### Đại Học Công Nghệ Thông Tin Khoa Hệ thống Thông tin

Môn học: Cơ Sở Dữ Liệu Thời lương: 45<sup>LT</sup>+30<sup>TH</sup>

Giảng viên: ThS. Thái Bảo Trân Email: trantb@uit.edu.vn

1

### **M**ŲC TIÊU

- Môn học nhằm cung cấp cho sinh viên kiến thức nền tảng về CSDL, đặc biệt CSDL quan hệ: mô hình dữ liệu quan hệ (Relational Data Model), các ngôn ngữ truy vấn,...
- Sinh viên sẽ tích lũy được những hiểu biết cơ bản để có thể sử dụng và khai thác CSDL quan hệ một cách thành thao.
- Sinh viên cũng sẽ được trang bị các kỹ năng khai báo, truy vấn một CSDL quan hệ với một hệ quản trị CSDL cụ thể (MS SQL Server) nhằm phục vụ cho nhiều môn học nâng cao về CSDL trong những học kỳ kế tiếp.

9

### Nội dung

- Tổng quan về CSDL (3t)
- Mô hình dữ liệu quan hệ (Codd) (6t)
- Ngôn ngữ đại số quan hệ (6t)
- Ngôn ngữ SQL (12t)
- Ràng buộc toàn vẹn (6t)
- Phụ thuộc hàm và các dạng chuẩn (9t)
- Ôn tập (3t)

3

Khoa HTTT- Đai học CNTT

### Tài Liệu Tham Khảo

- Slides môn Cơ sở dữ liệu, Khoa hệ thống thông tin, Đại học Công nghệ Thông tin, ĐHQG, HCM.
- Hector Garcia-Molina, Jeffrey D. Ullman, and Jennifer Widom, Database Systems - The Complete Book, Prentice Hall, ISBN: 0-13-031995-32002.
- Raghu Ramakrishnan and Johannes Gehrke, Database Management Systems, Third Edition, McGraw Hill, 2003. ISBN: 0-07-246563-8.
- C. J. Date, An Introduction to Database Systems, Eighth Edition, Addison Wesley, 2003. ISBN 0-321-19784-4

### Hình thức kiểm tra và đánh giá

- ❖ Giữa kỳ (Tự luận): 20%
- **❖** Thực hành : 30%
- ❖ Cuối kỳ (Tự luận): 50%
- ❖ Điểm cộng : ++2 (Thực hành)

5

### Chương 1: Tổng quan về Cơ Sở Dữ Liệu Thời lượng: 3 tiết

### **NỘI DUNG**

- 1. Giới thiệu
- 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
- 7. Các mô hình dữ liệu

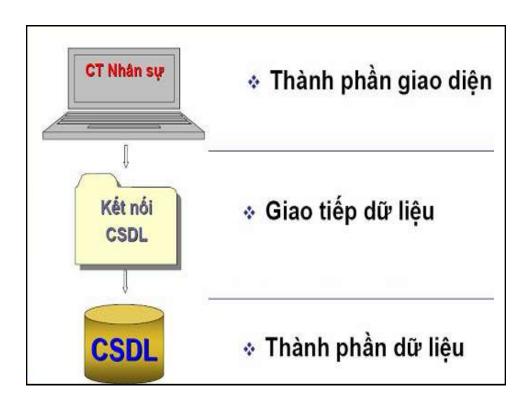
7

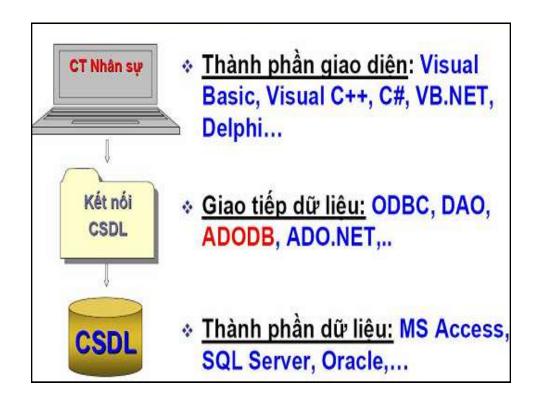
Khoa HTTT- Đại học CNTT

### 1. Giới thiệu

- Tại sao cần phải có một cơ sở dữ liệu?
- Một ứng dụng quản lý cần những gì ?
- · Giao diện: cửa sổ, nút điều khiển,...
- · Xử lý: tính toán, sắp xếp, tìm kiếm,...
- Lưu trữ: Tập tin (file), CSDL,...

Я





### 2. Hệ thống tập tin (1)

- 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)

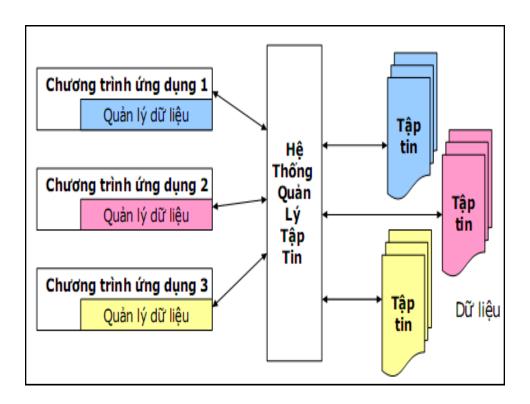
11

Khoa HTTT- Đại học CNTT

### 2. Hệ thống tập tin (2)

- Nhược điểm:
  - Dữ liệu bị trùng lắp→ dư thừa
  - Thiếu tính nhất quán giữa các dữ liệu
  - Chia sẻ dữ liệu bị hạn chế
  - Khó khăn trong việc truy suất đồng thời
  - Khó khôi phục
  - Vấn đề bảo mật và phân quyền kém

12

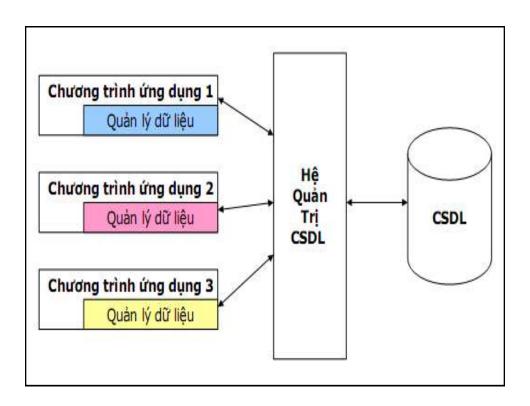


### 3. Cơ sở dữ liệu (1)

### • Định nghĩa:

- Cơ sở dữ liệu là một hệ thống các thông tin có cấu trúc, được lưu trữ trên các thiết bị lưu trữ thông tin thứ cấp như: băng từ, đĩa từ,...
- Có thể 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.

14



### 3. Cơ sở dữ liệu (2)

### • Ưu điểm:

- Giảm trùng lắp thông tin xuống mức thấp nhất và do đó 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.

16

### 3. Cơ sở dữ liệu (3)

- Những vấn đề cần giải quyết:
  - Tính chủ quyền dữ liệu
  - 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

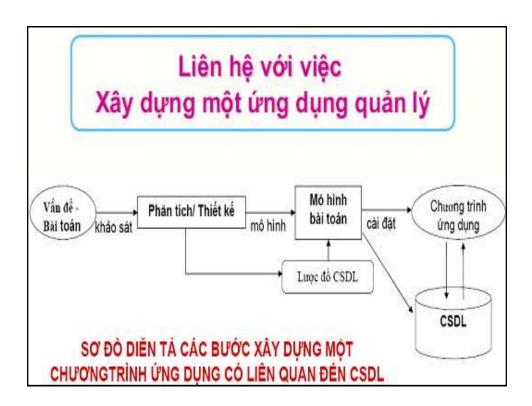
17

Khoa HTTT- Đại học CNTT

### 4. Các đối tượng sử dụng

- Người dùng cuối: sử dụng những công cụ hỗ trợ để khai thác CSDL.
- Người thiết kế CSDL: xây dựng các ứng dụng để phục vụ cho các mục đích quản lý.
- Người quản trị CSDL: tổ chức CSDL, bảo mật, cấp quyền, sao lưu, phục hồi dữ liệu, giải quyết các tranh chấp dữ liệu, ...

18



### 5. Hệ quản trị cơ sở dữ liệu (1)

- Hệ quản trị cơ sở dữ liệu:
- DataBase Management System DBMS
- Là hệ thống các phần mềm hỗ trợ tích cực cho các nhà phân tích, thiết kế và khai thác CSDL.
- Cung cấp cho người dùng và ứng dụng một môi trường thuận tiện và sử dụng hiệu quả tài nguyên dữ liệu.
- Các DBMS thông dụng: Visual FoxPro, Microsoft Access, SQL Server, DB2, Oracle, ...

20

### 5. Hệ quản trị cơ sở dữ liệu (2)

- Môt DBMS phải có:
- 1. Ngôn ngữ giao tiếp giữa người sử dụng và CSDL
- 2. Từ điển dữ liệu (Data Dictionary)
- 3. Có biện pháp bảo mật tốt khi có yêu cầu
- 4. Cơ chế giải quyết tranh chấp dữ liệu
- 5. Cơ chế sao lưu (Backup) và phục hồi (Restore) dữ liệu
- 6. Cung cấp một giao diện tốt, dễ sử dụng, dễ hiểu
- 7. Bảo đảm tính độc lập giữa dữ liệu và chương trình

21

Khoa HTTT- Đại học CNTT

### 5. Hệ quản trị cơ sở dữ liệu (3)

- 1. Ngôn ngữ giao tiếp giữa người sử dụng và CSDL
- 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) cho phép thay đổi cấu trúc bảng, khai báo bảo mật, cấp quyền cho người sử dụng.

22

### 5. Hệ quản trị cơ sở dữ liệu (4)

- 2. Từ điển dữ liêu (Data Dictionary)
- Dùng để mô tả các ánh xạ liên kết
- Ghi nhận các thành phần cấu trúc của CSDL, các CT ứng dụng, mật mã, các quyền hạn sử dụng,...
- 3. Có biên pháp bảo mật tốt khi có yêu cầu
- 4. Cơ chế giải quyết tranh chấp dữ liệu
- Cấp quyền ưu tiên cho người sử dụng (người quản trị CSDL thực hiên)
- Đánh dấu yêu cầu truy suất dữ liệu, phân chia thời gian, người nào có yêu cầu trước thì có quyền truy suất dữ liệu trước.

23

Khoa HTTT- Đại học CNTT

### 5. Hệ quản trị cơ sở dữ liệu (5)

- 5. Cơ chế sao lưu (Backup) và phục hồi (Restore) dữ kiệu khi có sư cố xảy ra.
- Định kỳ kiểm tra CSDL
- > Tạo nhật ký (LOG) thao tác CSDL
- 6. Cung cấp một giao diện tốt, dễ sử dụng, dễ hiểu
- 7. Bảo đảm tính độc lập giữa dữ liệu và chương trình: khi có thay đổi dữ liệu thì các CT ứng dụng đang chạy trên CSDL đó không cần phải viết lại và cũng không ảnh hưởng đến những NSD khác.

24

### 6. Các mức biểu diễn một CSDL (1)

### Một CSDL có 3 mức biểu diễn:

- Mức trong (Mức vật lý Physical)
- Mức quan niệm (Conception / Logical)
- Mức ngoài (Khung nhìn View)

25

Khoa HTTT- Đại học CNTT

### 6. Các mức biểu diễn một CSDL (2)

### Mức trong - Vật lý

- > Là mức lưu trữ CSDL
- Vấn đề cần giải quyết: 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 và người sử dụng chuyên môn.

26

### 6. Các mức biểu diễn một CSDL (2)

### Mức quan niệm - Logic

- CSDL cần phải lưu trữ bao nhiêu loại dữ liệu? là dữ liệu gì? mối quan hệ giữa các loại dữ liệu này ntn?
- Dành cho chuyên viên tin học khảo sát và phân tích cùng với những người quản trị CSDL xác định những loại thông tin cần thiết để đưa vào CSDL và mối quan hệ của chúng.
- CSDL mức quan niệm là một biểu diễn trừu tượng của CSDL mức vật lý, hay CSDL mức vật lý là cài đặt cụ thể của CSDL mức quan niêm.

27

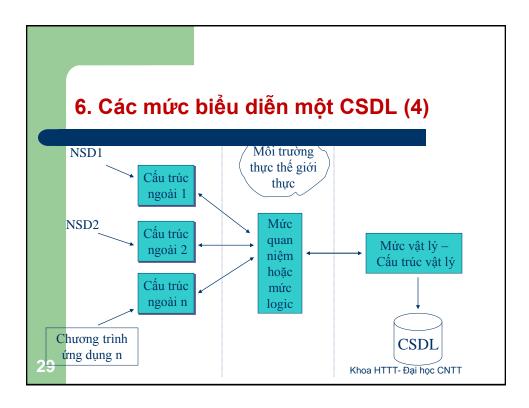
Khoa HTTT- Đại học CNTT

### 6. Các mức biểu diễn một CSDL (3)

### Mức ngoài - Khung nhìn

- Đây là mức của người sử dụng và các chương trình ứng dung
- > Được "nhìn" (View) CSDL theo một góc độ khác nhau.
- Có thể hoàn toàn không biết về cấu trúc tổ chức lưu trữ thông tin trong CSDL, thậm chí cả tên gọi các dữ liệu, thuộc tính.
- Họ chỉ làm việc trên một phần CSDL theo cách nhìn gọi là khung nhìn (View)

28



### 7. Các mô hình dữ liệu

- Mô hình dữ liệu là sự trừu tượng hóa của môi trường thực, biểu diễn dữ liệu ở mức logic. Phân loại các mô hình dữ liệu:
- Mô hình logic trên cơ sở mẫu tin
  - MH dữ liệu mạng
  - MH dữ liệu phân cấp
  - MH dữ liệu quan hệ
- Mô hình logic trên cơ sở đối tượng
  - MH dữ liệu thực thể kết hợp
  - MH dữ liệu hướng đối tượng, phân tán,...
- Mô hình vật lý (Physical Model)

### 7.1. Mô hình dữ liệu mạng (1)

- 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.
  - Mẫu tin (record)
  - Loại mẫu tin
  - Loại liên hệ (set type)
  - Bản số

31

### 7.1. Mô hình dữ liệu mạng (2)

- Mẫu tin: mô tả 1 đối tượng trong thế giới thực. ('NV001','Nguyen Van A','Nam','10/10/1970','Dong Nai')
- Loại mẫu tin: là 1 tập các mẫu tin có cùng tính chất.
   Ví du: 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
  - Ký hiệu:

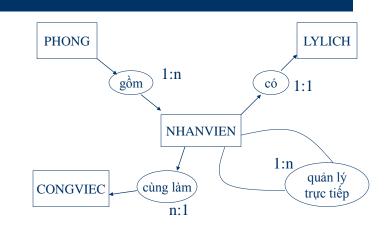
Tham gia

### 7.1. Mô hình dữ liệu mạng (3)

- Bản số: chỉ ra số lượng các mẫu tin tham gia trong mối 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.
  - (Recursive): một loại mẫu tin chủ cũng có thể đồng thời là loại mẫu tin thành viên với chính nó. Loại liên hệ này là Đệ quy

33

### 7.1. Mô hình dữ liệu mạng (4)



### 7.1. Mô hình dữ liệu mạng (5)

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

35

### 7.1. Mô hình dữ liệu mạng (6)

• Bài tập 1.1:

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

### 7.1. Mô hình dữ liệu mạng (7)

### • Gợi ý:

- KHACHHANG (MAKH, HOTEN, DCHI, SODT, NGSINH, DOANHSO, NGGN)
- > NHANVIEN (MANV, HOTEN, NGSINH, NGVL, HESO, MUCLUONG)
- > SANPHAM (MASP, TENSP, DVT, NUOCSX, GIA)
- > HOADON (SOHD, NGHD, MAKH, MANV, TRIGIA)
- > CTHD (SOHD, MASP, SL)

37

### 7.2. Mô hình thực thể-mối kết hợp

(Entity-Relationship Model viết tắt ER)

- Giới thiệu
- Các thành phần cơ bản
  - > Thực thể
    - Djnh nghĩa
    - □ Thuộc tính
    - Thuộc tính khóa
    - □ Loại thực thể yếu
- Lược đồ ER
  - Các bước tạo ER
  - Chuyển ER thành bảng

### Mối kết hợp

- □ Khái niệm
- □ Thuộc tính trên mối kết hợp
- Ràng buộc trên mối kết hợp
  - □ Mối kết hợp đệ quy

### Giới thiệu

- Mô hình thực thể mối kết hợp (ER) được CHEN giới thiệu năm 1976.
- Mô hình ER được sử dụng nhiều trong thiết kế dữ liệu ở mức quan niệm.
- Biểu diễn trừu tượng cấu trúc của CSDL

39

### Các thành phần cơ bản (1)

Loại thực thể (Entity Type)

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

SINHVIEN

LOP

### Các thành phần cơ bản (2)

Thực thể (Entity)

- Định nghĩa: Thực thể là một thể hiện hoặc một đối tượng của một loại thực thể.
- Ví dụ: Loại thực thể là SINHVIEN có các thực thể:
  - ('SV001', 'Nguyễn A', '1/2/1987','Nam')
  - ('SV002', 'Trần B', '13/2/1987', 'Nam')

41

### Các thành phần cơ bản (3)

Thuộc tính (Entity Attribute)

- Định nghĩa: Thuộc tính là những tính chất đặc trưng của loại thực thể cần lưu trữ.
- Ví dụ: Loại thực thể SINHVIEN có các thuộc tính: Mã sinh viên, họ tên, giới tính, ngày sinh, nơi sinh
- Ký hiệu:



### Các thành phần cơ bản (4)

Các loại thuộc tính (1)

 Đơn trị (Simple): mỗi thực thể chỉ có một giá trị ứng với các thuộc tính.

Ví du: MaSV, Hoten

 Đa trị (Multi-valued): thuộc tính có thể có nhiều giá trị đối với một thực thể.

Ví dụ: BANGCAP ký hiệu {BANGCAP}

 Đa hợp (Composite): thuộc tính có thể được tạo thành từ nhiều thành phần.

VÍ dụ: DIACHI(SONHA, DUONG, PHUONG, QUAN) hay thuộc tính HOTEN(HO, TENLOT, TEN).

43

### Các thành phần cơ bản (5)

Các loại thuộc tính (2)

- <u>Chú ý:</u> Các thuộc tính đa hợp và đa trị có thể lồng nhau tùy ý.
  - Ví dụ: thuộc tính BANGCAP của SINHVIEN là một thuộc tính đa hợp được ký hiệu bằng

{BANGCAP(TRUONGCAP,NAM,KETQUA, CHUYENNGANH)}

### Các thành phần cơ bản (6)

Khoá của loại thực thể (Entity type key)(1)

- 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.
- Ví dụ:
  - Mỗi sinh viên có một mã số duy nhất => Khoá của loại thực thể SINHVIEN là Mã sinh viên

45

### Các thành phần cơ bản (7)

Khoá của loại thực thể (Entity type key)(2)

### □ Chú ý:

- \* Mỗi tập thực thể phải có 1 khóa
- \* Một khóa có thể có 1 hay nhiều thuộc tính
- Có thể có nhiều khóa trong 1 tập thực thể, ta sẽ chọn ra 1 khóa làm khóa chính cho tập thực thể đó

### Các thành phần cơ bản (8)

Loại thực thể yếu (1)

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

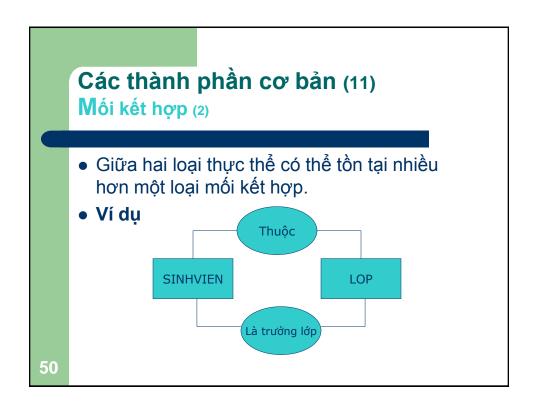
Thực thể

 Ví dụ: loại thực thể LANTHI có thuộc tính Lần và tham gia trong loại mối kết hợp Thi với loại thực thể SINHVIEN và MONHOC là loại thực thể yếu.

47

### Các thành phần cơ bản (9) Loại thực thể yếu (2) MONHOC (1,n) SINHVIEN LANTHI

# Các thành phần cơ bản (10) Mối kết hợp (Relationship) (1) • Định nghĩa: Mối kết hợp là sự kết hợp giữa hai hay nhiều loại thực thể • Ví dụ: giữa hai loại thực thể SINHVIEN và LOP có mối kết hợp "Thuoc" • Ký hiệu: bằng một hình oval hoặc hình thoi



### Các thành phần cơ bản (12)

Số ngôi của mối kết hợp (relationship degree)(1)

- Số ngôi của mối kết hợp là số loại thực thể tham gia vào loại mối kết hợp đó.
- Ví dụ 1: Loại mối kết hợp Thuộc kết hợp 2 loại thực thể HOCVIEN và LOP nên có số ngôi là 2.
- Ví dụ 2: Loại mối kết hợp Giảng dạy kết hợp 3 loại thực thể GIAOVIEN, MONHOC, LOP nên có số ngôi là 3.

51

### Các thành phần cơ bản (13) Số ngôi của mối kết hợp (relationship degree)(2)

### Các thành phần cơ bản (14)

Thuộc tính của mối kết hợp (1)

- Thuộc tính của mối kết hợp bao gồm các thuộc tính khóa của các loại thực thể tham gia vào loại mối kết hợp đó. Ngoài ra còn có thể có thêm những thuộc tính bổ sung khác.
- Ví dụ: Loại mối kết hợp Giảng dạy giữa ba loại thực thể GIAOVIEN, MONHOC và LOP có các thuộc tính là MaGV, Mamh, Malop, ngoài ra còn có thuộc tính riêng là Hocky, Nam

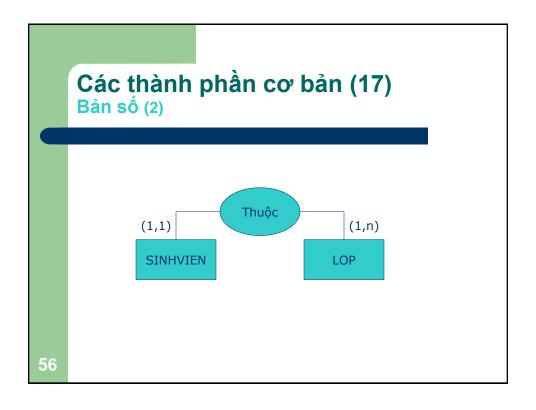
53

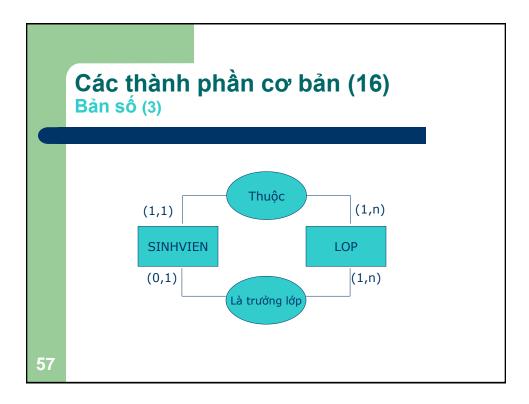
# Các thành phần cơ bản (15) Thuộc tính của mối kết hợp (2) LOP MONHOC Giang day Hocky Nam GIAOVIEN

### Các thành phần cơ bản (16)

Bản số (relationship cardinality)(1)

- Loại mối kết hợp thể hiện liên kết giữa các thực thể, mỗi liên kết được gọi là một nhánh.
- Định nghĩa: Bản số của nhánh là số lượng tối thiểu và số lượng tối đa các thực thể thuộc nhánh đó tham gia vào loại mối kết hợp.
- Ký hiệu: (bản số tối thiểu, bảng số tối đa)
- Ví dụ: Loại thực thể NhanVien và DeAn có loại mối kết hợp ThamGia.



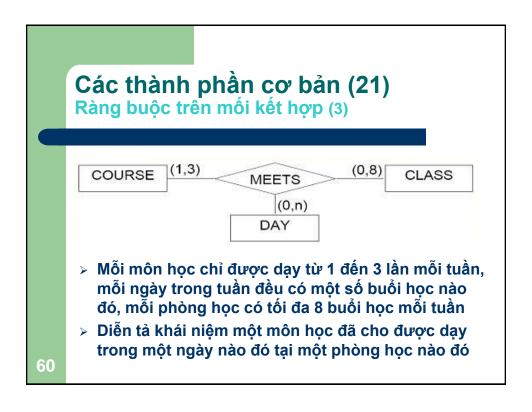


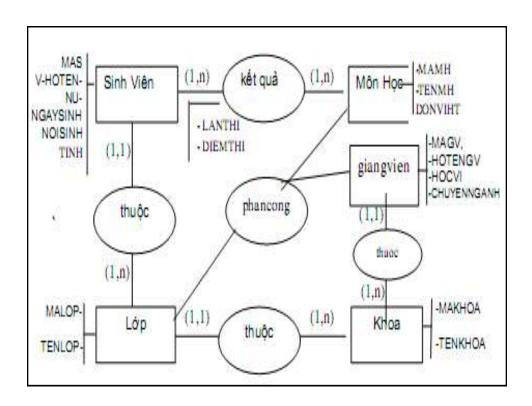
### Các thành phần cơ bản (19)

Ràng buộc trên mối kết hợp (1)

- Nhằm giới hạn khả năng có thể kết hợp của các thực thể tham gia
- > Xuất phát từ ràng buộc của thế giới thực
- > Có hai loại ràng buộc mối kết hợp chính
  - □ Ràng buộc dựa trên bản số
  - □ Ràng buộc dựa trên sự tham gia









### Lược đồ ER (1)

(Entity-Relationship Diagram viết tắt ERD)

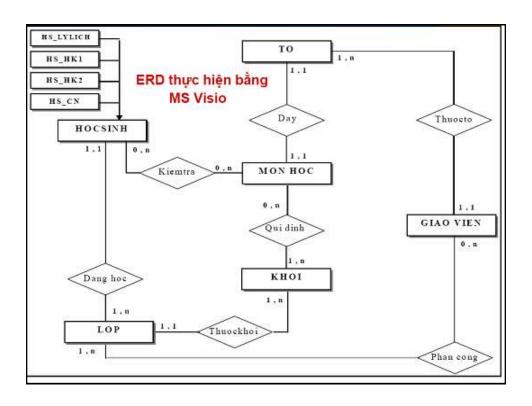
- ERD là đồ thị biểu diễn các tập thực thể, thuộc tính và mối quan hệ
- Đỉnh là tên tập thực thể, tên thuộc tính và tên quan hệ
- Cạnh là đường nối giữa:
  - Thực thể Thuộc tính
  - Mối quan hệ Thuộc tính
  - Thực thể Mối quan hệ

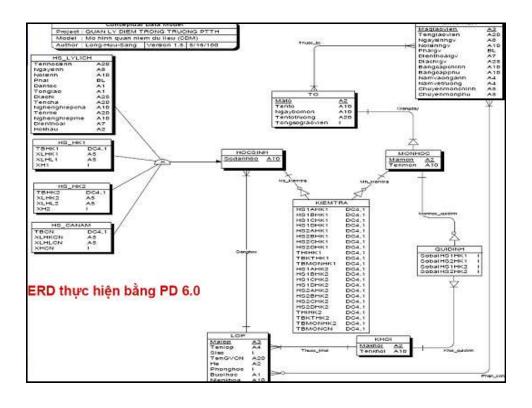
63

### Mô hình ER (2)

Các bước tạo ra ERD

- ❖ B1: Xác định thực thể, thuộc tính
- \* B2: Xác định mối kết hợp, thuộc tính
- ❖ B3: Xác định bản số
- ❖ B4: Vẽ mô hình bằng một số công cụ như:
  - MS Visio
  - **□** Power Designer





### Bài tập 1.2: Xây dựng mô hình ER

- Xây dựng mô hình ER cho CSDL quản lý giáo vụ gồm có các chức năng sau:
  - Lưu trữ thông tin: Học viên, giáo viên, môn học
  - 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.

67

### Bài tập 1.3: Xây dựng mô hình ER

• Vẽ sơ đồ ERD cho hệ thống thông tin sau Trường Đại học CNTT có nhu cầu xây dựng một hệ thống thông tin quản lý nhân sự nhằm mục đích đáp ứng ngày một tốt hơn nhu cầu cuộc sống của CB-CNV, đặt biệt là việc quan tâm, động viên kịp thời đối với con cái của mỗi CB-CNV của trường mỗi khi các cháu đạt thành tích trong học tập hay khen thưởng trong các đợt 1/6 hoặc hè. Các yêu cầu của hệ thống gồm:

### Bài tập 1.3: (tt) Xây dựng mô hình ER

- Quản lý thông tin về nhân thân của mỗi cán bộ công nhân viên (tên, tuổi, địa chỉ, điện thoại, vợ/chồng, ...)
- Quản lý thông tin về con cái của cán bộ công nhân viên (tên, ngày tháng năm sinh, giới tính, ...)
- Quản lý thông tin về thành tích trong học tập của các cháu (năm học, thành tích và các giải thưởng đặc biệt khác, ...)