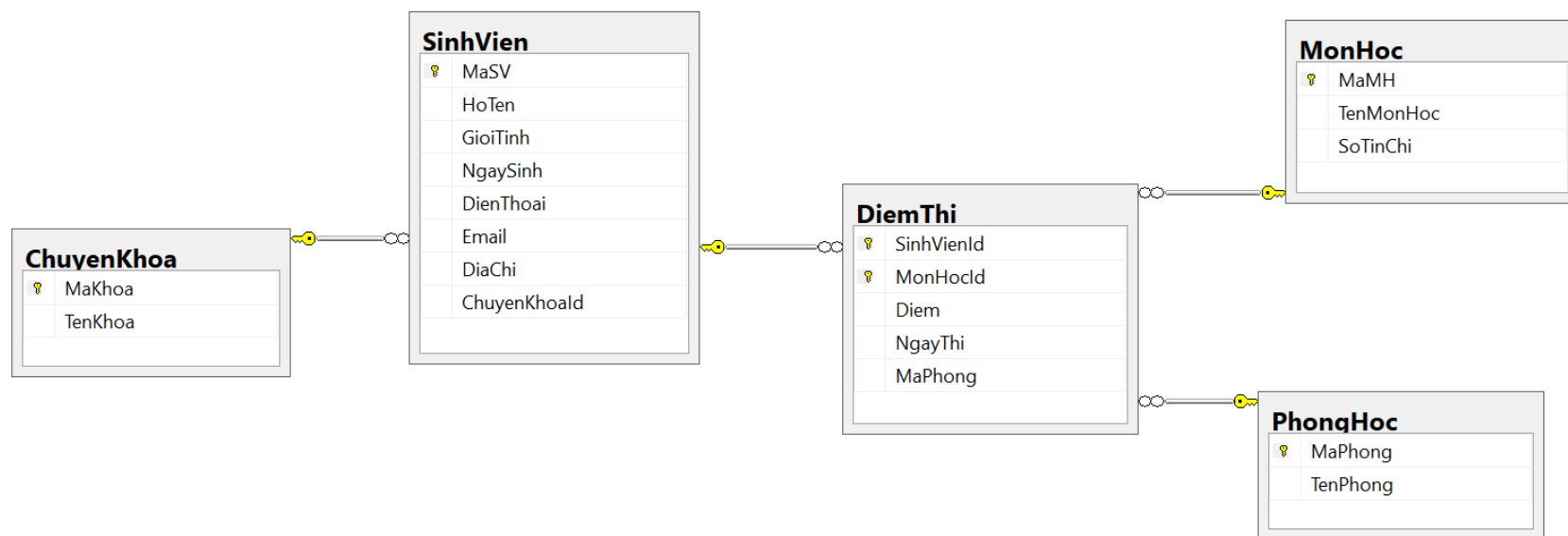


# MÔ HÌNH CƠ SỞ DỮ LIỆU QUAN HỆ

## ❖ Mô hình cơ sở dữ liệu quan hệ

- Mô hình CSDL quan hệ lần đầu tiên được E.F.Codd và tiếp sau đó được công ty IBM giới thiệu vào năm 1970. Ngày nay, hầu hết các tổ chức đã áp dụng CSDL quan hệ để quản lý dữ liệu trong đơn vị mình.
- **Mô hình cơ sở dữ liệu quan hệ:** Là cách thức biểu diễn dữ liệu dưới dạng bảng hay còn gọi là quan hệ, mô hình được xây dựng dựa trên cơ sở lý thuyết đại số quan hệ.
- **Cấu trúc dữ liệu:** Dữ liệu được tổ chức dưới dạng quan hệ hay còn gọi là bảng.

## ❖ Mô hình cơ sở dữ liệu quan hệ



## ❖ Các thành phần cơ bản của mô hình

### ▪ Một số khái niệm của mô hình quan hệ

- Mô hình quan hệ là cách thức biểu diễn dữ liệu dưới dạng các quan hệ (các bảng). Một quan hệ là một bảng dữ liệu 2 chiều (cột và dòng), mô tả một thực thể. Mỗi cột tương ứng với một thuộc tính của thực thể. Mỗi dòng chứa các giá trị dữ liệu của một đối tượng cụ thể thuộc thực thể.
- **Lược đồ quan hệ:**  $R(A_1, \dots, A_n)$ , trong đó  $R$  là tên quan hệ,  $A_i$  là các thuộc tính, mỗi  $A_i$  có miền giá trị tương ứng  $dom(A_i)$ .
- Lược đồ quan hệ được sử dụng để mô tả một quan hệ, bao gồm: Tên quan hệ, các thuộc tính và bậc của quan hệ (số lượng các thuộc tính)

## ❖ Các thành phần cơ bản của mô hình

### ▪ Một số khái niệm của mô hình quan hệ

Ví dụ: Quan hệ **NHANVIEN** trên tập các thuộc tính  $R=\{\text{MaNV}, \text{HoTen}, \text{NgaySinh}, \text{DiaChi}, \text{Luong}\}$  là một quan hệ 5 ngôi

MaNV	HoTen	NgaySinh	DiaChi	Luong	
001	Nguyễn Văn Nam	1994	Hà Nội	800	t1
002	Trần Thành Công	1993	Nam Định	650	t2
003	Lê Hải Hà	1993	Hà Nam	900	t3
004	Bùi Mạnh Hùng	1992	Hải Dương	1050	t4

$t1(001, 'Nguyễn Văn Nam', 1994, 'Hà Nội', 800)$  là 1 bộ của quan hệ **NHANVIEN**



## ❖ Các thành phần cơ bản của mô hình

### ■ Các tính chất của một quan hệ

- Giá trị đưa vào cột là đơn nhất
- Các giá trị trong cùng một cột phải thuộc cùng một miền giá trị (cùng kiểu)
- Thứ tự dòng cột tùy ý.

### ■ Các ràng buộc toàn vẹn trên quan hệ

- Ràng buộc là những quy tắc được áp đặt lên trên dữ liệu đảm bảo tính tin cậy và độ chính xác của dữ liệu. Các luật toàn vẹn được thiết kế để giữ cho dữ liệu phù hợp và đúng đắn.
- **Có 4 kiểu ràng buộc chính:** Ràng buộc miền giá trị, Ràng buộc khoá, Ràng buộc thực thể và Ràng buộc toàn vẹn tham chiếu.

## ❖ Các thành phần cơ bản của mô hình

### ▪ Các ràng buộc toàn vẹn trên quan hệ

- **Ràng buộc miền giá trị:** Là một hợp các kiểu dữ liệu và những giá trị giới hạn mà thuộc tính có thể nhận được. Thông thường việc xác định miền giá trị của các thuộc tính bao gồm một số các yêu cầu sau: *Tên thuộc tính, Kiểu dữ liệu, Độ dài dữ liệu, khuôn dạng của dữ liệu, các giá trị giới hạn cho phép, ý nghĩa, có duy nhất hay không, có cho phép giá trị rỗng hay không*
  - **Ràng buộc khoá:**
    - **Khóa chính:** là một (hoặc một tập) các thuộc tính đóng vai trò là nguồn của một phụ thuộc hàm mà đích lần lượt là các thuộc tính còn lại.
- => **Từ khóa chính có thể suy ra thông tin còn lại.**

## ❖ Các thành phần cơ bản của mô hình

### ■ Các ràng buộc toàn vẹn trên quan hệ

- Ràng buộc khoá:

- **Khóa dự tuyển:** Trong tập hợp các thuộc tính của một bảng, có thể có nhiều thuộc tính có thể dùng được làm khóa chính. Các thuộc tính đó được gọi là khóa dự tuyển.

**Khóa dự tuyển cần thỏa mãn 2 tính chất sau:**

- » Xác định duy nhất.
- » Không dư thừa: Khi xóa đi bất kỳ một thuộc tính nào của khóa đều phá hủy tính xác định duy nhất của khóa.

## ❖ Các thành phần cơ bản của mô hình

### ▪ Các ràng buộc toàn vẹn trên quan hệ

- Ràng buộc khoá:

- **Khóa ngoại (Foreign Key):** Trong nhiều trường hợp, khóa chính của một bảng được đưa sang làm thuộc tính bên bảng khác, thuộc tính đó gọi là khóa ngoại. Khóa ngoại đóng vai trò thể hiện liên kết giữa 2 bảng.
- **Khóa phụ (Second Key):** Đóng vai trò khi ta muốn sắp xếp lại dữ liệu trong bảng. Ví dụ: Ta có bảng SINHVIEN (MaSV, Hoten, GioiTinh, Diem). Muốn sắp xếp lại danh sách sinh viên theo thứ tự a, b, c.. của Họ tên. Khi đó thuộc tính Hoten được gọi là khóa phụ.

## ❖ Các thành phần cơ bản của mô hình

### ■ Các ràng buộc toàn vẹn trên quan hệ

- **Ràng buộc thực thể:** Mỗi một lược đồ quan hệ R, chúng ta phải xác định khoá chính của nó. Khoá chính trong lược đồ quan hệ được gạch chân ở phía dưới của thuộc tính.

EMPLOYEE

FNAME	MINIT	LNAME	<u>SSN</u>	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
-------	-------	-------	------------	-------	---------	-----	--------	----------	-----

DEPARTMENT

DNAME	<u>DNUMBER</u>	MGRSSN	MGRSTARTDATE
-------	----------------	--------	--------------

DEPT\_LOCATIONS

<u>DNUMBER</u>	<u>DLOCATION</u>
----------------	------------------

PROJECT

PNAME	<u>PNUMBER</u>	PLOCATION	DNUM
-------	----------------	-----------	------

WORKS\_ON

<u>ESSN</u>	<u>PNO</u>	HOURS
-------------	------------	-------

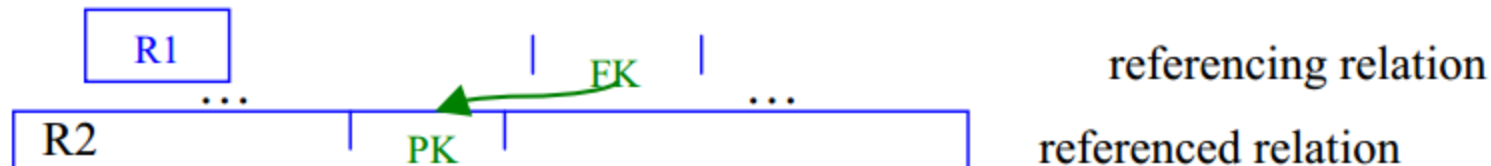
DEPENDENT

<u>ESSN</u>	<u>DEPENDENT_NAME</u>	SEX	BDATE	RELATIONSHIP
-------------	-----------------------	-----	-------	--------------

## ❖ Các thành phần cơ bản của mô hình

### ▪ Các ràng buộc toàn vẹn trên quan hệ

- **Ràng buộc toàn vẹn tham chiếu:** Một bộ giá trị trong một quan hệ tham chiếu tới một bộ giá trị đã tồn tại trong một quan hệ khác. Ràng buộc toàn vẹn tham chiếu phải xác định trên 2 quan hệ: quan hệ tham chiếu (referencing relation) và quan hệ được tham chiếu (referenced relation).



**Ví dụ:** Thuộc tính DNo của quan hệ EMPLOYEE tham chiếu tới thuộc tính DNumber của quan hệ DEPARTMENT.

## TẠI SAO PHẢI CHUẨN HÓA DỮ LIỆU

## ■ Chuẩn hóa dữ liệu

### ■ Tại sao phải chuẩn hóa dữ liệu ?

Xét ví dụ thông tin bảng điểm học tập của sinh viên như sau:

StNo	StName	SubNo	SubName	Credit	Mark
St01	Mai	Sub04	CSDL	3	9
St01	Mai	Sub01	TRR	5	10
St01	Mai	Sub02	PPS	4	8
St02	Vân	Sub04	CSDL	3	10
St02	Vân	Sub01	TRR	5	9
St03	Thanh	Sub07	Tiếng Anh	4	8



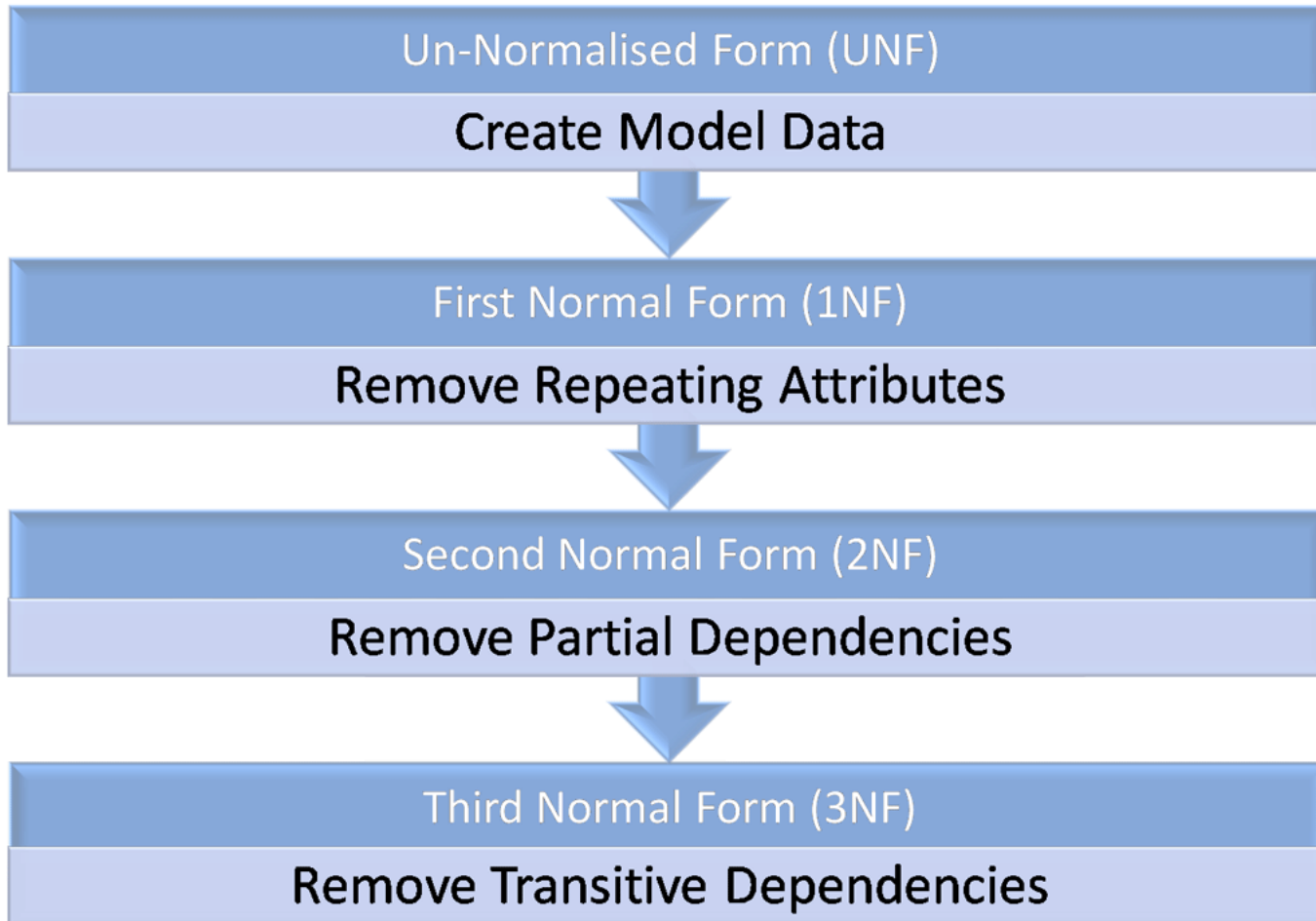
## ■ Chuẩn hóa dữ liệu

### ■ Tại sao phải chuẩn hóa dữ liệu ?

- **Dư thừa dữ liệu:** thông tin sinh viên và môn học bị lặp lại nhiều lần.
- **Không nhất quán:** Hệ quả của dư thừa dữ liệu. Sửa bản ghi thứ nhất tên sinh viên thành tên Lan, dữ liệu này sẽ không nhất quán với bản ghi thứ 2, 3 (do vẫn đang là tên Mai)
- **Dị thường khi thêm một bộ:** Nếu muốn thêm sinh viên mới nhập trường chưa có môn học thì không được do StNo, SubNo kết hợp là khóa chính
- **Dị thường khi xóa một bộ:** Ví dụ xóa bản ghi cuối thì thông tin môn học cũng bị mất.

## CHUẨN HÓA DỮ LIỆU

## ▪ Chuẩn hóa dữ liệu (Normal Form)



## ■ Chuẩn hóa dữ liệu

### ■ Dạng chuẩn 1 (First Normal Form)

Một quan hệ ở dạng chuẩn 1 nếu các giá trị của tất cả thuộc tính trong quan hệ là nguyên tử (tức là chỉ có 1 giá trị tại một thời điểm).

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

Vi phạm dạng chuẩn 1

- **Chuẩn hóa dữ liệu**
  - **Dạng chuẩn 1 (First Normal Form)**

DEPARTMENT

DNAME	<u>DNUMBER</u>	DMGRSSN	<u>DLOCATION</u>
Research	5	333445555	Bellaire
Research	5	333445555	Sugarland
Research	5	333445555	Houston
Administration	4	987654321	Stafford
Headquarters	1	888665555	Houston

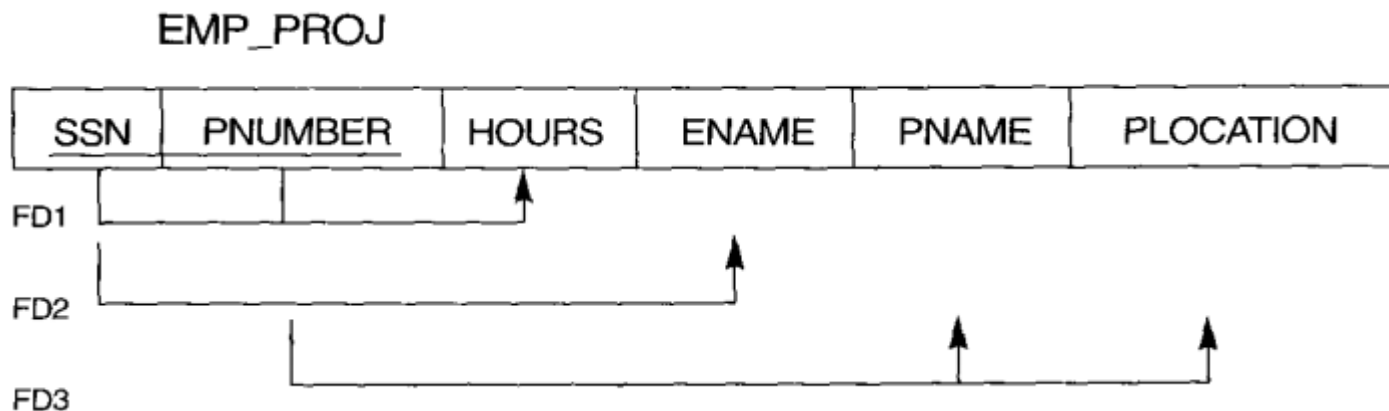
Dư thừa

## ■ Chuẩn hóa dữ liệu

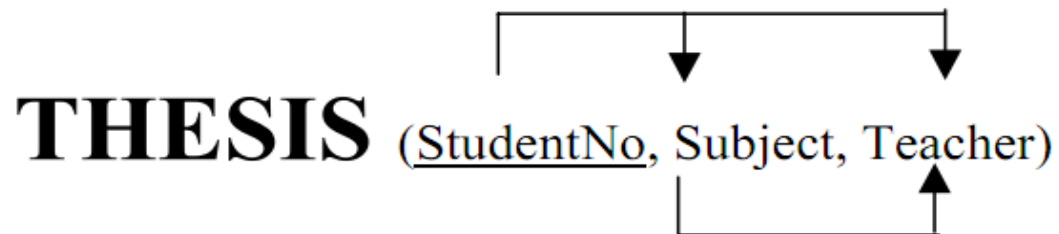
### ■ Dạng chuẩn 2 (Second Normal Form)

Một quan hệ ở dạng chuẩn 2 nếu:

- ✓ Quan hệ đó ở dạng chuẩn 1
- ✓ Các thuộc tính không phải là khóa **phụ thuộc đầy đủ vào khóa.**



- **Chuẩn hóa dữ liệu**
  - **Dạng chuẩn 2 (Second Normal Form)**



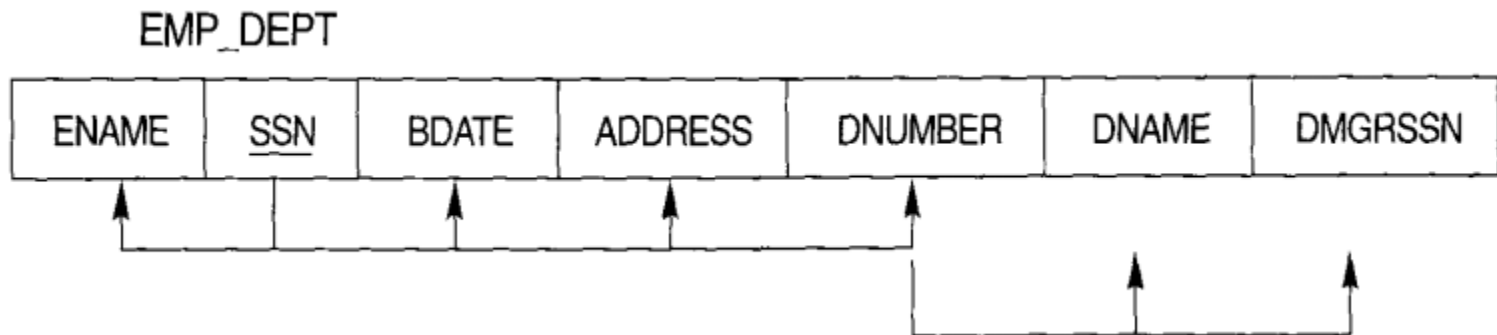
<u>StudentNo</u>	Subject	Teacher
SV01	1	Nguyễn Văn Hiệu
SV02	2	Ngô Lan Phương
SV03	1	Nguyễn Văn Hiệu
SV04	1	Nguyễn Văn Hiệu

## ■ Chuẩn hóa dữ liệu

### ■ Dạng chuẩn 3 (Third Normal Form)

Một quan hệ ở dạng chuẩn 3 nếu:

- Quan hệ ở dạng chuẩn 2
- Và không có chứa các phụ thuộc hàm phụ thuộc bắc cầu vào khoá.
- **Phụ thuộc hàm phụ thuộc bắc cầu:** Phụ thuộc hàm  $Y \rightarrow Z$  là phụ thuộc hàm bắc cầu nếu tồn tại hai phụ thuộc hàm:  $Y \rightarrow X$  và  $X \rightarrow Z$ .

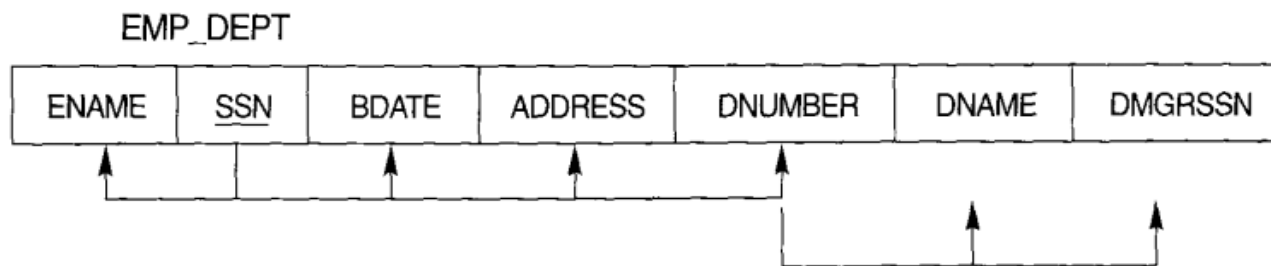




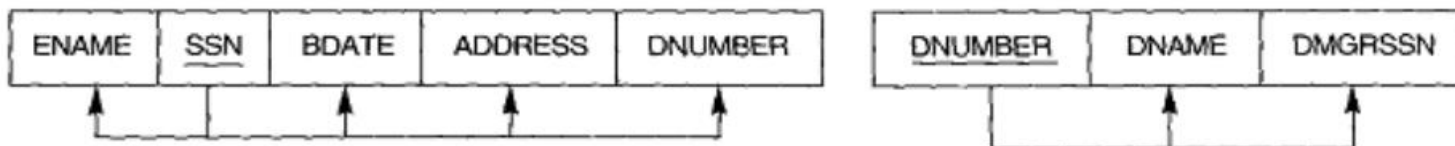
## ■ Chuẩn hóa dữ liệu

### ■ Dạng chuẩn 3 (Third Normal Form)

Ví dụ: Quan hệ EMP\_DEPT không phải ở dạng chuẩn 3 vì còn tồn tại phụ thuộc hàm **Ssn -> Dname** (Ssn -> DNumber, DNumber -> DName) là phụ thuộc hàm phụ thuộc bắc cầu vào khoá.



=> Tách thành 2 bảng lúc đó thành dạng chuẩn **3NF**

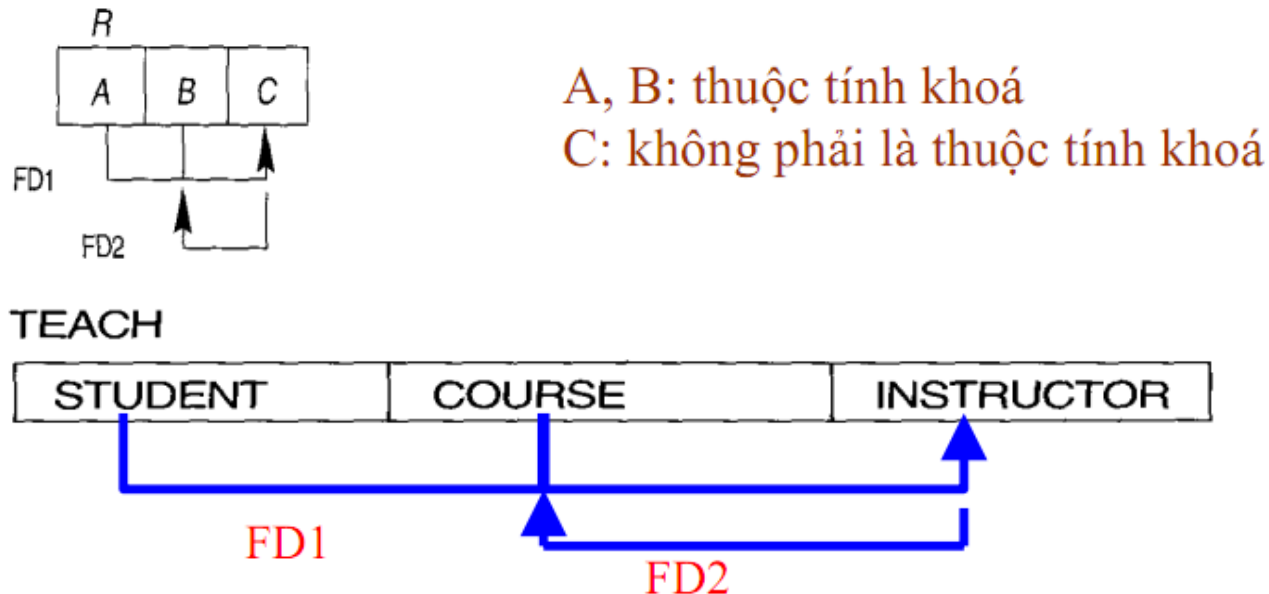


## ■ Chuẩn hóa dữ liệu

### ■ Dạng chuẩn Boyce\_Codd (Boyce-Codd Normal Form)

Quan hệ R ở dạng chuẩn **BCNF** khi tất cả các phụ thuộc hàm  $X \rightarrow A$  trong R đều phải có X là khoá của R.

**Ví dụ:** Quan hệ sau ở dạng **3NF** nhưng không phải **BCNF**



## ■ Chuẩn hóa dữ liệu

### ■ Nhận xét

- Trong một cơ sở dữ liệu tốt, các quan hệ nên được chuyển về dạng chuẩn 3.
- Tuy nhiên, dữ liệu vẫn có khả năng dư thừa khi quan hệ có hai tập khóa dự tuyển gói lẫn nhau, hoặc quan hệ có thuộc tính không khóa xác định một thuộc tính khóa.

**“HỌC ĐỂ LÀM VIỆC”**

**Thank You !**