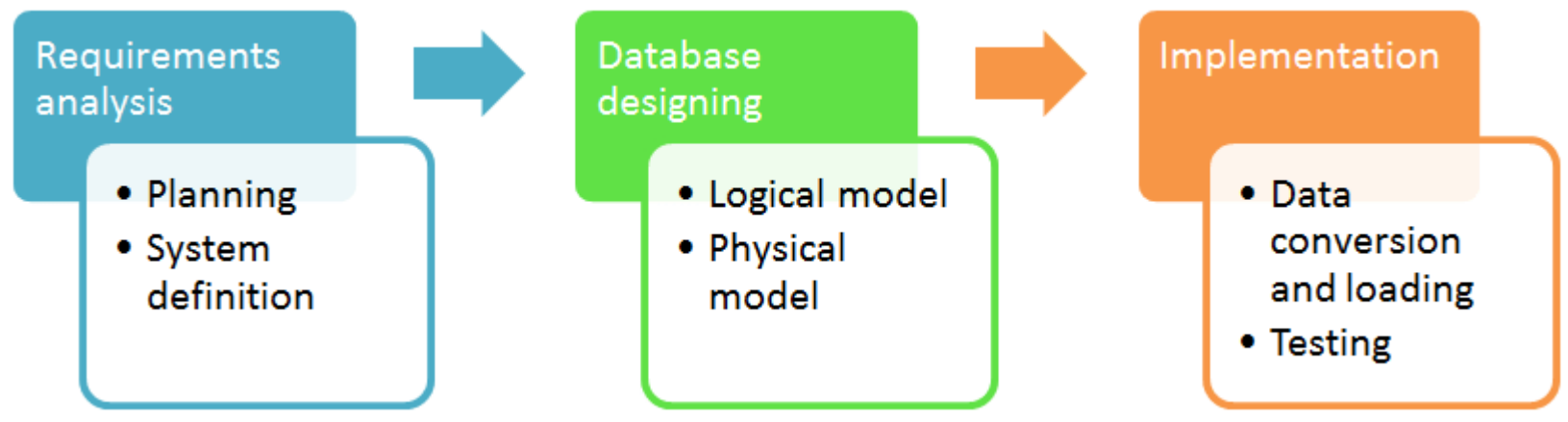


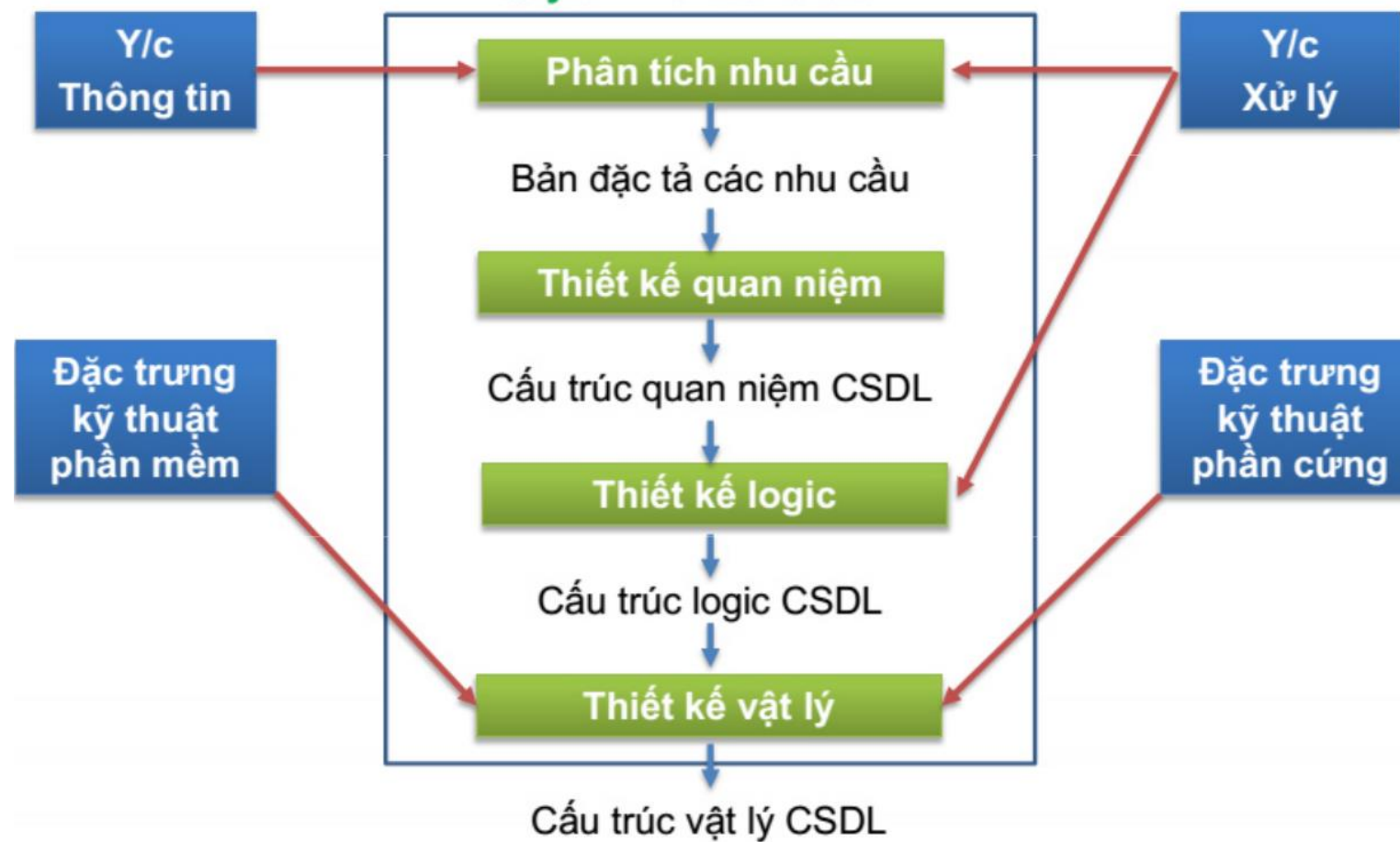
Chương 3: Thiết kế mức logic

- ▶ Các khái niệm cơ bản
- ▶ Các phương pháp thiết kế cơ sở dữ liệu logic
 - ▶ Top-down
 - ▶ Bottom-up
 - ▶ Phụ thuộc hàm
 - ▶ Bao đóng
 - ▶ Khoá
 - ▶ Chuẩn hoá

Thiết kế CSDL



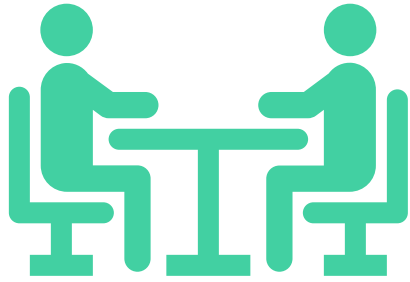
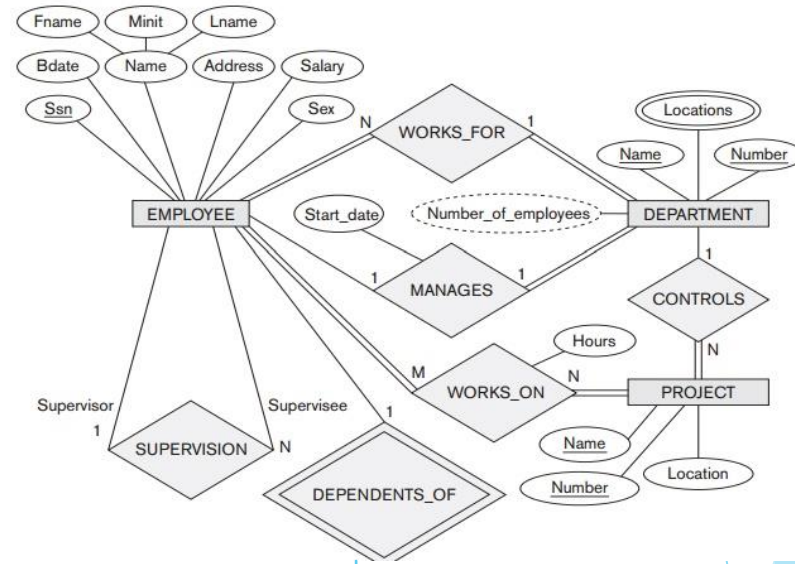
Quy trình thiết kế CSDL



Mô hình hoá dữ liệu

Figure 9.1

The ER conceptual schema diagram for the COMPANY database.



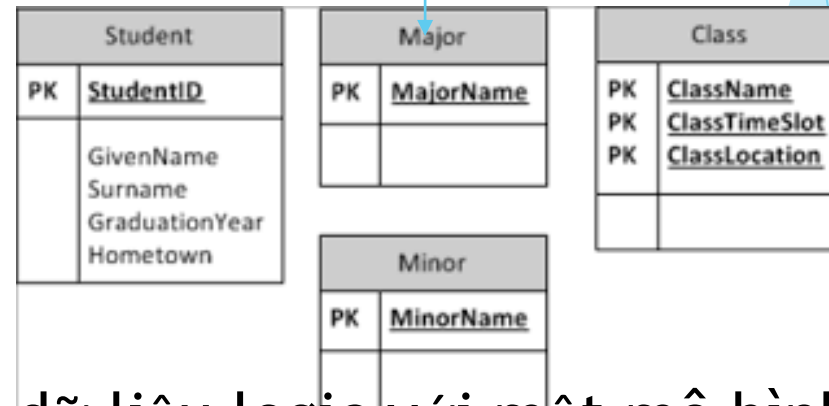
Mô tả công việc

Phân tích

Thiết kế



Triển khai



Mô tả dữ liệu logic với một mô hình dữ liệu (Sơ đồ quan hệ)

Thiết kế CSDL

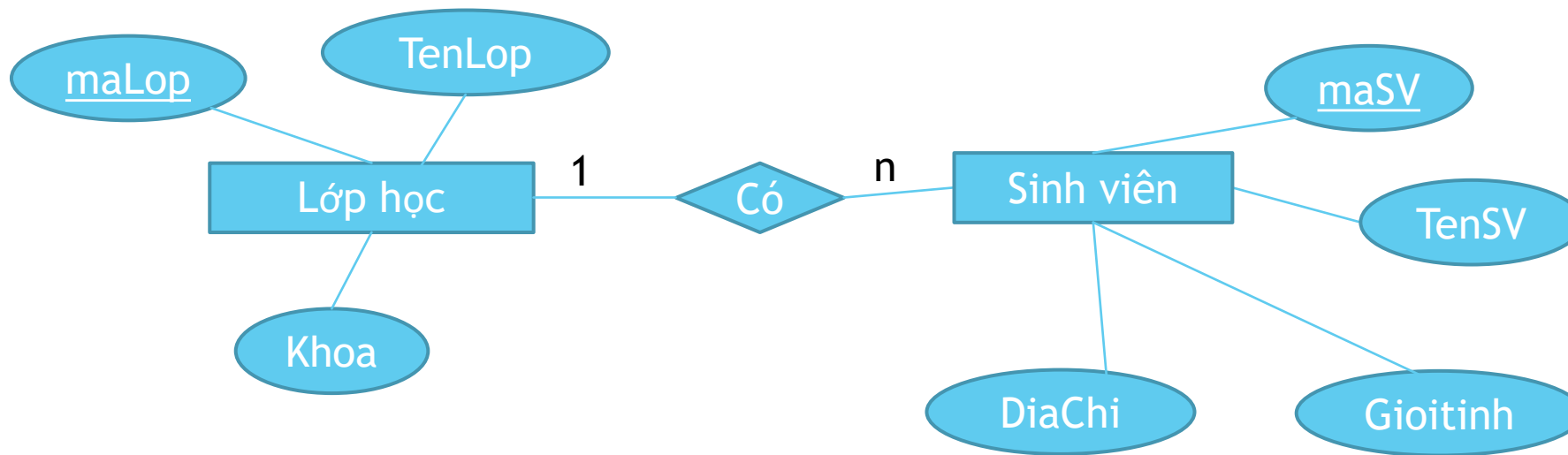
- ▶ Thu thập và phân tích yêu cầu
 - ▶ Các yêu cầu về CSDL
 - ▶ Các yêu cầu về chức năng thao tác (Thao tác trên CSDL)
- ▶ Thiết kế quan niệm và phân tích chức năng
 - ▶ Tạo một sơ đồ quan niệm (cấp cao), ví dụ thực thể liên kết
 - ▶ Đặc tả giao tác cấp cao tương ứng với các thao tác trên CSDL
- ▶ Thiết kế logic
 - ▶ Ánh xạ lược đồ quan niệm thành lược đồ logic (mô hình quan hệ)
- ▶ Thiết kế và triển khai các ứng dụng.

Nhắc lại: quá trình thiết kế mô hình dữ liệu quan niệm

- ▶ **Bước 1: Nhận dạng các kiểu thực thể**
- ▶ **Bước 2: Nhận dạng các kiểu liên kết**
- ▶ **Bước 3: Nhận dạng các thuộc tính của các kiểu thực thể và các mối liên kết**
- ▶ **Bước 4: Nhận dạng thuộc tính xác định cho mỗi kiểu thực thể**
- ▶ **Bước 5: Nhận dạng các cấu trúc siêu kiểu/kiểu con**
- ▶ **Bước 6: Vẽ sơ đồ ER**

Ví dụ:

► Vấn đề Lớp học và Sinh viên



Thiết kế mức logic

- ▶ Mô hình CSDL quan hệ hay ngắn gọn là mô hình quan hệ được E.F.Codd phát triển vào đầu những năm 1970.
- ▶ Các thành phần trong mô hình quan hệ gồm: Các quan hệ - các bộ - các thuộc tính.

Các kiểu thiết kế mức logic

- ▶ Trên xuống (Top-down)
- ▶ Dưới lên (bottom-up)
 - ▶ 1. Biểu diễn dữ liệu người dùng (biểu mẫu, báo cáo) dưới dạng các quan hệ
 - ▶ 2. Chuẩn hoá các quan hệ này
 - ▶ 3. Ghép các quan hệ có cùng khoá chính

Top-down

▶ Phương pháp Top-down

- ▶ Tiếp cận theo hướng mô hình liên kết thực thể, sau đó áp dụng các quy tắc chuyển đổi mô hình liên kết thực thể sang mô hình quan hệ.
- ▶ Các bước xây dựng mô hình liên kết thực thể
- ▶ Xác định các thực thể.
- ▶ Xác định các thuộc tính và gán thuộc tính vào mỗi liên kết.
- ▶ Xác định mỗi liên kết và loại mỗi liên kết giữa các thực thể.
- ▶ Gán thuộc tính vào mỗi liên kết (nếu có).
- ▶ Vẽ mô hình liên kết thực thể.

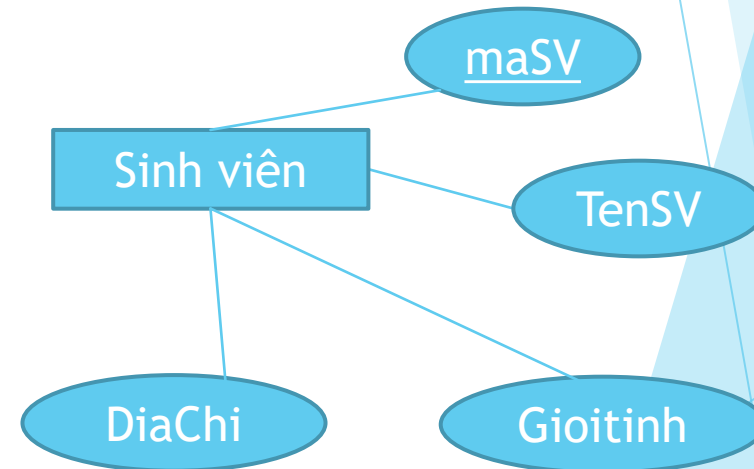
Top-down

- ▶ Phương pháp Top-down
- ▶ Chuyển đổi từ mô hình liên kết thực thể sang mô hình quan hệ
 - ▶ Tên thực thể -> Tên lược đồ quan hệ (Tên bảng)
 - ▶ Thuộc tính -> Thuộc tính
 - ▶ Thuộc tính khóa -> Khóa chính
 - ▶ Thuộc tính đa trị -> chuyển thành các thuộc tính đơn trị.
 - ▶ Thực thể yếu -> Lược đồ quan hệ (có khóa là thuộc tính khóa của bảng cha và một thuộc tính định danh của nó)

Các bước thiết kế logic theo Top-down

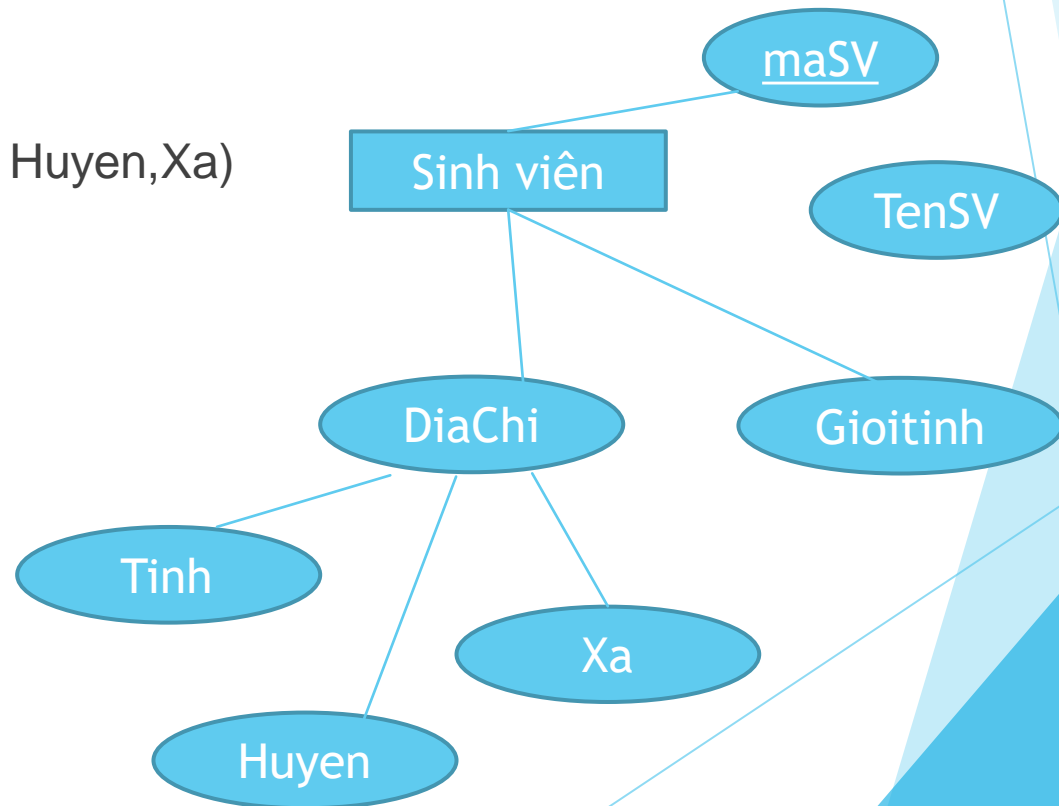
Chuyển đổi ERD thành các quan hệ

- ▶ Chuyển đổi kiểu thực thể thông thường (regular entity)
 - ▶ Mỗi kiểu thực thể thông thường được chuyển thành 1 quan hệ (bảng).
 - ▶ Khoá chính của kiểu thực thể trở thành khoá chính của bảng. Các thuộc tính của kiểu thực thể trở thành các cột của bảng
- ▶ SinhVien(MaSV, TenSV, GioiTinh, DiaChi)

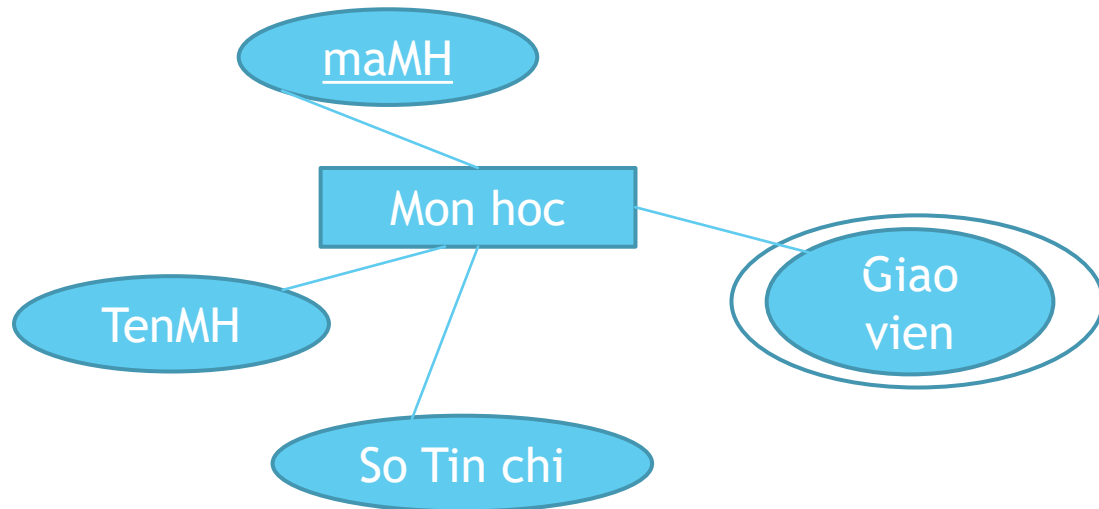


- ▶ Với thuộc tính phức hợp: các thành phần sẽ được chuyển thành các thuộc tính trong quan hệ

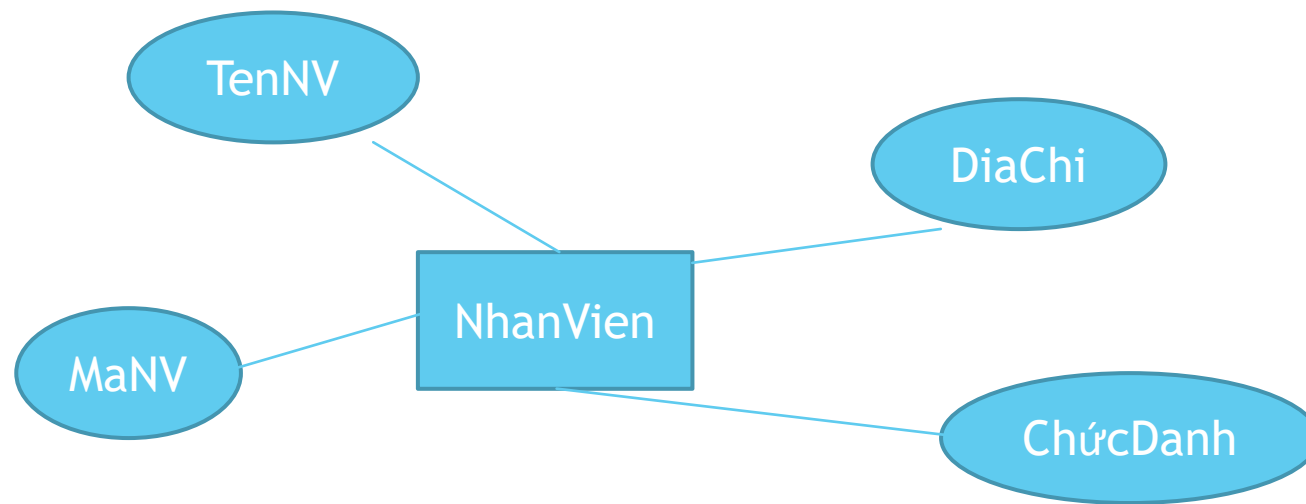
- ▶ Sinhvien(MaSV, TenSV, GioiTinh, Tinh, Huyen,Xa)



- ▶ Với mỗi thuộc tính đa trị
 - ▶ Thêm 1 quan hệ mới xác định bởi thuộc tính đa trị và khoá của tập thực thể tương ứng:
 - ▶ Monhoc(MaMH, TenMH, SoTinChi)
MH_GV(maMH, giao_vien)

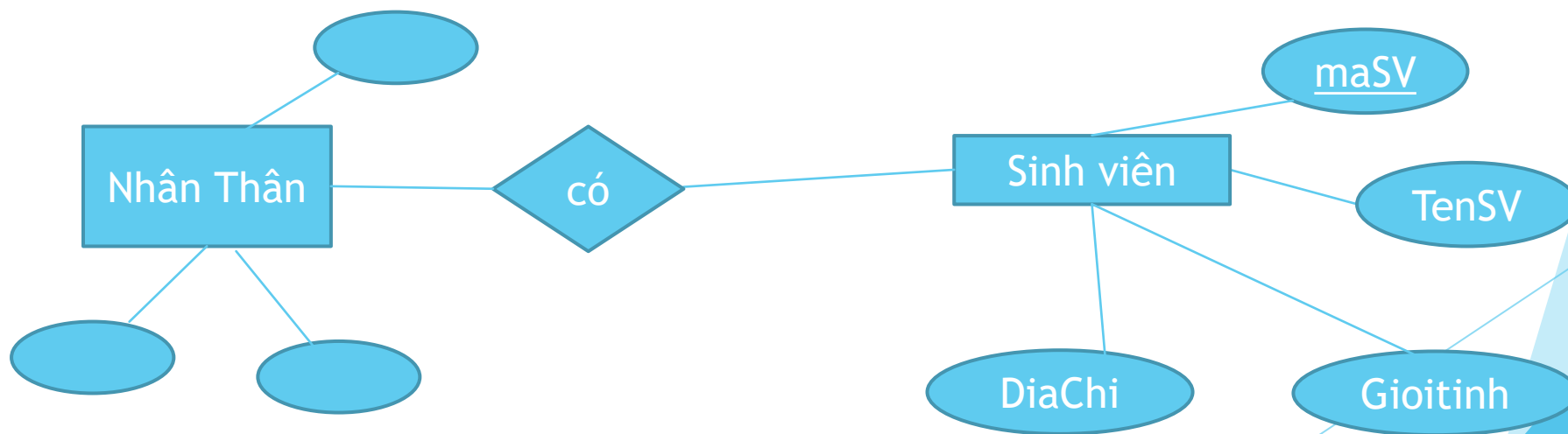


MaMH	MaGV
CSDL	PU001
CSDL	PU002
Datamining	PU001
Datamining	PU002



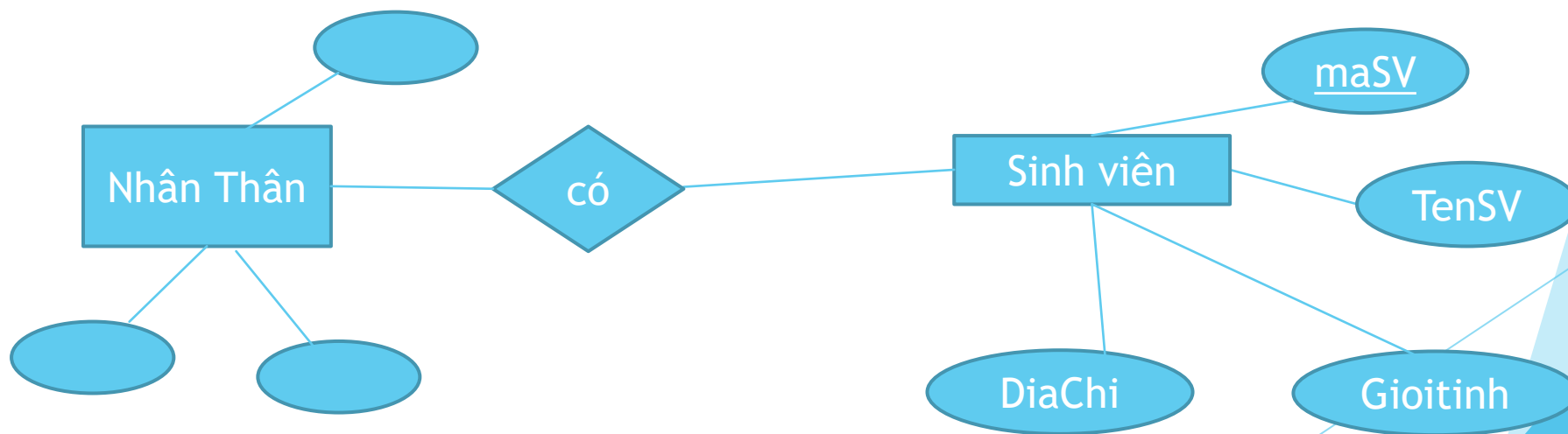
- ▶ `NhanVien(maNV, TenNV, DiaChi)`
- ▶ `NhanVien_ChucDan(MaNV, ChucDanh)`

- ▶ **Chuyển đổi kiểu thực thể yếu (weak entity)**
- ▶ Mỗi thực thể yếu chuyển thành 1 quan hệ. Thuộc tính xác định của kiểu thực thể chủ được đưa vào như là 1 thuộc tính của quan hệ.
- ▶ Khóa chính của quan hệ = thuộc tính xác định của thực thể chủ + thuộc tính xác định của thực thể yếu



Thực thể yếu

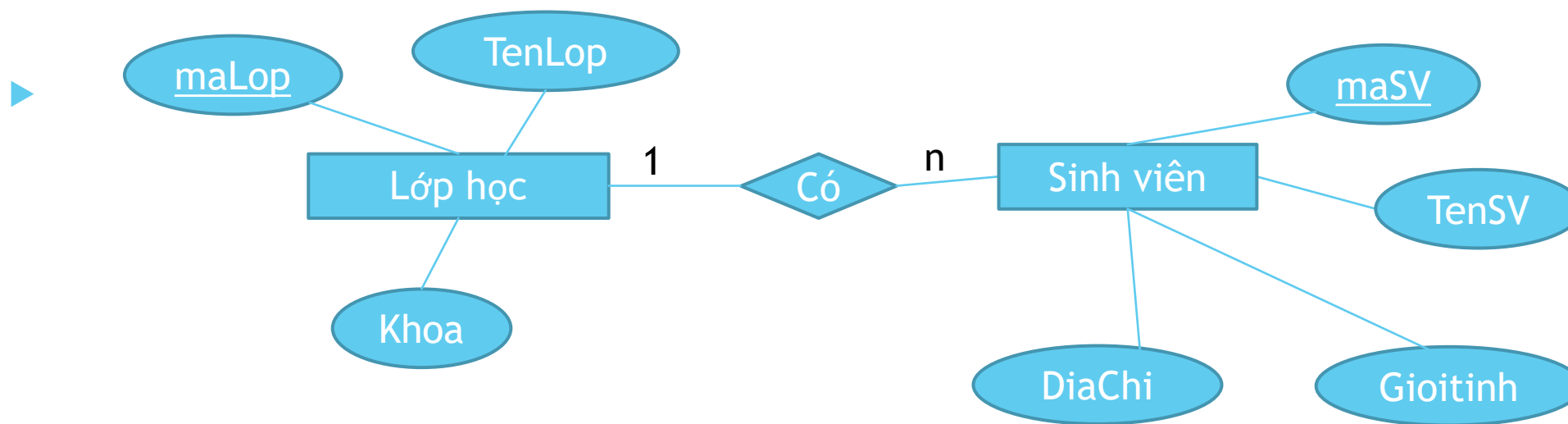
- ▶ SinhVien_NhanThan(MaSV, MaNhanThan, Ten Nhan Than, Gioi Tinh, MoiQuanHe)



Chuyển đổi liên kết

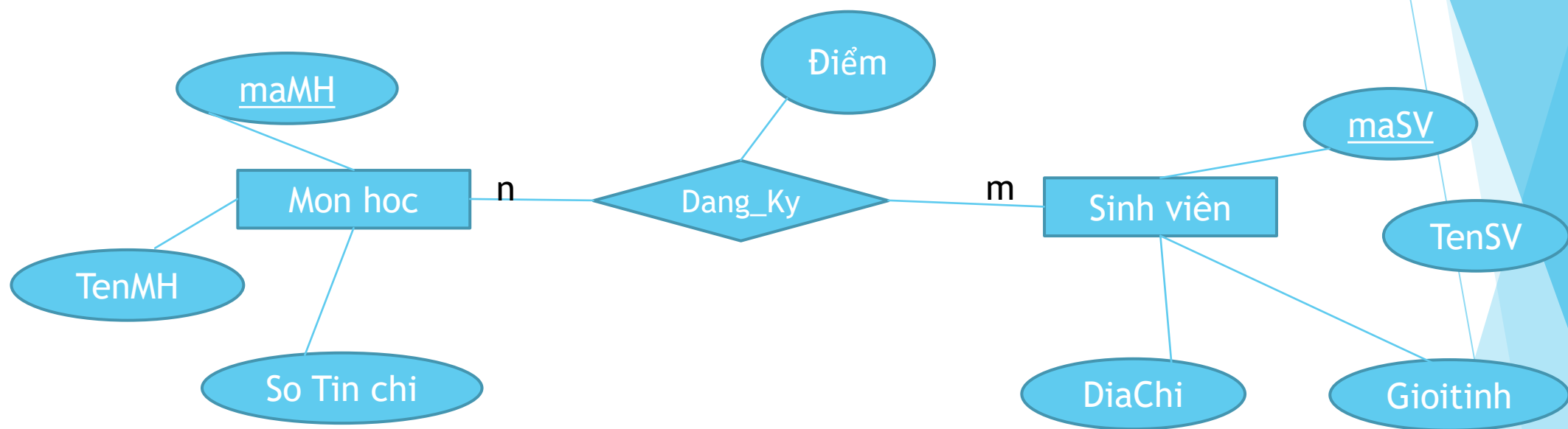
► Liên kết 1-n

- Với mỗi liên kết 1-N: Thuộc tính xác định của kiểu thực thể phía 1 được đưa vào làm khoá ngoại của kiểu thực thể phía N.



- Dùng khoá ngoại: thêm khoá chính của quan hệ 1 (lớp học) vào quan hệ bên n (Sinh viên) làm khoá ngoại.
- SinhVien(MaSV, TenSV, GioiTinh, DiaChi, MaLop)

► Liên kết n-m



Thêm 1 quan hệ mới xác định bởi các thuộc tính nằm trong khóa của các thực thể có liên quan và các thuộc tính của liên kết DANG_KY(maSV,maMH, diem)

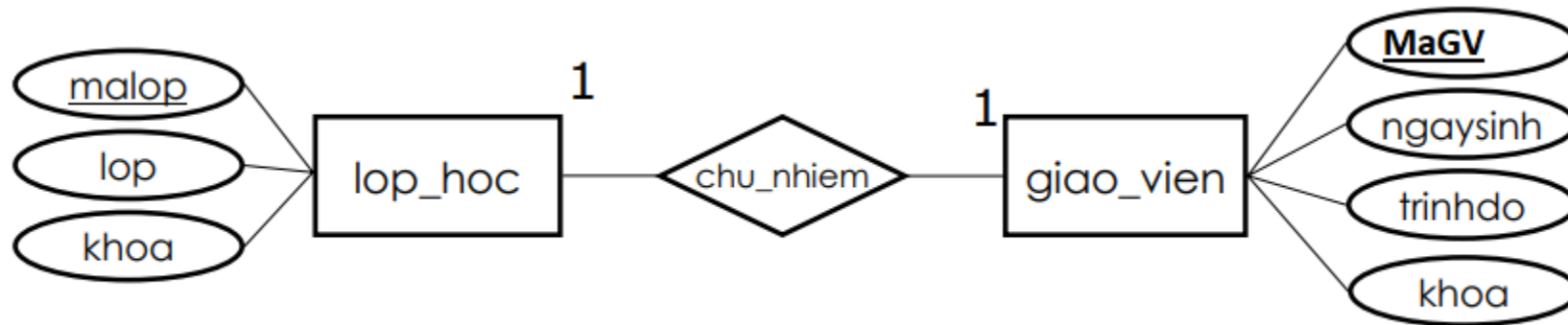
MonHoc(MaMH, TenMH, SoTinChi)

SinhVien(MaSV, TenSV, GioiTinh, DiaChi)

DangKy(SinhVien, MaMH, Diem)

Biến đổi các liên kết

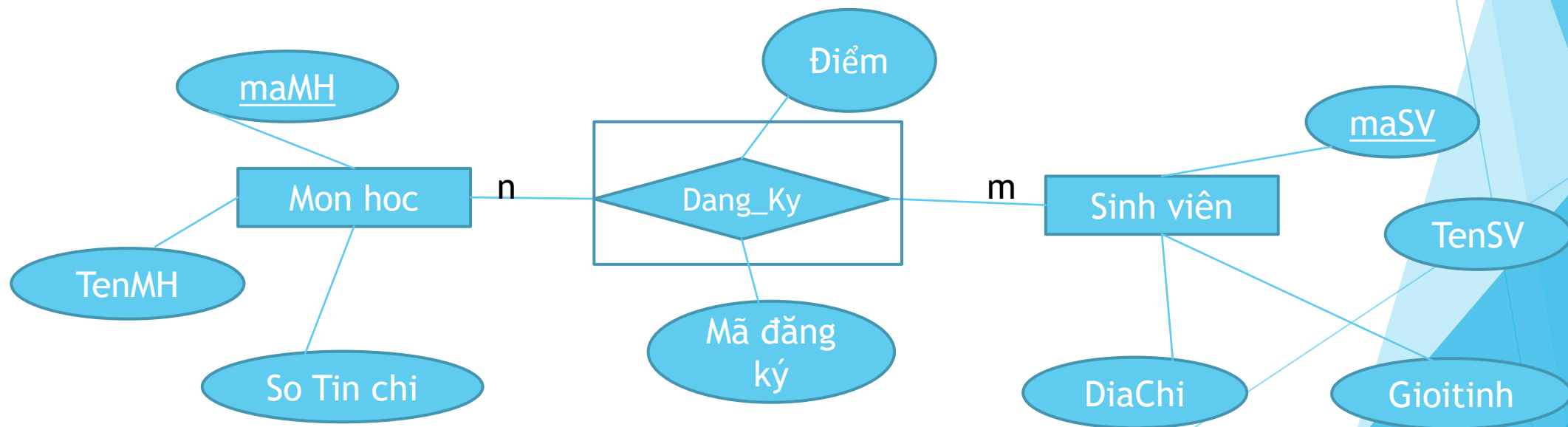
▶ Liên kết 1-1



- ▶ Dùng khoá ngoài LOP_HOC(malop, lop, khoa, maGV)
- ▶ Hoặc GiaoVien(MaGV,TenGV,MaLop)

- ▶ **Chuyển đổi kiểu thực thể kết hợp**
- ▶ Một quan hệ mới được tạo ra để diễn tả cho kiểu thực thể kết hợp
 - ▶ Nếu kiểu thực thể kết hợp không có thuộc tính xác định riêng, khoá chính của quan hệ mới bao gồm khoá chính của 2 quan hệ tương ứng với 2 kiểu thực thể có liên quan
 - ▶ Nếu kiểu thực thể kết hợp có thuộc tính xác định riêng thì thuộc tính này được dùng làm khoá chính cho quan hệ mới.

- ▶ Trường hợp 1: DangKy(MaSV, MaMH, Diem)
- ▶ Trường hợp 2: DangKy(MaDK, Diem, MaSV, MaMH)

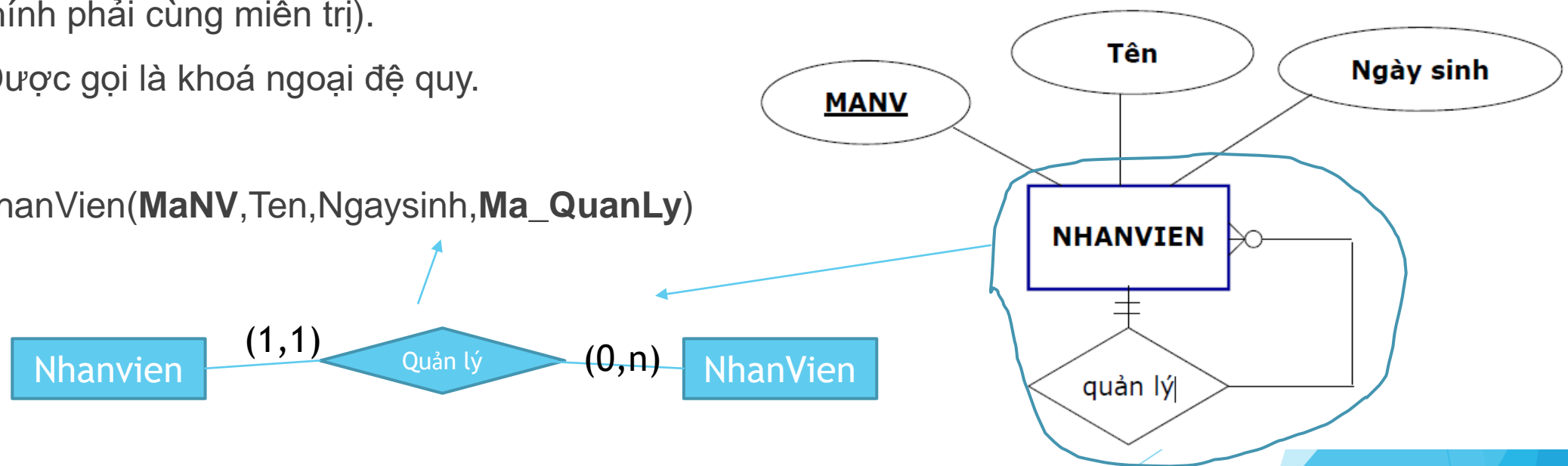


► **Chuyển đổi mỗi liên kết 1 ngôi**

► Với liên kết 1 ngôi 1-M: Thêm 1 khoá ngoại vào quan hệ của mỗi liên kết này để tham chiếu đến các trị của khoá chính của quan hệ (khóa ngoại và khóa chính phải cùng miền trị).

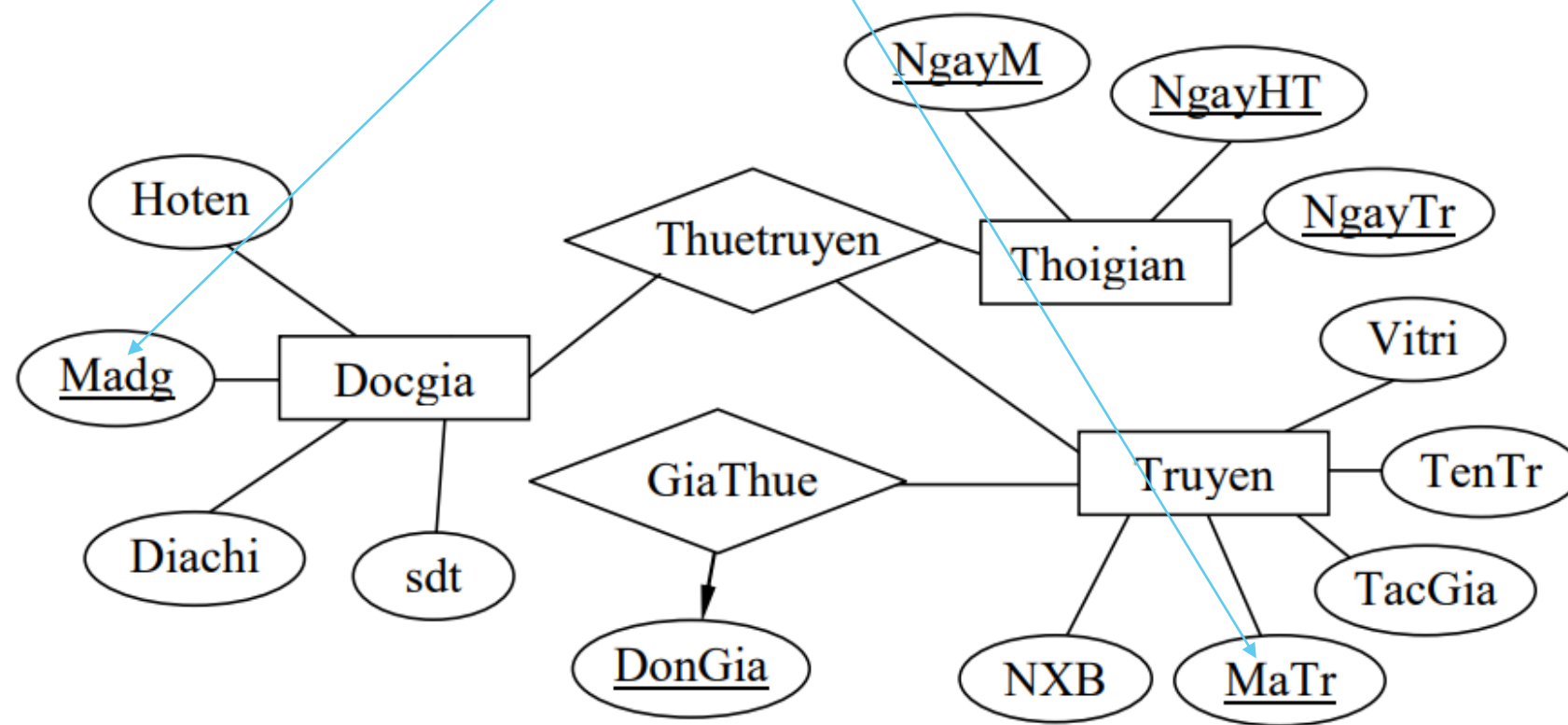
► Được gọi là khóa ngoại đệ quy.

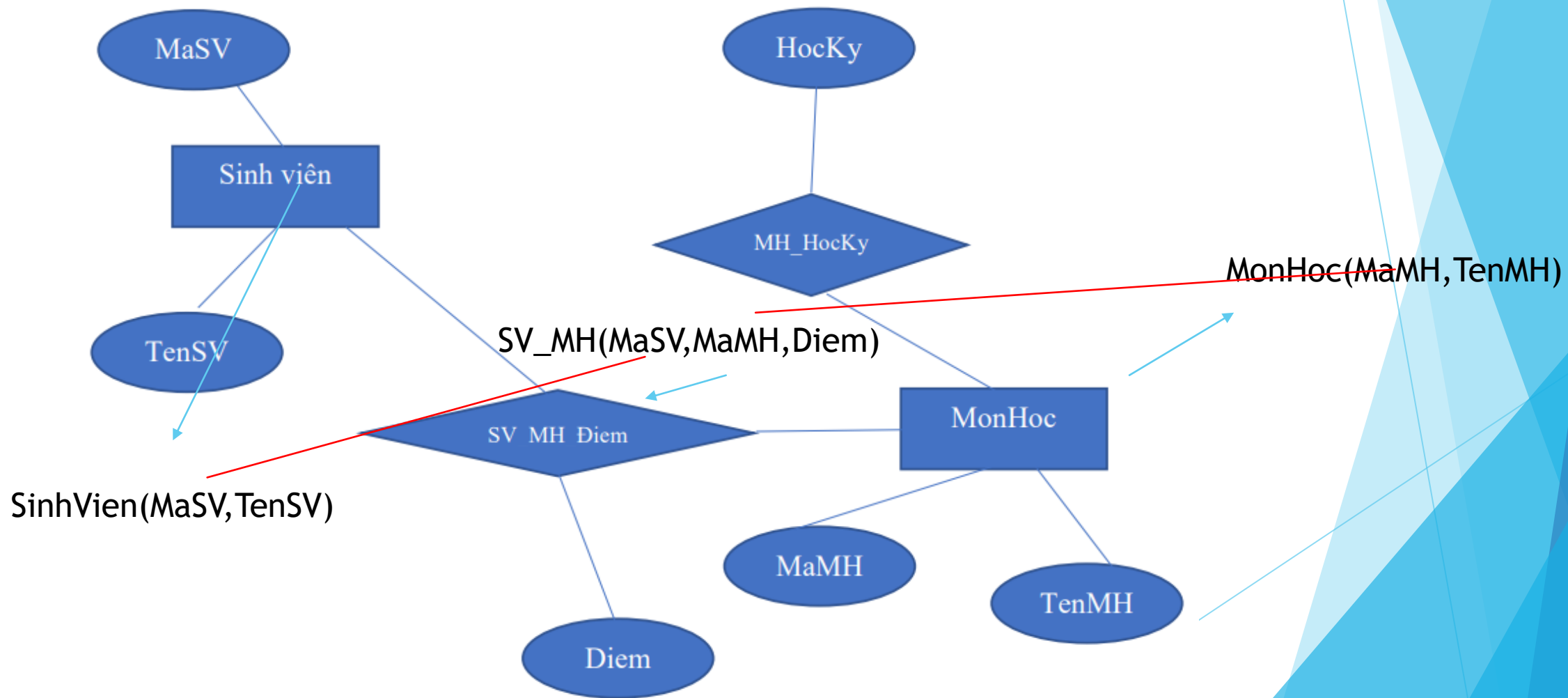
► NhanVien(**MaNV**,Ten,Ngaysinh,**Ma_QuanLy**)

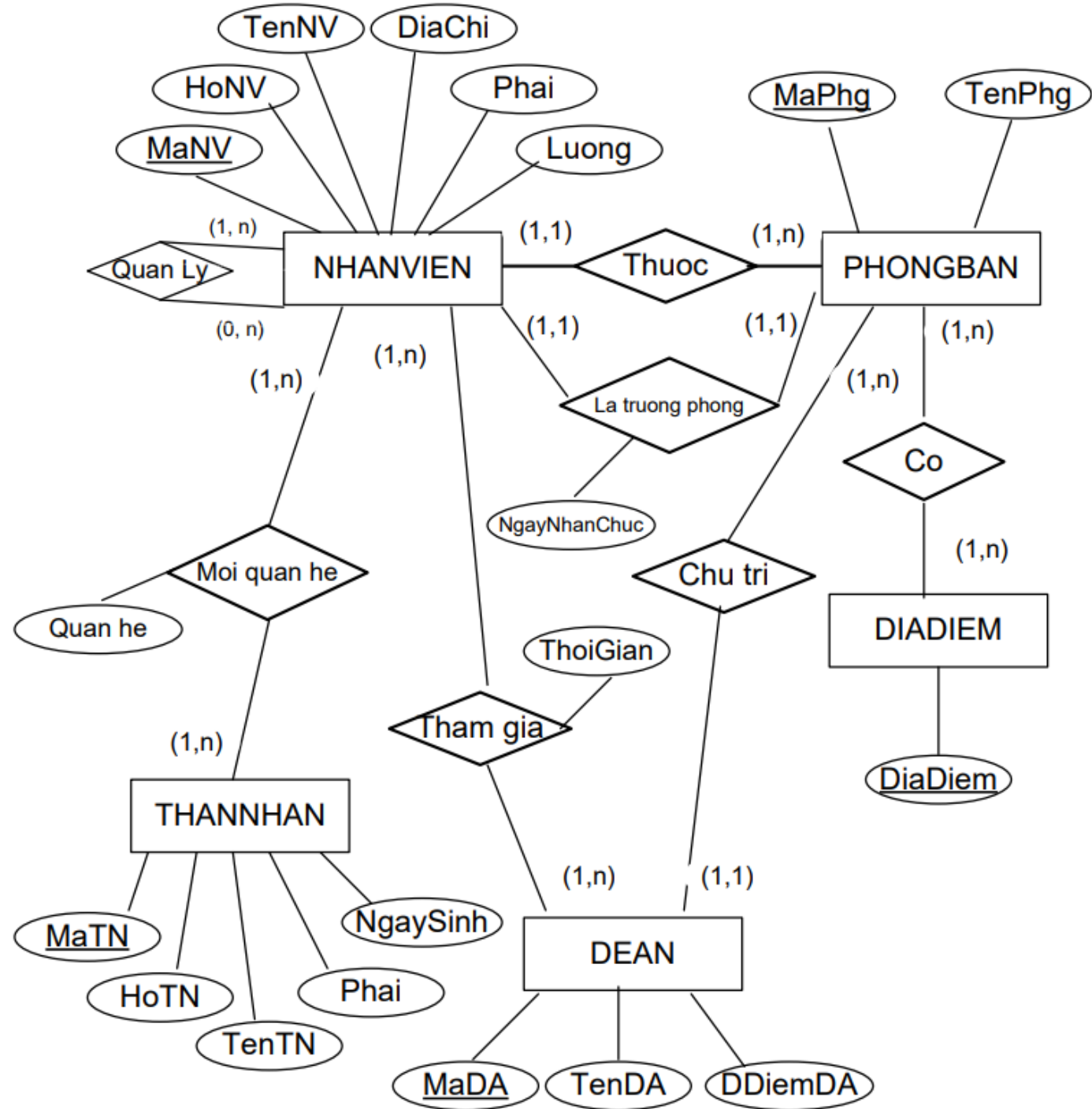


Làm nhanh!

ThueTruyen(**MaDG**, **MaTruyen**, NgayM, NgayHT, NgayTR)







Bottom-up

- ▶ Vấn đề???
 - ▶ Dư thừa
 - ▶ Cấu trúc không liên kết
- ▶ Dưới lên (bottom-up)
 - ▶ 1. Biểu diễn dữ liệu người dùng (biểu mẫu, báo cáo) dưới dạng các quan hệ
 - ▶ 2. Chuẩn hoá các quan hệ này
 - ▶ 3. Ghép các quan hệ có cùng khoá chính
- ▶ Nhóm tất cả các thuộc tính liên quan của hệ thống vào trong một quan hệ.
- ▶ Áp dụng các luật chuẩn hoá để tách quan hệ đó thành các quan hệ có cấu trúc tốt hơn, giảm bớt dư thừa dữ liệu.

Sự dư thừa

- ▶ Ví dụ: Xét quan hệ NhanVien(TenNV, HoNV, NS, Dchi, GT, Luong, BangCap)

TENNV	HONV	NS	DCHI	GT	LUONG	BANGCAP
Tung	Nguyen	12/08/1955	638 NVC Q5	Nam	40000	Trung học
Nhu	Le	06/20/1951	291 HVH QPN	Nu	43000	Trung học
Nhu	Le	06/20/1951	291 HVH QPN	Nu	43000	Đại học
Hung	Nguyen	09/15/1962	Ba Ria VT	Nam	38000	Thạc sỹ

- ▶ Quan hệ SinhVien(MASV, TenSV, NS, DiaChi, MonHoc)

- ▶ Giải pháp:

- ▶ Tách, xóa
- ▶ Tạo bảng mới

MaSV	TenSV	NS	DiaChi	MonHoc
K140001	Nguyen T	8/8/2002	HaNoi	CSDL
K140001	Nguyen T	8/8/2002	HaNoi	Web
K140002	Tran T	18/8/2002	HaNam	CSDL

Sự dư thừa

- ▶ Một số quy tắc
 - ▶ Rõ ràng về mặt ngữ nghĩa, tránh các phụ thuộc giữa các thuộc tính với nhau
 - ▶ Tránh sự trùng lặp về nội dung -> đảm bảo tránh được các dị thường khi nhập dữ liệu
 - ▶ Tránh đặt các thuộc tính có nhiều giá trị NULL
 - ▶ Ảnh hưởng tới các phép kết nối
 - ▶ Thiết kế các lược đồ quan hệ sao cho chúng có thể được nối với điều kiện bằng trên các thuộc tính là khoá chính hoặc khoá ngoại theo cách đảm bảo không sinh ra các bộ “giả”

Lý thuyết về chuẩn hoá dựa trên các phụ thuộc hàm: sẽ là nền tảng cơ sở để thực hiện việc phân tích và chuẩn hoá lược đồ quan hệ

Mục đích của chuẩn hoá

- ▶ Xác định được 1 tập các lược đồ quan hệ cho phép tìm kiếm thông tin một cách dễ dàng, đồng thời tránh được dư thừa dữ liệu
- ▶ Hướng tiếp cận: Tách các lược đồ quan hệ “có vấn đề” thành những lược đồ quan hệ “chuẩn hơn”

Phụ thuộc hàm

- ▶ Xét quan hệ **R**: SinhVien(**MaSV**,Ho,Ten,NS,GT,NoiSinh,MaLop)
- ▶ MaSV là khoá chính:
- ▶ Từ MaSV có thể xác định được các thuộc tính còn lại hay Các thuộc tính PHỤ THUỘC HÀM vào MaSV
- ▶ Ký hiệu:
 - ▶ MaSV-> Ho
 - ▶ MaSV->Ten
 - ▶ MaSV->NgS
 - ▶ MaSV->GT
 - ▶ MaSV->NoiSinh

MASV	HO	TEN	NGSINH	GIOITINH	NOISINH	MALOP
K1101	Nguyen Van	A	27/1/1986	Nam	TpHCM	K11
K1102	Tran Ngoc	Han	14/3/1986	Nu	Kien Giang	K11
K1103	Ha Duy	Lap	18/4/1986	Nam	Nghe An	K11
K1104	Tran Ngoc	Linh	30/3/1986	Nu	Tay Ninh	K11
K1105	Tran Minh	Long	27/2/1986	Nam	TpHCM	K11
K1106	Le Nhat	Minh	24/1/1986	Nam	TpHCM	K11

Phụ thuộc hàm

- ▶ Định nghĩa phụ thuộc hàm (PTH) trong 1 quan hệ R như sau:
- ▶ Cho $R(U)$ là một quan hệ, U là tập các thuộc tính
- ▶ X, Y là tập con của U : $X, Y \subseteq U$
- ▶ X xác định hàm Y hay Y phụ thuộc hàm vào X nếu
 - ▶ Với \forall quan hệ r xác định trên $R(U)$ và với 2 bộ t_1 và t_2 bất kỳ mà $t_1[X] = t_2[X]$ thì $t_1[Y] = t_2[Y]$

Phụ thuộc hàm

► Ví dụ:

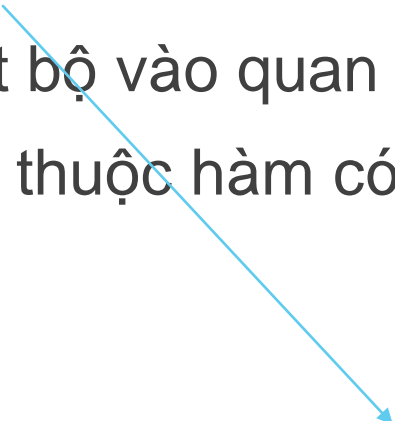
Quan Hệ R: NhanVien

X: HoTen có thể suy diễn ra tất cả các thuộc tính khác

HoTen	NgaySinh	MaPB	TenPB
Nguyễn Văn A	1/1/1980	PB01	Hành chính
Nguyễn Văn B	20/2/1981	PB02	Tổng hợp
Trần C	13/6/1981	PB03	Dự án

Nếu thêm một bộ vào quan hệ NhanVien:

Tính chất phụ thuộc hàm có đúng không?



Trần C	10/2/1982	PB01	Hành chính
--------	-----------	------	------------

Làm nhanh!

- ▶ Cho quan hệ R như sau:
- ▶ Trong các phụ thuộc hàm sau, PTH nào không thoả mãn

- ▶ $A \rightarrow D$, $A \rightarrow C$, $B \rightarrow A$, $C \rightarrow D$, $D \rightarrow C$, $D \rightarrow A$

A	B	C	D
x	u	x	y
y	x	z	x
z	y	y	y
y	z	w	z

Hệ tiên đề Amstrong

- ▶ Amstrong đã đưa ra hệ luật tiên đề còn được gọi là hệ luật dẫn Amstrong sau:
- ▶ Với $X, Y, Z, W \subseteq U$, U là tập thuộc tính. Phụ thuộc hàm có các tính chất sau:
 - ▶ (1) Tính phản xạ (reflexivity)
 - ▶ Nếu $Y \subseteq X$ thì $X \rightarrow Y$
 - ▶ (2) Tính tăng trưởng (augmentation)
 - ▶ Nếu $X \rightarrow Y$ thì $XZ \rightarrow YZ$
 - ▶ (3) Tính bắc cầu (transitivity)
 - ▶ Nếu $X \rightarrow Y$ và $Y \rightarrow Z$ thì $X \rightarrow Z$

Hệ tiên đề Amstrong

- ▶ Từ hệ tiên đề Amstrong suy ra một số tính chất sau:
- ▶ (1) Tính phân rã sau
 - ▶ Nếu $X \rightarrow YZ$ thì $X \rightarrow Y$ và $X \rightarrow Z$
- ▶ (2) Tính tách:
 - ▶ Nếu $X \rightarrow Y$ và $Z \subseteq Y$ thì $X \rightarrow Z$
- ▶ (3) Tính kết hợp
 - ▶ Nếu $X \rightarrow Y$ và $X \rightarrow Z$ thì $X \rightarrow YZ$
- ▶ (4) Tính tựa bắc cầu
 - ▶ Nếu $X \rightarrow Y$ và $YZ \rightarrow W$ thì $XZ \rightarrow W$

- ▶ (1) Tính phản xạ (reflexivity)
 - ▶ Nếu $Y \subseteq X$ thì $X \rightarrow Y$
- ▶ (2) Tính tăng trưởng (augmentation)
 - ▶ Nếu $X \rightarrow Y$ thì $XZ \rightarrow YZ$
- ▶ (3) Tính bắc cầu (transitivity)
 - ▶ Nếu $X \rightarrow Y$ và $Y \rightarrow Z$ thì $X \rightarrow Z$

Làm nhanh!

- ▶ Cho lược đồ quan hệ $R(A, B, C, D, E, G, H)$ và tập phụ thuộc hàm
- ▶ $FD = \{AB \rightarrow C, B \rightarrow D, CD \rightarrow E, CE \rightarrow GH, G \rightarrow A\}$
- ▶ Chứng tỏ $AB \rightarrow E$ được suy diễn từ phụ thuộc hàm FD.

▶ Từ hệ tiên đề Armstrong suy ra một số tính chất sau:

- ▶ (1) Tính phân rã sau
 - ▶ Nếu $X \rightarrow YZ$ thì $X \rightarrow Y$ và $X \rightarrow Z$
 - ▶ (2) Tính tách:
 - ▶ Nếu $X \rightarrow Y$ và $Z \subseteq Y$ thì $X \rightarrow Z$
 - ▶ (3) Tính kết hợp
 - ▶ Nếu $X \rightarrow Y$ và $X \rightarrow Z$ thì $X \rightarrow YZ$
 - ▶ (4) Tính tựa bắc cầu
 - ▶ Nếu $X \rightarrow Y$ và $YZ \rightarrow W$ thì $XZ \rightarrow W$
-
- ▶ $AB \rightarrow C$ (1)
 - ▶ $AB \rightarrow AB$ (2) Tính phản xạ
 - ▶ $AB \rightarrow B$ (3) Tính phân rã
 - ▶ $B \rightarrow D$ (4) lấy từ F
 - ▶ $AB \rightarrow D$ (5) Tính bắc cầu 3 + 4
 - ▶ $AB \rightarrow CD$ (6) Tính kết hợp 1 + 5
 - ▶ $CD \rightarrow E$ (7) Lấy từ F
 - ▶ $AB \rightarrow E$ Tính bắc cầu 6 + 7

Làm nhanh!

- ▶ Cho lược đồ quan hệ $R(A, B, C, D, E, G, H)$ và tập phụ thuộc hàm
- ▶ $\mathbf{FD} = \{AB \rightarrow C, B \rightarrow D, CD \rightarrow E, CE \rightarrow GH, G \rightarrow A\}$
- ▶ Chứng tỏ $AB \rightarrow E$ được suy diễn từ phụ thuộc hàm \mathbf{FD} .
 $CD \rightarrow G$

- ▶ (1) Tính phản xạ (reflexivity)

- ▶ Nếu $Y \subseteq X$ thì $X \rightarrow Y$

- ▶ (2) Tính tăng trưởng (augmentation)

- ▶ Nếu $X \rightarrow Y$ thì $XZ \rightarrow YZ$

- ▶ (3) Tính bắc cầu (transitivity)

- ▶ Nếu $X \rightarrow Y$ và $Y \rightarrow Z$ thì $X \rightarrow Z$

- ▶ Từ hệ tiên đề Amstrong suy ra một số tính chất sau:

- ▶ (1) Tính phân rã sau

- ▶ Nếu $X \rightarrow YZ$ thì $X \rightarrow Y$ và $X \rightarrow Z$

- ▶ (2) Tính tách:

- ▶ Nếu $X \rightarrow Y$ và $Z \subseteq Y$ thì $X \rightarrow Z$

- ▶ (3) Tính kết hợp

- ▶ Nếu $X \rightarrow Y$ và $X \rightarrow Z$ thì $X \rightarrow YZ$

- ▶ (4) Tính tựa bắc cầu

- ▶ Nếu $X \rightarrow Y$ và $YZ \rightarrow W$ thì $XZ \rightarrow W$

Làm nhanh!

- ▶ Cho lược đồ quan hệ $R(A, B, C, D, E, G, H)$ và tập phụ thuộc hàm
- ▶ $FD = \{AB \rightarrow C, B \rightarrow D, CD \rightarrow E, CE \rightarrow GH, G \rightarrow A\}$
- ▶ Chứng tỏ $AB \rightarrow E$ được suy diễn từ phụ thuộc hàm FD .

- ▶ **Ví dụ 1:** cho lược đồ quan hệ $R(U)$ và tập phụ thuộc hàm F như sau:

- ▶ $R(A, B, C, D, E)$
- ▶ $F = \{AB \rightarrow C, AC \rightarrow B, BC \rightarrow DE\}$
- ▶ Tìm khóa K

- ▶ $AB \rightarrow C$ (1)
- ▶ $AB \rightarrow AB$ (2) Tính phản xạ
- ▶ $AB \rightarrow B$ (3) Tính phân rã
- ▶ $B \rightarrow D$ (4) lấy từ F
- ▶ $AB \rightarrow D$ (5) Tính bắc cầu 3 +4
- ▶ $AB \rightarrow CD$ (6) Tính kết hợp 1 +5
- ▶ $CD \rightarrow E$ (7) Lấy từ F
- ▶ $AB \rightarrow E$ Tính bắc cầu 6+ 7

Bao đóng F^+ của tập phụ thuộc hàm F

- ▶ **Đ/N :** Bao đóng của tập phụ thuộc hàm F là tập lớn nhất các phụ thuộc hàm có thể được suy diễn logic từ F
- ▶ Ký hiệu : F^+
- ▶ Ví dụ
- ▶ $R = \{A, B, C, D\}$
- ▶ $F = \{A \rightarrow B, B \rightarrow C, A \rightarrow D, B \rightarrow D\}$
- ▶ $F^+ = \{A \rightarrow B, B \rightarrow C, A \rightarrow D, B \rightarrow D, A \rightarrow C, A \rightarrow BD, A \rightarrow BCD, A \rightarrow BC, A \rightarrow CD, B \rightarrow CD\}$
- ▶ **Các tính chất của F^+**
- ▶ Tính phản xạ: $F \subseteq F^+$
- ▶ b. Tính đơn điệu: $F \subseteq G \Rightarrow F^+ \subseteq G^+$
- ▶ c. Tính lũy đẳng: $F^{++} = F^+$

Bao đóng thuộc tính X^+

- ▶ Định nghĩa: Bao đóng của tập thuộc tính X đối với tập phụ thuộc hàm F , ký hiệu là X^+ :
 - ▶ X^+ là tập tất cả các thuộc tính T có thể suy dẫn từ X nhờ tập bao đóng của các phụ thuộc hàm F^+
 - ▶ $X^+ = \{T \in U \mid X \rightarrow T \in F^+\}$: U là các thuộc tính trong R

Example:

$R = \{A, B, C, D\}$

$F = \{A \rightarrow B, B \rightarrow C, A \rightarrow D, B \rightarrow D\}$

$F^+ = \{A \rightarrow B, B \rightarrow C, A \rightarrow D, B \rightarrow D, A \rightarrow C, A \rightarrow BD, A \rightarrow BCD, A \rightarrow BC, A \rightarrow CD, B \rightarrow CD\}$

Thuật toán tìm bao đóng của tập thuộc tính

- ▶ Thuật toán tìm bao đóng của tập thuộc tính
- ▶ Cho $R = \{A_1, \dots, A_n\}$. F là PTH trên R . X là tập thuộc tính
- ▶ Ta xây dựng tập X^0, \dots, X^k như sau:
- ▶ Bước 1: $X^0 = X$
- ▶ Bước 2: $X^{(i+1)} = X^i Z^i$ với $Z^i = \{A : A \in R \mid X^i \text{ và } X^i \rightarrow A \in F^+\}$
- ▶ Tập $X^0, X^1 \dots X^k$ là tập tăng dần và tập R là hữu hạn nên sau một một số bước thuật toán sẽ kết thúc.
- ▶ Bước 3: Nếu $X^k = X^{k+1}$ hoặc $X^k = U$ thì X^k là tập X^+ , nếu không quay lại bước 2.

Làm nhanh!

- ▶ Ví dụ: Cho lược đồ quan hệ $R(A, B, C, D, E, G, H)$ và tập phụ thuộc hàm:

$$F = \{ \begin{array}{l} f1: B \rightarrow A, \\ f2: DA \rightarrow CE, \\ f3: D \rightarrow H, \\ f4: GH \rightarrow C, \\ f5: AC \rightarrow D \end{array} \}$$

$K0 = ABCDEGH;$

$K1 = BCDEGH$; do $B \rightarrow A$ (f1);

- ▶ Tìm AC^+

$F = \{ \begin{array}{l} f1: B \rightarrow A, \\ f2: DA \rightarrow CE, \\ f3: D \rightarrow H, \\ f4: GH \rightarrow C, \\ f5: AC \rightarrow D \end{array} \}$

- ▶ Bước 1: $X^0 = AC$
- ▶ Bước 2: Từ $f1$ đến $f4$ không thỏa, $f5$ thỏa nên $X^1 = AC \cup D = ACD$
- ▶ Lặp lại bước 2:
 - ▶ $f1$ không thỏa, $f2$ thỏa nên $X^2 = ACD \cup CE = ACDE$
 - ▶ $f3$ thỏa nên $X^3 = ACDE \cup H = ACDEH$
 - ▶ $f4$ không thỏa, $f5$ đã thỏa
- ▶ Lặp lại bước 2: $f2$, $f3$ và $f5$ đã thỏa, $f1$ và $f4$ không thỏa. Nên $X^4 = X^3 = ACDEH$
- ▶ Vậy $AC^+ = ACDEH$

- ▶ Vậy $AC \rightarrow E$ có thuộc F^+ ????
- ▶ Rõ ràng $AC^+ = ACDEH$
- ▶ Vì $E \in AC^+$ nên $AC \rightarrow E \in F^+$

Do đâu có phụ thuộc hàm?

- ▶ Tính khóa (duy nhất)” của các thuộc tính.
- ▶ Các ràng buộc trên miền giá trị và ràng buộc ứng dụng.
- ▶ Các ràng buộc của thế giới thực.
 - ▶ MaID → Ten Nguoi
- ▶ Phụ thuộc hàm giúp chúng ta định nghĩa “khóa”

Khóa của sơ đồ quan hệ

- ▶ Trong lược đồ quan hệ, một số thuộc tính đóng vai trò quan trọng, và từ các thuộc tính này có thể xác định các thuộc tính khác.
- ▶ Khái niệm sơ đồ quan hệ
- ▶ Cho lược đồ quan hệ $R=\{A^1, A^2, \dots, A^n\}$, F là tập các phụ thuộc hàm trên R . Sơ đồ quan hệ là cặp R, F như trên, ký hiệu SĐQH là $W=<R, F>$
- ▶ Sơ đồ quan hệ là một lược đồ quan hệ và tập phụ thuộc hàm trên nó.

Khóa

- ▶ Cho lược đồ quan hệ $R(U)$ với $U = \{ A_1, \dots, A_n \}$ có tập các phụ thuộc hàm F và $\emptyset \neq K \subseteq U$.

K là một khoá của R nếu:

- ▶ $K \subseteq U \in F^+$;
 - ▶ Không tồn tại $X \subset K$ mà $X \rightarrow U \in F^+$.
- ▶ Điều này tương đương với
- ▶ $K^+ = U$;
 - ▶ $X \subset K$ thì $X^+ \neq U$.
- ▶ Hay
- ▶ $K^+ = U$; (1)
 - ▶ và $\forall A \in K$ thì $(K - A)^+ \neq U$. (2)

- ▶ Nếu K chỉ thoả mãn điều kiện (1) $K^+ = U$ thì K được gọi là khoá bao hàm.
- ▶ Thoả mãn cả 2 điều kiện thì K là khoá tối thiểu.

Khóa

- ▶ Thuộc tính ***A*** được gọi là ***thuộc tính khóa*** nếu $A \in K$ với K là khóa bất kỳ của R . Ngược lại A được gọi là ***thuộc tính không khóa***.
- ▶ Một lược đồ quan hệ có thể có nhiều khóa và tập thuộc tính không khóa

Khóa

► Định lý:

- Nếu K là khoá của quan hệ $W = \langle R, F \rangle$, r là quan hệ trên R thì mọi cặp phần tử khác nhau t_1, t_2 của r ta luôn có $t_1.k \neq t_2.k$
- Ngược lại nếu K là tập tối thiểu và với mọi quan hệ r trên R mà mọi cặp t_1, t_2 của r mà $t_1.k \neq t_2.k$ thì K là khoá của quan hệ W

Thuật toán tìm khoá

- ▶ Do bao đóng của k là $R_k(U)$ nên coi ban đầu $k = U$. Sau đây loại trừ các phần tử và kiểm tra xem $k^+ = U$
- ▶ Input $W = \langle R, F \rangle$
- ▶ Output: k - khoá của W

Thuật toán:

Bước 1: Đặt $k = U$

Bước 2: Lặp lại quá trình loại khỏi các phần tử A mà $(k - A)^+ = U$

Làm nhanh!

- ▶ **Ví dụ 1:** cho lược đồ quan hệ $R(U)$ và tập phụ thuộc hàm F như sau:
 - ▶ $R(A,B,C,D,E)$
 - ▶ $F = \{AB \twoheadrightarrow C, AC \rightarrow B, BC \rightarrow DE\}$
 - ▶ Tìm khóa K

▶ **Ví dụ 1:** cho lược đồ quan hệ $R(U)$ và tập phụ thuộc hàm F như sau:

- ▶ $R(A, B, C, D, E)$
- ▶ $F = \{AB \twoheadrightarrow C, AC \rightarrow B, BC \rightarrow DE\}$
- ▶ Tìm khóa K

▶ **Giải:**

- ▶ B1: $K^+ = U \Rightarrow K^+ = ABCDE$
- ▶ B2: $(K \setminus A)^+ \Rightarrow (BCDE)^+ = BCDE \neq U \Rightarrow K = ABCDE$
- ▶ B3: $(K \setminus B)^+ \Rightarrow (ACDE)^+ = ABCDE = U \Rightarrow K = ACDE$
- ▶ B4: $(K \setminus C)^+ \Rightarrow (ADE)^+ = ADE \neq U \Rightarrow K = ACDE$
- ▶ B5: $(K \setminus D)^+ \Rightarrow (ACE)^+ = ACEBD = U \Rightarrow K = ACE$
- ▶ B6: $(K \setminus E)^+ \Rightarrow (AC)^+ = ACBDE = U \Rightarrow \mathbf{K = AC}$

Chuẩn hoá

- ▶ Mục đích việc thiết kế CSDL
 - ▶ Xác định được 1 tập các lược đồ quan hệ cho phép tìm kiếm thông tin một cách dễ dàng, đồng thời tránh được dư thừa dữ liệu.

Chuẩn hoá

- ▶ Vấn đề đặt ra:
 - ▶ Có cần phải tinh chỉnh thiết kế nữa hay không?
 - ▶ Thiết kế đã là tốt hay chưa?
- ▶ Mỗi dạng chuẩn là một tập các điều kiện trên lược đồ nhằm đảm bảo các tính chất của nó (liên quan tới dư thừa và bất thường trong cập nhật)
- ▶ Chuẩn hóa dữ liệu: quá trình phân tích lược đồ quan hệ dựa trên phụ thuộc hàm (FD) và các khóa chính để đạt được:
 - ▶ Cực tiểu sự dư thừa
 - ▶ Cực tiểu các phép cập nhật bất thường

Chuẩn hóa

▶ Thủ tục chuẩn hóa cung cấp:

- ▶ Một cơ cấu hình thức để phân tích các lược đồ quan hệ dựa trên **các khóa** của nó và **các phụ thuộc hàm giữa các thuộc tính** của nó
- ▶ Một loạt các **kiểm tra dạng chuẩn** có thể thực hiện trên các lược đồ quan hệ riêng rẽ sao cho CSDL có thể được chuẩn hóa đến một mức cần thiết.

▶ Tính chất

- ▶ Nói không mất mát (hoặc nói không phụ thêm)
- ▶ Bảo toàn sự phụ thuộc: Nó bảo đảm rằng từng phụ thuộc hàm sẽ được biểu hiện trong các quan hệ riêng rẽ nhận được sau khi tách.

Các dạng chuẩn

- ▶ Mục đích: Mỗi dạng chuẩn đảm bảo ngăn ngừa (giảm thiểu) một số các dạng dư thừa hay dị thường dữ liệu
 - ▶ Các dạng chuẩn hay sử dụng:
 - ▶ Dạng chuẩn 1 (1NF)
 - ▶ Dạng chuẩn 2 (2NF)
 - ▶ Dạng chuẩn 3 (3NF) -> Dạng chuẩn Boye-Code (BCNF)
 - ▶ Dạng chuẩn 4NF, 5NF do tính đa trị và phụ thuộc hàm nối

Một số khái niệm liên quan đến các chuẩn hoá

▶ Thuộc tính khoá:

- ▶ Cho lược đồ quan hệ $R(A_1, A_2, \dots, A_n)$, thuộc tính K được gọi là **thuộc tính khoá** nếu K là một thuộc tính thành phần trong một khoá nào đó của R , ngược lại được gọi là **thuộc tính không khoá**
- ▶ Ví dụ: $R(A, B, C, D)$, $F=\{AB \rightarrow C, B \rightarrow D, BC \rightarrow A\}$
- ▶ Trong ví dụ trên, lược đồ R có 2 khoá là AB, BC . Khi đó A, B, C là thuộc tính khoá, D là thuộc tính không khoá.

▶ Giá trị nguyên tố (đơn)

- ▶ Giá trị nguyên tố là giá trị không phân nhỏ được nữa.
- ▶ Ví dụ: ChiTietLanThi: CSDL Thuc Hanh 7, Ly Thuyet 7
- ▶ Phân tiếp ra được là MônThi, Loại HìnhThi, Điểm số

Dạng chuẩn 1. 1NF

- ▶ Lược đồ R ở dạng chuẩn 1 nếu mọi thuộc tính đều mang giá trị nguyên tố.
- ▶ Trong bài toán xét dạng chuẩn, dạng chuẩn thấp nhất là dạng chuẩn 1.
- ▶ Ví dụ: Quan hệ SinhVien_Thi

MaSV	MaMH	ChiTietThiMon
K140001	CSDL	Thuc Hanh 7, Ly Thuyet 7
K140002	CSDL	Thuc Hanh 6, Ly Thuyet 7
K140003	CSDL	Thuc Hanh 7, Ly Thuyet 6
K140004	CSDL	Thuc Hanh 5, Ly Thuyet 7

- ▶ Nhận thấy rằng ChiTietThiMon không mang giá trị nguyên tố, do đó SinhVien_Thi không đạt dạng chuẩn 1.

Dạng chuẩn 2. 2NF

- ▶ **Phụ thuộc hàm đầy đủ:** Một phụ thuộc hàm $X \rightarrow Y$ là một phụ thuộc hàm đầy đủ nếu loại bỏ bất kỳ thuộc tính A nào ra khỏi X thì phụ thuộc hàm không còn đúng nữa.
 - ▶ $\forall A, A \in X, (X - \{A\}) \rightarrow Y$: là sai.
- ▶ **Phụ thuộc hàm bộ phận:** Một phụ thuộc hàm $X \rightarrow Y$ là phụ thuộc bộ phận nếu có thể bỏ một thuộc tính $A \in X$, ra khỏi X phụ thuộc hàm vẫn đúng, điều đó có nghĩa là với
 - ▶ $\exists A \in X, (X - \{A\}) \rightarrow Y$

Dạng chuẩn 2. 2NF

- ▶ Lược đồ R ở dạng chuẩn 2 nếu thoả:
 - ▶ (1) R đạt dạng chuẩn 1
 - ▶ (2) Mọi thuộc tính không khóa của R đều phụ thuộc đầy đủ vào khóa
- ▶ Cách làm:
 - ▶ Với các quan hệ có thuộc tính khóa đơn thì ko phải xét
 - ▶ Chỉ kiểm tra các lược đồ có chứa phụ thuộc hàm bộ phận

- ▶ Xét quan hệ MONHOC (MaMH, TenMH, STC, Loai)

MaMH	TenMH	STC	Loai
CT101	Nhập môn tin học	4	BB
CT102	TH kỹ năng máy tính	5	BB
TN311	Xác suất thống kê	3	BB
CT103	Thiết kế Web	4	TC
CT110	Nguyên lý lập trình 2	5	BB

- ▶ Xét quan hệ SinhVien_MonHoc(MaSV, TenSV, MaMH, TenMH, Diem, GioiTinh, DiaChi):
- ▶ Khoá chính là: MaSV, MaMH
 - ▶ Thuộc tính khoá: MaSV, MaMH
- ▶ MaSV-> TenSV, GioiTinh, DiaChi
- ▶ MaMH-> TenMH

MaSV	TenSV	MaMH	TenMH	Diem	GioiTinh	DiaChi
K140001	Nguyễn Văn A	CSDL	Cơ Sở dữ liệu	7	Nam	Hà nội, Hà đông,
K140001	Nguyễn Văn A	Web	Web design	8	Nam	Hà nội, Hà đông,
K140002	Trần Thị B	CSDL	Cơ Sở dữ liệu	7	Nữ	Đông hưng, Thái bình,
K140002	Trần Thị B	Web	Web design	8	Nữ	Đông hưng, Thái bình,
K140003	Lê Thanh C	CSDL	Cơ Sở dữ liệu	8	Nam	Vinh, Nghệ An

- ▶ Create table SinhVien(
 - ▶ maSV char primary key,
 - ▶ tenSV Char not null
- ▶);
- ▶ Create table MonHoc(
 - ▶ maMH char primary key,
 - ▶ tenMH Char not null
- ▶);
- ▶ Create table SV_MH(
 - ▶ CodeSV char not null,
 - ▶ MaMH char not null,
 - ▶ Diem float,
 - ▶ CONSTRAINT xxxx FOREIGN KEY (CodeSV) REFERENCES SinhVien(maSV)
- ▶);

MaSV	TenSV	MaMH	TenMH	Diem	GioiTinh	DiaChi
K140001	Nguyễn Văn A	CSDL	Cơ Sở dữ liệu	7	Nam	Hà nội, Hà đông,
K140001	Nguyễn Văn A	Web	Web design	8	Nam	Hà nội, Hà đông,
K140002	Trần Thị B	CSDL	Cơ Sở dữ liệu	7	Nữ	Đông hưng, Thái bình,
K140002	Trần Thị B	Web	Web design	8	Nữ	Đông hưng, Thái bình,
K140003	Lê Thanh C	CSDL	Cơ Sở dữ liệu	8	Nam	Vinh, Nghệ An

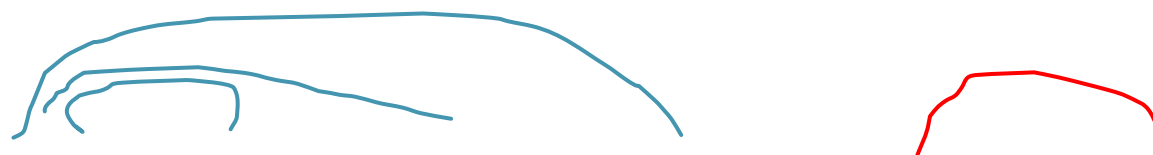
MaSV	TenSV	GioiTinh	DiaChi
K140001	Nguyễn Văn A	Nam	Hà nội, Hà đông,
K140002	Trần Thị B	Nữ	Đông hưng, Thái bình,
K140003	Lê Thanh C	Nam	Vinh, Nghệ An

MaSV	MaMH	Diem
K140001	CSDL	7
K140001	Web	8
K140002	CSDL	7
K140002	Web	8
K140003	CSDL	8

MaMH	TenMH
CSDL	Cơ Sở dữ liệu
Web	Web design

Dạng chuẩn 3 (3NF)

- ▶ Đ/N: Một sơ đồ quan hệ R được coi là ở dạng chuẩn 3 nếu
 - ▶ Sơ đồ quan hệ này ở 2NF
 - ▶ Mọi thuộc tính không khóa đều không phụ thuộc bắc cầu vào khóa chính
- ▶ Xét ví dụ: quan hệ SinhVien_Khoa



- ▶ SinhVien_Khoa(MaSV, TenSV, GioiTinh, DiaChi, MaKhoa, TenKhoa)
- ▶ Tất cả các thuộc tính phải phụ thuộc vào thuộc tính khóa
 - ▶ Một vài thuộc tính phụ thuộc vào thuộc tính không phải là khóa
 - ▶ Để chuẩn hóa : Tách nhóm các thuộc tính đó thành quan hệ mới



▶ SinhVien_Khoa(MaSV,TenSV,GioiTinh,DiaChi, MaKhoa,TenKhoa)



▶ SinhVien(MaSV,TenSV,GioiTinh,DiaChi) ; Khoa(MaKhoa,TenKhoa)

Tóm tắt chuẩn 1NF-3NF

NF	Nhận biết	Chuẩn Hoá
1	Quan hệ ko có thuộc tính đa trị và quan hệ lặp	Chuyển tất cả quan hệ lặp hoặc đa trị thành 1 quan hệ mới
2	Phụ thuộc 1 phần vào thuộc tính khóa	Tách thuộc tính phụ thuộc 1 phần thành lược đồ mới, đảm bảo quan hệ với lược đồ liên quan
3	Phụ thuộc ẩn, tồn tại phụ thuộc hàm giữa các thuộc tính ko phải là khóa	Tách các thuộc tính đó thành lược đồ mới

Dạng chuẩn Boyce-Codd

- ▶ Một lược đồ quan hệ R được gọi là ở dạng chuẩn Boyce-Codd (BCNF) nếu nó
 - ▶ Thỏa mãn dạng chuẩn 3NF
 - ▶ Không có thuộc tính phụ thuộc hàm vào thuộc tính không khóa.
- ▶ Xét ví dụ:
- ▶ R (A₁, A₂, A₃, A₄, A₅) Với các phụ thuộc hàm:
 - ▶ A₁, A₂ → A₃, A₄, A₅
 - ▶ A₄ → A₂
- ▶ Đây không phải là chuẩn Boyce-Codd, vẫn còn thuộc tính khóa, và thuộc tính phụ thuộc hàm.

Dạng chuẩn Boyce-Codd

- ▶ Nếu một lược đồ quan hệ không thoả mãn điều kiện BCNF, thủ tục chuẩn hóa bao gồm:
 - ▶ Loại bỏ các thuộc tính khóa phụ thuộc hàm vào thuộc tính không khóa ra khỏi quan hệ
 - ▶ Tách chúng thành một quan hệ riêng có khoá chính là thuộc tính không khóa gây ra phụ thuộc.
- ▶ $R(\underline{A_1}, \underline{A_2}, A_3, A_4, A_5)$ Với các phụ thuộc hàm:
 - ▶ $A_1, A_2 \rightarrow A_3, A_4, A_5$
 - ▶ $A_4 \rightarrow A_2$
- ▶ Lược đồ được tách ra như sau:
 - ▶ $R1(\underline{A_4}, A_2)$
 - ▶ $R2(\underline{A_1}, \underline{A_4}, A_3, A_5)$

▶ Ví dụ: Xét quan hệ Sinh viên, môn học, Giáo Viên:

▶ $SV_MH_GV(\underline{MaSV}, MaMH, MaGV)$

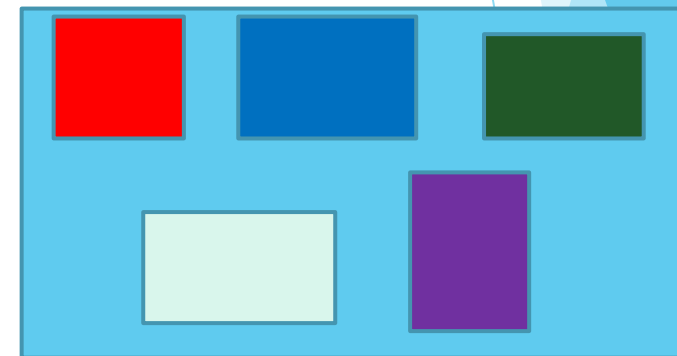
▶ Phụ thuộc hàm $MaSV, MaMH \rightarrow GV$; $MaGV \rightarrow MaMH$

▶ Ta tách:

▶ $SV_MH(\underline{MaSV}, \underline{MaMH})$; $GV_MH(\underline{MaGV}, MaMH)$

Phép tách các lược đồ quan hệ

- ▶ Mục đích:
- ▶ Thay thế một sơ đồ quan hệ $R(A^1, A^2, \dots, A^n)$ bằng một tập các sơ đồ con $\{R_1, R_2, \dots, R_k\}$ trong đó $R_i \subseteq R$ và $R = R_1 \cup R_2 \cup \dots \cup R_k$
- ▶ Yêu cầu của phép tách:
 - ▶ Bảo toàn thuộc tính, ràng buộc
 - ▶ Bảo toàn dữ liệu
- ▶ Tách theo chiều dọc
- ▶ Tách theo chiều ngang



MaSV	TenSV	MaMH	TenMH	Diem	GioiTinh	DiaChi
K140001	Nguyễn Văn A	CSDL	Cơ Sở dữ liệu	7	Nam	Hà nội, Hà đông,
K140001	Nguyễn Văn A	Web	Web design	8	Nam	Hà nội, Hà đông,
K140002	Trần Thị B	CSDL	Cơ Sở dữ liệu	7	Nữ	Đông hưng, Thái bình,
K140002	Trần Thị B	Web	Web design	8	Nữ	Đông hưng, Thái bình,
K140003	Lê Thanh C	CSDL	Cơ Sở dữ liệu	8	Nam	Vinh, Nghệ An

<u>MaSV</u>	<u>TenSV</u>	<u>MaMH</u>	<u>TenMH</u>
K140001	<u>Nguyễn Văn A</u>	CSDL	Cơ <u>Sở dữ liệu</u>
K140001	<u>Nguyễn Văn A</u>	Web	Web design
K140002	<u>Trần Thị B</u>	CSDL	Cơ <u>Sở dữ liệu</u>
K140002	<u>Trần Thị B</u>	Web	Web design
K140003	Lê Thanh C	CSDL	Cơ <u>Sở dữ liệu</u>

<u>Diem</u>	<u>GioiTinh</u>	<u>DiaChi</u>
7	Nam	<u>Hà nội, Hà đông,</u>
8	Nam	<u>Hà nội, Hà đông,</u>
7	<u>Nữ</u>	<u>Đông hưng, Thái bình</u>
8	<u>Nữ</u>	<u>Đông hưng, Thái bình</u>
8	Nam	<u>Vinh, Nghệ An</u>