

CHƯƠNG 06: THIẾT KẾ DỮ LIỆU

Nội dung chi tiết

- Thiết kế luận lý dữ liệu
 - Cấp cao
 - Cấp thấp
- Thiết kế mã
- Bài tập cuối chương

Nội dung chi tiết

- **Thiết kế luận lý dữ liệu**
 - Cấp cao
 - Cấp thấp
- Thiết kế mã
- Bài tập cuối chương

Thiết kế luận lý dữ liệu

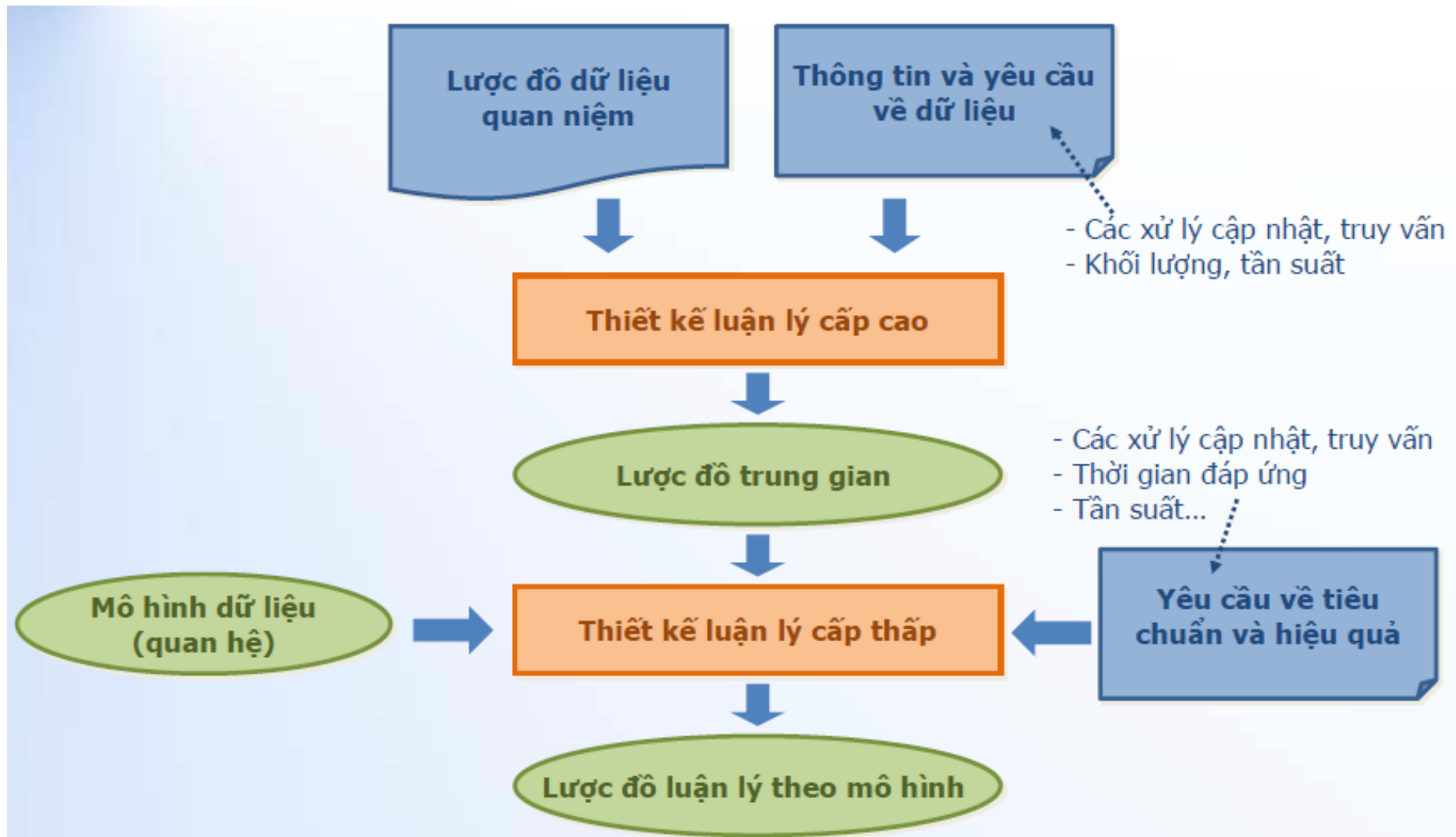
■ Cấp cao

- ◆ Độc lập với mô hình cài đặt
- ◆ Dùng chung cho nhiều loại mô hình dữ liệu

■ Cấp thấp

- ◆ Chuyển đổi lược đồ kết quả của thiết kế luận lý cấp cao sang 1 mô hình dữ liệu nhất định

Thiết kế luận lý dữ liệu



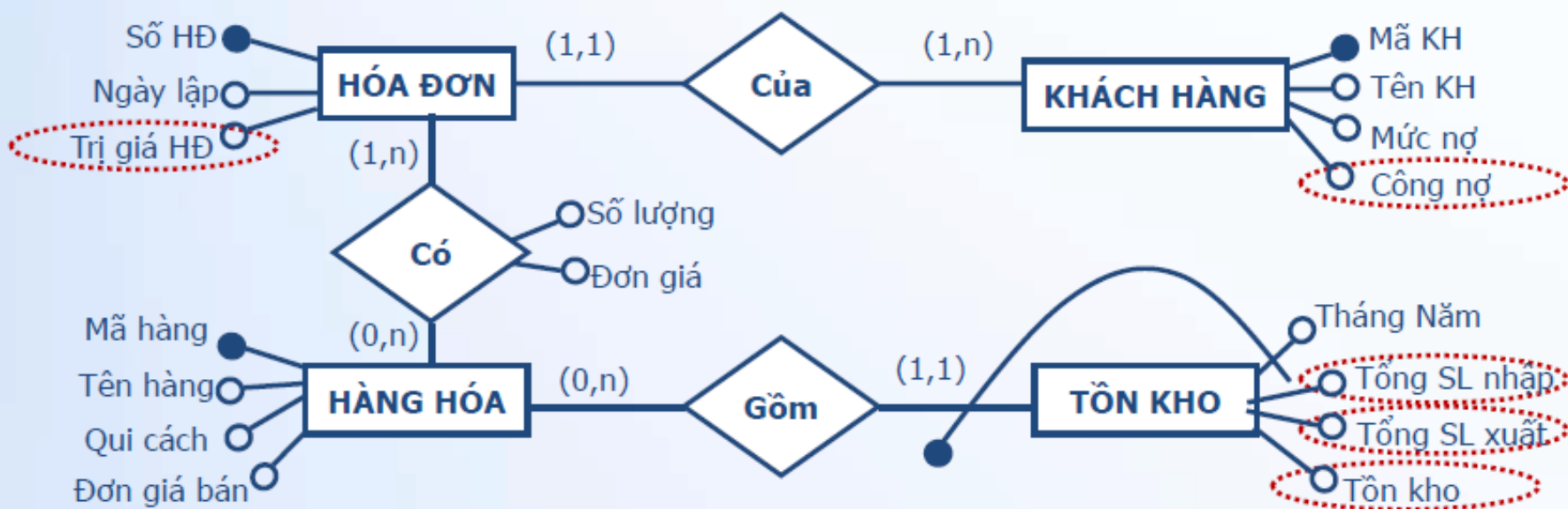
Thiết kế luận lý dữ liệu cấp cao

- Quyết định dữ liệu suy diễn
- Chuyển đổi tổng quát hóa và tập con
- Chọn lựa khóa chính

Quyết định dữ liệu suy diễn

■ Dữ liệu suy diễn

- ◆ Là những thuộc tính mà giá trị của nó có thể tính toán số học từ những thuộc tính khác



Quyết định dữ liệu suy diễn

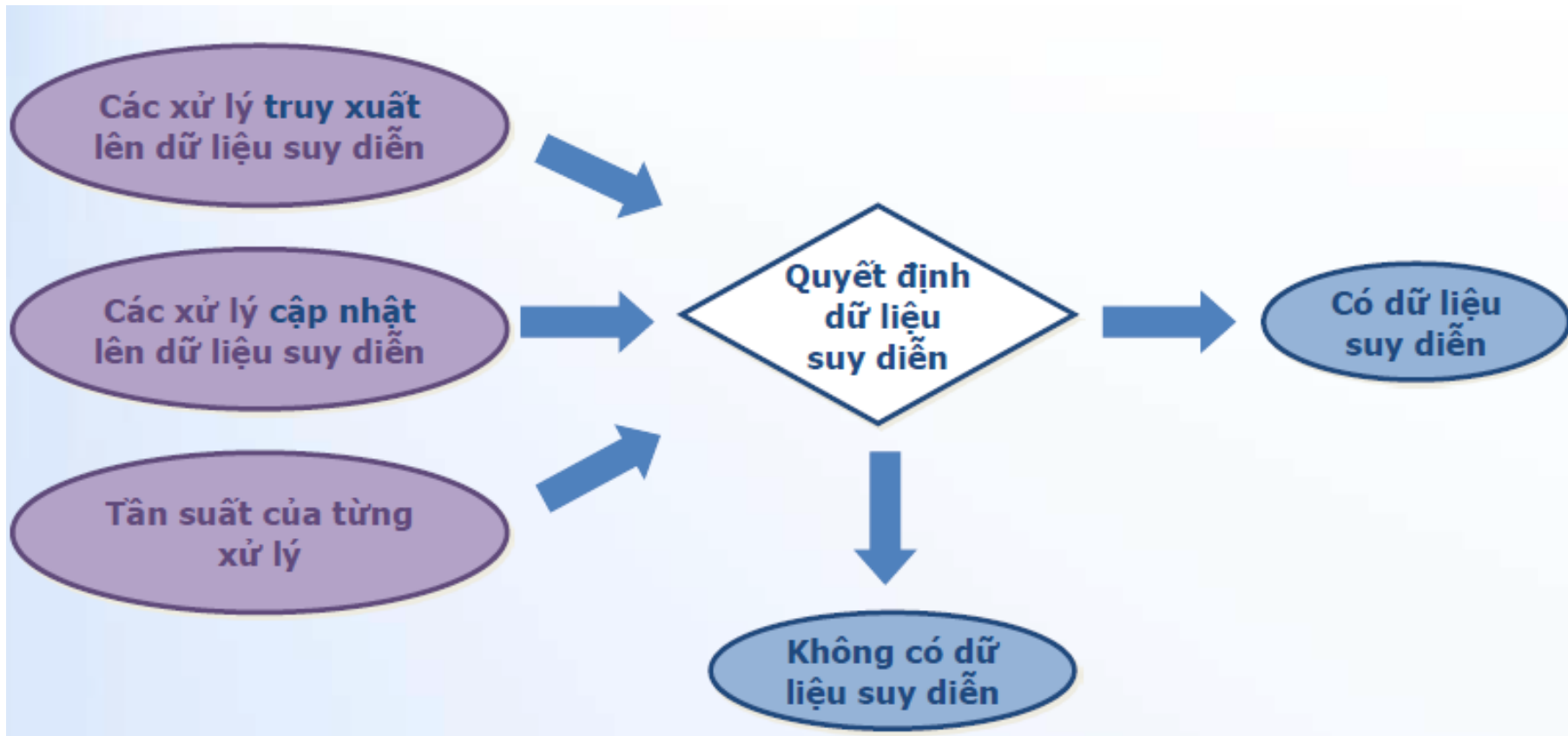
■ Ưu điểm

- ◆ Tăng tốc độ truy vấn
 - Không tính toán lại giá trị tại thời điểm truy vấn

■ Khuyết điểm

- ◆ Khi cập nhật phải kiểm tra tính nhất quán với các dữ liệu liên quan → giảm tốc độ cập nhật
- ◆ Tăng dung lượng lưu trữ
- ◆ Phải lập trình để bảo đảm ràng buộc

Quyết định dữ liệu suy diễn



Ví dụ



Ví dụ

■ Có dữ liệu suy diễn (A)

Tên tác vụ	Khái niệm	Loại	Đọc/Ghi	Tần suất (Ngày)
Mở tài khoản	Tài khoản	Thực thể	Ghi	100
	Khách hàng	Thực thể	Ghi	100
	Cửa	Mỗi kết hợp	Ghi	100
Đọc tổng số dư khách hàng	Khách hàng	Thực thể	Đọc	3000
Gửi tiền	Tài khoản	Thực thể	Đọc	2000
			Ghi	2000
	Khách hàng	Thực thể	Đọc	2000
			Ghi	2000
Rút tiền	Tài khoản	Thực thể	Đọc	1000
			Ghi	1000
	Khách hàng	Thực thể	Đọc	1000
			Ghi	1000
Khách hàng			Đọc	3000
			Ghi	3000

Ví dụ

■ Không có dữ liệu suy diễn (B)

Tên tác vụ	Khái niệm	Loại	Đọc/Ghi	Tần suất (Ngày)
Mở tài khoản	Tài khoản	Thực thể	Ghi	100
	Khách hàng	Thực thể	Ghi	100
	Của	Mỗi kết hợp	Ghi	100
Đọc tổng số dư khách hàng	Khách hàng	Thực thể	Đọc	3000
	Tài khoản	Thực thể	Đọc	3000x2
	Của	Mỗi kết hợp	Đọc	3000x2
Gửi tiền	Tài khoản	Thực thể	Đọc	2000
			Ghi	2000
Rút tiền	Tài khoản	Thực thể	Đọc	1000
			Ghi	1000
Tài khoản			Đọc	6000
Của			Đọc	6000

Ví dụ

- Nếu $A \gg B$
 - ◆ Không sử dụng thuộc tính suy diễn
- Nếu $A \ll B$
 - ◆ Sử dụng thuộc tính suy diễn
- Chú ý
 - ◆ Có những nhu cầu phát sinh 1 cách định tính
→ không thể quyết định bằng phương pháp định lượng

Chuyển đổi tổng quát hóa & tập con

■ Cần thiết khi

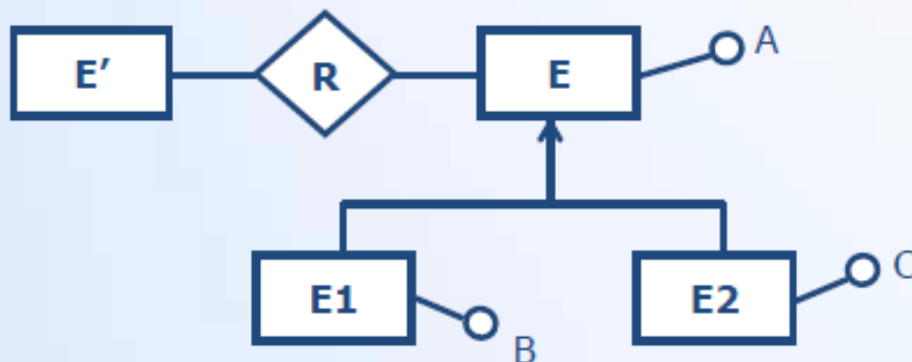
- ◆ Mô hình cài đặt không hỗ trợ
 - Mô hình quan hệ, mạng, phân cấp...

■ Xem xét

- ◆ Các đặc trưng kế thừa của thực-thể-chuyên-biệt từ thực-thể-tổng-quát
 - Thuộc tính, định danh, mối kết hợp
- ◆ Mối kết hợp tổng quát hóa

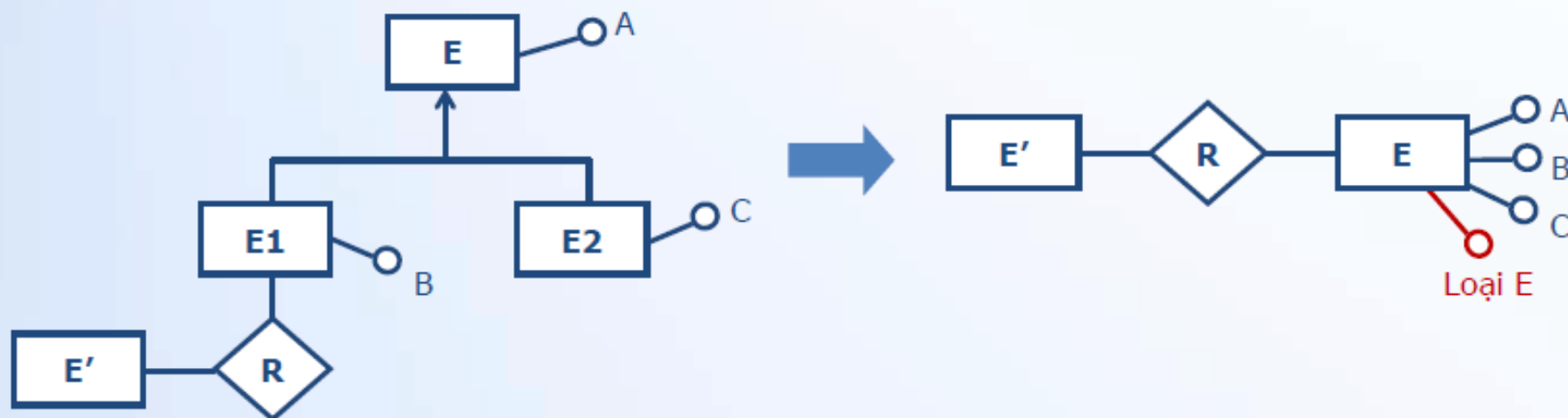
Chuyển đổi tổng quát hóa & tập con

- Phương án chọn lựa
 - ◆ Dùng thực thể tổng quát
 - ◆ Dùng thực thể chuyên biệt
 - ◆ Dùng mối kết hợp

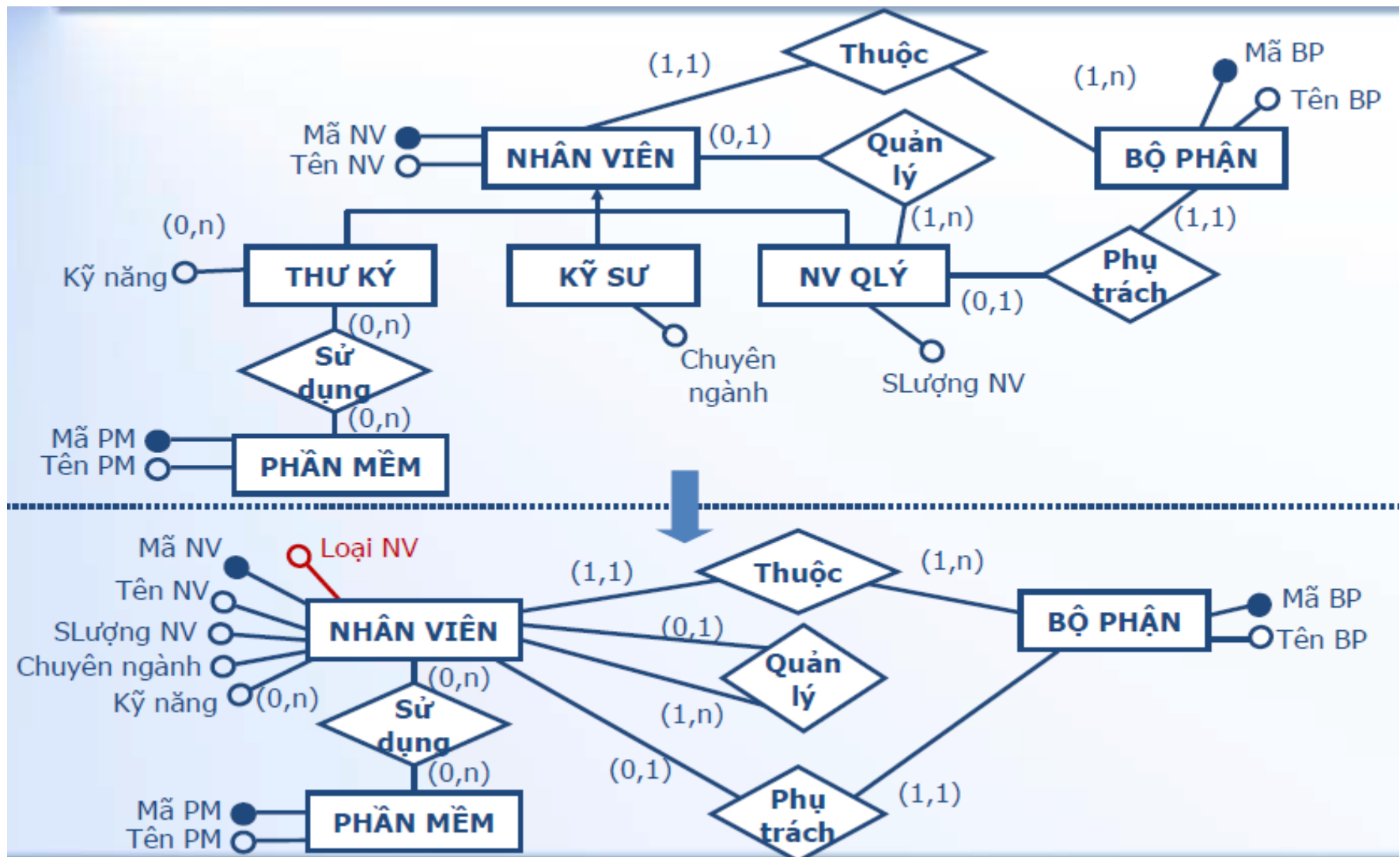


Dùng thực thể tổng quát

- Các đặc trưng của thực-thể-chuyên-biệt sẽ được chuyển sang thực-thể-tổng-quát
- Thêm thuộc tính LoạiE
 - ◆ $\text{MGT}(\text{LoạiE}) = \{E, E1, E2, E1E2\}$



Ví dụ



Dùng thực thể tổng quát

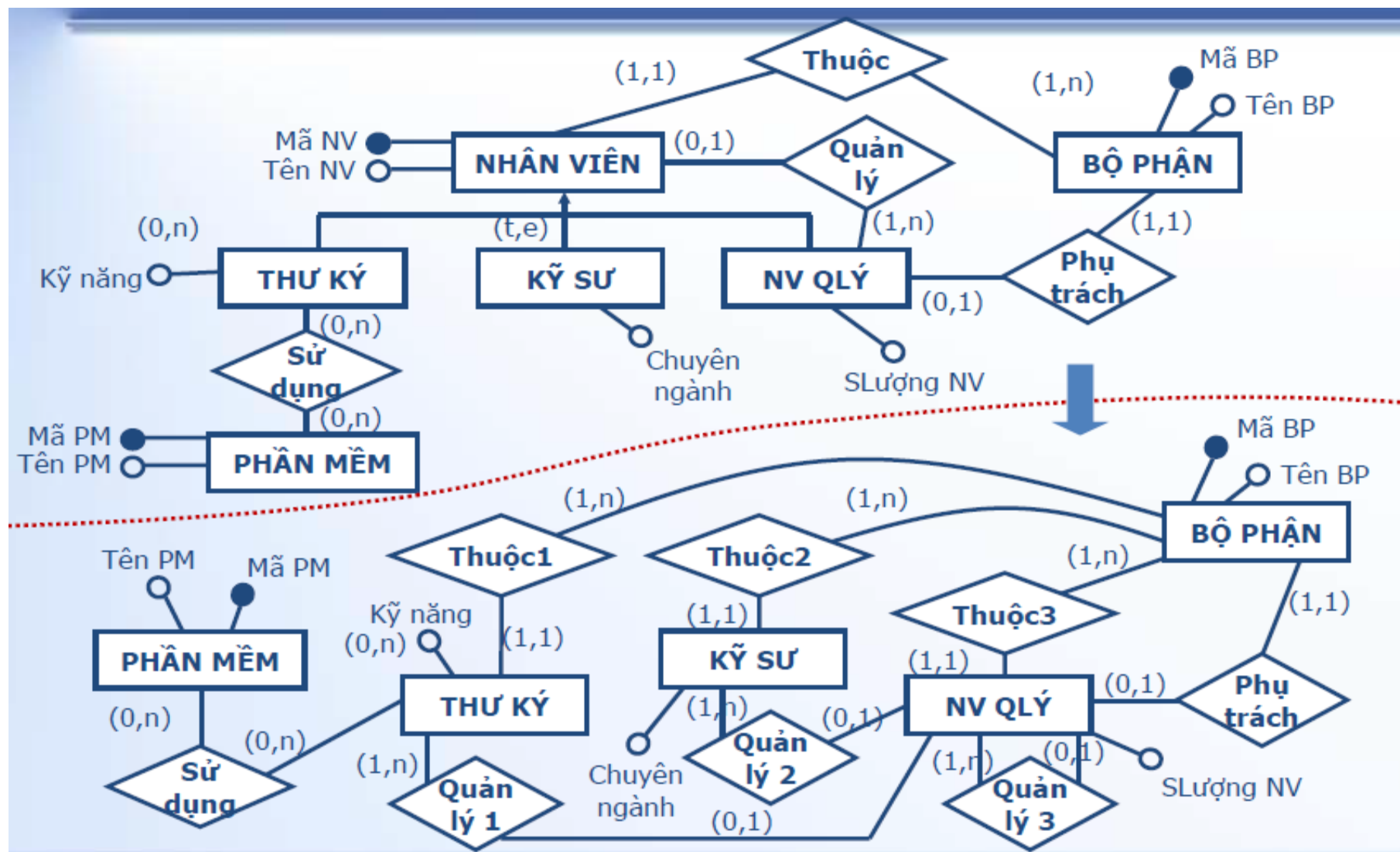
Ưu điểm	Khuyết điểm
Giải pháp đơn giản nhất, không phát sinh thêm các mối kết hợp	Có thể phát sinh ra 1 số lớn các giá trị rỗng cho các thuộc tính mà chỉ dùng cho một loại thực thể tập con mà thôi
Áp dụng cho tất cả các cấu trúc tổng quát hóa như toàn bộ (t) và bán phần (p), chồng lấp (o) và riêng biệt (e)	<p>Tất cả các tác vụ muốn truy cập đến một thực thể tập con phải truy cập toàn bộ tất cả các thực thể tập con</p> <p>Phát sinh thêm một số RBTV cần phải kiểm tra</p>

Dùng thực thể chuyên biệt

- Chuyển đổi tất cả đặc trưng của thực-thể-tổng-quát xuống các thực-thể-chuyên-biệt



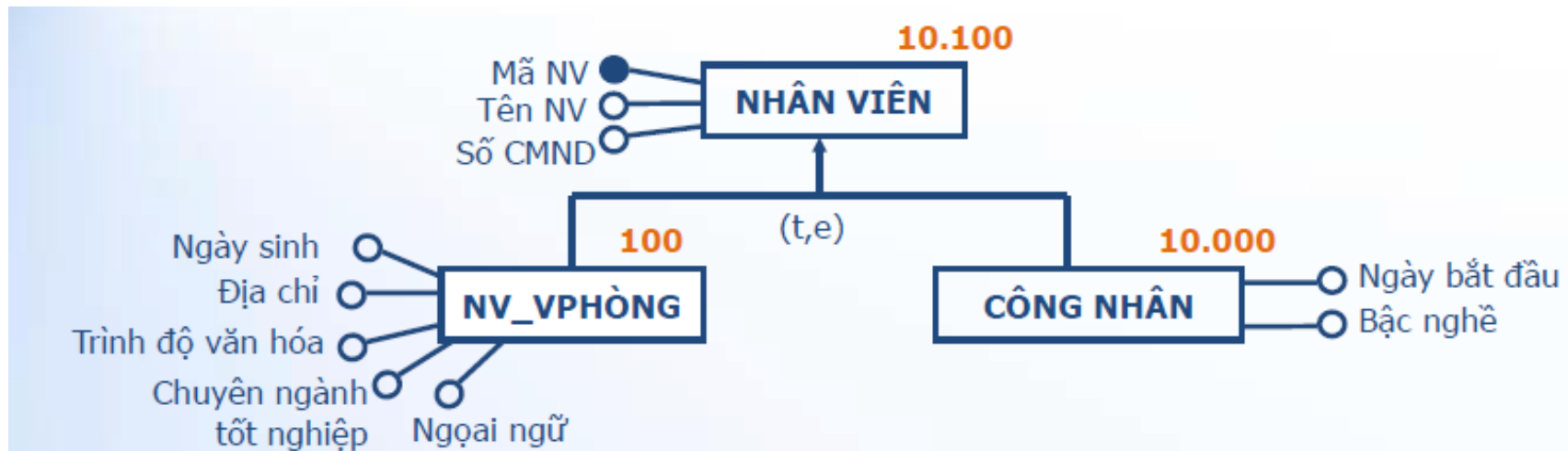
Ví dụ



Dùng thực thể chuyên biệt

Ưu điểm	Khuyết điểm
Thuộc tính riêng của thực thể tập con chỉ biểu diễn riêng cho loại thực thể tập con đó	Cách chọn lựa này không áp dụng được cho cấu trúc tổng quát hóa loại chổng chéo (o) và bán phần (p), chỉ dùng được cho toàn phần (t) và riêng biệt (e)
Các tác vụ liên quan đến một loại thực thể tập con chỉ truy cập đến loại thực thể tập con đó	<p>Quan niệm các thực thể tập con trước đây cùng là chuyên biệt hóa của một thực thể tổng quát không còn nữa. Quan niệm này có thể là chủ yếu liên quan đến một số xử lý</p> <p>Nếu thuộc tính của thực thể tổng quát là đáng kể thì sự lập lại trong lược đồ đáng xem xét lại</p> <p>Các tác vụ trước đây thao tác lên thực thể tổng quát nay phải thao tác lên tất cả các thực thể tập con</p>

Lựa chọn chuyên biệt | tổng quát



<i>Xử lý</i>	<i>Tần suất</i>	<i>Chọn lựa tối ưu</i>
(o1) Tính lương sản phẩm cho công nhân phân xưởng	2/ tháng	Tách (gộp)
(o2) Tính lương cho nhân viên văn phòng	1/tháng	Tách
(o3) Tìm kiếm thông tin về công nhân	1000/ngày	Tách (gộp)
(o4) Tổng hợp danh sách chung của toàn bộ nhân viên	5/tháng	Gộp
(o5) Truy xuất thông tin nhân viên văn phòng	20/tháng	Tách

Lựa chọn chuyên biệt | tổng quát

■ Phụ thuộc

◆ Dung lượng

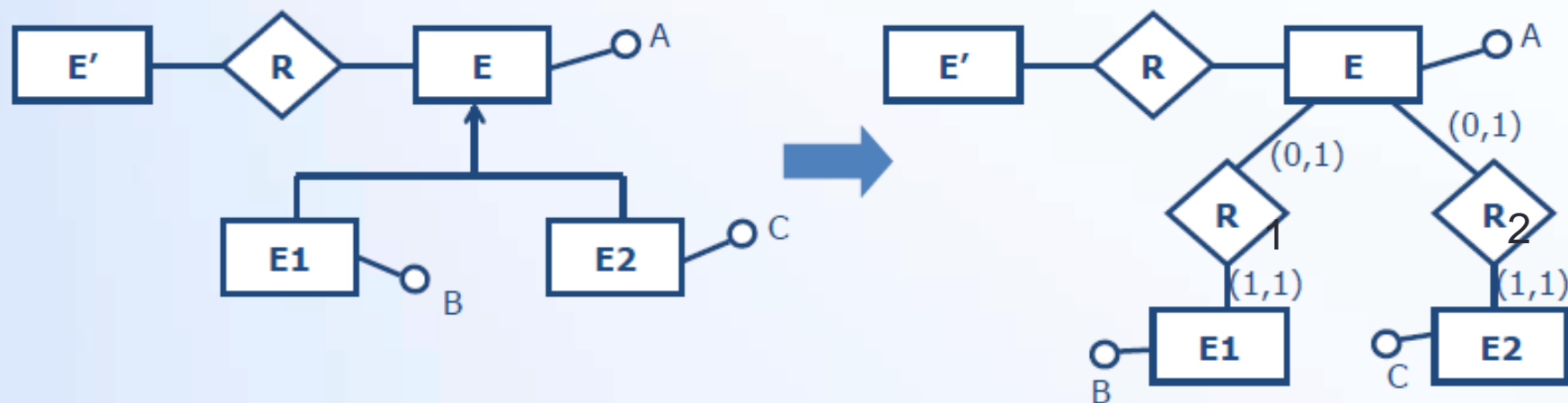
- Gộp: dung lượng lớn
- Tách: dung lượng tối ưu

◆ Xử lý

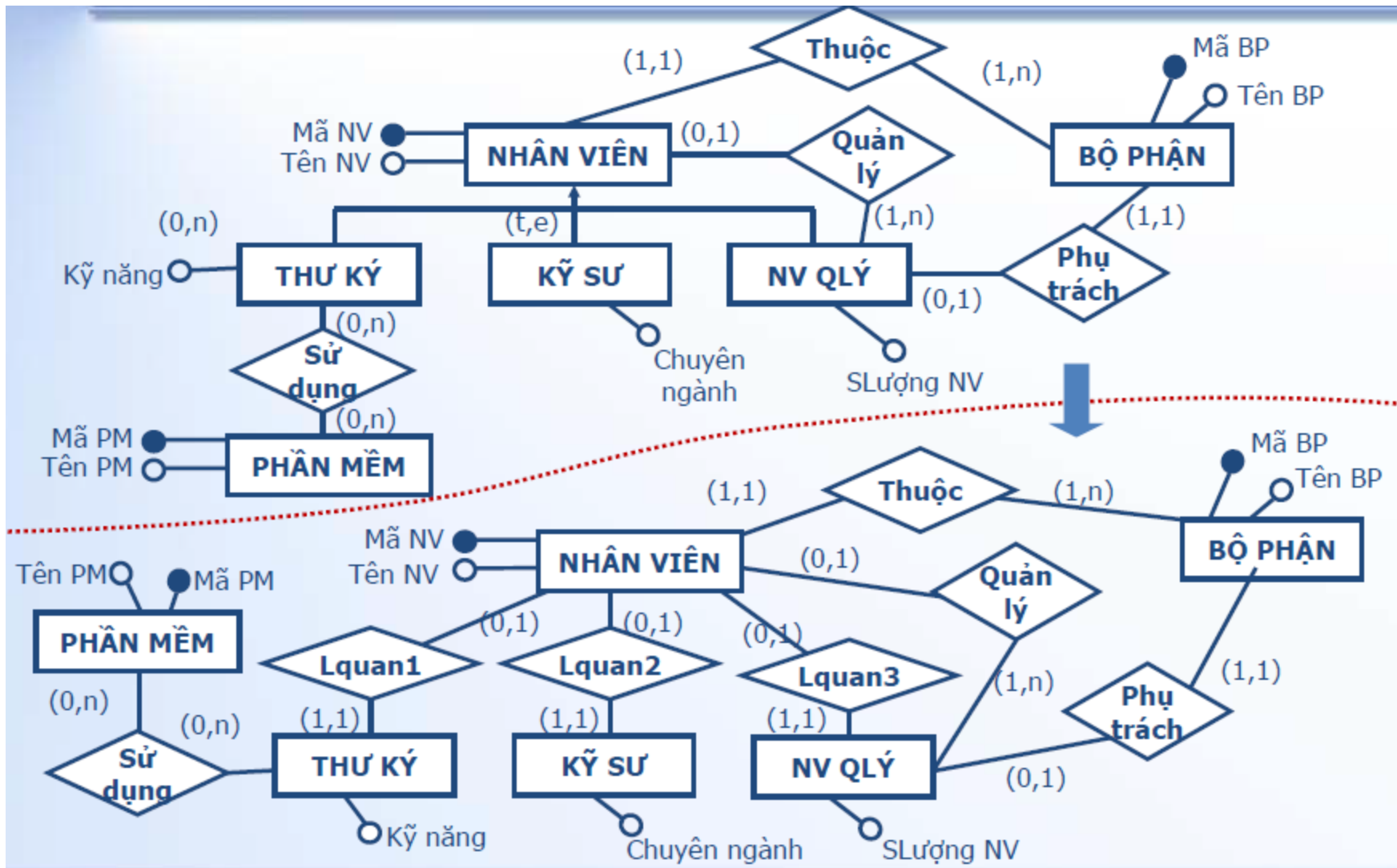
- Ưu tiên các xử lý có tần suất cao
- Ví dụ
 - Nếu ưu tiên o4 : gộp
 - Nếu ưu tiên o2, o3: tách

Dùng mối kết hợp

- Bỏ tổng quát, chuyên biệt
- Tạo mối kết hợp từ thực-thể-tổng-quát đến thực-thể-chuyên-biệt



Ví dụ



Dùng mối kết hợp

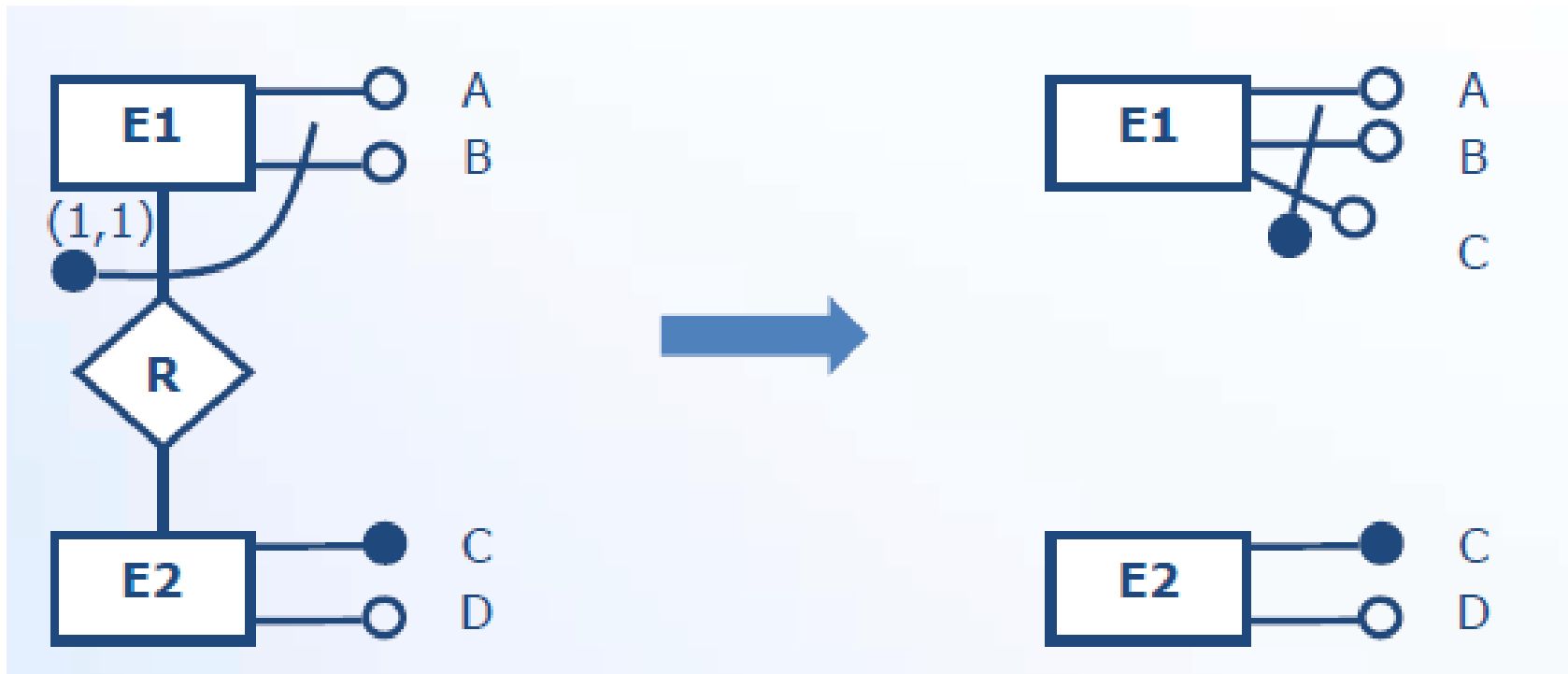
Ưu điểm	Khuyết điểm
Có thể mô hình tất cả các loại cấu trúc tổng quát hóa toàn bộ / bán phần và chồng chéo / riêng biệt	Lược đồ kết quả khá phức tạp. Ví dụ như thêm một thể hiện cho một thực thể tập con phải thêm mới một thể hiện cho quan hệ và một thể hiện cho thực thể tổng quát hóa
Rất uyển chuyển khi thay đổi yêu cầu của ứng dụng	Phải chấp nhận sự dư thừa khi biểu diễn mối liên kết IS-A thành mối kết hợp

Thiết kế luận lý dữ liệu cấp thấp

- Chuẩn bị
 - ◆ Loại bỏ định danh ngoài
 - ◆ Loại bỏ thuộc tính đa trị và thuộc tính kết hợp
- Chuyển đổi mô hình thực thể kết hợp sang mô hình quan hệ

(ERM → LDM)

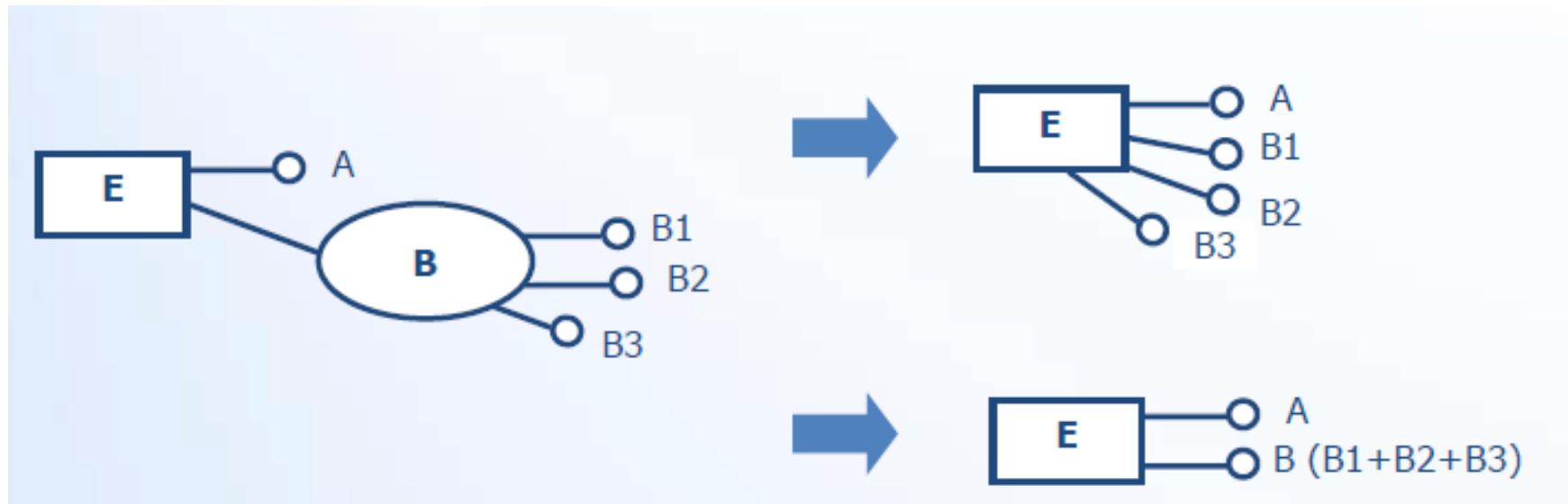
Loại bỏ định danh bên ngoài



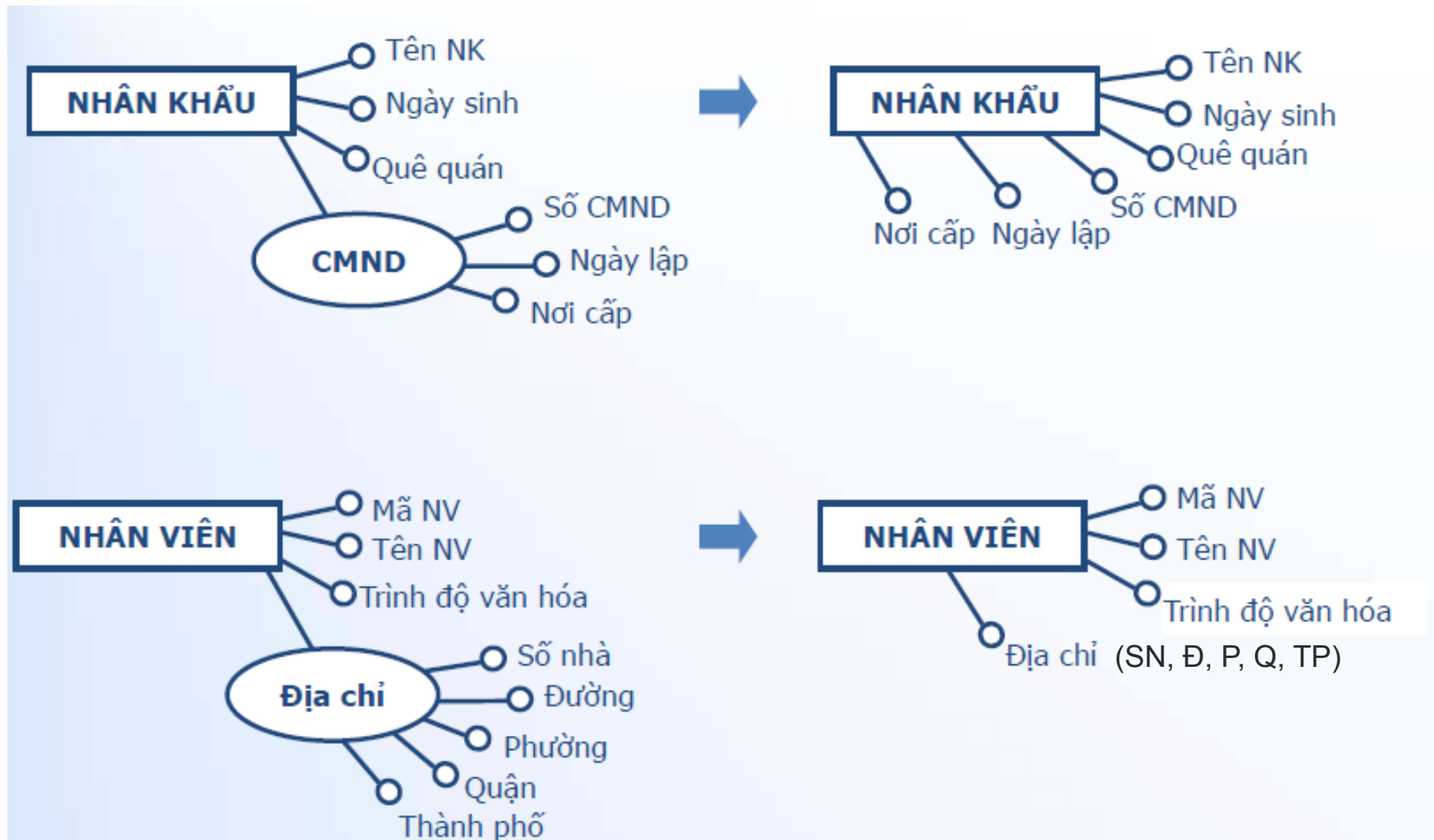
Ví dụ



Loại bỏ thuộc tính kết hợp



Ví dụ

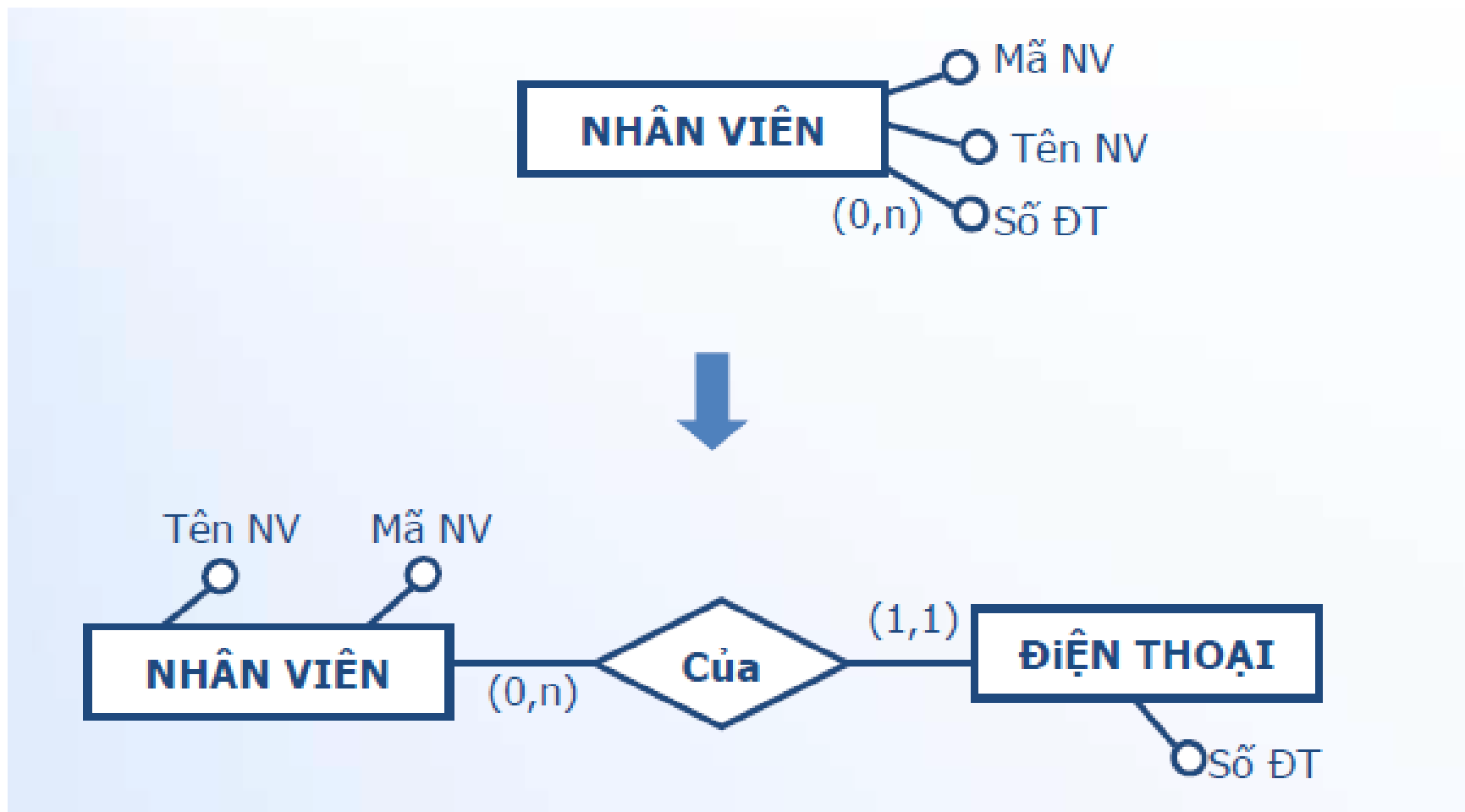


Loại bỏ thuộc tính đa trị

■ Thuộc tính đa trị ở thực thể

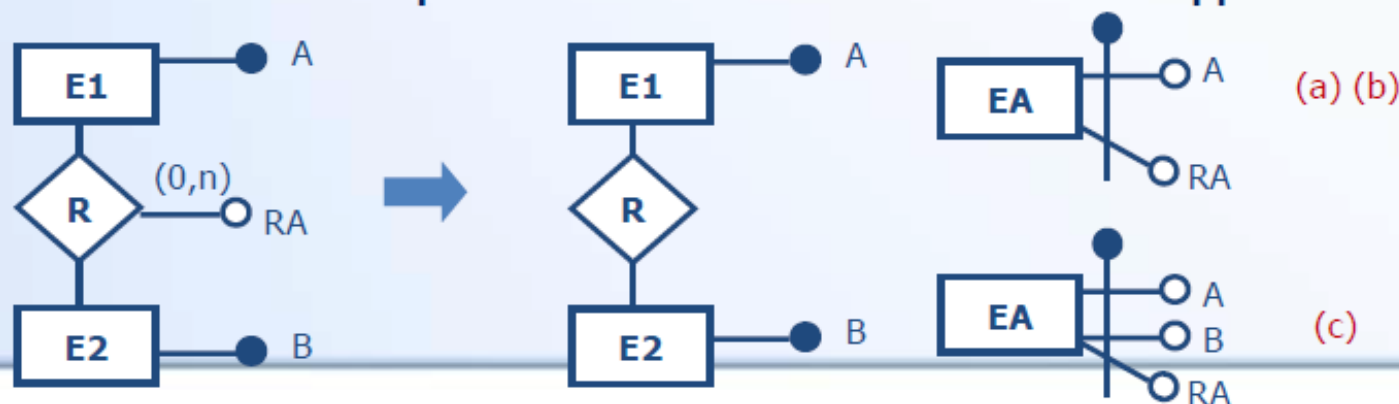


Ví dụ:

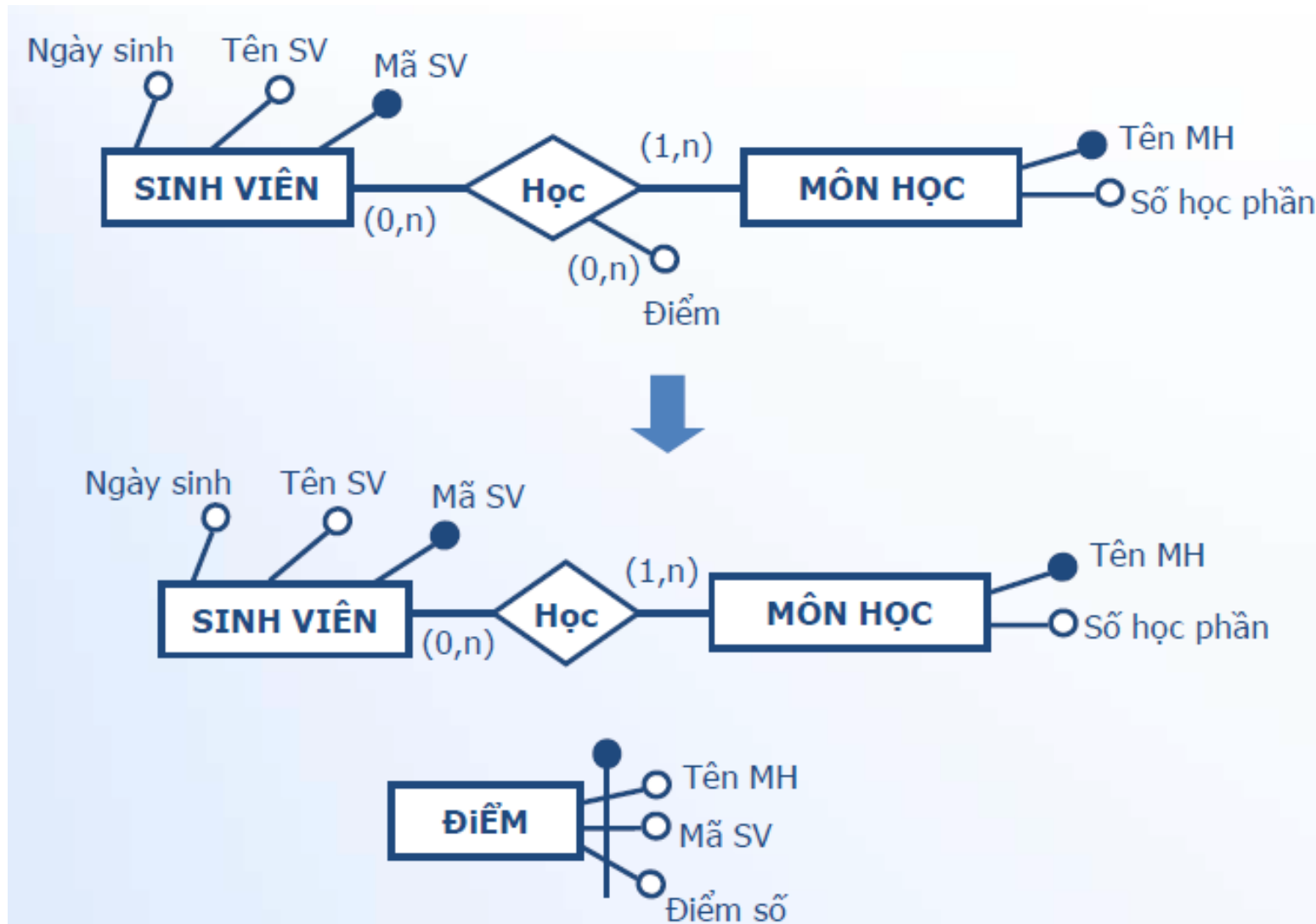


Loại bỏ thuộc tính đa trị (tt)

- Thuộc tính đa trị ở mỗi kết hợp
 - ◆ (a) R là mỗi kết hợp 1-1
 - EA sẽ có định danh của E1 hoặc E2 kết hợp với RA
 - ◆ (b) R là mỗi kết hợp 1-N
 - EA sẽ có định danh của E1 kết hợp với RA
 - ◆ (c) R là mỗi kết hợp N-N
 - EA sẽ có định danh của E1 và E2 kết hợp với RA



Ví dụ:



Chuyển mô hình thực thể kết hợp (ERM) → Mô hình dữ liệu mức luận lý (LDM)

- Chuyển đổi thực thể
- Chuyển đổi mối kết hợp

Chuyển đổi thực thể

- 1 thực thể → 1 quan hệ
- Thuộc tính → thuộc tính
- Định danh → khóa chính



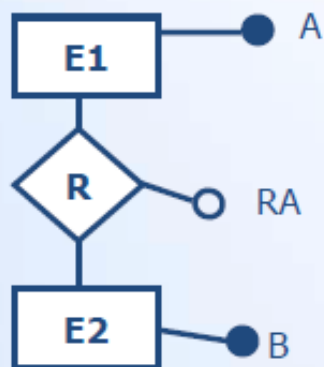
SINH_VIÊN(MA_SV, TEN_SV, NGÀY_SINH)



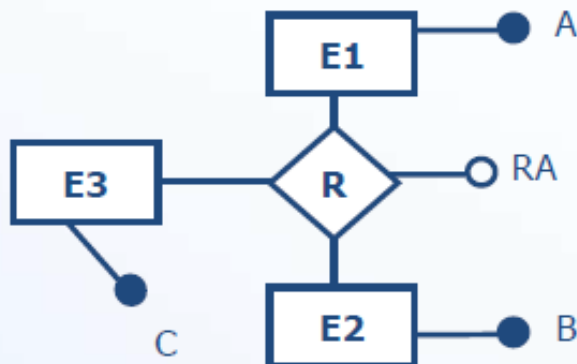
MÔN_HỌC(TÊN_MH, SỐ_HỌC_PHẦN)

Chuyển đổi mỗi kết hợp

■ Quy tắc chung



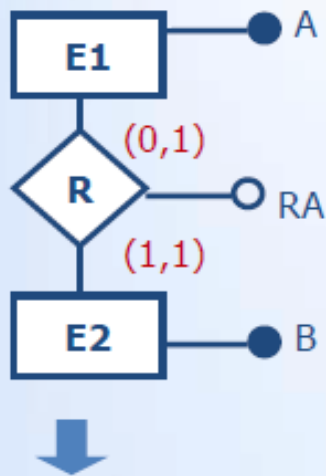
R(A, B, RA)



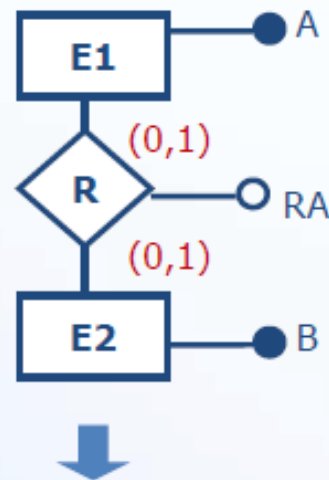
R(A, B, C, RA)

Chuyển đổi mỗi kết hợp

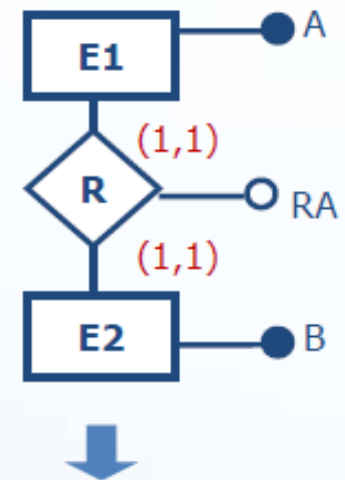
■ Mỗi kết hợp 1-1



$E1(\underline{A}, \dots)$
 $E2(\underline{B}, \dots, A) : A \text{ là khóa}$

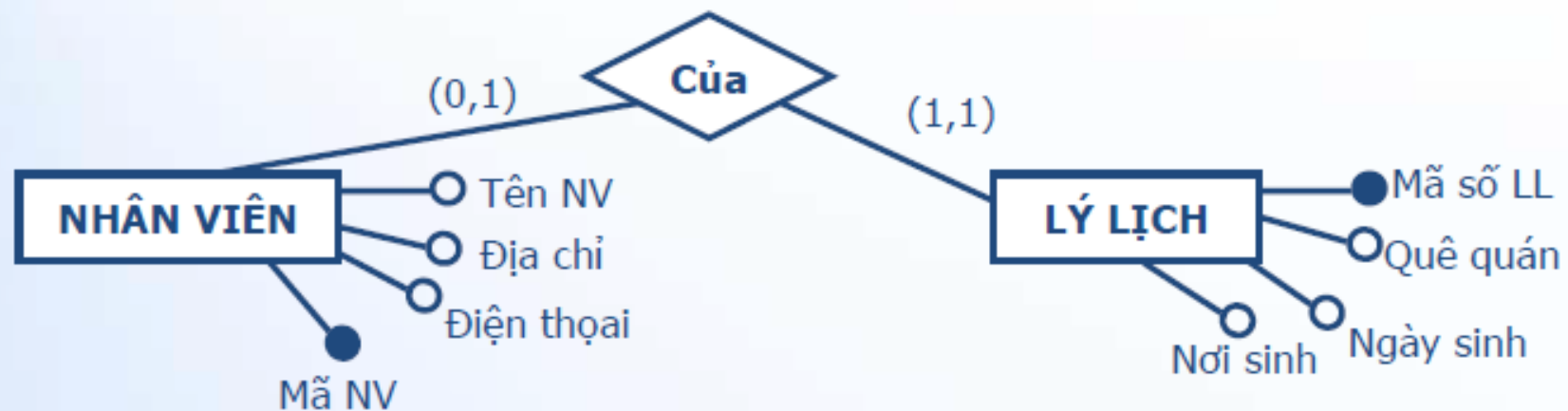


$E1(\underline{A}, \dots)$
 $E2(\underline{B}, \dots, A) : A \text{ là khóa}$
 Hoặc
 $E1(\underline{A}, \dots, B) : B \text{ là khóa}$
 $E2(\underline{B}, \dots)$



$E1E2(\underline{A}, B, \dots) : B \text{ là khóa}$
 Hoặc
 $E1E2(A, \underline{B}, \dots) : A \text{ là khóa}$

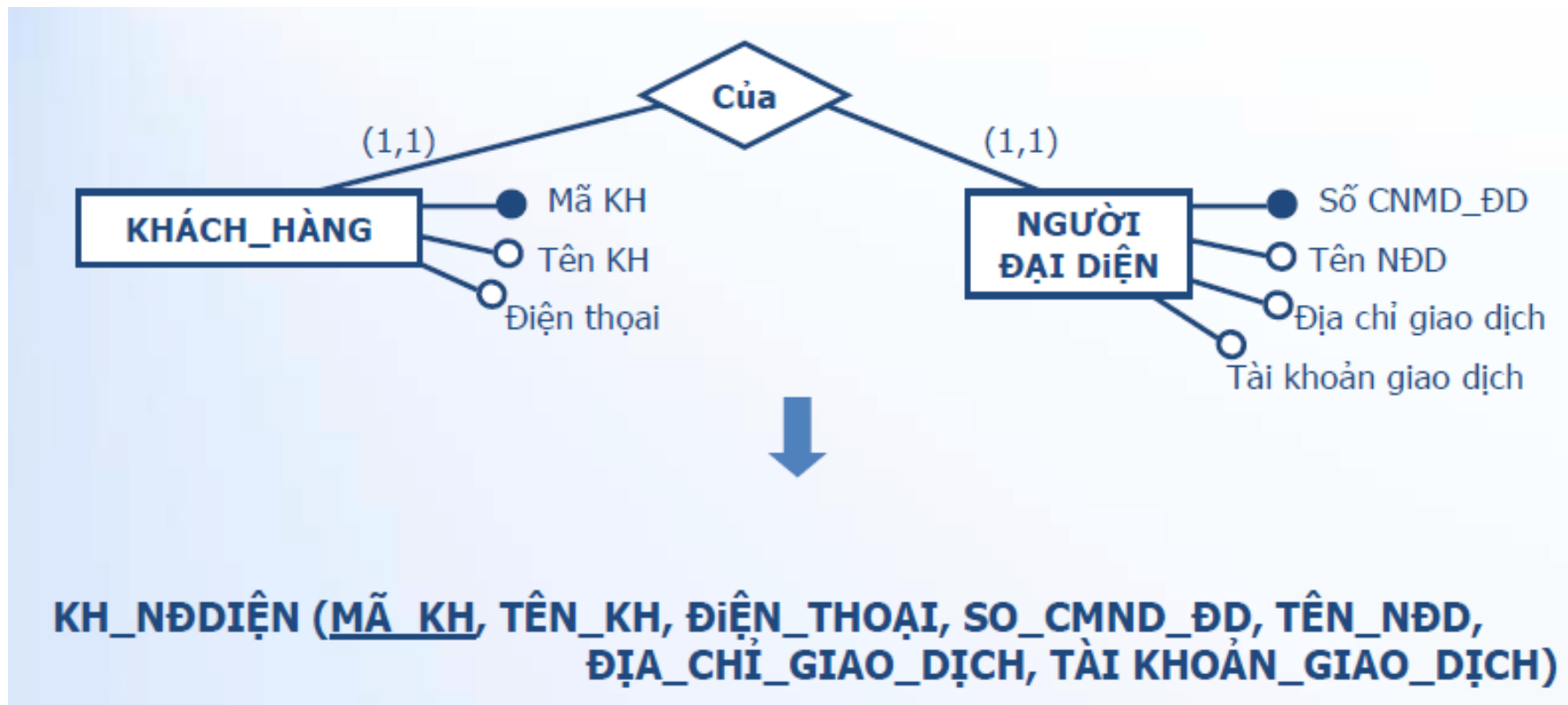
Ví dụ:



NHÂN_VIÊN(MA_NV, TÊN_NV, ĐỊA_CHỈ, ĐIỆN_THOẠI)

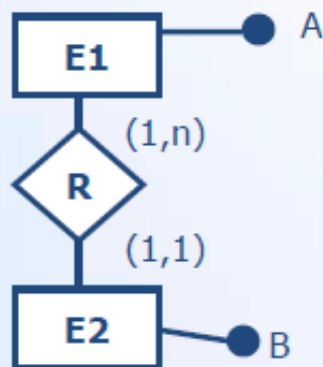
LÝ_LỊCH(MÃ_SỐ_LL, QUÊ_QUÁN, NGÀY_SINH, NƠI_SINH, MA_NV)

Ví dụ



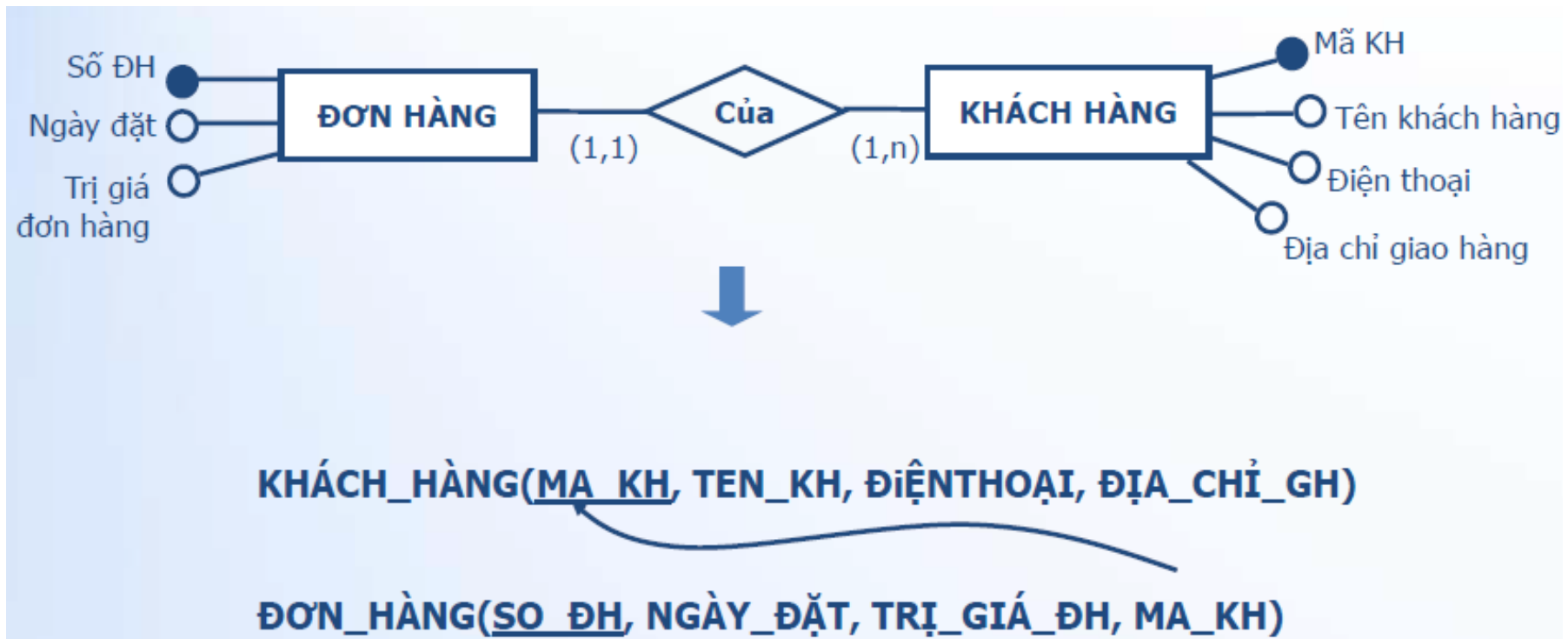
Chuyển đổi mối kết hợp

■ Mối kết hợp 1-N



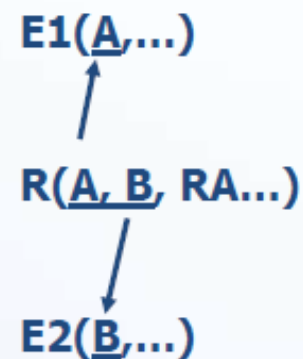
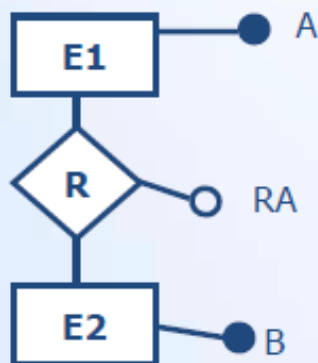
$E1(\underline{A}, \dots)$
 $E2(\underline{B}, \dots, A)$

Ví dụ:

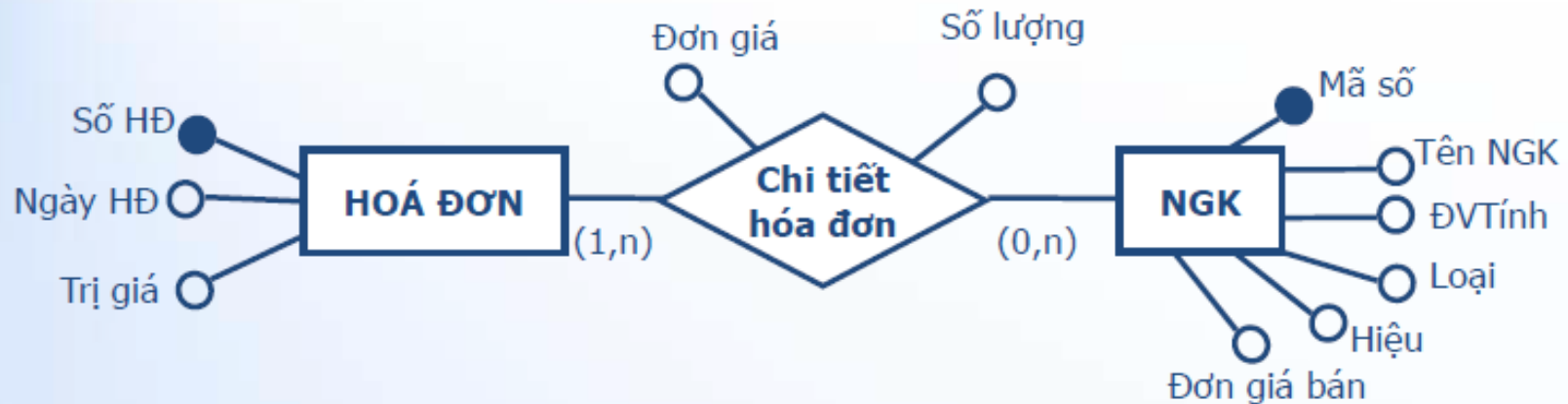


Chuyển đổi mối kết hợp (tt)

■ Mối kết hợp N-N



Ví dụ



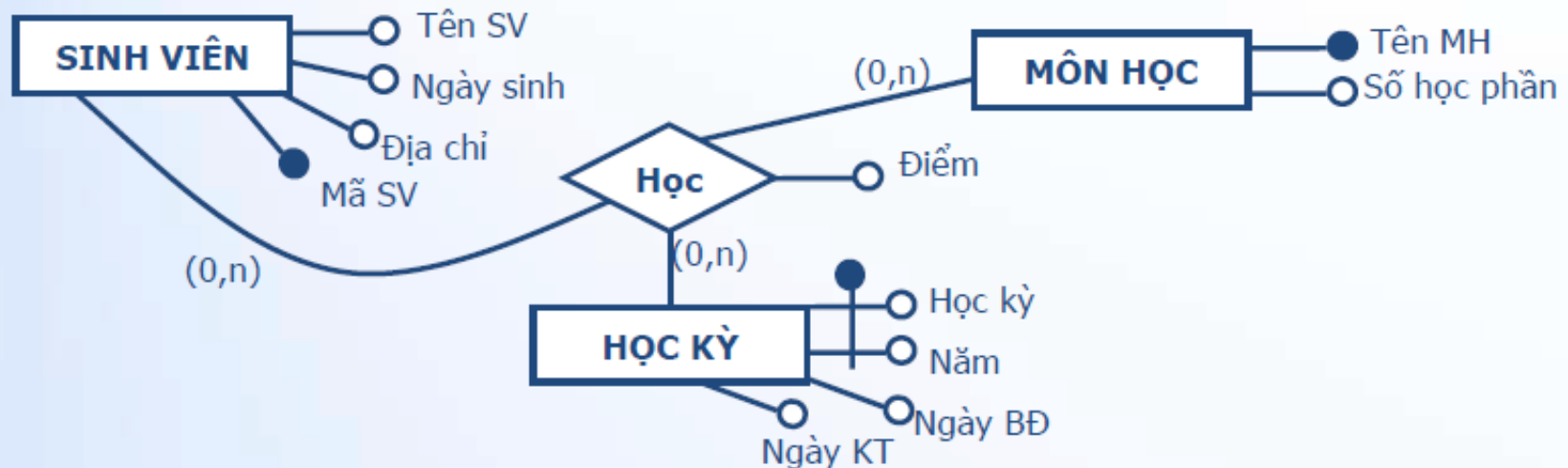
HÓA ĐƠN(SỐ HD, NGÀY_HD, TRỊ GIÁ)

CHITIET_HD(SO HD, MÃ SỐ, SỐ LƯỢNG, ĐƠN GIÁ)

NGK(MÃ SỐ, TÊN_NGK, ĐVTÍNH, LOẠI, HIỆU, ĐƠN GIÁ_BÁN)

Chuyển đổi mối kết hợp (tt)

■ Mối kết hợp đa phân



SINH_VIÊN(MÃ_SV, TÊN_SV, NGÀY_SINH, ĐỊA_CHỈ)

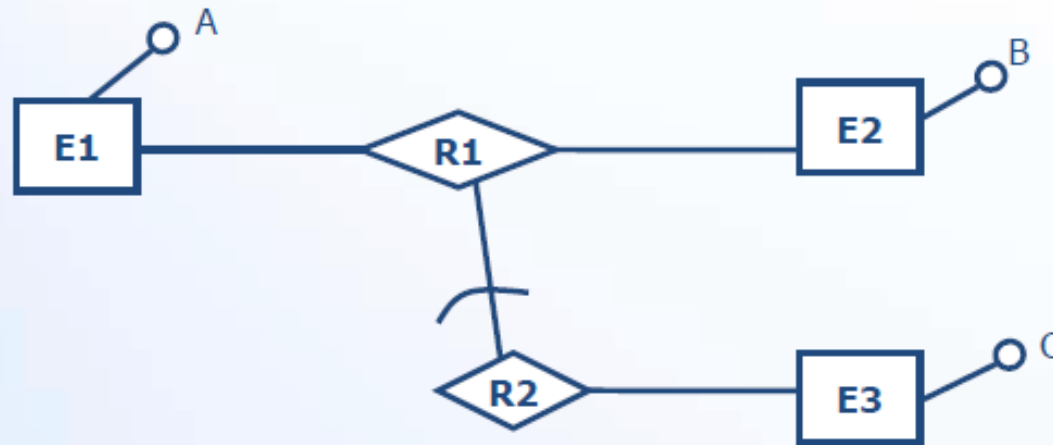
MÔN_HỌC(TÊN_MH, SỐ_HỌC_PHẦN)

HỌC_KỲ(HỌC_KỲ, NĂM, NGÀY_BĐ, NGÀY_KT)

HỌC(MÃ_SV, TÊN_MH, HỌC_KỲ, NĂM, ĐIỂM)

Chuyển đổi mối kết hợp (tt)

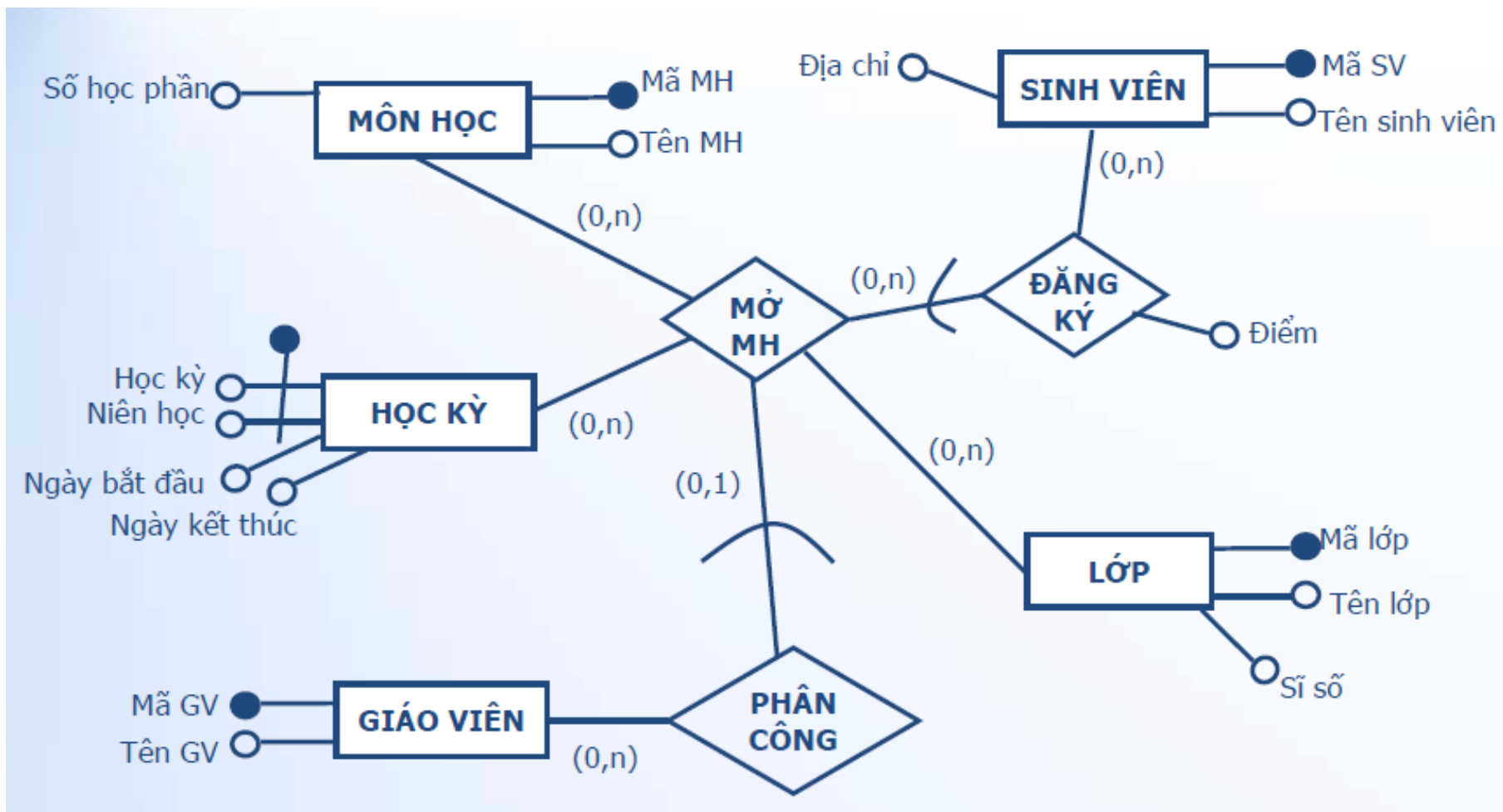
■ Mối kết hợp mở rộng



R1(A, B, ...)

R2(A, B, C, ...)

Ví dụ:



Ví dụ

MÔN_HỌC(MA_MH, TÊN_MH, SỐ_HP)

HỌC_KỲ(HOC_KỲ, NIÊN_HOC, NGÀY_BĐ, NGÀY_KT)

GIÁO_VIÊN(MÃ_GV, TÊN_GV)

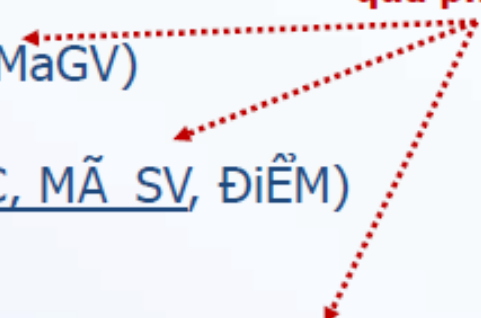
SINH_VIÊN(MÃ_SV, TÊN_SV)

LỚP(MÃ_LỚP, TÊN_LỚP, SĨ_SỐ)

MỞ_MH(MÃ_MH, MÃ_LỚP, HOC_KỲ, NIÊN_HOC, MãGV)

ĐĂNG_KÝ(MÃ_MH, MÃ_LỚP, HOC_KỲ, NIÊN_HOC, MÃ_SV, ĐIỂM)

Khóa chính
quá phức tạp



Nội dung chi tiết

- Thiết kế luận lý dữ liệu
 - Cấp cao
 - Cấp thấp
- **Thiết kế mã**
- Bài tập cuối chương

Thiết kế mã

- Tìm các hình thức mã hóa
 - ◆ Mã quận, mã bưu điện
 - ◆ Mã xe
 - ◆ Mã sinh viên
 - ◆ ...
- Dễ dàng cho việc
 - ◆ Quản lý và xử lý thông tin trong máy tính
 - ◆ Nhận dạng và phân loại dữ liệu
- Phải được thiết kế trên cơ sở uyển chuyển và có khả năng mở rộng

Thiết kế mã

■ Ngữ nghĩa của mã hóa

◆ Duy nhất

– Mã phải duy nhất để nhận dạng đối tượng dữ liệu

◆ Phân loại

– Phân loại dữ liệu thành các nhóm khác nhau, dựa trên mã phân biệt được nhóm

◆ Sắp xếp

– Mã hóa thể hiện thứ tự của dữ liệu

◆ Kiểm tra

– Kiểm soát dữ liệu có được nhập đúng hay không

Thiết kế mã

■ Một số đặc điểm

- ◆ Vùng mã hóa: phạm vi mà mã được sử dụng
 - Có liên kết với các hệ thống bên ngoài, mã nên là chuẩn công nghiệp chung
 - Trong cùng 1 đơn vị, mã có phạm vi toàn công ty
- ◆ Chu kỳ sử dụng
 - Ước lượng 1 lượng mã cần đủ trong chu kỳ
- ◆ Ngữ nghĩa
 - Có thể hiểu mã bởi các thành phần liên quan

Ví dụ vùng mã hoá

- Mã môn học được thiết kế khác nhau trong cùng 1 trường

TH201 - Thiết kế Cơ Sở Dữ Liệu

Khoa CNTT

CNTT201 - Thiết kế Cơ Sở Dữ Liệu

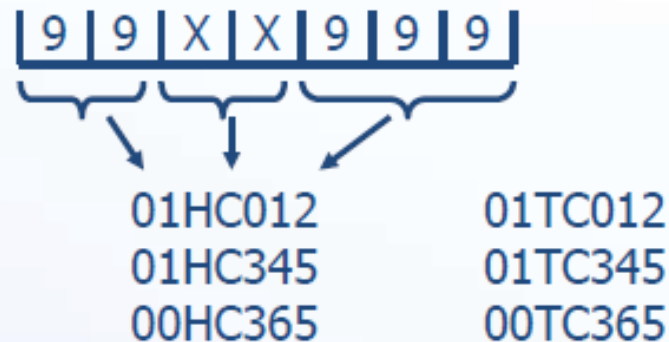
Phòng Đào Tạo

Gây ra 1 số hoạt động không cần thiết như : chuyển đổi mã liên quan đến kết quả học tập trên môn học, ...

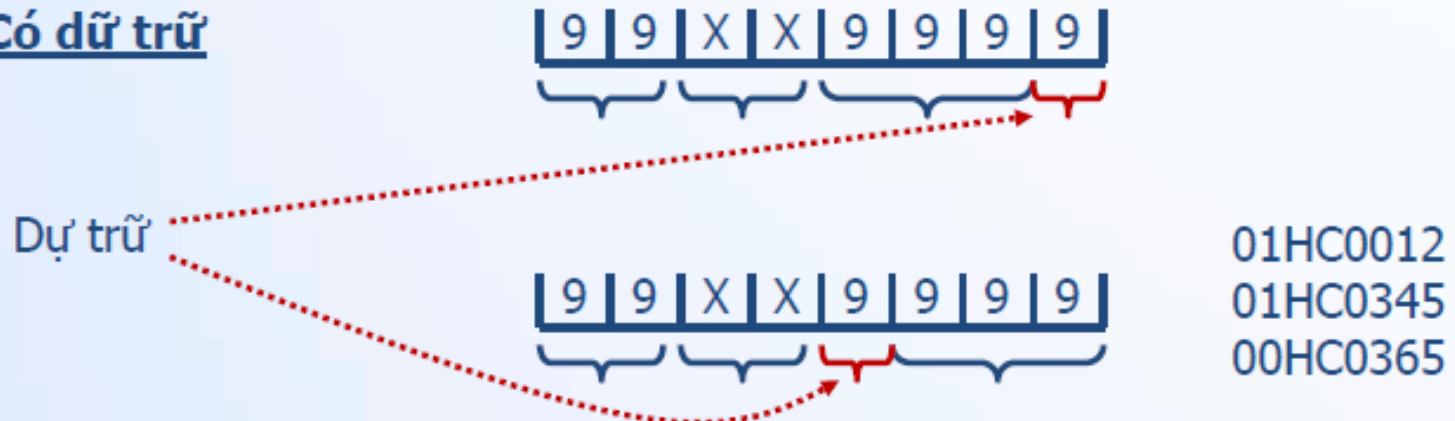
Ví dụ chu kỳ sử dụng

■ Mã số sinh viên

Không dữ trữ

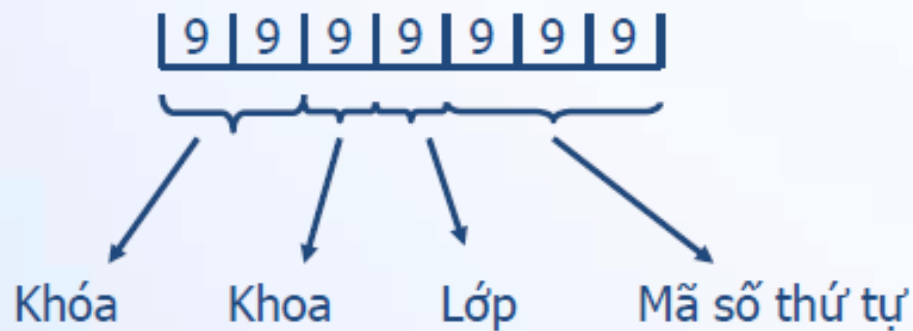


Có dữ trữ



Ví dụ ngữ nghĩa

■ Mã số sinh viên



0011037

0011456

0111230

Một số loại mã

Tên mã	Mô tả	Ví dụ
Mã tuần tự	Một số được gán một cách tuần tự	<u>Mã số xe của tỉnh:</u> 50 Thành phố HCM 60 Đồng Nai 62 Long An 63 Tiền Giang
Mã khối	Một số được gán cho mỗi khối. Rồi một số sẽ được gán tuần tự trong khối đó → thuận tiện cho phân loại dữ liệu	<u>Mã khách hàng của từng chi nhánh được qui định như sau:</u> 0001 khách hàng tại trung tâm 1000 khách hàng chi nhánh A 2000 khách hàng chi nhánh B 3000 khách hàng chi nhánh C

Một số loại mã

Tên mã	Mô tả	Ví dụ
Mã thập phân	Các đối tượng sẽ được mã hóa từ 0 đến 9, rồi đến lượt thành viên của mỗi đối tượng này cũng sẽ được mã hóa từ 0 đến 9,	<u>Mã hóa phòng ban trong một công ty:</u> 00 Ban giám đốc 1 Bộ phận kinh doanh 1 Nhóm 1 2 Nhóm 2 2 Bộ phận phát triển 10 Chi nhánh A 20 Chi nhánh B 30 Chi nhánh C
Mã theo ký số	Mỗi ký số của mã sẽ gán cho một ngữ nghĩa	<u>Mã SV của trường ĐHKHTN TPHCM</u> ① ② ③ ④ 01 1 1 100 01 2 1 120 ① Năm ② Khoa ③ Lớp ④ số thứ tự

Một số loại mã

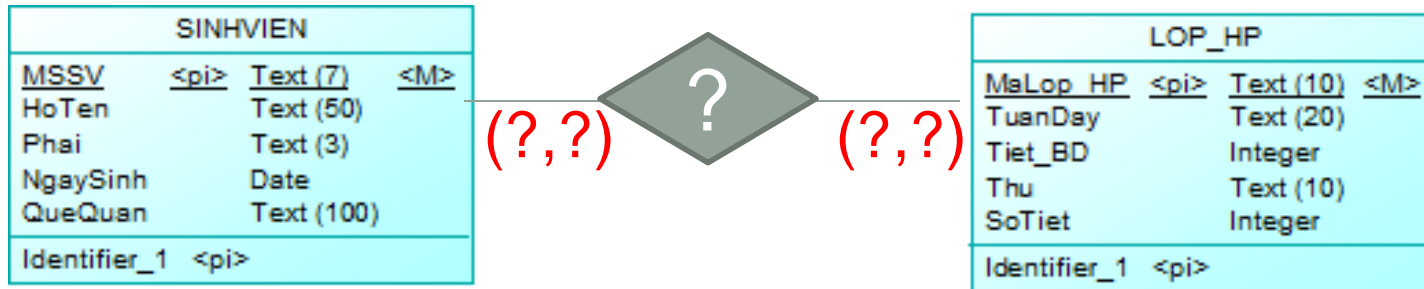
Tên mã	Mô tả	Ví dụ
Mã gợi nhớ	Chữ viết tắt hoặc biểu tượng của đối tượng được dùng để mã hóa	<u>Mã hóa lớp học của trường đại học như sau:</u> 00TC Lớp tại chức khóa 2000 01TC Lớp tại chức khóa 2001 00HC Lớp hoàn chỉnh khóa 2000 01HC Lớp hoàn chỉnh khóa 2001
Mã kiểm tra	Dùng một ký số kiểm tra thêm vào sau mỗi mã số	

Nội dung chi tiết

- Thiết kế luận lý dữ liệu
 - Cấp cao
 - Cấp thấp
- Thiết kế mã
- **Bài tập cuối chương**

Bài tập cuối chương

- **Bài tập 01:**
- Gắn bản số cho quan hệ:

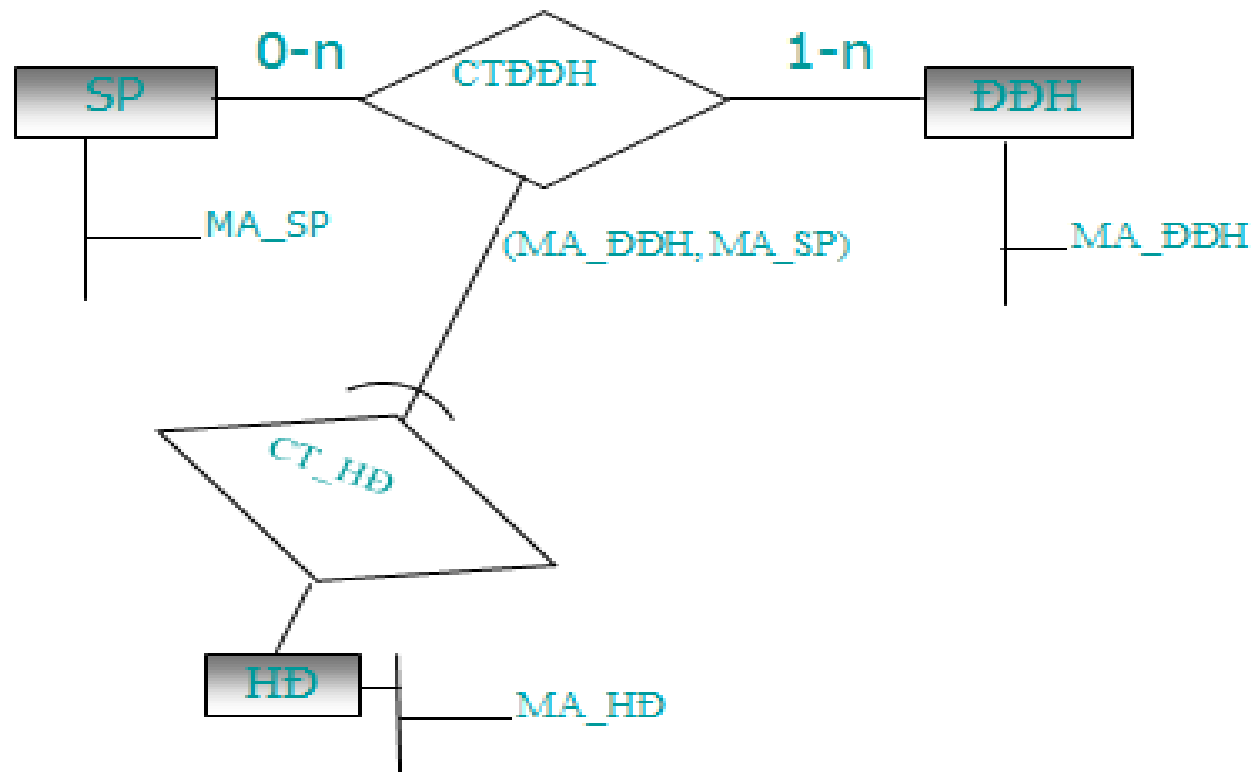


- Chuyển sang LDM và vẽ tham chiếu?

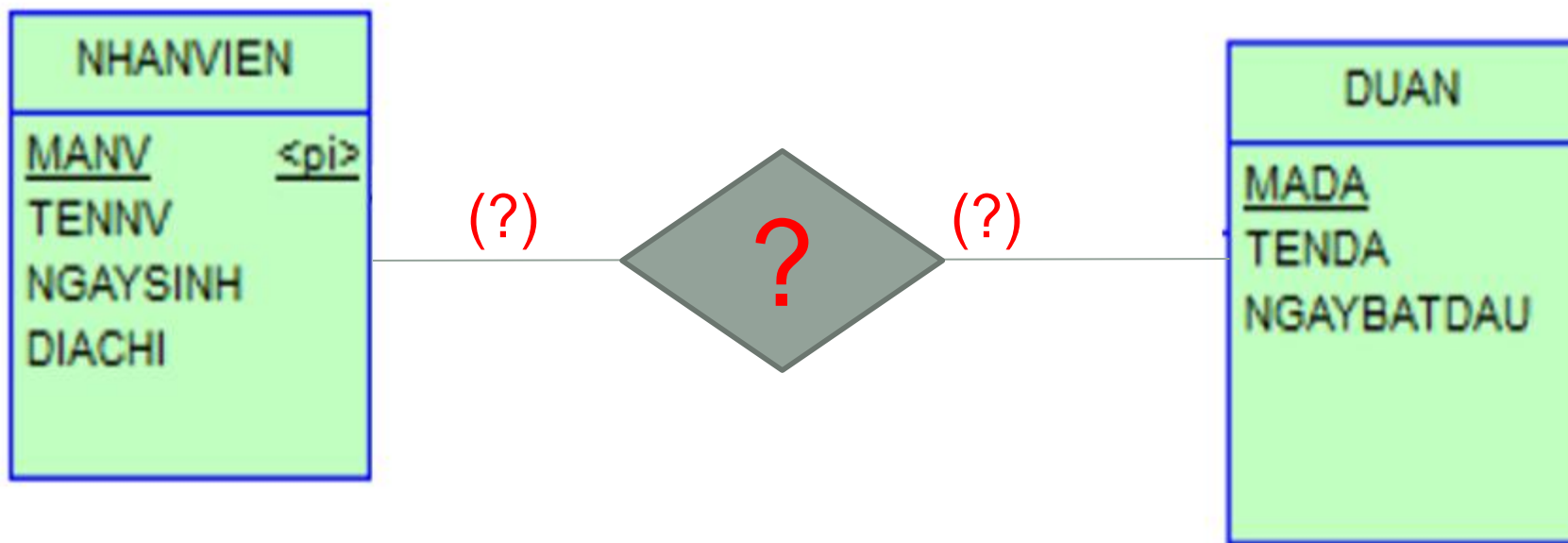
Bài tập cuối chương

- **Bài tập 02**: Cho biết ngữ nghĩa mã số sv **CST115486**?
- **Bài tập 03**: Ý nghĩa mã các môn học **COS301, COS302**?

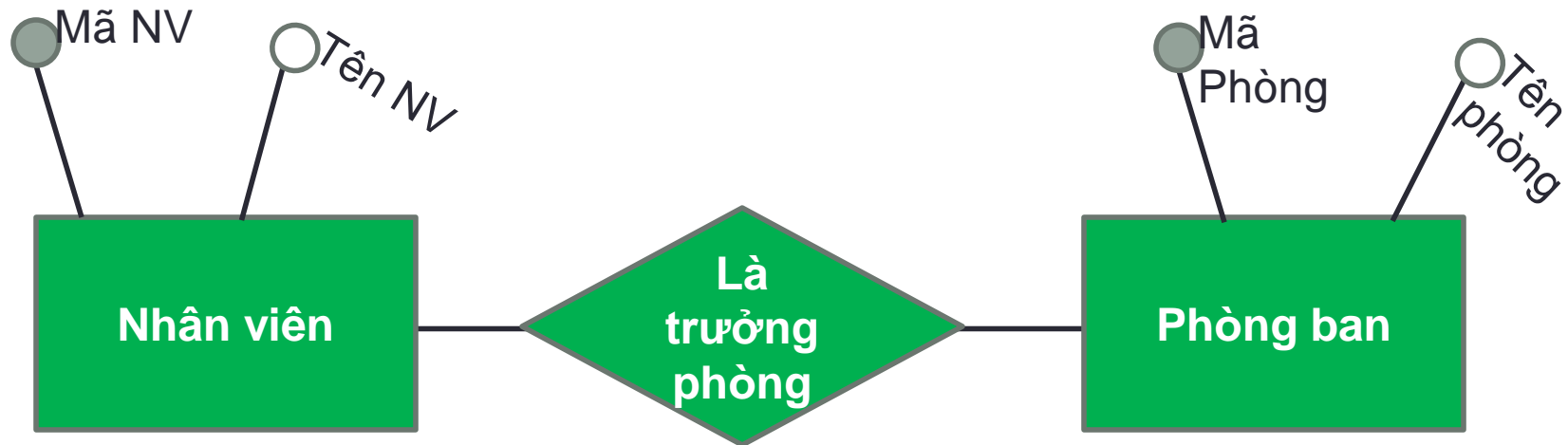
- **Bài tập 04:**
- Chuyển sang LDM và vẽ tham chiếu:



- **Bài tập 05:**
- Chuyển sang LDM và vẽ tham chiếu:



- **Bài tập 06:**
- Gắn bản số cho quan hệ
- Chuyển sang LDM và vẽ tham chiếu:



- **Bài tập 07:**
- Gắn bản số cho quan hệ, tên mỗi kết hợp
- Chuyển sang LDM và vẽ tham chiếu:

