



CƠ SỞ DỮ LIỆU

CHƯƠNG 1: TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU

ThS. LÊ NGÔ THỰC VI



MỤC TIÊU



1. Vai trò và tầm quan trọng của CSDL.
2. Hiểu rõ khái niệm, đối tượng sử dụng CSDL, hệ quản trị CSDL
3. Xác định được các dữ kiện, thiết kế mô hình thực thể mối kết hợp (ERD)



NỘI DUNG

1. Giới thiệu
2. Quá trình phát triển
3. Khái niệm CSDL
4. Các đối tượng sử dụng CSDL
5. Hệ quản trị CSDL
6. Các mức biểu diễn một CSDL
7. Các mô hình dữ liệu



GIỚI THIỆU

1

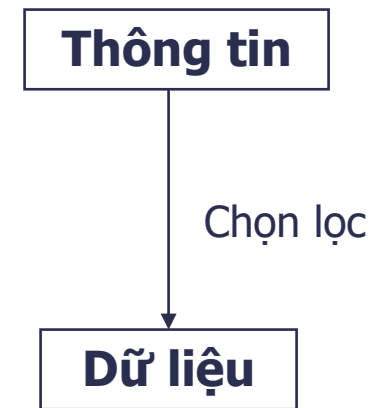
1. Giới thiệu





1. Giới thiệu - Dữ liệu

- Dữ liệu (Data)
 - Một mô tả hình thức về những sự kiện, khái niệm
 - Tên, địa chỉ, số điện thoại của khách hàng
 - Báo cáo doanh thu
 - Đăng ký học phần
 - ...



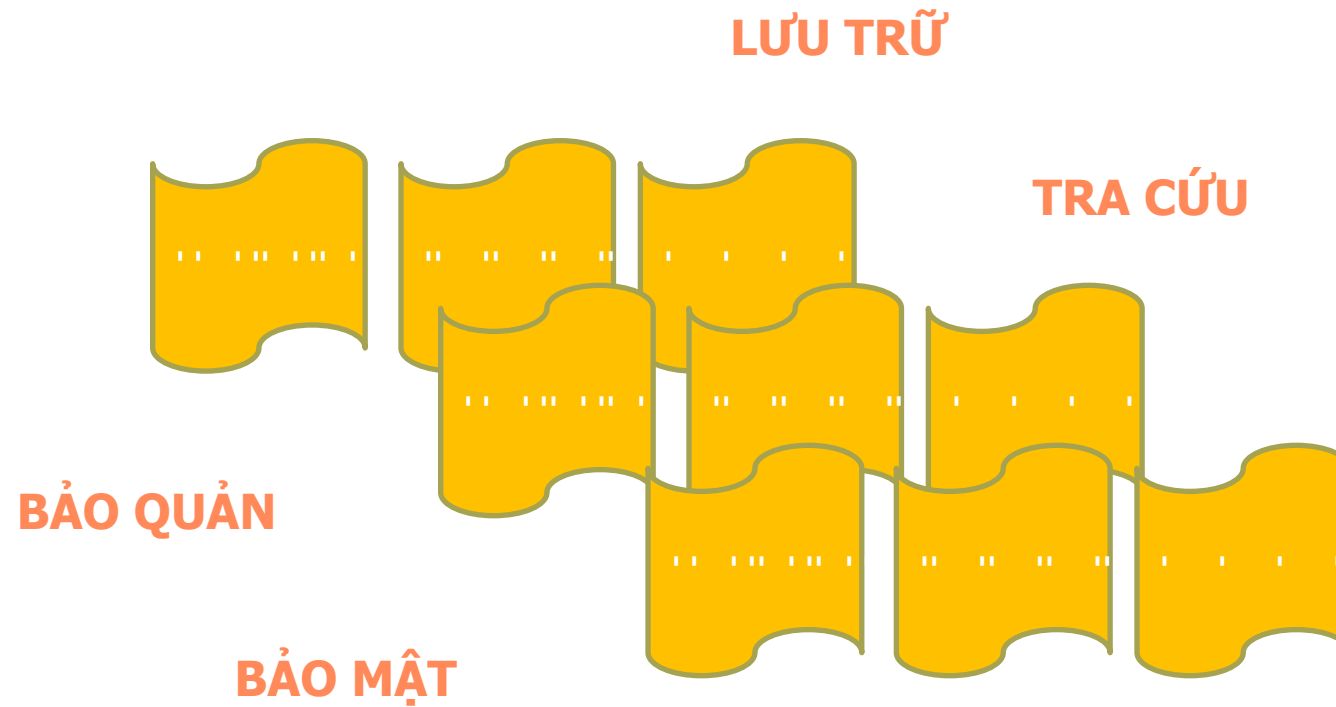


QUÁ TRÌNH PHÁT TRIỂN

2

2. Quá trình phát triển

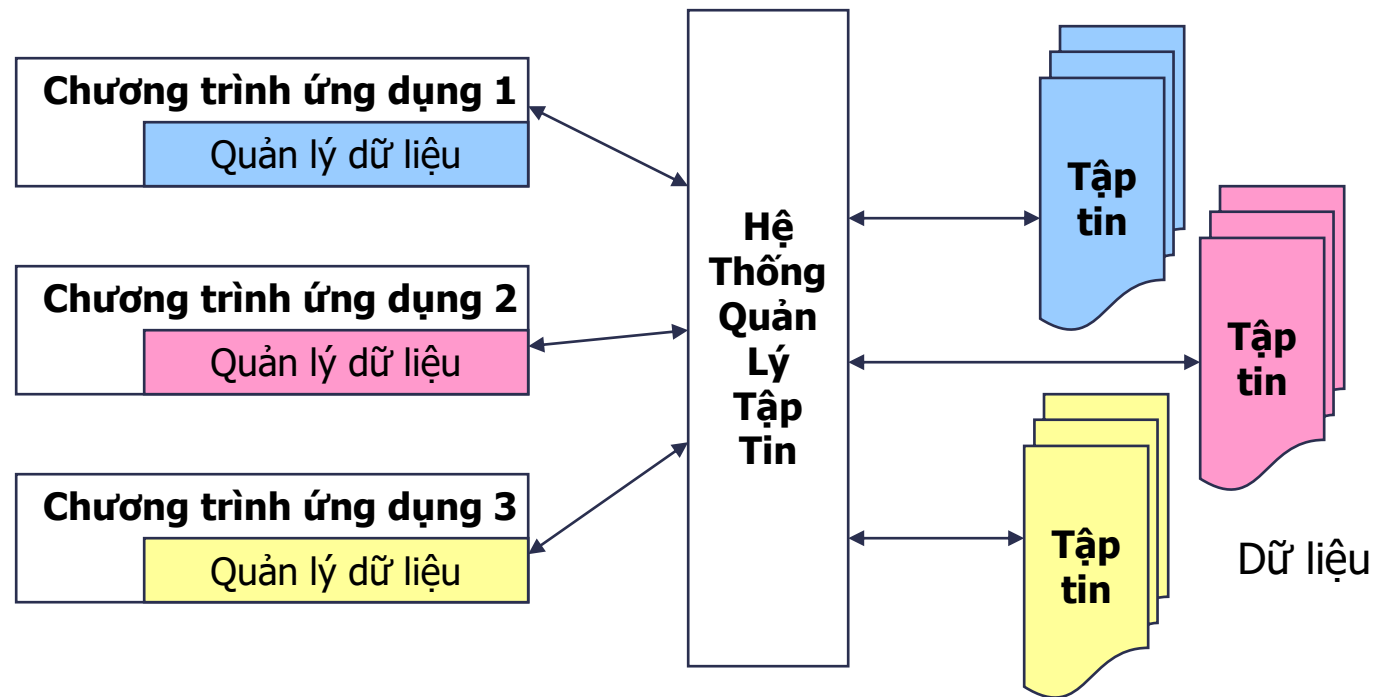
- Tài liệu giấy



2. Quá trình phát triển (tt)

- **Hệ thống tập tin (File)**

- Là tập hợp các tập tin **riêng lẻ** phục vụ cho một mục đích sử dụng nhất định





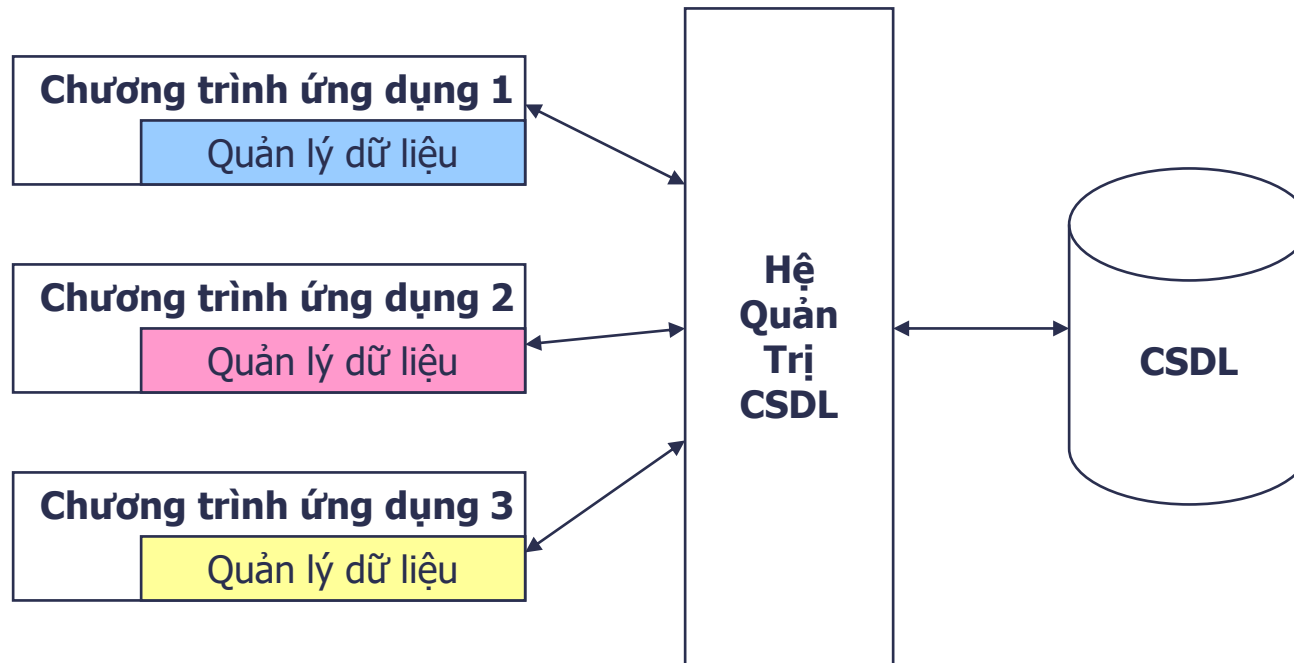
2. Quá trình phát triển (tt)

- **Hệ thống tập tin (File)**

- 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
 - Dữ liệu trùng lặp và dư thừa → lãng phí
 - Thiếu tính nhất quán giữa các dữ liệu
 - Khó khăn trong việc truy xuất
 - Chia sẻ dữ liệu bị hạn chế
 - Khó khôi phục
 - Không đáp ứng nhu cầu khai thác thông tin đồng thời

2. Quá trình phát triển (tt)

- Tiếp cận CSDL





KHÁI NIỆM CƠ SỞ DỮ LIỆU

3



3. Cơ sở dữ liệu

- **Cơ sở dữ liệu** (Database) 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.
- Một CSDL biểu diễn 1 phần của thế giới thực (*danh sách sinh viên, danh mục môn học...*)
- Tập ngẫu nhiên dữ liệu không được xem là một cơ sở dữ liệu



3. Cơ sở dữ liệu (tt)

- **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 (tt)

- **Những vấn đề cần giải quyết:**
 - Tính chủ quyền dữ liệu.
 - Tính bảo mật và quyền khai thác thông tin của người sử dụng
 - Đảm bảo dữ liệu khi có sự cố



CÁC ĐỐI TƯỢNG SỬ DỤNG

4

4. Các đối tượng sử dụng CSDL

- **Người dùng cuối (end user)**

- Ít khi truy cập CSDL, nhưng cần những thông tin khác nhau trong mỗi lần truy cập và dùng những câu truy vấn phức tạp (VD: Người quản lý)
- Thường xuyên truy vấn và cập nhật CSDL nhờ vào một số các chức năng đã được xây dựng sẵn (VD: Nhân viên)
- Thông thạo về HQT CSDL, tự xây dựng những truy vấn phức tạp cho công việc (VD: Nhà khoa học, người phân tích tư vấn kinh doanh)



4. Các đối tượng sử dụng CSDL (tt)

- **Chuyên viên tin học:**

- Chịu trách nhiệm xây dựng các ứng dụng phục vụ cho quản lý
 - Lựa chọn cấu trúc phù hợp để lưu trữ dữ liệu
 - Quyết định những dữ liệu nào cần được lưu trữ
- Liên hệ với người dùng để nắm bắt được những yêu cầu và đưa ra một thiết kế CSDL thỏa yêu cầu này
- Các quản trị viên sẽ quản lý các CSDL sau khi việc thiết kế hoàn tất

4. Các đối tượng sử dụng CSDL (tt)

- **Quản trị viên (Database Administrator):**
 - Am hiểu về CSDL và hệ quản trị CSDL
 - Tổ chức CSDL; Xử lý an toàn, bảo mật, cấp quyền; Sao lưu, phục hồi dữ liệu; Giải quyết các tranh chấp dữ liệu





HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU

(Database Management System - DBMS)

5

5. Hệ quản trị CSDL (DBMS)

- Tập hợp các chương trình cho phép người sử dụng tạo ra và duy trì CSDL
- Một phần mềm hệ thống cho phép định nghĩa, xây dựng và xử lý dữ liệu
 - Định nghĩa – khai báo bộ khung dữ liệu cùng với các mô tả chi tiết về dữ liệu
 - Xây dựng – lưu trữ dữ liệu lên bộ nhớ phụ
 - Xử lý – truy vấn, cập nhật và phát sinh báo cáo

5. Hệ quản trị CSDL (DBMS) (tt)

- Các DBMS thông dụng:

- Visual FoxPro,
- Microsoft Access,
- SQL Server,
- DB2, Oracle ...

hầu hết các DBMS hiện nay đều dựa trên mô hình quan hệ.



5. Hệ quản trị CSDL (DBMS) (tt)

- **Các tính năng của DBMS:**

- Kiểm soát được tính dư thừa của dữ liệu
 - Tích hợp các nhu cầu dữ liệu của người dùng để xây dựng một CSDL thống nhất
- Chia sẻ dữ liệu
 - Trong môi trường đa người dùng, các HQT phải cho phép truy xuất dữ liệu đồng thời
- Cấp quyền, hạn chế truy cập không cho phép
 - Từng người dùng và nhóm người dùng có một tài khoản và mật mã để truy xuất dữ liệu

5. Hệ quản trị CSDL (DBMS) (tt)

- **Các tính năng của DBMS:**

- Đảm bảo các ràng buộc toàn vẹn (Integrity Constraints)
 - RBTV là những qui định cần được thỏa mãn để đảm bảo dữ liệu luôn phản ánh đúng ngữ nghĩa của thế giới thực hoặc yêu cầu thiết kế
 - Một số RB được khai báo với HQT và HQT sẽ tự động kiểm tra.
- Khả năng sao lưu dự phòng khi gặp sự cố
 - Có khả năng khôi phục dữ liệu khi có sự hư hỏng về phần cứng hoặc phần mềm
- Đảm bảo tính độc lập giữa dữ liệu và chương trình



5. Hệ quản trị CSDL (DBMS) (tt)

- **Các tính năng của DBMS:**

- Ngôn ngữ giao tiếp giữa người sử dụng và CSDL
 - Ngôn ngữ định nghĩa dữ liệu (DDL – Data Definition Language)
 - Cho phép khai báo cấu trúc CSDL, mối liên hệ của dữ liệu
 - Được dùng để định nghĩa các lược đồ mức quan niệm và lược đồ trong
 - Ngôn ngữ thao tác dữ liệu (DML – Data Manipulation Language)
 - Cho phép thêm, xóa, sửa dữ liệu



5. Hệ quản trị CSDL (DBMS) (tt)

- **Các tính năng của DBMS:**
 - Ngôn ngữ giao tiếp giữa người sử dụng và CSDL
 - 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)
 - Cho phép khai báo bảo mật, cấp quyền cho người sử dụng

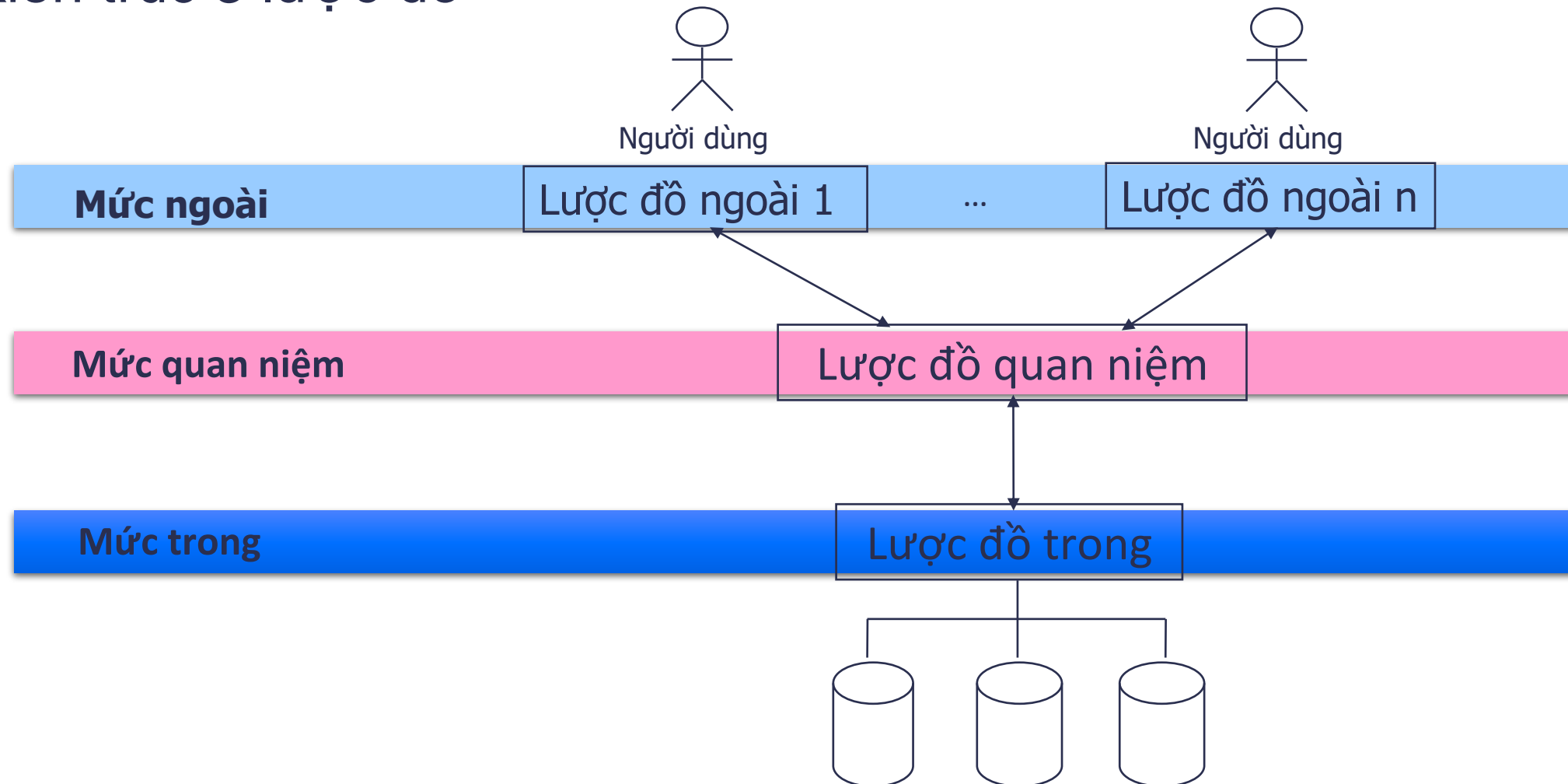


CÁC MỨC BIỂU DIỄN MỘT CSDL

6

6. Các mức biểu diễn một CSDL

- Kiến trúc 3 lược đồ



6. Các mức biểu diễn một CSDL (tt)

- **Mức ngoài**

- Còn gọi là mức khung nhìn (view)
- Mô tả một phần của CSDL mà 1 nhóm người dùng quan tâm đến và che dấu phần còn lại của CSDL đối với nhóm người dùng đó

- **Mức quan niệm (Conception hay Logical)**

- Mức nhìn tổng thể về CSDL
- Mô tả cấu trúc của toàn thể CSDL gồm thực thể, kiểu dữ liệu, mối liên hệ và ràng buộc

- **Mức trong (mức vật lý – Physical)**

- Mô tả cấu trúc lưu trữ vật lý CSDL: Dữ liệu gì? Lưu trữ thế nào? Ở đâu?...
- Góc nhìn dành cho người phát triển CSDL và người quản trị



CÁC MÔ HÌNH DỮ LIỆU

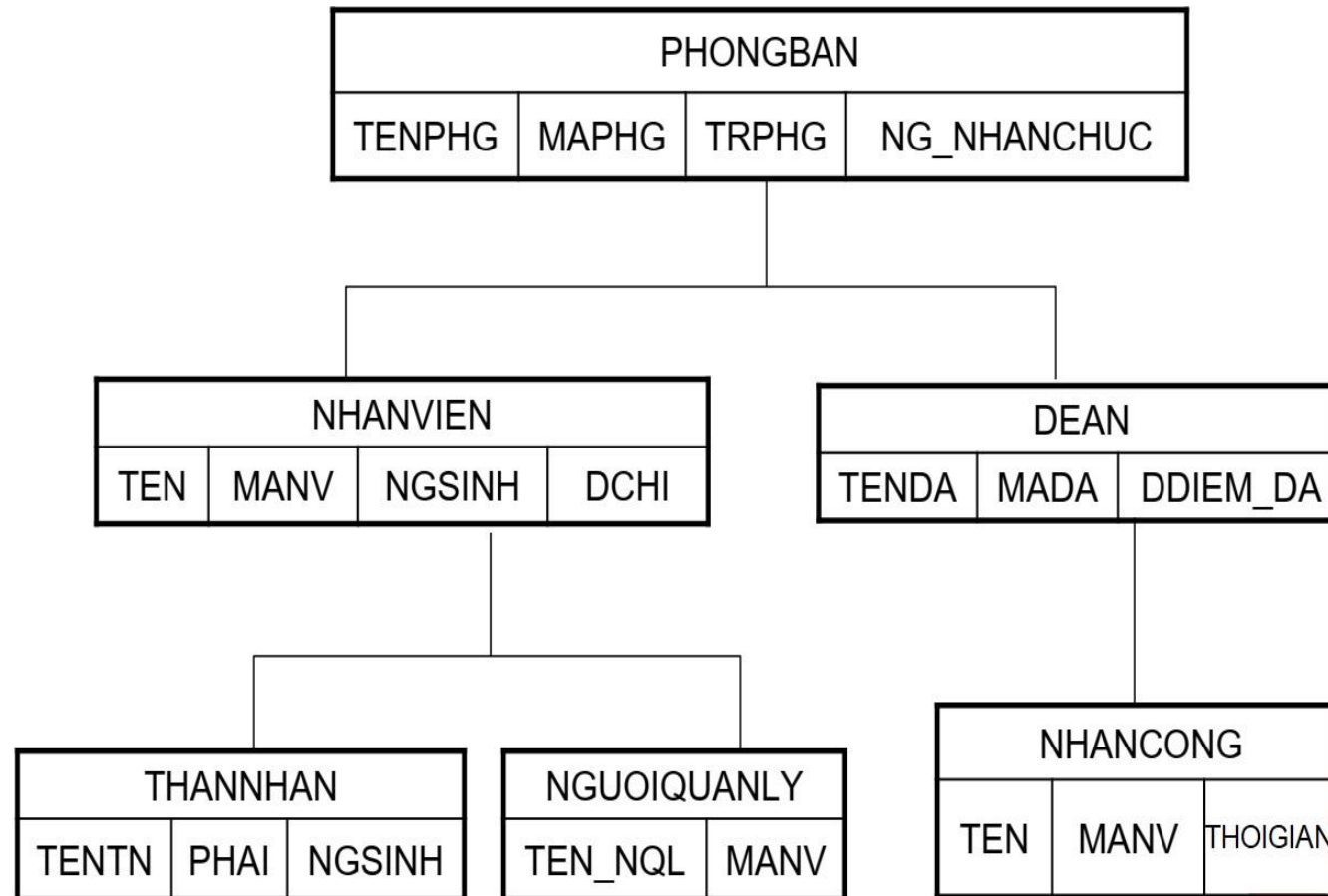
7

7. Các mô hình dữ liệu

- Mô hình dữ liệu (Data model) là sự trừu tượng hóa của môi trường thực để biểu diễn dữ liệu.
- Một số mô hình
 - Mô hình dữ liệu phân cấp
 - Mô hình dữ liệu mạng
 - Mô hình dữ liệu thực thể mối kết hợp
 - Mô hình dữ liệu quan hệ
 - Mô hình dữ liệu hướng đối tượng

7. Các mô hình dữ liệu (tt)

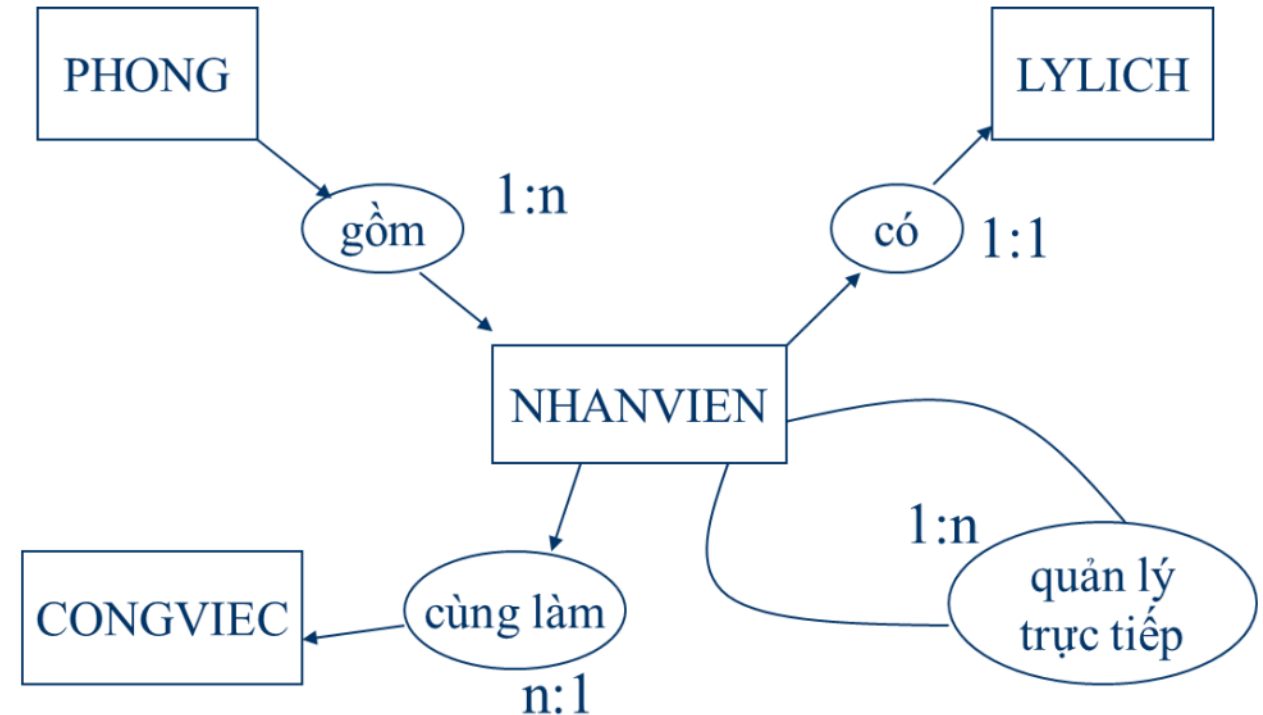
- Mô hình dữ liệu phân cấp



7. Các mô hình dữ liệu (tt)

- **Mô hình dữ liệu mạng** (*Network Data Model*) còn gọi tắt là mô hình mạng hoặc mô hình lưới là mô hình được biểu diễn bởi một đồ thị có hướng.

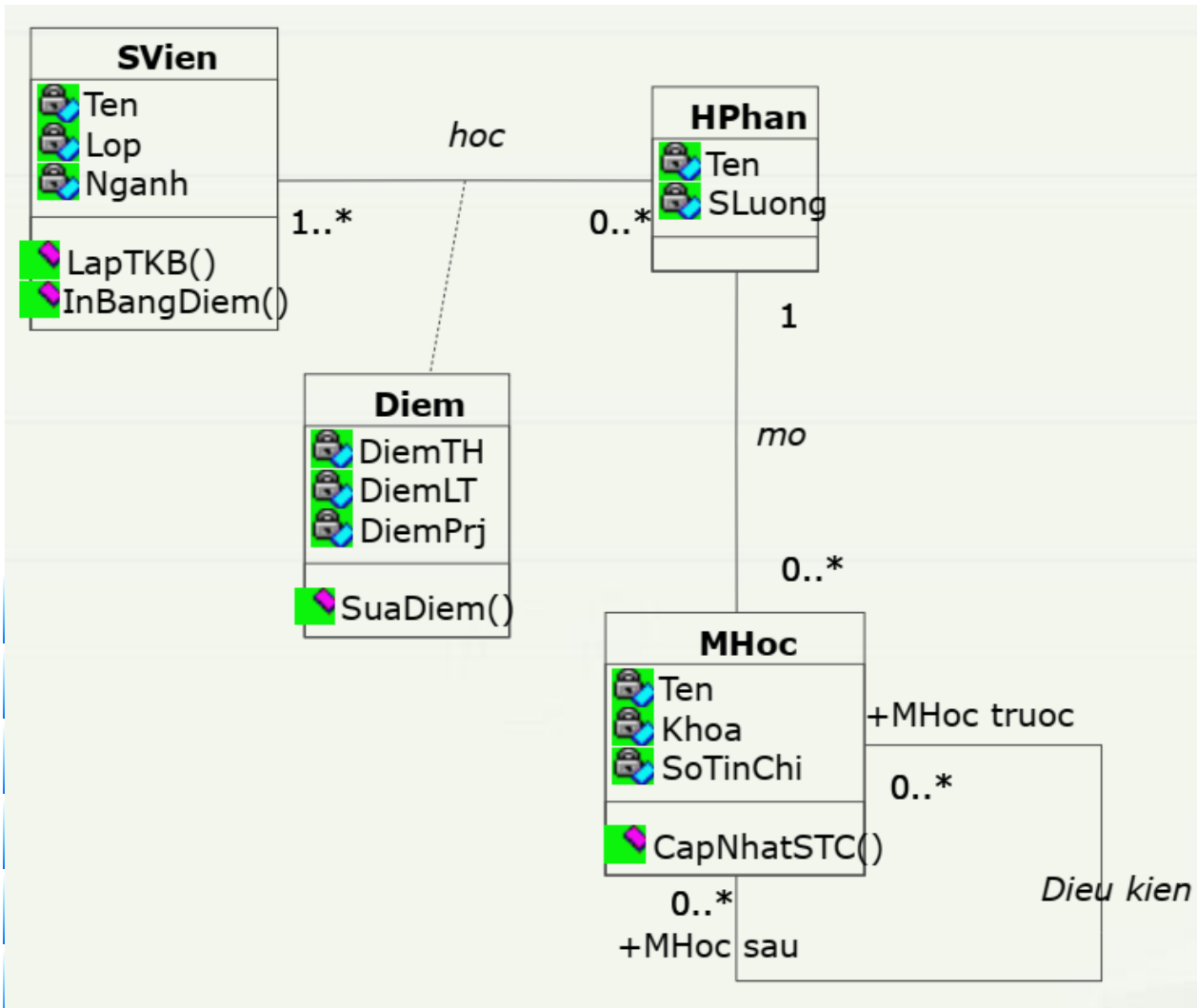
- Tương đối đơn giản
- Dễ sử dụng
- Không thích hợp biểu diễn CSDL có quy mô lớn
- Khả năng diễn đạt ngữ nghĩa kém



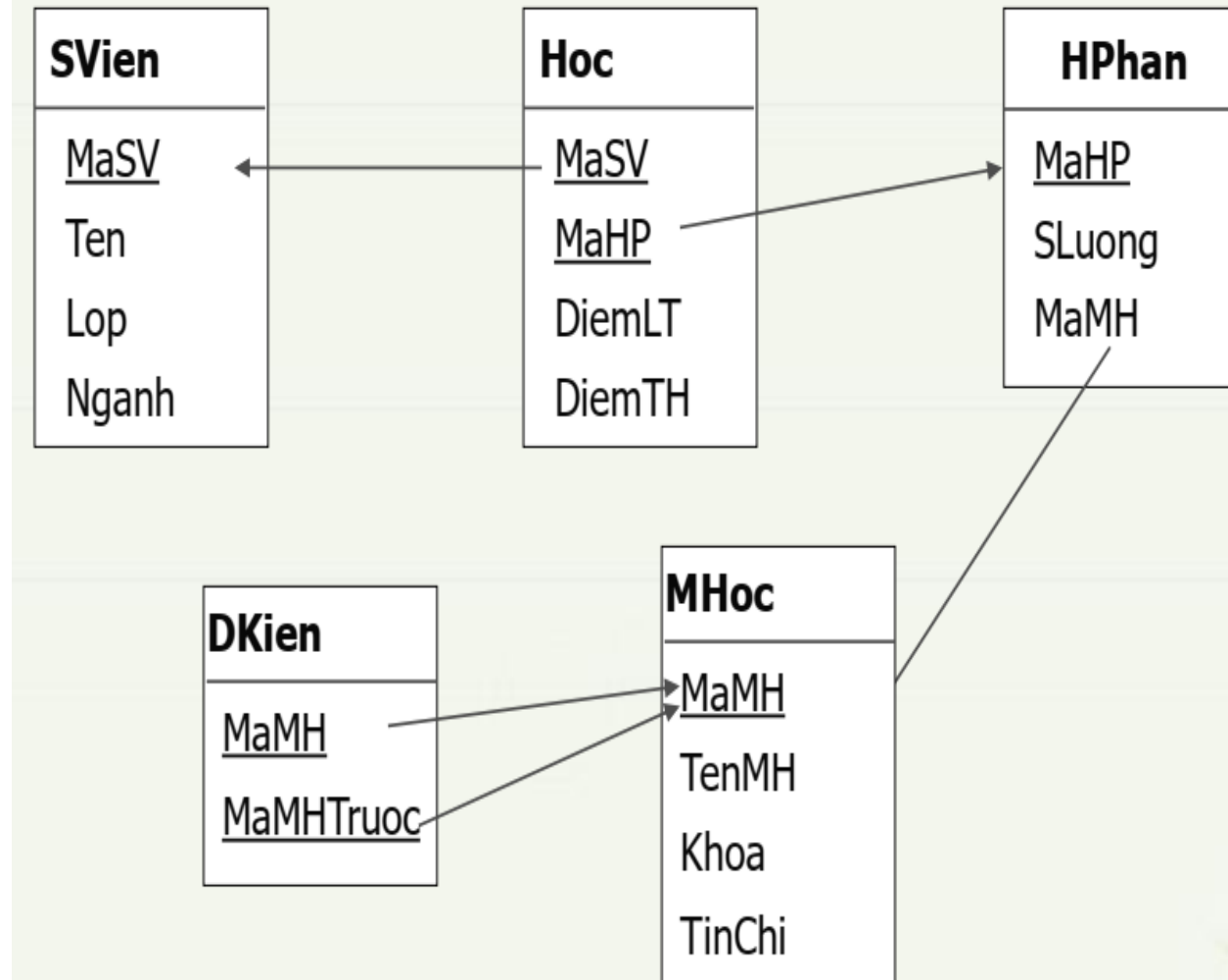


7. Các mô hình dữ liệu (tt)

Mô hình dữ liệu hướng đối tượng

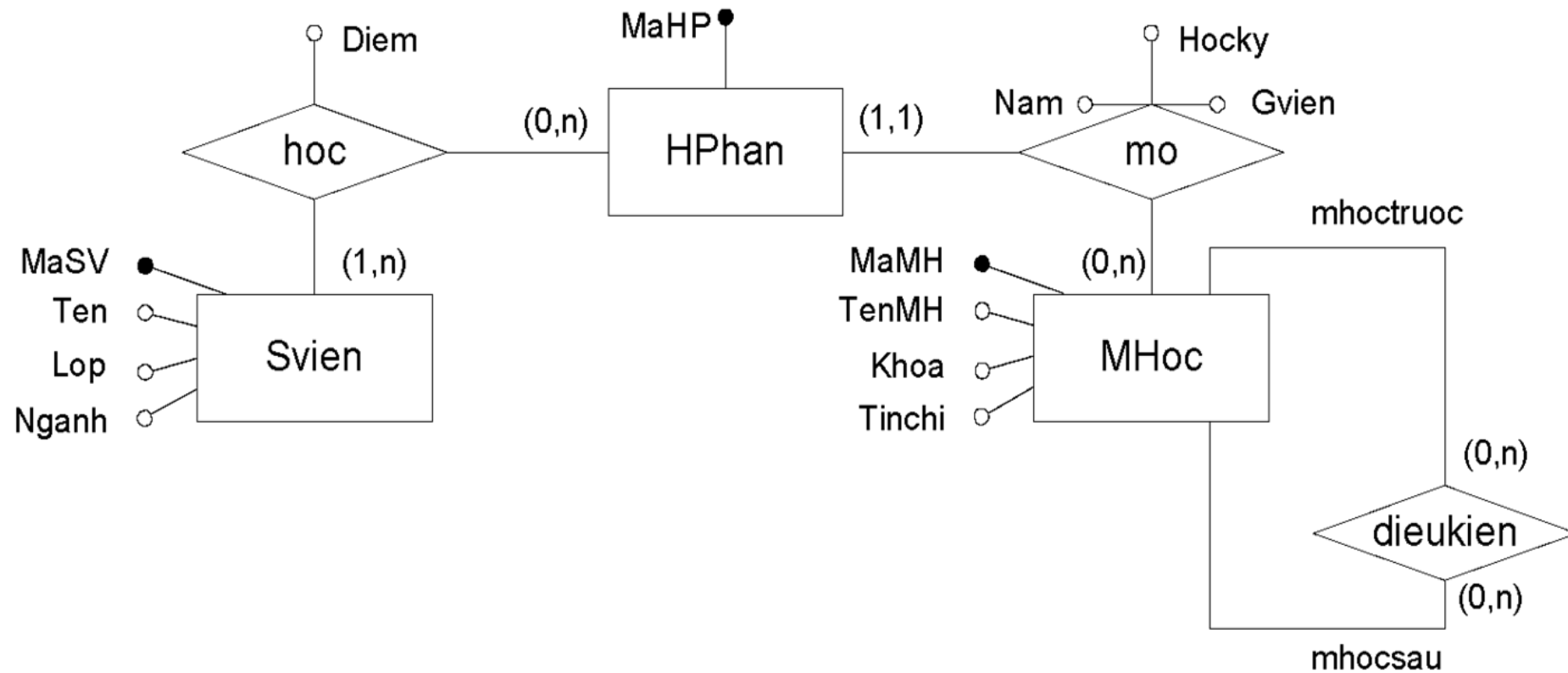


Mô hình dữ liệu quan hệ



7. Các mô hình dữ liệu (tt)

- Mô hình thực thể mối kết hợp**





CÁC MÔ HÌNH DỮ LIỆU (tt)

MÔ HÌNH THỰC THỂ MỐI KẾT HỢP (ER)

7




Khái niệm (ERD)

- Mô hình thực thể mối kết hợp được CHEN giới thiệu năm 1976.
- Được dùng để thiết kế CSDL ở mức quan niệm
- Biểu diễn trừu tượng *cấu trúc* của CSDL
- Lược đồ thực thể kết hợp (**E**ntity-**R**elationship **D**iagram)
 - Tập thực thể (Entity Sets)/ Thực thể (Entity)
 - Thuộc tính (Attributes)
 - Mối quan hệ (Relationship)



ERD (tt) – Loại thực thể (tập thực thể)

- **Định nghĩa:** loại thực thể (Entity Type, còn gọi Tập thực thể) là những loại đối tượng hay sự vật của thế giới thực tồn tại cụ thể cần được quản lý.
 - Một thực thể là một đối tượng của thế giới thực
 - Tập hợp các thực thể giống nhau tạo thành 1 loại thực thể
 - Ví dụ : HOCVIEN, LOP, MONHOC, ...
- **Ký hiệu:** 



ERD (tt) – Thực thể (Entity)

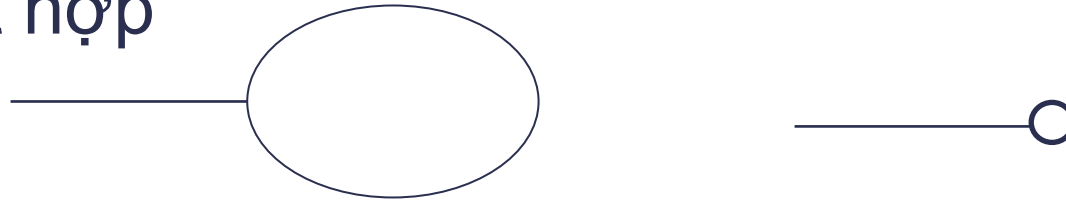
Ví dụ:

- Một học viên là một thực thể
- Tập hợp các học viên tạo thành loại thực thể
- Loại thực thể HOCVIEN có các thực thể:
 - ('HV001', 'Nguyen Nam', '1/2/1987', 'Nam')
 - ('HV002', 'Trần Nam', '13/2/1987', 'Nam')

ERD (tt) – Thuộc tính

- **Định nghĩa:** thuộc tính là những **đặc tính riêng biệt** của tập thực thể hoặc mối kết hợp

- **Ký hiệu:**



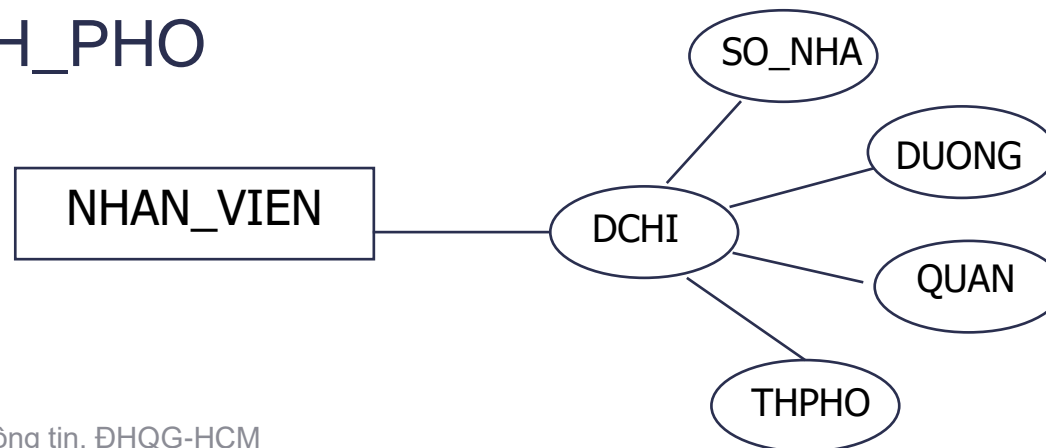
- **Ví dụ:** Loại thực thể HOCVIEN có các thuộc tính

- Mã học viên
- Họ tên
- Giới tính
- Ngày sinh
- ...



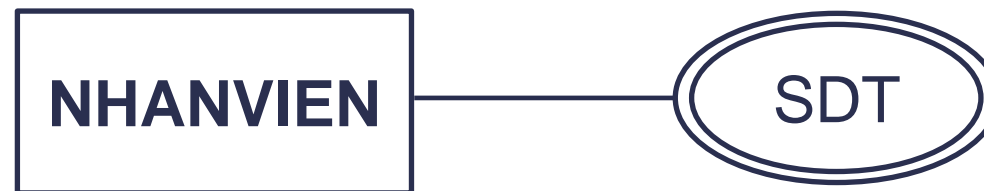
ERD (tt) – Các loại Thuộc tính

- **Thuộc tính đơn:** không thể chia làm nhiều thành phần nhỏ hơn
 - VD: Tập thực thể NHAN_VIEN có thuộc tính LUONG, HO_TEN, GIOI_TINH
- **Thuộc tính kết hợp:** gồm nhiều thành phần nhỏ hơn
 - VD: Tập thực thể NHAN_VIEN có thuộc tính DIA_CHI
 - DIA_CHI có thể chia thành các thành phần: SO_NHA, DUONG, QUAN, THANH_PHO



ERD (tt) – Các loại Thuộc tính (tt)

- **Thuộc tính đơn trị:** nhận một giá trị đơn đối với một thực thể cụ thể
 - VD: Thuộc tính HO_TEN của tập thực thể NHAN_VIEN là đơn trị, mỗi nhân viên chỉ có một họ tên
- **Thuộc tính đa trị:** nhận nhiều giá trị đối với một thực thể cụ thể
 - VD: Thuộc tính DIEN_THOAI của tập thực thể NHAN_VIEN là đa trị, mỗi nhân viên có thể có nhiều số điện thoại



- **Thuộc tính suy diễn:** giá trị được tính từ giá trị của thuộc tính khác



ERD (tt) – Khóa của loại thực thể

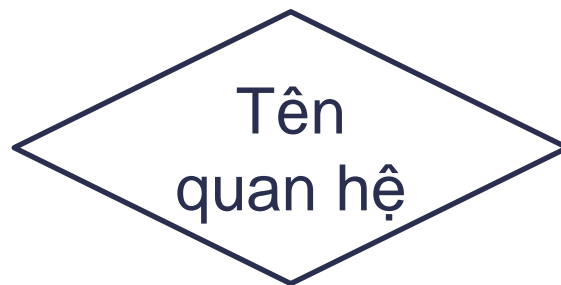
- Các thực thể trong tập thực thể cần phải được phân biệt
- **Khóa K** của tập thực thể E là một hay nhiều thuộc tính sao cho:
 - Lấy ra 2 thực thể bất kỳ e1, và e2 trong E
 - Thì e1 và e2 **không thể có các giá trị giống nhau tại các thuộc tính trong K**
- Lưu ý
 - Mỗi loại thực thể phải có ít nhất 1 khóa
 - Có thể có nhiều khóa trong 1 loại thực thể, ta sẽ chọn ra 1 khóa làm **khóa chính** cho tập thực thể đó
 - Một khóa có thể có 1 hay nhiều thuộc tính



ERD (tt) – Mỗi quan hệ/ mỗi kết hợp

- **Định nghĩa:** Diễn tả mỗi **quan hệ ngữ nghĩa** giữa ít nhất 2 tập thực thể khác nhau
- **Ví dụ:** giữa tập thực thể NHAN_VIEN và PHONG_BAN có các mối quan hệ sau
 - Một nhân viên thuộc một phòng ban nào đó
 - Một phòng ban có một nhân viên làm trưởng phòng

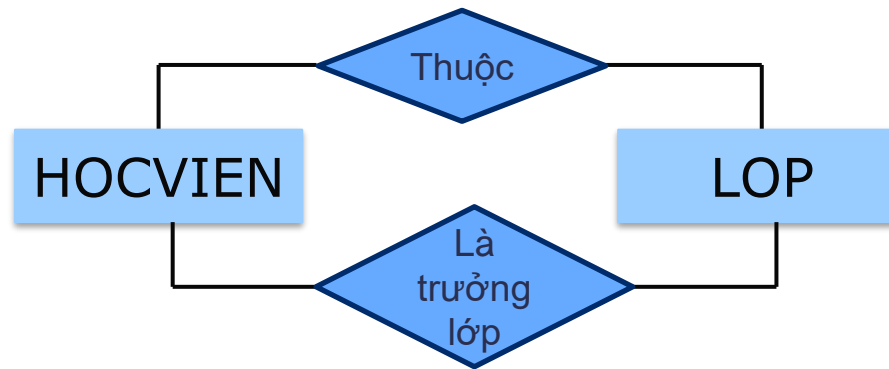
- **Ký hiệu:**



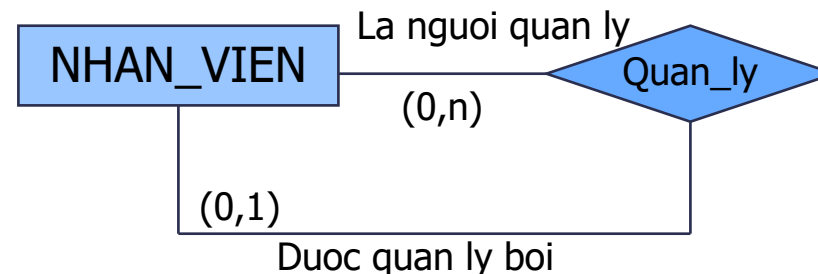
ERD (tt) – Mỗi quan hệ/ mỗi kết hợp

- Giữa hai loại thực thể có thể tồn tại nhiều hơn một quan hệ

- Ví dụ



- Vai trò: Một loại thực thể có thể tham gia nhiều lần vào một quan hệ với nhiều vai trò khác nhau





ERD (tt) – Mỗi quan hệ/ mỗi kết hợp – Thể hiện

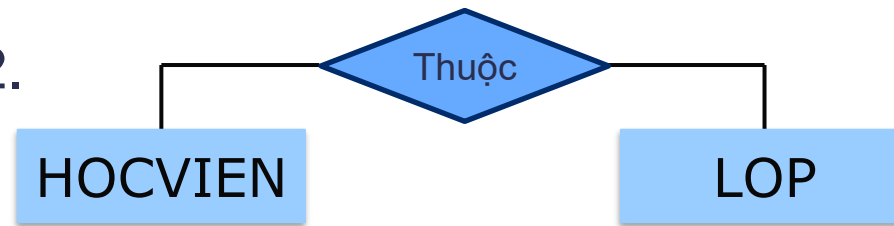
- Một CSDL được mô tả bởi lược đồ ER sẽ chứa đựng những **dữ liệu cụ thể** gọi là **thể hiện CSDL**
- Thể hiện CSDL còn chứa các mối quan hệ cụ thể.
 - Cho mỗi quan hệ R kết nối n tập thực thể E_1, E_2, \dots, E_n
 - Thể hiện của R là tập hữu hạn các danh sách (e_1, e_2, \dots, e_n)
 - Trong đó e_i là các giá trị được chọn từ các tập thực thể E_n
- Ví dụ:



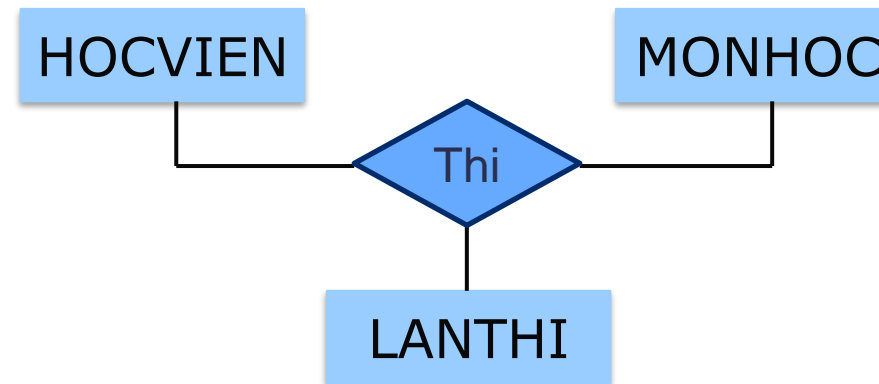
ERD (tt) – Số ngôi

- **Định nghĩa:** Số ngôi của loại mỗi kết hợp là **số loại thực thể** tham gia vào loại mỗi kết hợp đó.

- Ví dụ 1: Loại mỗi kết hợp **Thuộc** kết hợp 2 loại thực thể HOCVIEN và LOP nên có số ngôi là 2.

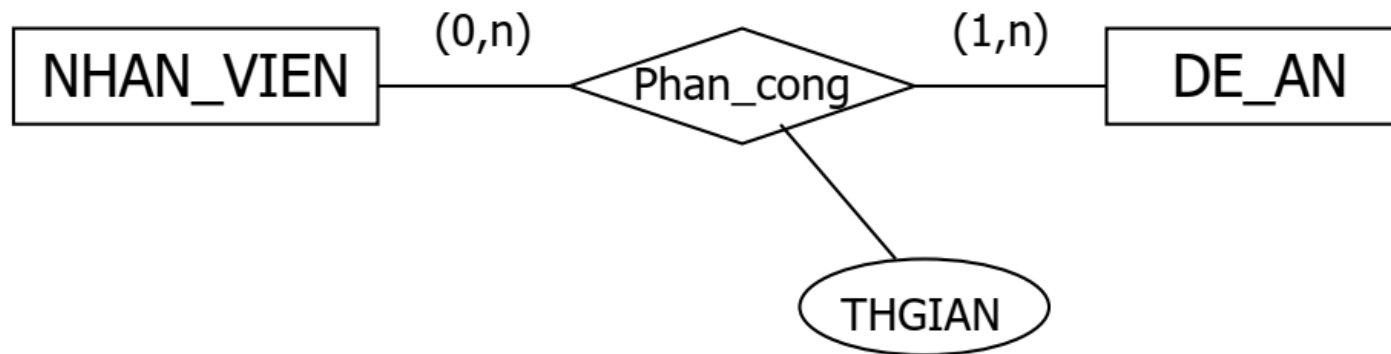


- Ví dụ 2: Loại mỗi kết hợp **Thi** kết hợp 3 loại thực thể LANTHI, HOCVIEN, và MONHOC nên có số ngôi là 3.



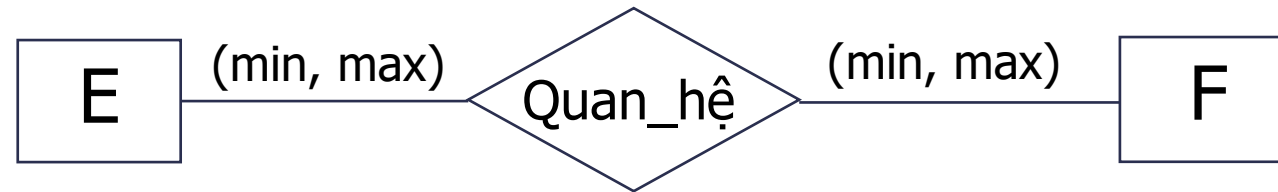
ERD (tt) – Thuộc tính của mỗi quan hệ

- Thuộc tính trên mỗi quan hệ mô tả tính chất cho mỗi quan hệ đó.
- Thuộc tính này không thể gắn liền với những thực thể tham gia vào mỗi quan hệ.



ERD (tt) – Bản số (relationship cardinality)

- **Định nghĩa:** Bản số (min, max) chỉ định mỗi thực thể $e \in E$ tham gia ít nhất và nhiều nhất vào thể hiện của R



- **Ký hiệu:** (bản số tối thiểu, bản số tối đa)
 - (0,1) – không hoặc 1
 - (1,1) – duy nhất 1
 - (0,n) – không hoặc nhiều
 - (1,n) – một hoặc nhiều



ERD (tt) – Bản số (relationship cardinality)

- Một phòng ban có nhiều nhân viên



- Một nhân viên chỉ thuộc một phòng ban



- Một đề án có thể có nhiều nhân viên tham gia hoặc không có nhân viên nào tham gia



- Một nhân viên có thể là trưởng phòng ban hoặc không là trưởng phòng ban nào



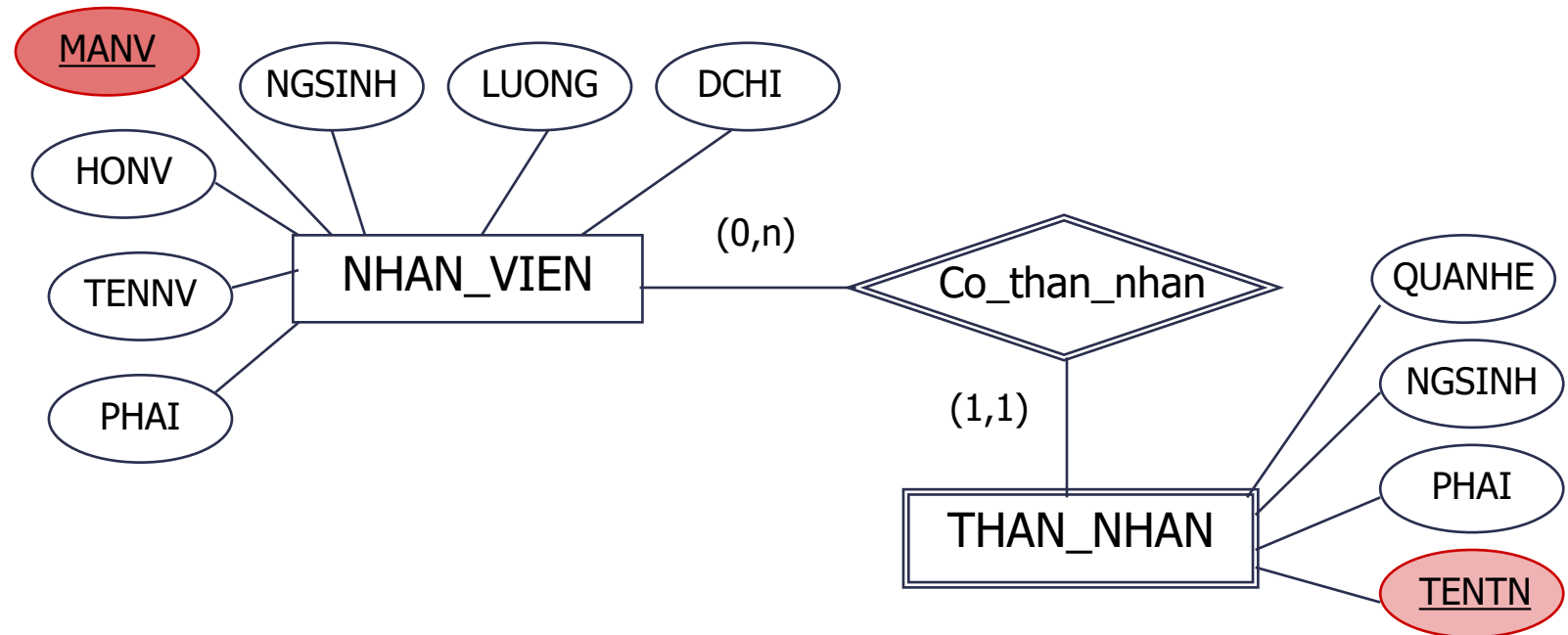
ERD (tt) – Loại thực thể yếu

- **Định nghĩa:**

- Là thực thể mà Khóa có được từ những thuộc tính của tập thực thể khác
- Phải tham gia trong một mối quan hệ xác định, trong đó có một tập thực thể chủ gọi là thực thể mạnh

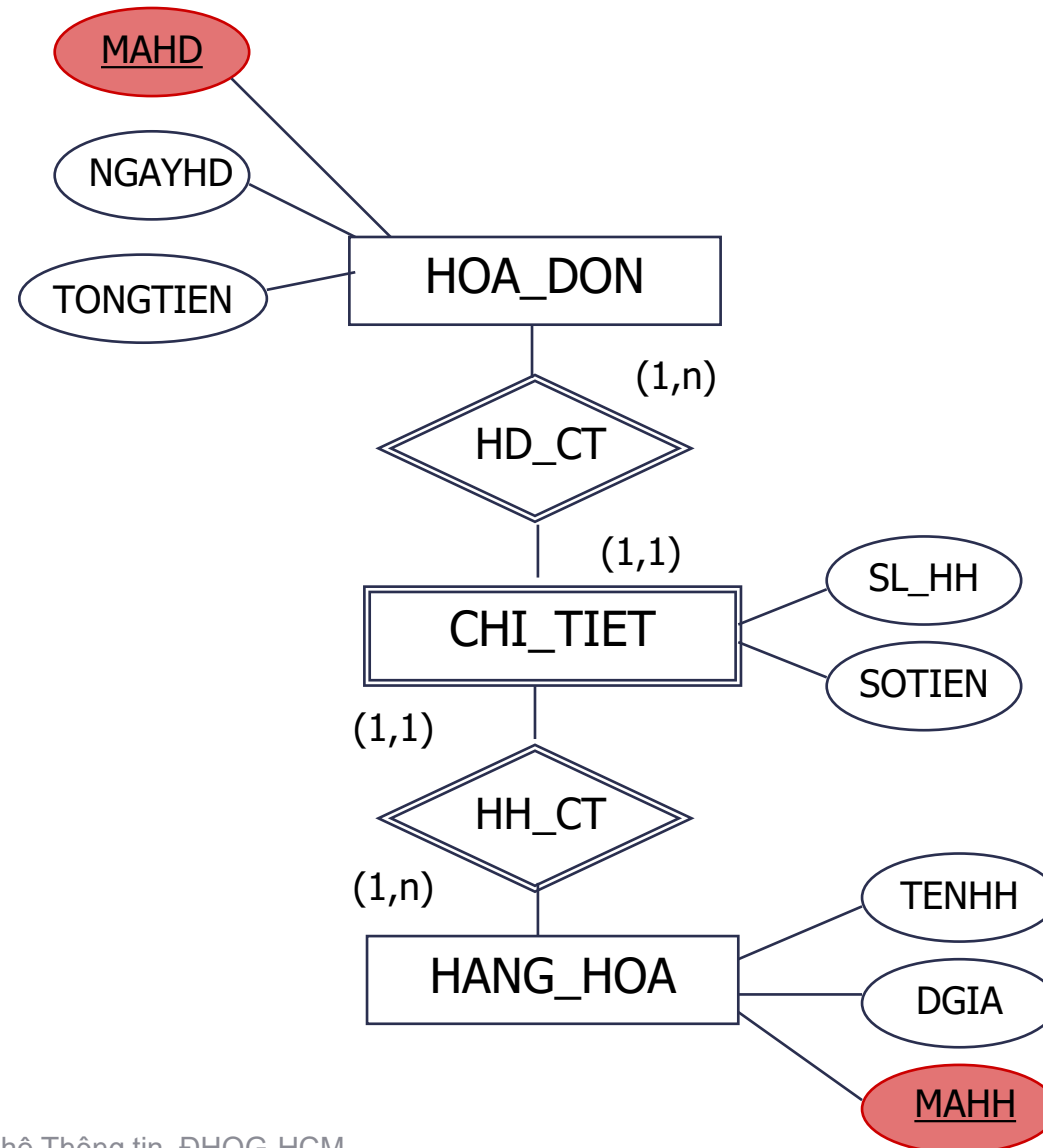
- **Ký hiệu:**

Tập thực thể



ERD (tt) – Loại thực thể yếu

- Ví dụ khác:





ERD (tt) – Biểu diễn lược đồ ER

Mô hình thực thể mối kết hợp (ERD) là **đồ thị** biểu diễn các tập thực thể (loại thực thể), thuộc tính và mối kết hợp, trong đó:

- **Đỉnh**

Tên tập thực thể

Tập thực thể

Tên thuộc tính

Thuộc tính

Tên quan hệ

Quan hệ

- **Cạnh** là đường nối giữa

- ❖ Tập thực thể và thuộc tính
- ❖ Mối quan hệ và tập thực thể
- ❖ Mối quan hệ và thuộc tính



ERD (tt) – Biểu diễn lược đồ ER

Các bước thiết kế:

- (1) Xác định tập thực thể
- (2) Xác định mối quan hệ
- (3) Xác định thuộc tính và gán thuộc tính cho tập thực thể và mối quan hệ
- (4) Quyết định thuộc tính khóa
- (5) Quyết định (min, max) cho mối quan hệ



Bài tập

8



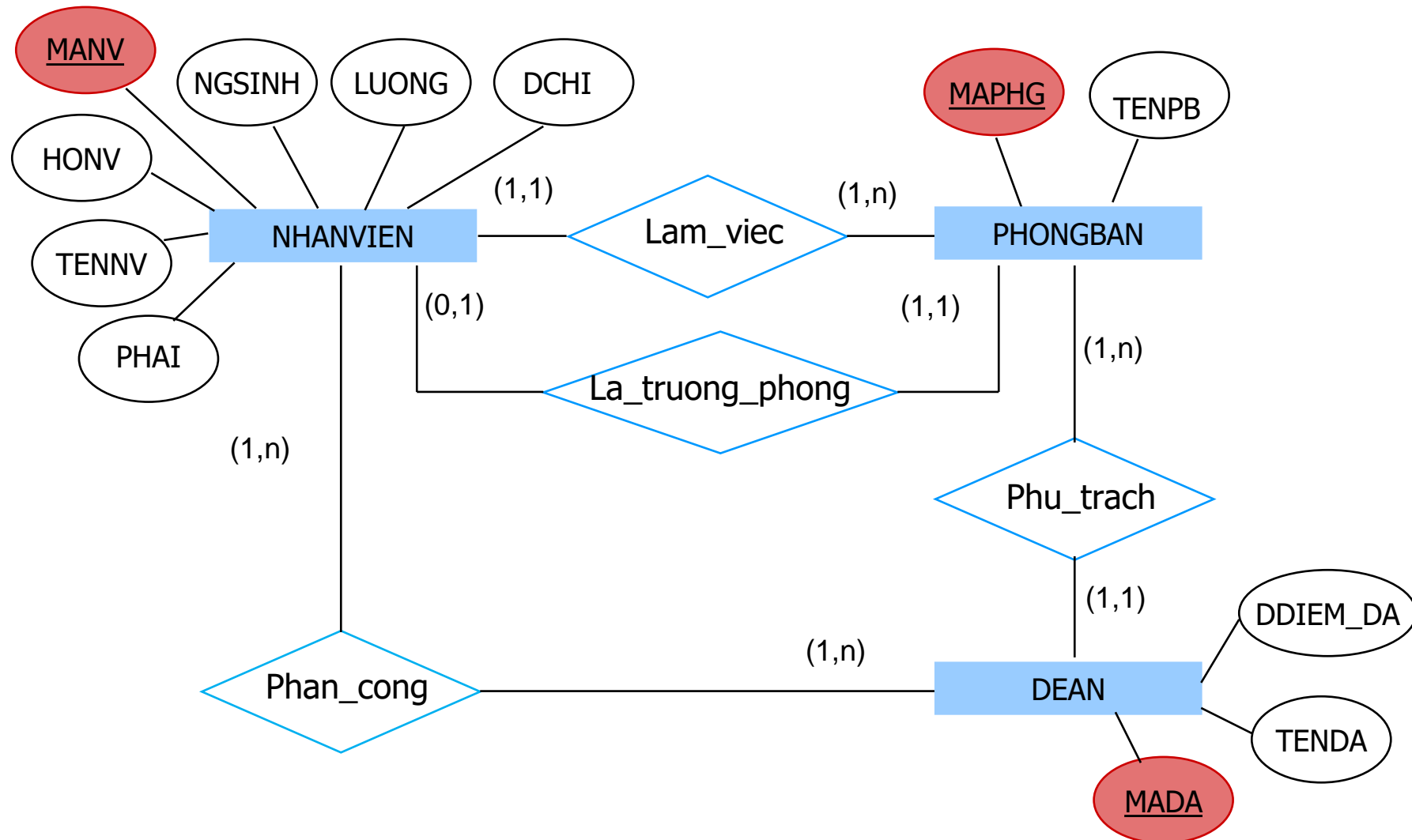
Bài tập 1

CSDL đề án công ty theo dõi các thông tin liên quan đến nhân viên, phòng ban và đề án

- Công ty có nhiều phòng ban, mỗi phòng ban có tên duy nhất, mã phòng duy nhất, một trưởng phòng và ngày nhận chức. Mỗi phòng ban có thể ở nhiều địa điểm khác nhau.
- Đề án có tên duy nhất, mã duy nhất, do 1 một phòng ban chủ trì và được triển khai ở 1 địa điểm.
- Nhân viên có mã số, tên, địa chỉ, ngày sinh, phái và lương. Mỗi nhân viên làm việc ở 1 phòng ban, tham gia vào các đề án với số giờ làm việc khác nhau. Mỗi nhân viên đều có một người quản lý trực tiếp.
- Một nhân viên có thể có nhiều thân nhân. Mỗi thân nhân có tên, phái, ngày sinh và mối quan hệ với nhân viên đó.



ERD (tt) – Biểu diễn lược đồ ER– Ví dụ





Bài tập 2

Xây dựng mô hình ERD cho CSDL quản lý giáo vụ gồm có các chức năng sau:

- Lưu trữ thông tin: Học viên (có mã học viên duy nhất, họ tên, giới tính, NTNS) , giáo viên (mã giáo viên duy nhất, họ tên, giới tính), môn học (mã môn học duy nhất, tên môn học, số tín chỉ lý thuyết, số tín chỉ thực hành), lớp học (mã lớp duy nhất, tên lớp, sĩ số dự kiến).
- Xếp lớp cho học viên, chọn lớp trưởng cho lớp
- Phân công giảng dạy: giáo viên dạy lớp nào với môn học gì, ở học kỳ, năm học nào.
- Lưu trữ kết quả thi: học viên thi môn học nào, lần thi thứ mấy, điểm thi bao nhiêu.



THẢO LUẬN

