

CHƯƠNG 1

TỔNG QUAN

NỘI DUNG

- Phần 1: Các khái niệm
 1. Đặt vấn đề
 2. Hệ thống tập tin (File System)
 3. Định nghĩa một CSDL
 4. Các đối tượng sử dụng CSDL
 5. Hệ quản trị CSDL
 6. Các mức của một CSDL
- Phần 2: Các mô hình dữ liệu
 1. Giới thiệu
 2. Mô hình dữ liệu mạng
 3. Mô hình dữ liệu phân cấp
 4. Mô hình thực thể mối kết hợp
 5. Mô hình dữ liệu quan hệ
 6. Mô hình dữ liệu hướng đối tượng

CÁC KHÁI NIỆM

1. Đặt vấn đề



1. Đặt vấn đề

- Vấn đề Quản lý dữ liệu
- Cơ sở dữ liệu là gì?

2. Hệ thống tập tin

- Là tập hợp các tập tin riêng lẻ phục vụ cho một mục đích của đơn vị sử dụng.
- **Ưu điểm**
 - Triển khai ứng dụng nhanh
 - Khả năng đáp ứng nhanh chóng, kịp thời (vì chỉ phục vụ cho mục đích hạn hẹp)
- **Khuyết điểm**
 - Trùng lặp dữ liệu → lãng phí, dữ liệu không nhất quán
 - Chi phí cao
 - Chia sẻ dữ liệu kém

3. Cơ sở dữ liệu (1)

- Cơ sở dữ liệu là một hệ thống các thông tin có cấu trúc, được lưu trữ trên các thiết bị lưu trữ nhằm thỏa mãn yêu cầu khai thác thông tin đồng thời của nhiều người sử dụng hay nhiều chương trình ứng dụng với những mục đích khác nhau

3. Cơ sở dữ liệu (2)

● Ưu điểm

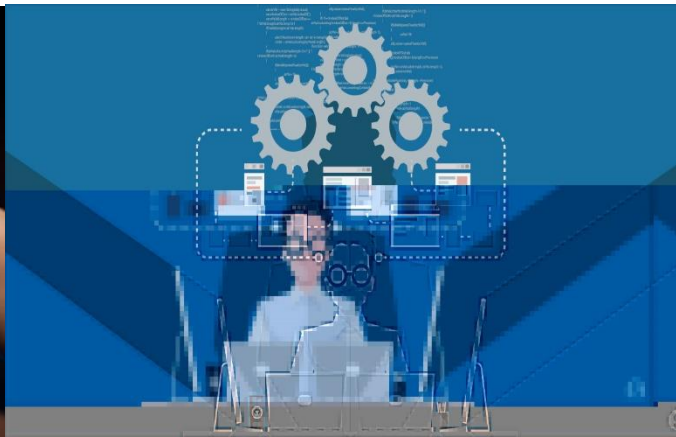
- Giảm trùng lặp thông tin xuống mức thấp nhất, đảm bảo tính nhất quán và toàn vẹn dữ liệu.
- Đảm bảo dữ liệu được truy xuất theo nhiều cách khác nhau.
- Khả năng chia sẻ thông tin cho nhiều người, nhiều ứng dụng khác nhau.

3. Cơ sở dữ liệu (3)

- Những vấn đề cần giải quyết
 - Dữ liệu phải được cập nhật
 - Tính bảo mật và quyền khai thác thông tin của người sử dụng
 - Tranh chấp dữ liệu
 - Đảm bảo dữ liệu khi có sự cố

4. Các đối tượng sử dụng CSDL

- Người sử dụng không chuyên về CNTT, CSDL: công cụ khai thác
- Chuyên viên tin học xây dựng các ứng dụng để phục vụ cho các mục đích quản lý
- Quản trị CSDL: tổ chức CSDL, bảo mật, phân quyền, sao lưu, phục hồi dữ liệu, giải quyết các tranh chấp dữ liệu ...



5. Hệ quản trị cơ sở dữ liệu (1)

- Hệ quản trị cơ sở dữ liệu (DBMS – DataBase Management System) là hệ thống các phần mềm hỗ trợ tích cực cho các nhà phân tích, thiết kế và khai thác CSDL.
- Các DBMS thông dụng: Visual FoxPro, Microsoft Access, SQL Server, DB2, My SQL, Oracle ... hầu hết các DBMS hiện nay đều dựa trên mô hình quan hệ.

5. Hệ quản trị cơ sở dữ liệu (2)

- Một DBMS phải có:
 - Ngôn ngữ giao tiếp giữa người sử dụng và CSDL
 - Từ điển dữ liệu (Data Dictionary)
 - Có biện pháp bảo mật khi có yêu cầu
 - Cơ chế giải quyết tranh chấp dữ liệu
 - Có cơ chế sao lưu (backup), phục hồi (restore)
 - Đảm bảo tính độc lập giữa dữ liệu và chương trình

5. Hệ quản trị cơ sở dữ liệu (3)

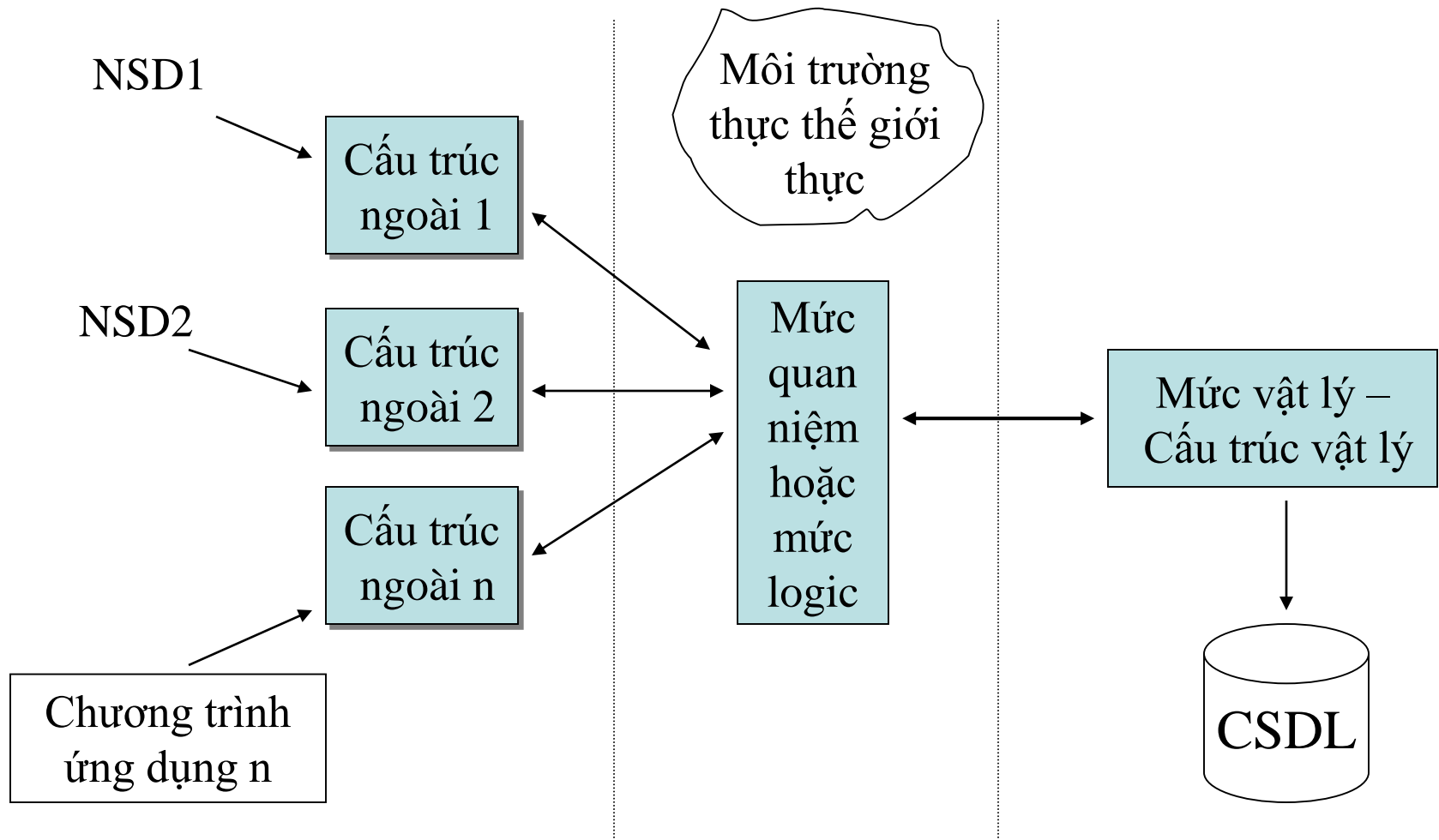
● Ngôn ngữ giao tiếp

- Ngôn ngữ mô tả dữ liệu (**DDL–Data Definition Language**): cho phép khai báo cấu trúc CSDL, các mối liên hệ của dữ liệu, các quy định, ràng buộc dữ liệu.
- Ngôn ngữ thao tác dữ liệu (**DML–Data Manipulation Language**): cho phép thực hiện thao tác thêm, xóa, sửa dữ liệu.
- Ngôn ngữ truy vấn có cấu trúc (**SQL–Structured Query Language**): cho phép người khai thác sử dụng để truy vấn thông tin cần thiết.
- Ngôn ngữ quản lý dữ liệu (**DCL–Data Control Language**) quản lý bảo mật, phân quyền cho người sử dụng.

5. Các mức biểu diễn 1 CSDL

- **Mức trong:** (mức vật lý – Physical) là mức lưu trữ CSDL
 - Cần giải quyết vấn đề gì?
 - Dữ liệu gì?
 - Lưu trữ như thế nào? ở đâu?
 - Cần các chỉ mục gì?
 - Truy xuất tuần tự hay ngẫu nhiên.
 - **Dành cho người quản trị CSDL.**
- **Mức quan niệm:** (Conception hay Logical)
 - Cần phải lưu trữ bao nhiêu loại dữ liệu?
 - Là dữ liệu gì? mối quan hệ giữa chúng?
- **Mức ngoài:** của người sử dụng và các chương trình ứng dụng

Các mức biểu diễn 1 CSDL



NỘI DUNG

- Phần 1: Các khái niệm của một hệ CSDL

1. Đặt vấn đề
2. Hệ thống tập tin (File System)
3. Định nghĩa một CSDL
4. Các đối tượng sử dụng CSDL
5. Hệ quản trị CSDL
6. Các mức của một CSDL

- Phần 2: Các mô hình dữ liệu

1. Giới thiệu
2. Mô hình dữ liệu mạng
3. Mô hình dữ liệu phân cấp
4. Mô hình thực thể mối kết hợp
5. Mô hình dữ liệu quan hệ
6. Mô hình dữ liệu hướng đối tượng

CÁC MÔ HÌNH DỮ LIỆU

1. Giới thiệu

- Mô hình dữ liệu là sự trừu tượng hóa của môi trường thực. Một số mô hình:
 - Mô hình dữ liệu mạng
 - Mô hình dữ liệu phân cấp
 - Mô hình dữ liệu thực thể mối kết hợp
 - Mô hình dữ liệu quan hệ
 - Mô hình dữ liệu hướng đối tượng

2. Mô hình dữ liệu mạng (1)

- Mô hình dữ liệu mạng (Network Data Model) còn gọi tắt là mô hình mạng (Network Model) là mô hình được biểu diễn bởi một đồ thị có hướng.
- Được xây dựng bởi Honeywell vào năm 1964-1965
- Một số khái niệm:
 - Mẫu tin (record)
 - Loại mẫu tin
 - Loại liên hệ (set type)
 - Bản số

2. Mô hình dữ liệu mạng (2)

- **Mẫu tin:**

- Mô tả 1 đối tượng trong thế giới thực
- ('NV001','Nguyen Lam','Nam','10/10/1970','Dong Nai')

- **Loại mẫu tin:**

- Là 1 tập các mẫu tin có cùng tính chất
- Ví dụ: nhân viên (NHANVIEN)

- Ký hiệu:

NHANVIEN

CONGVIEC

- **Loại liên hệ:**

- Mô tả sự liên kết giữa 1 loại mẫu tin chủ và 1 loại mẫu tin thành viên

- Ví dụ: tham gia (ThamGia)

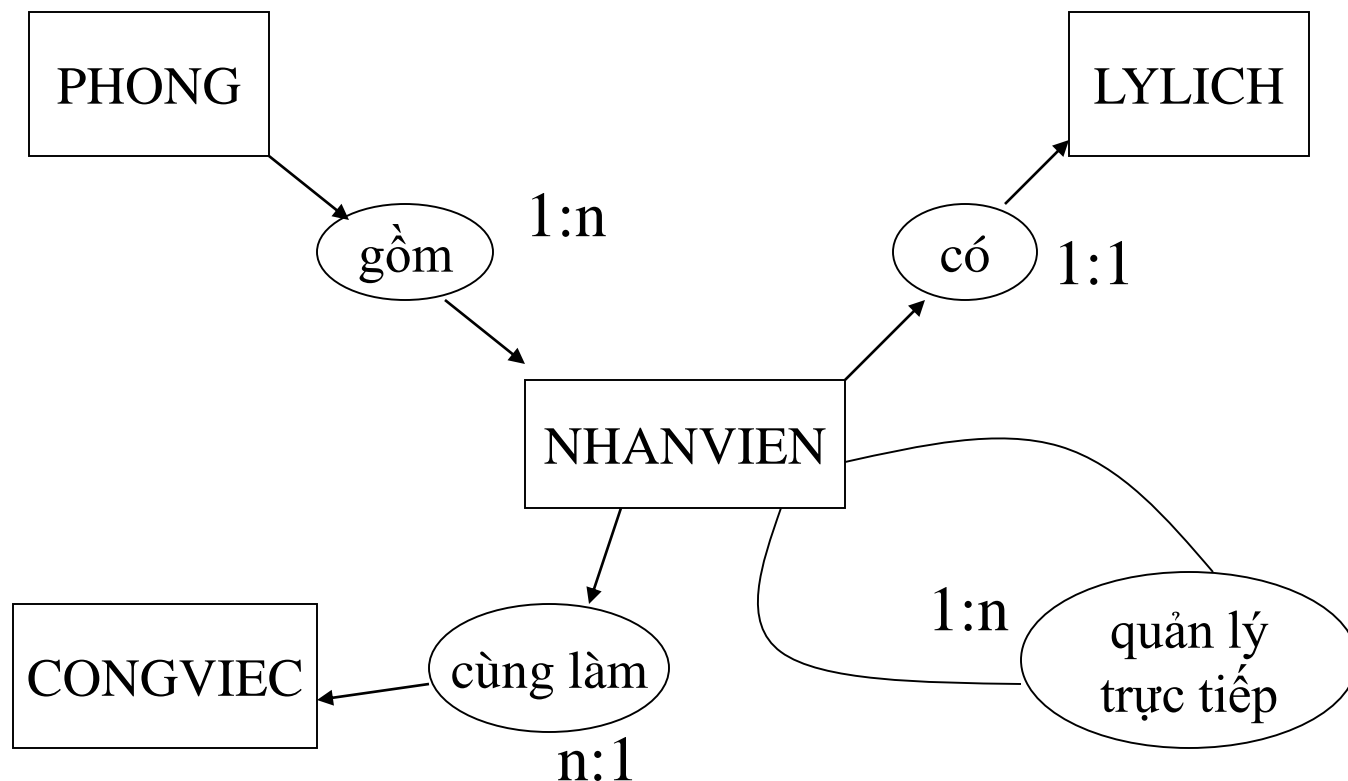
- Ký hiệu:

Tham gia

2. Mô hình dữ liệu mạng (3)

- **Bản số:** chỉ ra số lượng các mẫu tin tham gia trong mỗi liên hệ
 - **(1:1)** (one-to-one): mỗi mẫu tin của loại mẫu tin chủ kết hợp với đúng 1 mẫu tin của loại mẫu tin thành viên.
 - **(1:n)** (one-to-many): mỗi mẫu tin của loại mẫu tin chủ kết hợp với 1 hay nhiều mẫu tin thành viên.
 - **(n:1)** (many-to-one): nhiều mẫu tin của loại mẫu tin chủ kết hợp với đúng 1 mẫu tin của loại mẫu tin thành viên.
 - **Đệ qui** (Recursive): một loại mẫu tin chủ cũng có thể đồng thời là loại mẫu tin thành viên với chính nó.

2. Mô hình dữ liệu mạng (4)



2. Mô hình dữ liệu mạng (5)

● Nhận xét

- Tương đối đơn giản
- Dễ sử dụng
- Không thích hợp biểu diễn CSDL có quy mô lớn
- Khả năng diễn đạt ngữ nghĩa kém

2. Mô hình dữ liệu mạng (6)

- **Bài tập:**

Xây dựng mô hình dữ liệu mạng cho cơ sở dữ liệu quản lý bán hàng trong một siêu thị

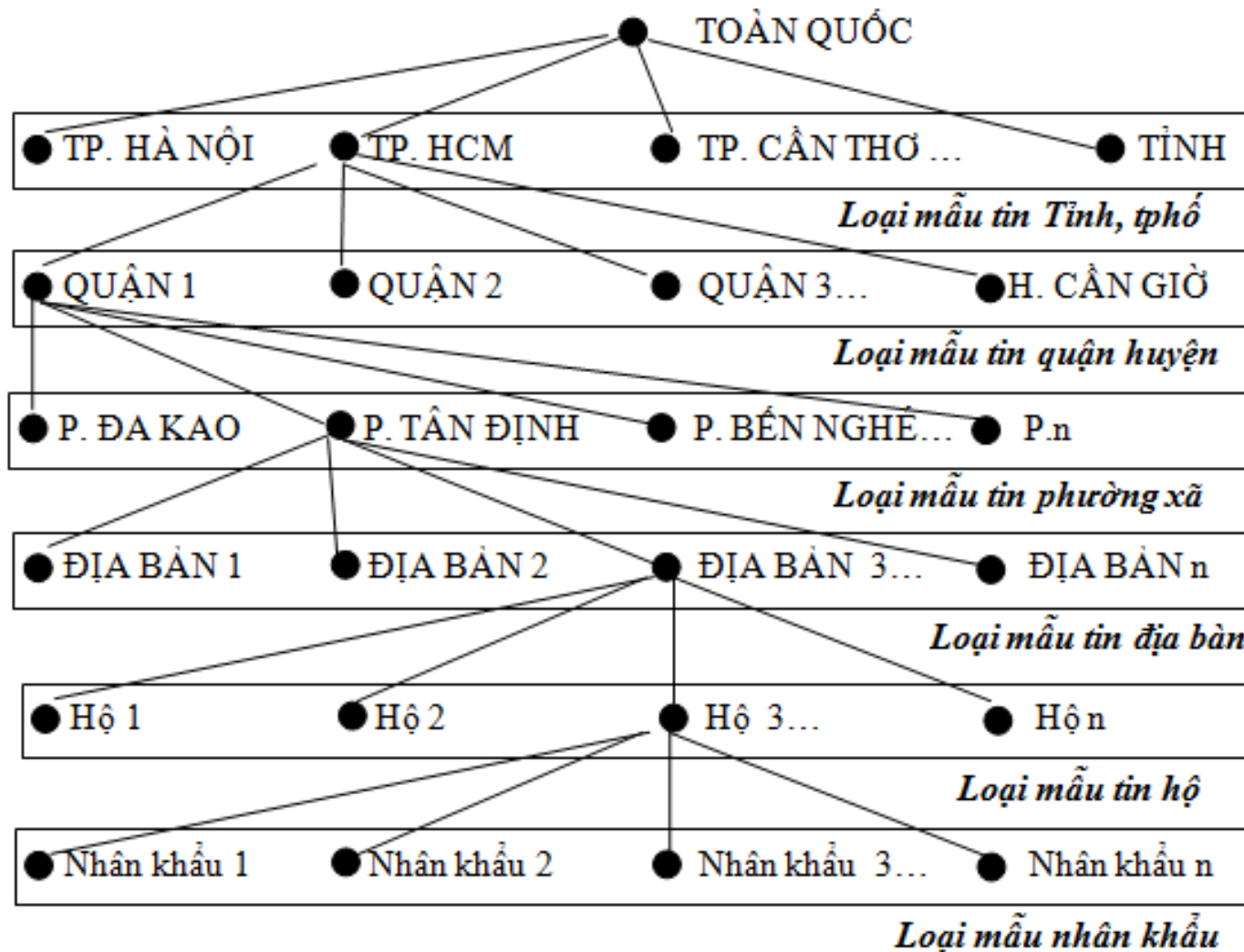
3. Mô hình dữ liệu phân cấp (1)

- Mô hình dữ liệu phân cấp (Hierarchical Data Model) còn gọi tắt là mô hình phân cấp (Hierarchical Model) dạng cây phân cấp, trong đó mỗi nút của cây biểu diễn một thực thể, giữa nút con với nút cha được liên hệ với nhau theo một mối quan hệ xác định.
- Được xây dựng bởi IBM và North American Rockwell vào khoảng năm 1965.
- Một số khái niệm:
 - Loại mẫu tin: giống mô hình dữ liệu mạng
 - Loại mối liên hệ: phân cấp

3. Mô hình dữ liệu phân cấp (2)

- Giữa hai loại mẫu tin **chỉ tồn tại một mối quan hệ duy nhất**. Như vậy, mỗi liên hệ từ mẫu tin chủ tới các mẫu tin thành viên là $1 \div n$ và từ mẫu tin thành viên với mẫu tin chủ là $1 \div 1$

3. Mô hình dữ liệu phân cấp (3)



3. Mô hình dữ liệu phân cấp (4)

● Nhận xét

- Đơn giản
- Không thích hợp biểu diễn CSDL có quy mô lớn
- Khả năng diễn đạt ngữ nghĩa kém

4. Mô hình thực thể mối kết hợp (ER – Entity Relationship Model)

- Giới thiệu
- Loại thực thể, thực thể
- Thuộc tính của loại thực thể
- Khoá của loại thực thể
- Loại mối kết hợp, mối kết hợp
- Thuộc tính của loại mối kết hợp
- Bản số
- Mô hình ER mở rộng

Giới thiệu

- Mô hình thực thể mối kết hợp (**Entity-Relationship Model** viết tắt ER) được Peter Pin-Shan CHEN giới thiệu năm 1976.
- Nhìn thể giới thực như là một tập các đối tượng căn bản được gọi là **các tập thực thể**, và **các mối quan hệ** giữa các đối tượng này.
- Mô hình ER được sử dụng nhiều trong thiết kế dữ liệu ở mức quan niệm.

Loại thực thể (Entity Type)

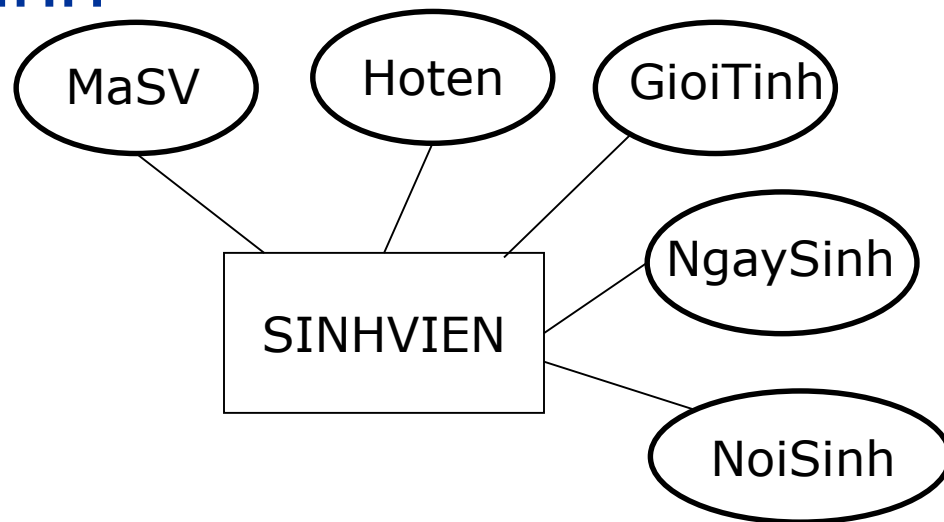
- Loại thực thể là những loại đối tượng hay sự vật của thế giới thực tồn tại cụ thể cần được quản lý.
- NHANVIEN, PHONGBAN, LOP, MONHOC, SINHVIEN...
- Ký hiệu:

Tên loại thực thể

NHANVIEN

Thuộc tính của loại thực thể (Entity Attribute)

- Là những tính chất đặc trưng của loại thực thể.
- Loại thực thể SINHVIEN có các thuộc tính: Mã sinh viên, họ tên, giới tính, ngày sinh, nơi sinh
- Ký hiệu:



Thực thể

- Thực thể là một thể hiện của một loại thực thể.
- Loại thực thể là SINHVIEN có các thực thể:
 - ('SV001', 'Nguyễn Minh', 'Nam', '1/2/1987', 'Tp. Hồ Chí Minh')
 - ('SV002', 'Trần Năm', 'Nam', '13/2/1987', 'Đồng Nai')

Các loại thuộc tính (1)

- Đơn trị (Simple)

- Chỉ có một giá trị, không thể chia nhỏ.
- Ví dụ: MaSV, TenMonHoc, SoTC,...

- Đa hợp (Composite) (hay còn gọi là kết hợp)

- Thuộc tính có thể được tạo thành từ nhiều thành phần.
- Ví dụ: DiaChi (SoNha, Duong, Phuong, Quan, Tp_Tinh) hay HoTen (Ho,TenLot,Ten).

→ Lưu ý về mặt ý nghĩa của Thuộc tính trong từng thiết kế

- Đa trị (Multi-valued)

- Thuộc tính có thể có nhiều giá trị đối với một thực thể.
- BANGCAP ký hiệu {BANGCAP}

Các loại thuộc tính (2)

- Các thuộc tính đa hợp và đa trị có thể lồng nhau.
- Ví dụ: thuộc tính BangCap của HOCVIEN là một thuộc tính đa hợp được ký hiệu bằng {BangCap (NoiCap, NamCap, KetQua, ChuyenNganh)}

● Thuộc tính suy ra

- Giá trị của thuộc tính được suy ra từ các thuộc tính khác.
- Ví dụ:
 - ◆ Giá trị của thuộc tính lương (Luong) được suy ra từ giá trị thuộc tính Hệ số lương (HSL)
 - ◆ Giá trị của thuộc tính kết quả thi (KetQuaThi) sẽ được suy ra giá trị thuộc tính điểm thi (Diem)

● Thuộc tính có giá trị rỗng

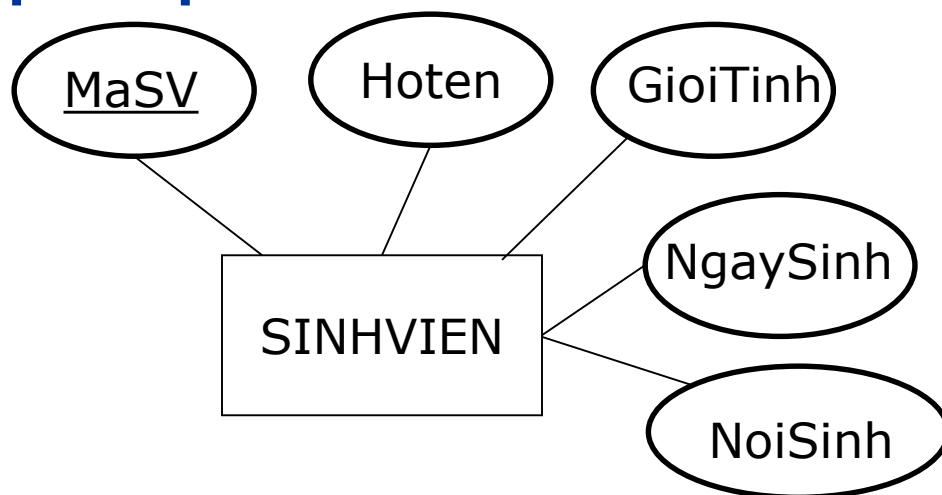
- Giá trị của thuộc tính là rỗng (NULL)
- Ý nghĩa: giá trị của thuộc tính là không có hoặc chưa được xác định

Bài tập nhóm

- Nhóm: 5 sinh viên
- Thời gian: 10 phút
- Yêu cầu:
 - Tìm ít nhất 3 loại thực thể (tên loại thực thể, các thuộc tính, xác định khóa)
 - Tìm Thể hiện của các loại thực thể tìm được

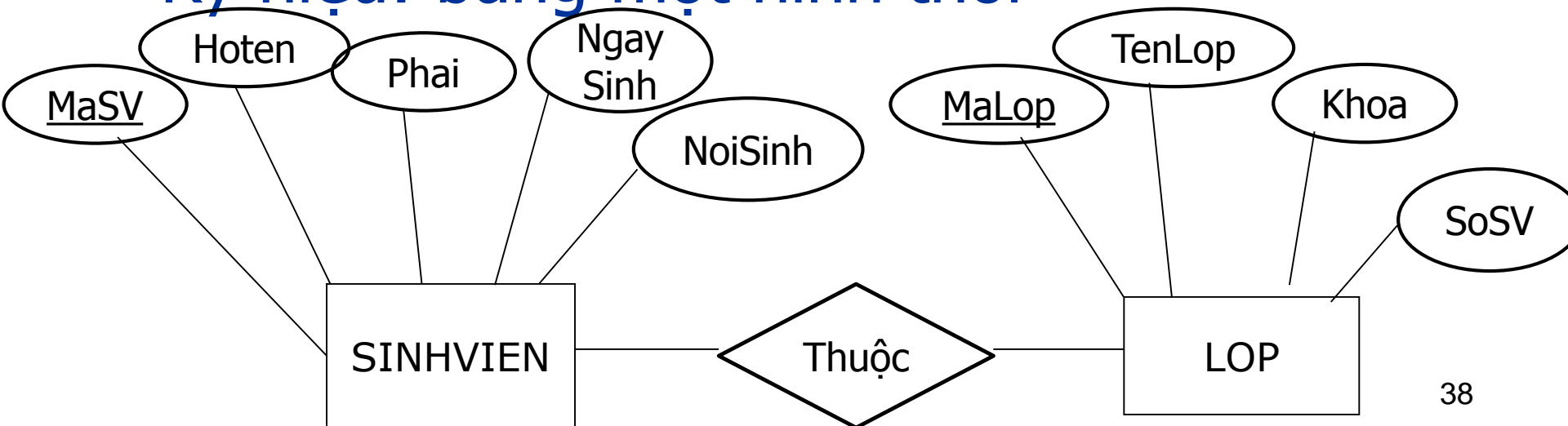
Khoá của loại thực thể

- Khóa của loại thực thể là thuộc tính nhận diện thực thể.
- Căn cứ vào giá trị của khóa có thể xác định **duy nhất** một thực thể.
- Mỗi sinh viên có một mã số duy nhất → Khóa của loại thực thể **SINHVIEN** là **MaSV**



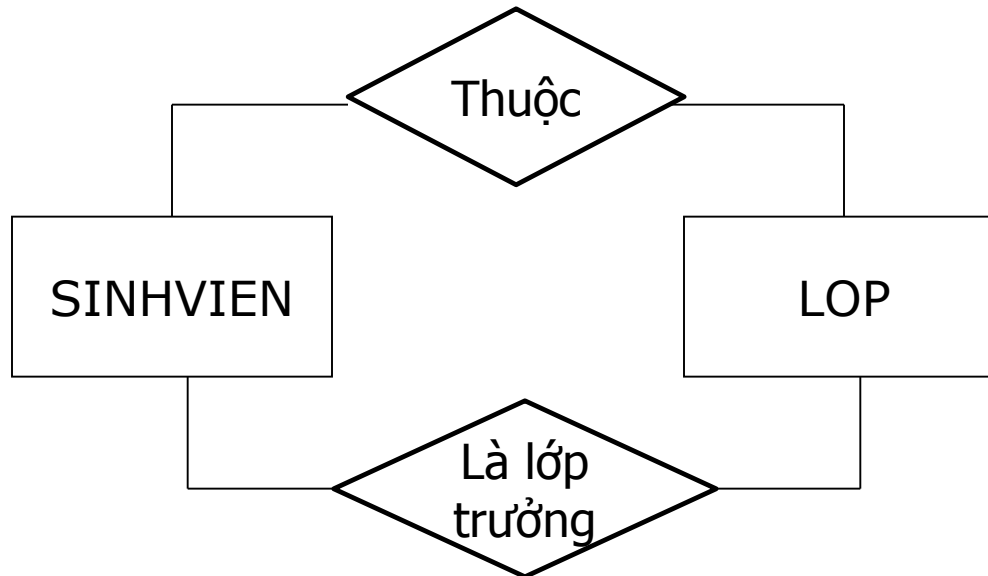
Loại mỗi kết hợp (1) (relationship type)

- Loại mỗi kết hợp là sự liên kết giữa hai hay nhiều loại thực thể
- Giữa hai loại thực thể SINHVIEN và LOP có loại mỗi kết hợp THUOC
- Ký hiệu: bằng một hình thoi



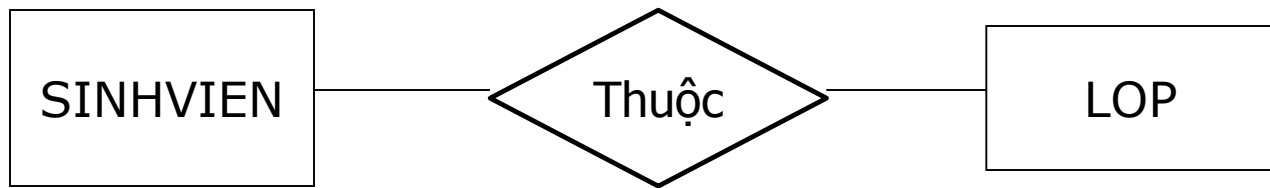
Loại mỗi kết hợp (2)

- Giữa hai loại thực thể có thể tồn tại nhiều hơn một loại mỗi kết hợp.



Số ngôi của loại mỗi kết hợp (relationship degree)

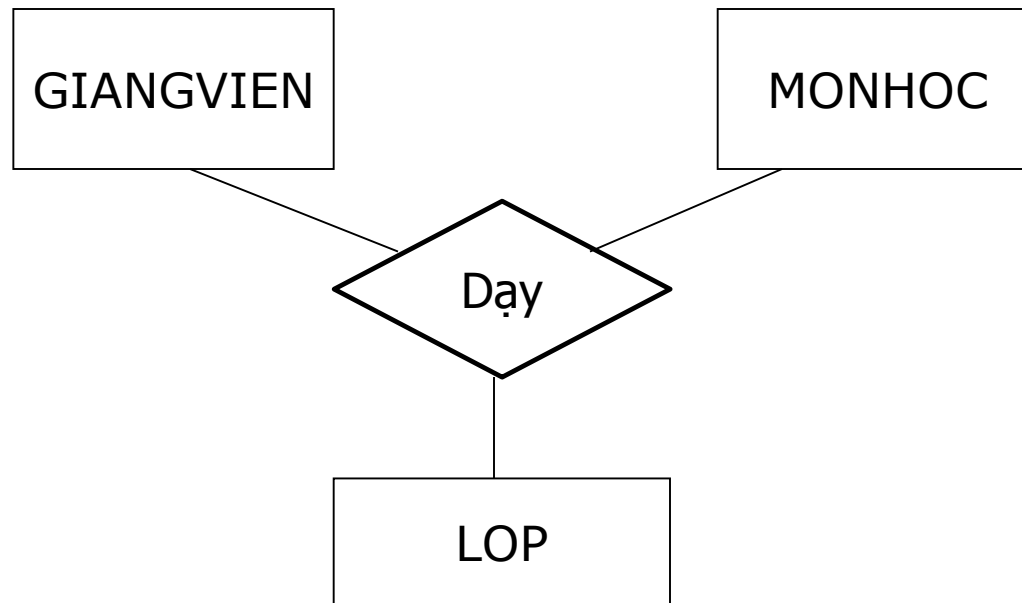
- Là số loại thực thể tham gia vào loại mỗi kết hợp đó.



- Số ngôi của loại MKH Thuộc là 2

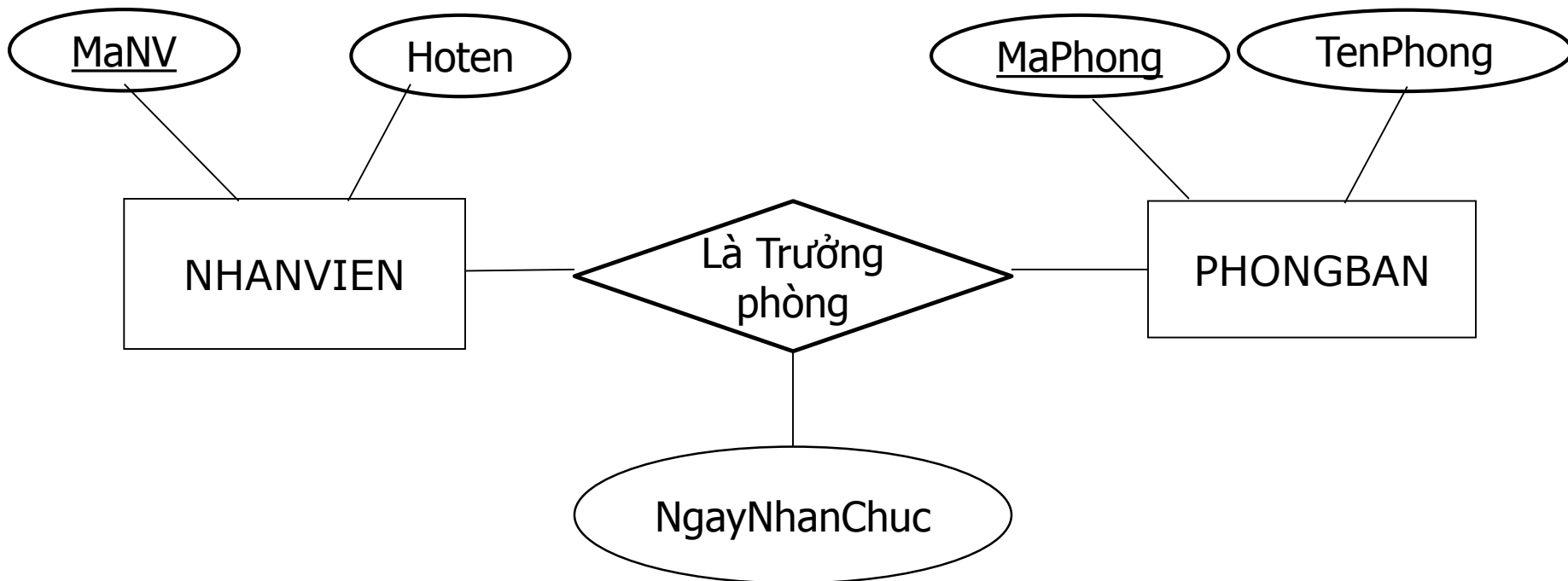
Số ngôi của loại mỗi kết hợp

- Số ngôi của loại MKH Dạy là 3

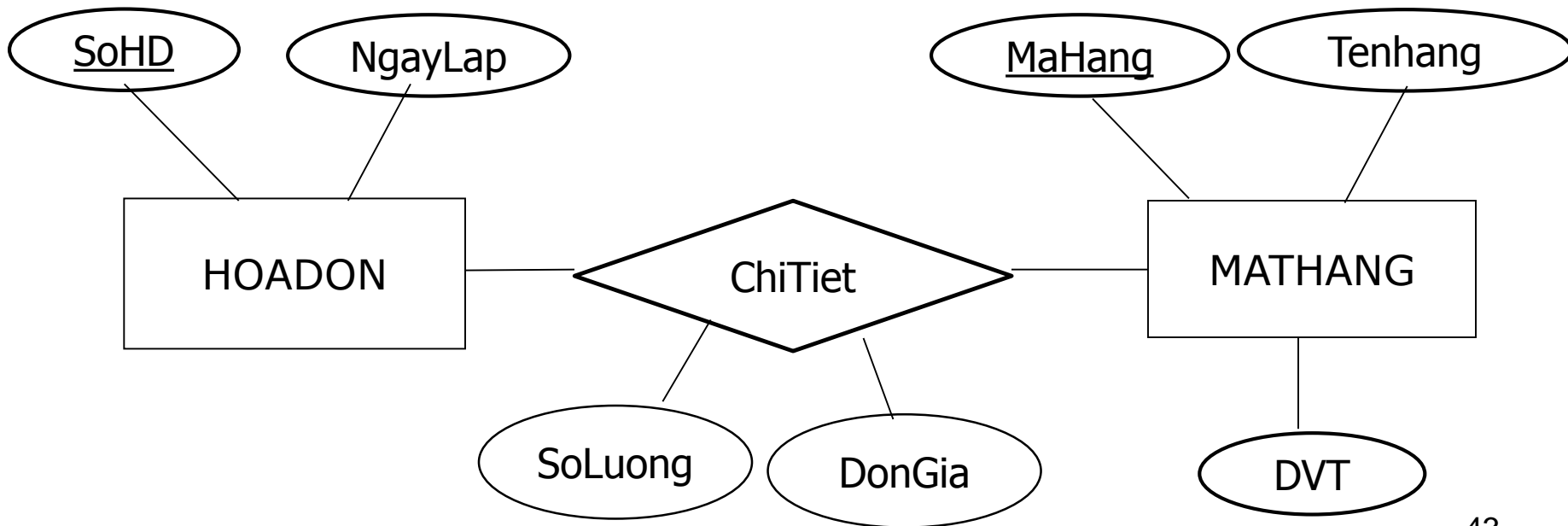


Thuộc tính của loại mỗi kết hợp (relationship type attribute)

- Gồm các thuộc tính khoá của các loại thực thể tham gia vào loại mỗi kết hợp đó, và còn có thể có thêm những thuộc tính bổ sung khác

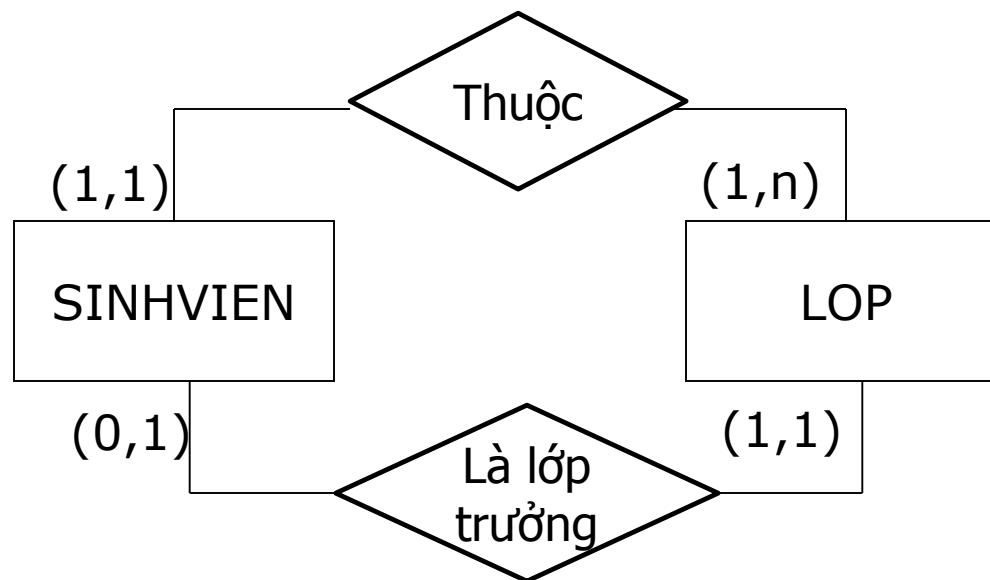


Thuộc tính của loại mỗi kết hợp

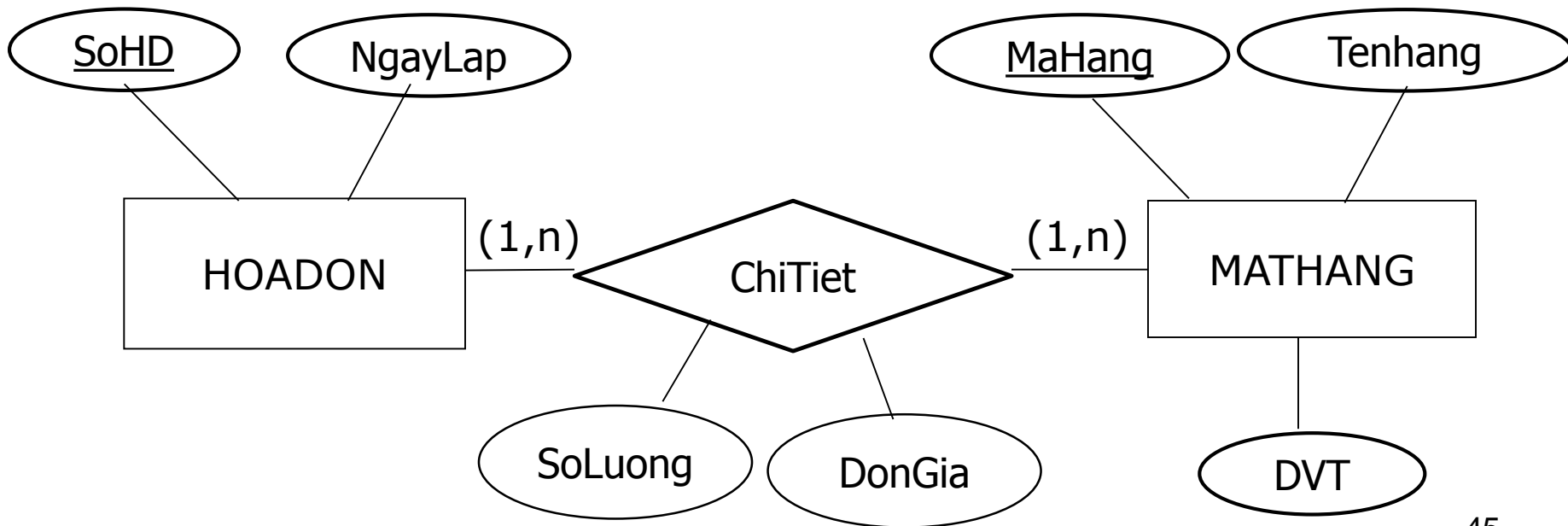


Bản số (relationship cardinality)

- Bản số thể hiện số lượng tối thiểu và số lượng tối đa các thực thể tham gia vào loại mỗi kết hợp.
- Ký hiệu: (số lượng tối thiểu, số lượng tối đa)



Bản số (relationship cardinality)

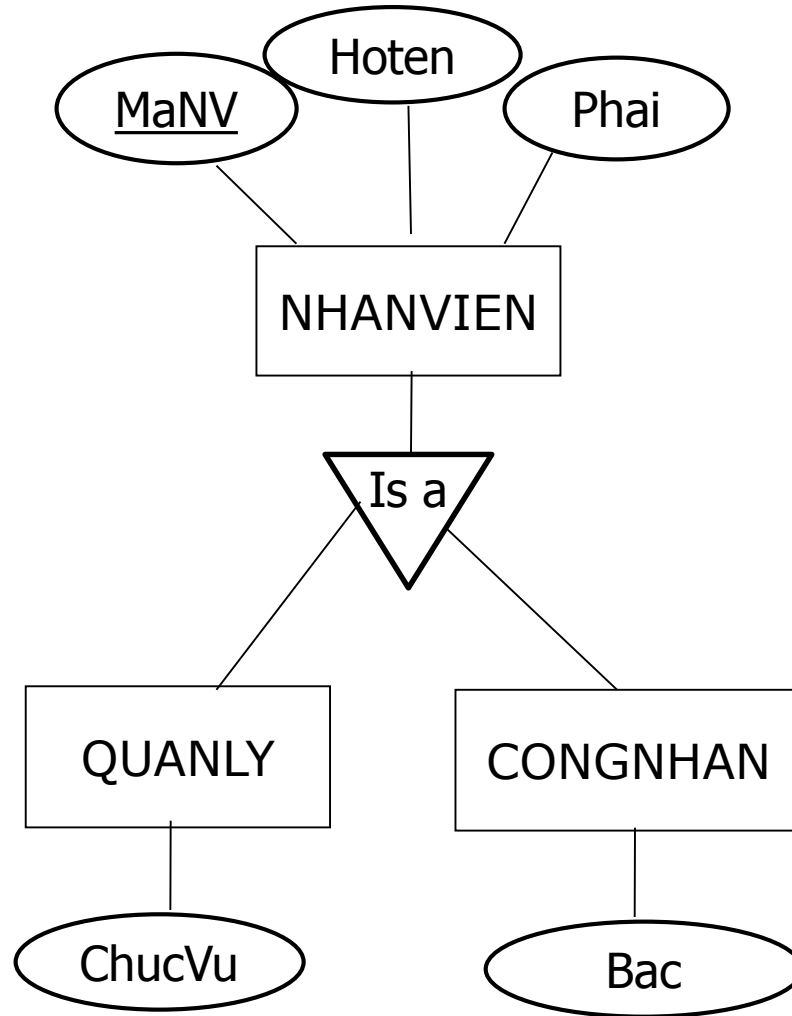


45

Mô hình ER mở rộng

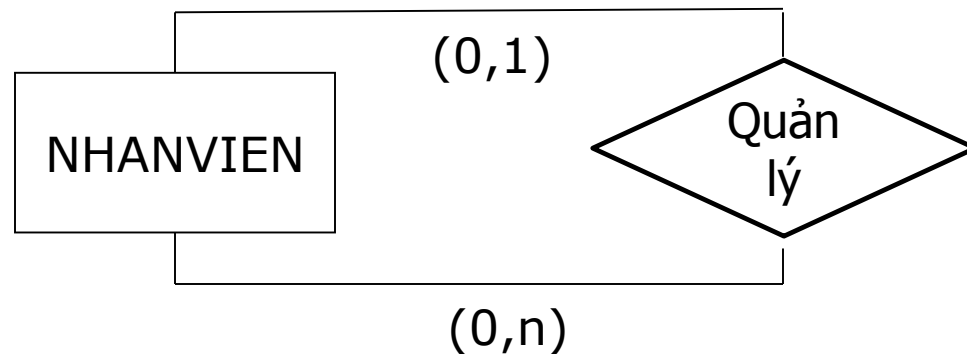
- Chuyên biệt hoá, tổng quát hóa
- Mỗi kết hợp đệ quy
- Loại thực thể yếu
- Mỗi kết hợp mở rộng

Chuyên biệt hóa, tổng quát hóa



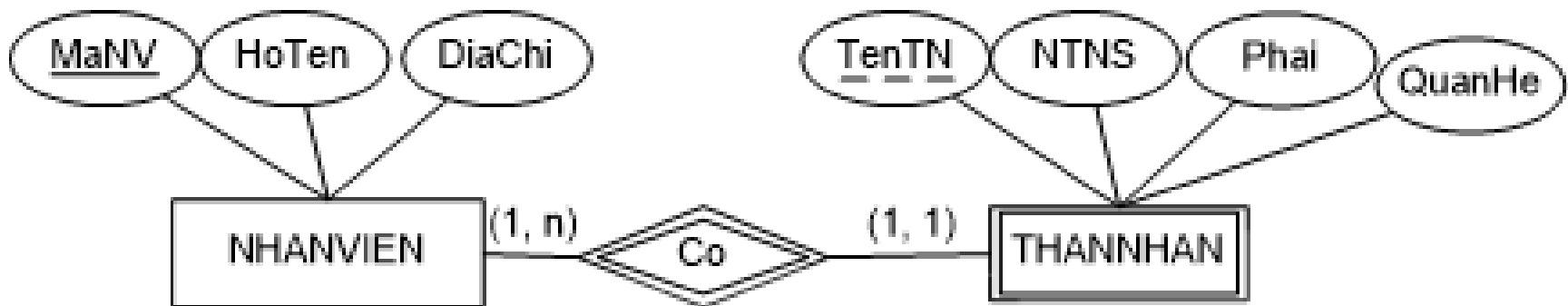
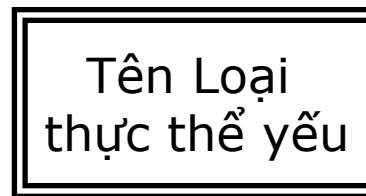
Mối kết hợp đệ quy

- Là loại mối kết hợp được tạo thành từ cùng một loại thực thể (hay một loại thực thể có loại mối kết hợp với chính nó)
- Mỗi nhân viên có một người quản lý trực tiếp và người quản lý đó cũng là một nhân viên

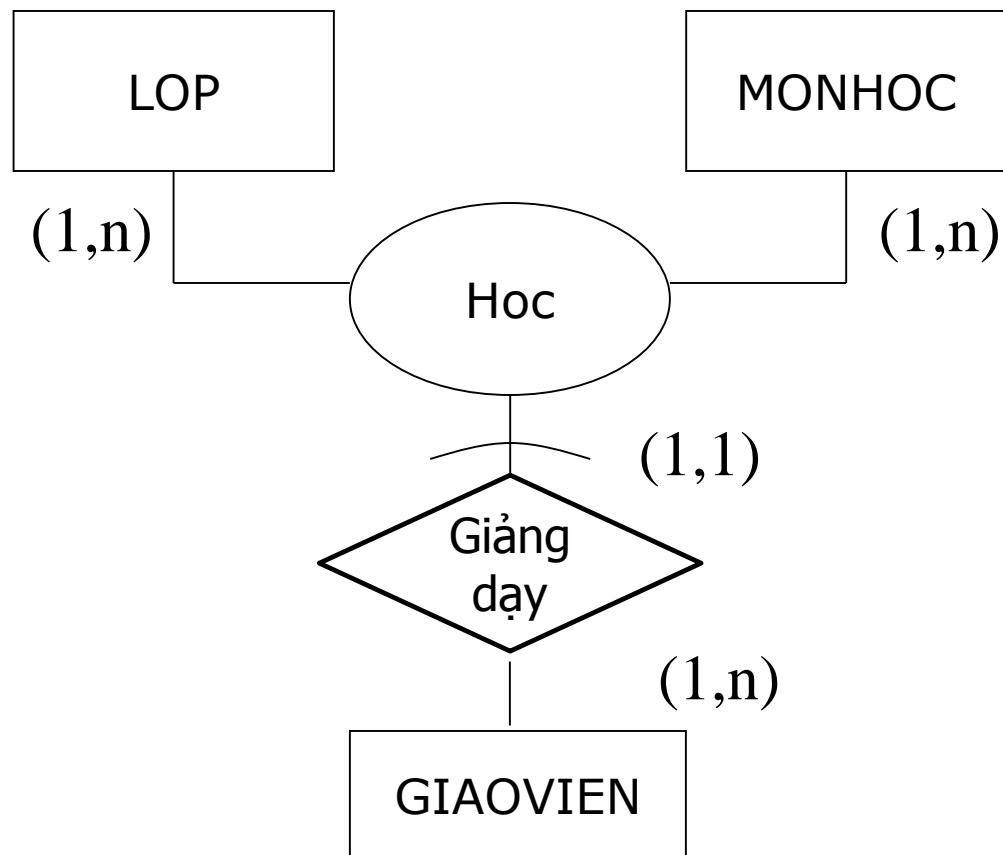


Loại thực thể yếu

- Là loại thực thể không có khóa.
- Phải tham gia trong một loại mối kết hợp xác định trong đó có một loại thực thể chủ.
- Ký hiệu:



Mối kết hợp mở rộng



Bài tập về nhà

Mô tả tình huống cho các bài tập được đăng ở <https://courses.uit.edu.vn/>

- Bài 1: Xây dựng mô hình ER cho tình huống quản lý dự án của một công ty.
- Bài 2: Xây dựng mô hình ER cho tính huống quản lý công việc ghi danh và quản lý lớp học tại trung tâm ngoại ngữ.

Mô hình dữ liệu hướng đối tượng

- Mô hình dữ liệu hướng đối tượng (Object Oriented Data Model) ra đời vào cuối những năm 80 và đầu những năm 90
- Dựa trên cách tiếp cận hướng đối tượng.
- Các khái niệm
 - lớp (class), thuộc tính
 - sự kế thừa (inheritance)
 - ..
- Đặc trưng
 - tính đóng gói (encapsulation)
 - tính đa hình (polymorphism)
 - tính tái sử dụng (reusability).

NỘI DUNG

- Phần 1: Các khái niệm của một hệ CSDL
 1. Đặt vấn đề
 2. Hệ thống tập tin (File System)
 3. Định nghĩa một CSDL
 4. Các đối tượng sử dụng CSDL
 5. Hệ quản trị CSDL
 6. Các mức của một CSDL
- Phần 2: Các mô hình dữ liệu
 1. Giới thiệu
 2. Mô hình dữ liệu mạng
 3. Mô hình dữ liệu phân cấp
 4. Mô hình thực thể mối kết hợp
 5. Mô hình dữ liệu quan hệ
 6. Mô hình dữ liệu hướng đối tượng

