

Đại học Công Nghệ Thông Tin

Môn: Cơ Sở Dữ Liệu
Thời lượng: 45LT+30TH

Nội dung

- ◆ Bài 1: Các khái niệm của một hệ CSDL
- ◆ Bài 2: Các mô hình CSDL
- ◆ Bài 3: Mô hình dữ liệu quan hệ (của Codd)
- ◆ Bài 4: Ngôn ngữ đại số quan hệ
- ◆ Bài 5: Ngôn ngữ SQL
- ◆ Bài 6: Ngôn ngữ tân từ
- ◆ Bài 7: Ràng buộc toàn vẹn trong một CSDL
- ◆ Bài 8: Tối ưu hóa câu hỏi bằng đại số quan hệ

Bài 1: Các khái niệm của một hệ CSDL

Bài 1: Các khái niệm của một hệ CSDL

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

1. Giới thiệu

- ♦ Tại sao cần phải có một cơ sở dữ liệu?

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)

◆ 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 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:
 - 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

4. Các đối tượng sử dụng

- ◆ Người sử dụng CSDL không chuyên về lĩnh vực tin học và CSDL -> cần công cụ để họ có thể khai thác CSDL khi cần.
- ◆ 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, cấp 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, 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)

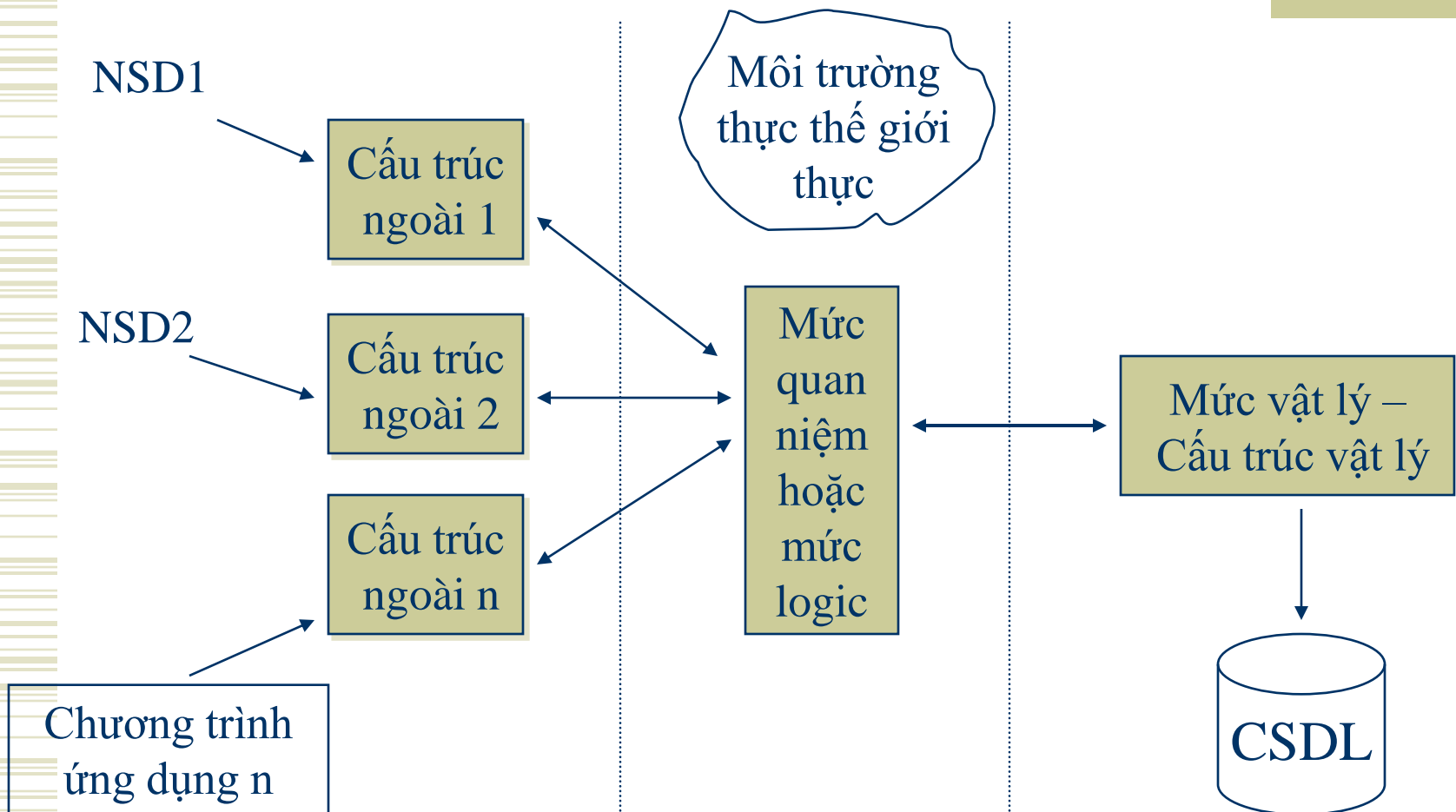
Ngôn ngữ giao tiếp:

- ♦ 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*) cho phép thay đổi cấu trúc bảng, khai báo bảo mật, cấp quyền cho người sử dụng.

6. 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 và người sử dụng chuyên môn.
- ♦ **Mức quan niệm:** (Conception hay Logical) cần phải lưu trữ bao nhiêu loại dữ liệu? là dữ liệu gì? mối quan hệ
- ♦ **Mức ngoài:** của người sử dụng và các chương trình ứng dụng

6. Các mức biểu diễn 1 CSDL



Bài 2: Các mô hình dữ liệu

Nội dung

1. Giới thiệu
2. Mô hình dữ liệu mạng
3. Mô hình thực thể mối kết hợp

1. Giới thiệu

Mô hình dữ liệu là sự trừu tượng hóa của môi trường thực, biểu diễn dữ liệu ở mức quan niệm. Giới thiệu một số mô hình như:

- 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 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

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: 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

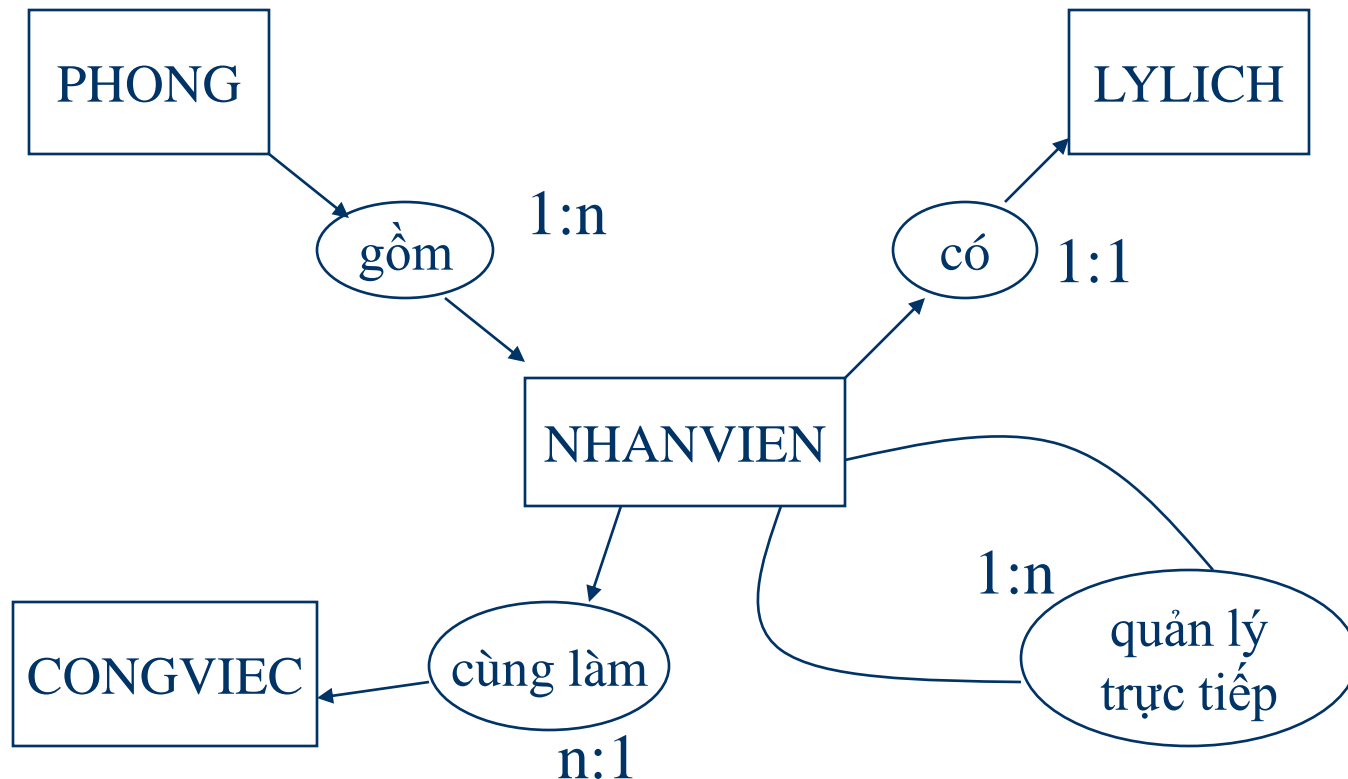
- Ký hiệu:

Tham gia

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.
 - **(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**

2. Mô hình dữ liệu mạng (4)



2. 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

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 (giáo viên mô tả hiện trạng, xác định các yêu cầu)

3. Mô hình thực thể mối kết hợp

3.1 Giới thiệu

3.2 Loại thực thể, thực thể

3.3 Thuộc tính của loại thực thể

3.4 Khoá của loại thực thể

3.5 Loại mối kết hợp, mối kết hợp

3.6 Thuộc tính của loại mối kết hợp

3.7 Bản số

3.8 Mô hình ER mở rộng

3.1 Giới thiệu

- ♦ Mô hình thực thể mối kết hợp (*Entity-Relationship Model* viết tắt ER) được CHEN giới thiệu năm 1976.
- ♦ Mô hình ER được sử dụng nhiều trong thiết kế dữ liệu ở mức quan niệm.

3.2 Loại thực thể

- ♦ **Định nghĩa:** loại thực thể (Entity Type) là những loại đối tượng hay sự vật của thế giới thực tồn tại cụ thể cần được quản lý.
- ♦ **Ví dụ :** HOCVIEN, LOP, MONHOC, ...
- ♦ **Ký hiệu:**

HOCVIEN

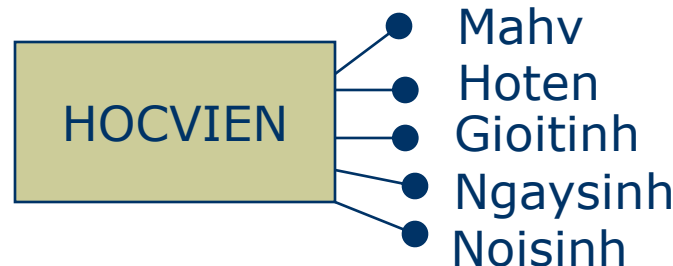
LOP

3.2 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 HOCVIEN có các thực thể:
 - ('HV001', 'Nguyen Nam', '1/2/1987', 'Nam')
 - ('HV002', 'Trần Nam', '13/2/1987', 'Nam')

3.3 Thuộc tính của loại thực thể² (Entity Attribute)

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



3.3 Các loại thuộc tính (1)

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

Ví dụ: Mahv,Hoten

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

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

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

Ví dụ: BANGCAP ký hiệu {BANGCAP}

3.3 Các loại thuộc tính (2)

- ◆ Tóm lại, các thuộc tính đa hợp và đa trị có thể lồng nhau tùy ý.
 - **Ví dụ:** thuộc tính BANGCAP của HOCVIEN là một thuộc tính đa hợp được ký hiệu bằng

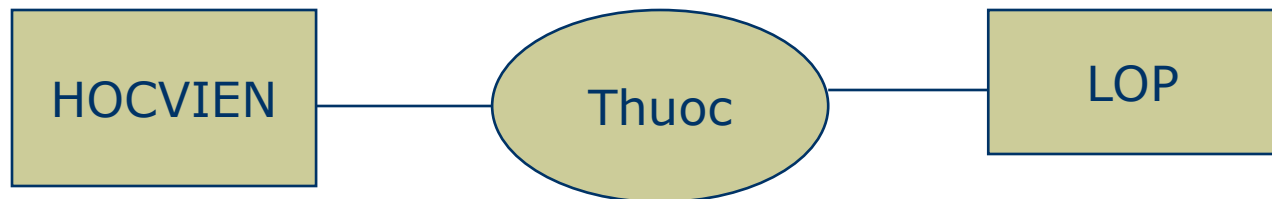
{BANGCAP(TRUONGCAP,NAM,KETQUA,
CHUYENNGANH)}

3.4. Khoá của loại thực thể² (entity type key)

- ◆ Khoá của loại thực thể là thuộc tính nhận diện thực thể.
- ◆ Căn cứ vào giá trị của khoá có thể xác định **duy nhất** một thực thể.
- ◆ Ví dụ:
 - Mỗi học viên có một mã số duy nhất => Khoá của loại thực thể **HOCVIEN** là **Mã học viên**

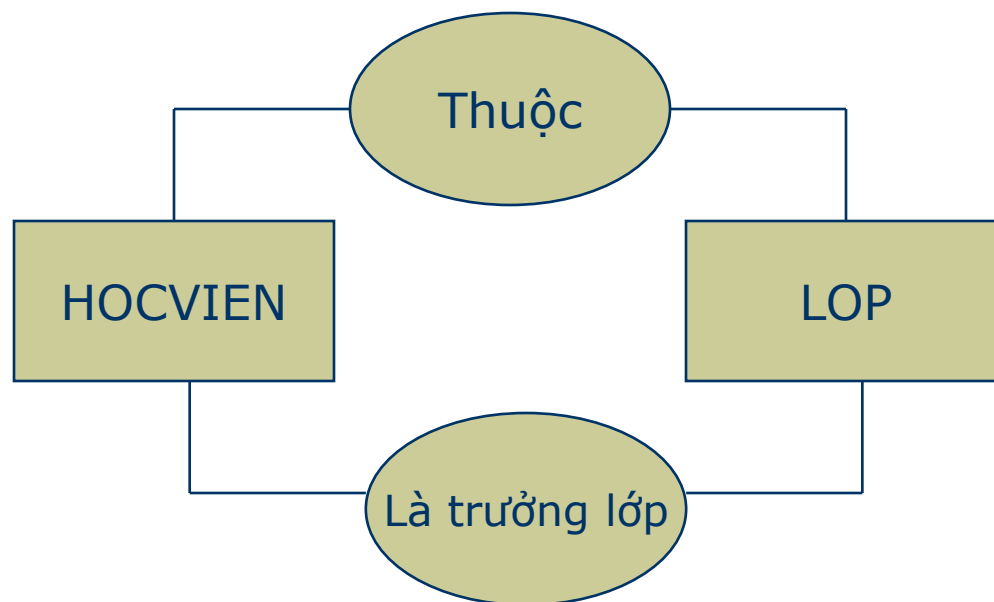
3.5 Loại mối kết hợp (1) (relationship type)

- ♦ **Định nghĩa:** loại mối kết hợp là sự kết hợp giữa hai hay nhiều loại thực thể
- ♦ **Ví dụ:** giữa hai loại thực thể HOCVIEN và LOP có loại mối kết hợp THUOC
- ♦ **Ký hiệu:** bằng một hình oval hoặc hình thoi



3.5 Loại mối kết hợp (2)

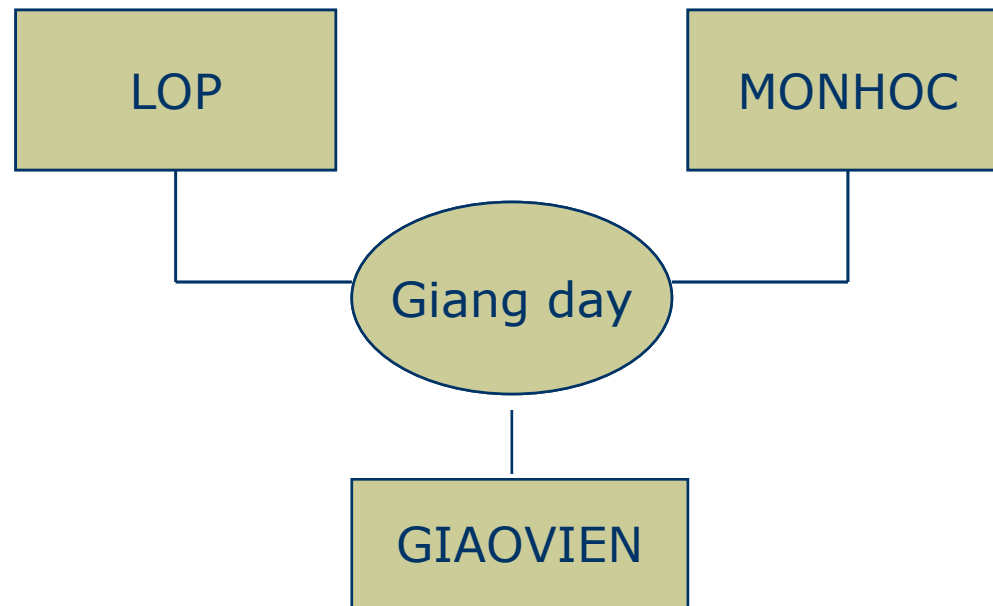
- ◆ Giữa hai loại thực thể có thể tồn tại nhiều hơn một loại mối kết hợp.
- ◆ Ví dụ



3.5 Số ngôi của loại mỗi kết hợp (relationship degree)

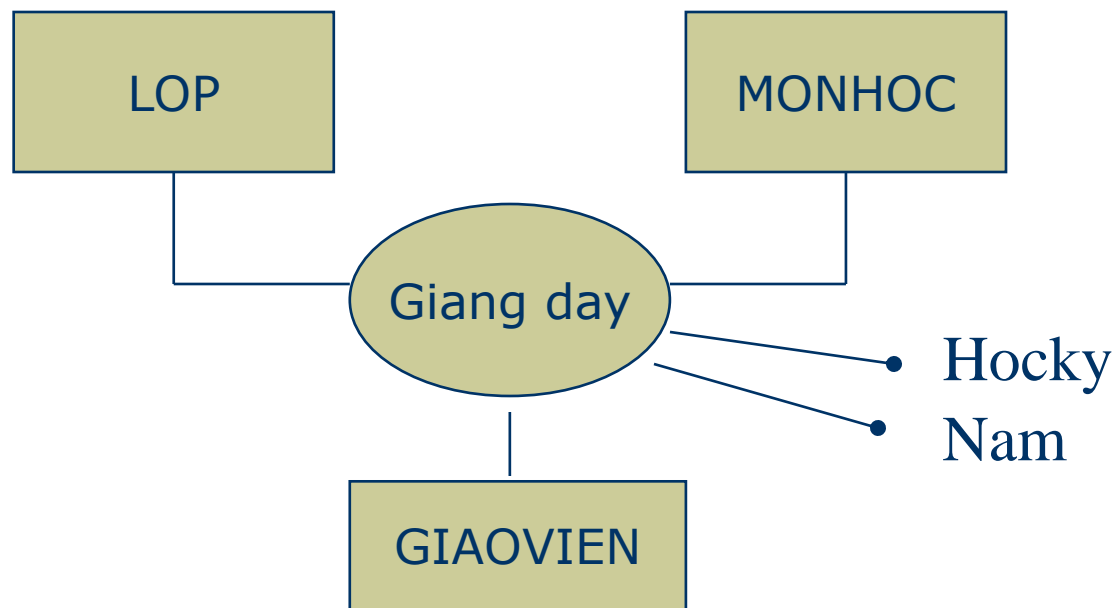
- ♦ Số ngôi của loại mỗi kết hợp là số loại thực thể tham gia vào loại mỗi kết hợp đó.
- ♦ **Ví dụ 1:** Loại mỗi kết hợp **Thuộc** kết hợp 2 loại thực thể HOCVIEN và LOP nên có số ngôi là 2.
- ♦ **Ví dụ 2:** Loại mỗi kết hợp **Giảng dạy** kết hợp 3 loại thực thể GIAOVIEN, MONHOC, LOP nên có số ngôi là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

3.5 Số ngôi của loại mối kết hợp



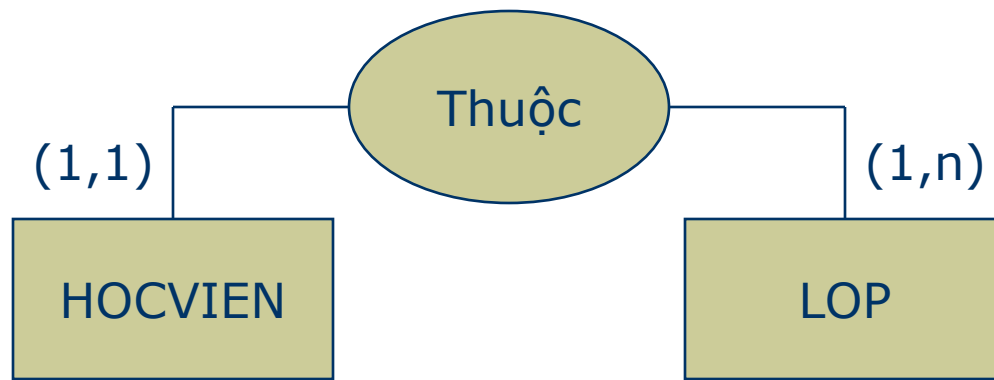
3.6 Thuộc tính của loại mối kết hợp (relationship type attribute)

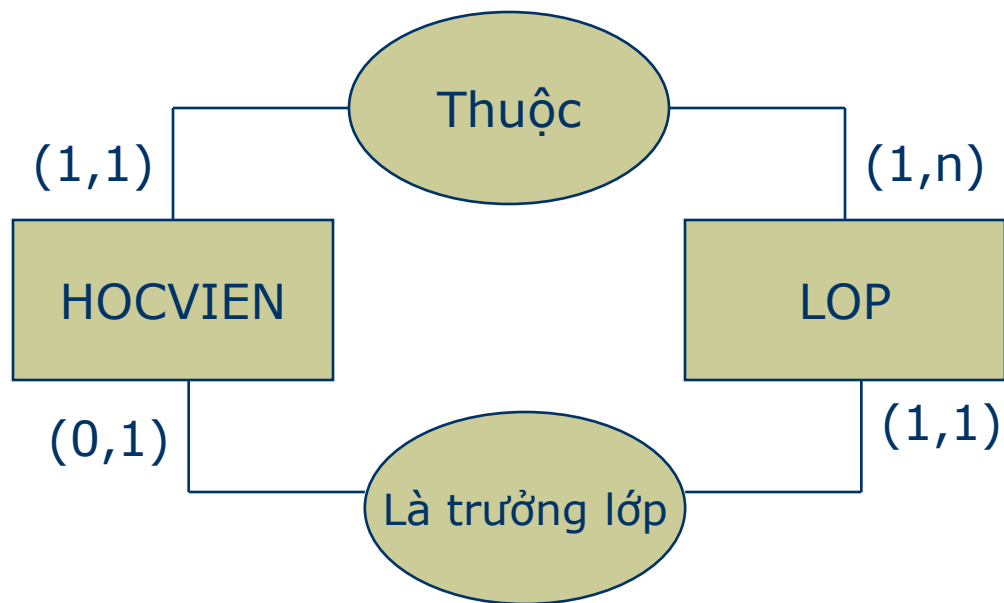
- ◆ Thuộc tính của loại mối kết hợp bao gồm các thuộc tính khoá của các loại thực thể tham gia vào loại mối kết hợp đó. Ngoài ra còn có thể có thêm những thuộc tính bổ sung khác.
- ◆ **Ví dụ:** Loại mối kết hợp **Giảng dạy** giữa ba loại thực thể GIAOVIEN, MONHOC và LOP có các thuộc tính là Magv, Mamh,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



3.7 Bản số (relationship cardinality)

- ♦ 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, bản số tối đa)
- ♦ **Ví dụ:** Loại thực thể NhanVien và DeAn có loại mối kết hợp ThamGia.

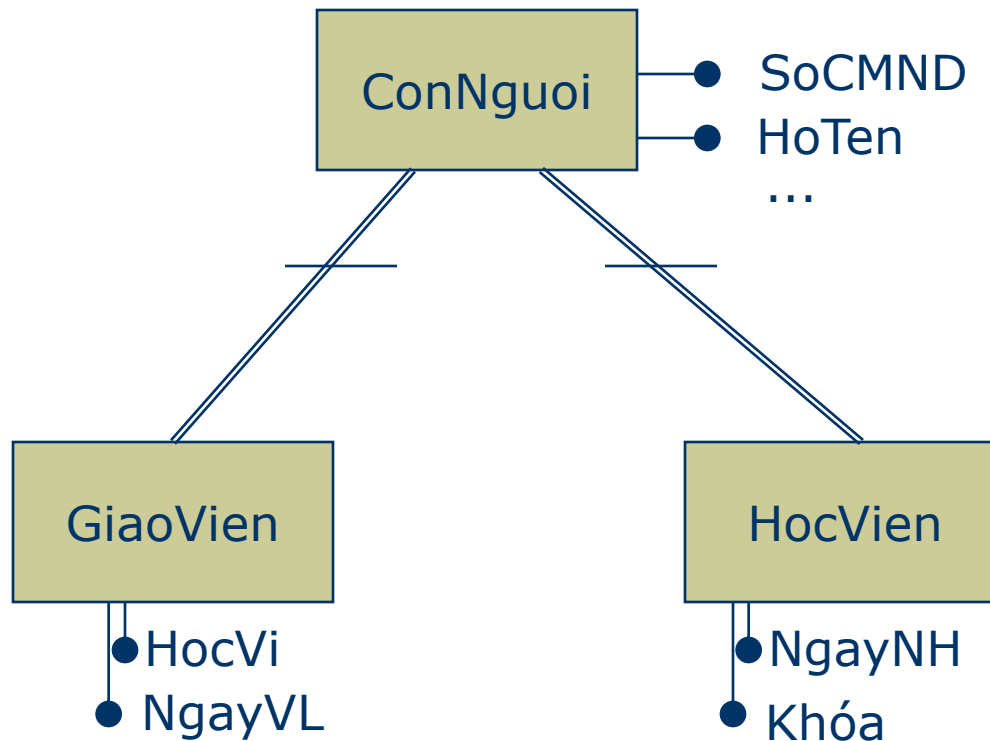




3.7 Mô hình ER mở rộng

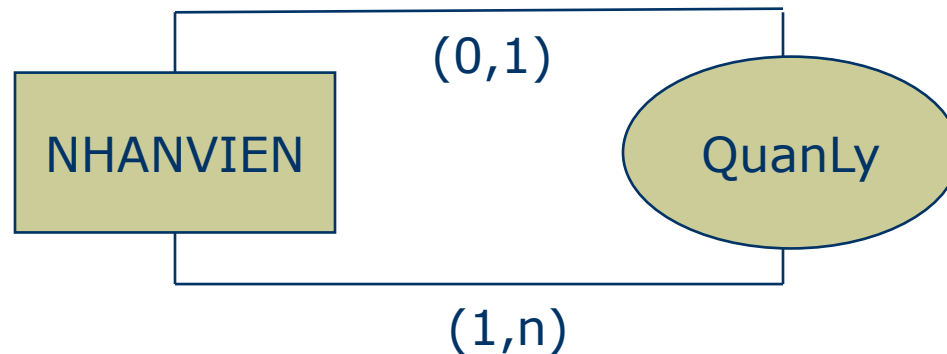
- 3.7.1 Chuyên biệt hoá / Tổng quát hóa
- 3.7.2 Mỗi kết hợp đệ quy
- 3.7.3 Loại thực thể yếu

3.7.1 Chuyên biệt hóa (tổng quát hóa)



3.7.2 Mỗi kết hợp đệ quy

- ♦ **Định nghĩa:** 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)
- ♦ **Ví dụ:** Mỗi nhân viên có một người quản lý trực tiếp và người quản lý đó cũng là một nhân viên



3.7.3 Loại thực thể yếu

♦ Định nghĩa:

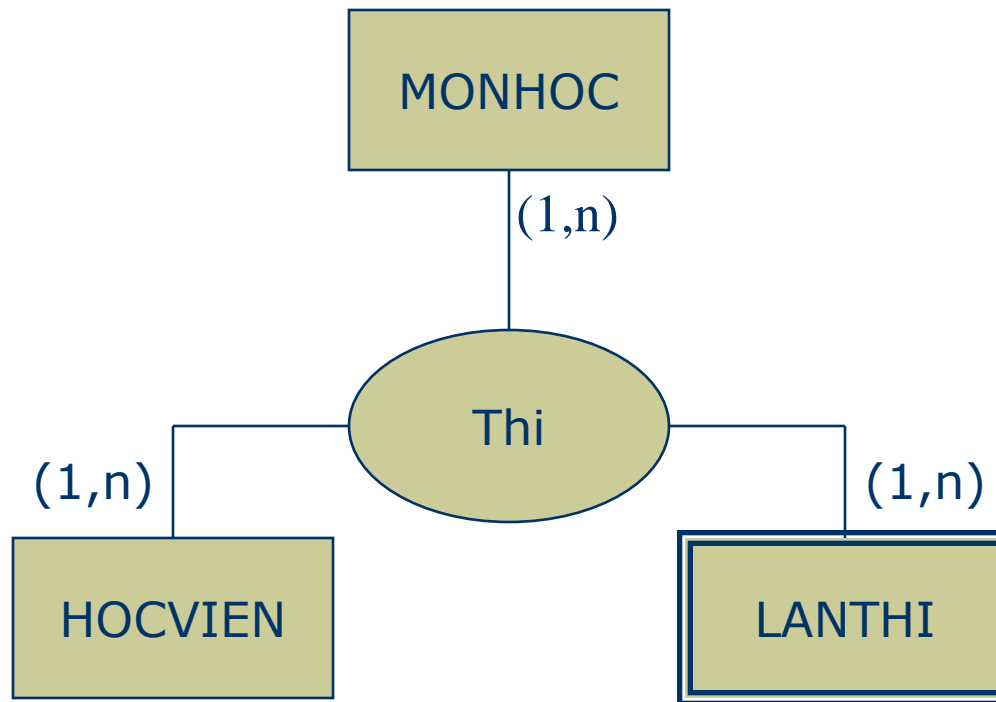
- Là loại thực thể không có thuộc tính khóa
- Phải tham gia trong một loại mối kết hợp xác định trong đó có một loại thực thể chủ.

♦ Ký hiệu:

Thực thể

- ♦ Ví dụ: loại thực thể LANTHI có thuộc tính Làn và tham gia trong loại mối kết hợp **Thi** với loại thực thể HOCVIEN và MONHOC là loại thực thể yếu.

3.7.3 Loại thực thể yếu



Bài tập

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.

liệu quan hệ (Relational Data Model)

Nội dung

1. Giới thiệu
2. Các khái niệm
 - 2.1 Thuộc tính
 - 2.2 Quan hệ
 - 2.3 Bộ giá trị
 - 2.4 Thể hiện của quan hệ
 - 2.5 Tân từ
 - 2.6 Phép chiếu
 - 2.7 Khóa
 - 2.8 Lược đồ quan hệ và lược đồ CSDL
 - 2.9 Hiện thực mô hình ER bằng mô hình dữ liệu quan hệ.

1. Giới thiệu

- ♦ Mô hình Dữ liệu Quan hệ (*Relational Data Model*) dựa trên khái niệm quan hệ.
- ♦ *Quan hệ* là khái niệm toán học dựa trên nền tảng hình thức về lý thuyết tập hợp.
- ♦ Mô hình này do TS. E. F. Codd đưa ra năm 1970.

2.1 Thuộc tính (attribute)

- ◆ Thuộc tính:
 - *Tên gọi*: dãy ký tự (gợi nhớ)
 - *Kiểu dữ liệu*: Số, Chuỗi, Thời gian, Luận lý, OLE.
 - *Miền giá trị*: tập giá trị mà thuộc tính có thể nhận. Ký hiệu miền giá trị của thuộc tính A là **Dom**(A).
- ◆ Ví dụ: GIOITINH kiểu dữ liệu là Chuỗi, miền giá trị
$$\text{Dom}(\text{GIOITINH}) = (\text{'Nam'}, \text{'Nu'})$$
- Tại một thời điểm, một thuộc tính không có giá trị hoặc chưa xác định được giá trị \Rightarrow giá trị Null

2.2 Quan hệ (relation)

- ♦ **Định nghĩa:** quan hệ là một tập hữu hạn các thuộc tính.
 - **Ký hiệu:** $Q(A_1, A_2, \dots, A_n)$
 - Trong đó Q là tên quan hệ, $Q^+ = \{A_1, A_2, \dots, A_n\}$ là tập các thuộc tính của quan hệ Q
 - **Ví dụ:**
HOCVIEN (Mahv, Hoten, Ngsinh, Gioitinh, Noisinh, Malop)
LOP (Malop, Tenlop, Siso, Trglop, Khoa)

2.3 Bộ (tuple)

- ♦ **Định nghĩa:** Bộ là các thông tin của một đối tượng thuộc quan hệ, được gọi là mẫu tin (record), dòng.
- ♦ Quan hệ là một bảng (table) với các cột là các thuộc tính và mỗi dòng được gọi là bộ.
- ♦ Một bộ của quan hệ $Q(A_1, A_2, \dots, A_n)$ là $Q(A_1, A_2, \dots, A_n)$ với $\forall a_i \in Dom(A_i)$
- ♦ **Ví dụ:** HOCVIEN(Mahv, Hoten, Ngsinh, Noisinh) có q=(1003, Nguyen Van Lam, 1/1/1987, Dong Nai) nghĩa là học viên có mã số là 1003, họ tên là Nguyen Van Lam, sinh ngày 1/1/1987 ở Dong Nai

2.4 Thể hiện của quan hệ (instance)

- ♦ **Định nghĩa:** thể hiện của một quan hệ là tập hợp các bộ giá trị của quan hệ tại một thời điểm nhất định.
- ♦ **Ký hiệu:** thể hiện của quan hệ Q là T_Q
- ♦ **Ví dụ:** T_{HOCVIEN} là thể hiện của quan hệ HOCVIEN tại thời điểm hiện tại gồm có các bộ như sau:

| HOCVIEN | | | | |
|---------|----------------|----------|------------|-------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Kinh | Nam | Tay Ninh | K11 |

2.5 Tân từ

- ♦ **Định nghĩa:** tân từ là một quy tắc dùng để mô tả một quan hệ.
- ♦ **Ký hiệu:** $||Q||$
- ♦ **Ví dụ:** THI (Mahv, Mamh, Lanthi, Diem)
 $||THI||$: mỗi học viên được phép thi một môn học nhiều lần, mỗi lần thi lưu trữ học viên nào thi môn gì? lần thi thứ mấy? và điểm là bao nhiêu?

2.6 Phép chiếu (1)

- ♦ **Phép chiếu** : Dùng để trích giá trị của một số thuộc tính trong danh sách các thuộc tính của quan hệ.
- ♦ **Ký hiệu**: phép chiếu của quan hệ R lên tập thuộc tính A là **$R[A]$ hoặc $R.A$** .
- ♦ **Ví dụ**:

| HOCVIEN | | | | |
|---------|----------------|----------|------------|-------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Linh | Nu | Tay Ninh | K11 |

- $h_{v_1} =$
- $h_{v_2} =$
- $h_{v_3} =$

2.6 Phép chiếu (2)

- ♦ Phép chiếu của quan hệ HOCVIEN lên thuộc tính NoiSinh của quan hệ HOCVIEN:

$HOCVIEN[Noisinh] = \{ 'Nghe An', 'Kien Giang', 'Tay Ninh' \}$

| HOCVIEN | | | | |
|---------|----------------|----------|------------|-------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Linh | Nu | Tay Ninh | K11 |

2.6 Phép chiếu (3)

- ◆ Phép chiếu lên 1 tập thuộc tính
 $K = \{ \text{Hoten}, \text{Noisinh} \}$ của quan hệ HOCVIEN

$\text{HOCVIEN}[\text{Hoten}, \text{Noisinh}] = \{ \{ \text{'Ha Duy Lap'}, \text{'Nghe An'} \}, \{ \text{'Tran Ngoc Han'}, \text{'Kien Giang'} \}, \{ \text{'Tran Ngoc Linh'}, \text{'Tay Ninh'} \} \}$

| HOCVIEN | | | | |
|---------|----------------|----------|------------|-------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Linh | Nu | Tay Ninh | K11 |

2.6 Phép chiếu (4)

- ♦ **Chiếu của một bộ lên tập thuộc tính:** dùng để trích chọn các giá trị cụ thể của bộ giá trị đó theo các thuộc tính được chỉ ra trong danh sách thuộc tính của một quan hệ.
- ♦ **Ký hiệu:** chiếu của một bộ giá trị t lên tập thuộc tính A của quan hệ R là $t_R[A]$ hoặc $t[A]$. Nếu A có 1 thuộc tính $t_R.A$
- ♦ **Ví dụ:** cho quan hệ HOCVIEN với tập thuộc tính $HOCVIEN^+ = \{Mahv, Hoten, Gioitinh, Noisinh, Malop\}$, chứa 3 bộ giá trị hv_1, hv_2 và hv_3

2.6 Phép chiếu (5)

- ♦ Phép chiếu 1 bộ lên 1 thuộc tính
 - $h\nu_1[\text{Hoten}] = \{\text{'Ha Duy Lap'}\}$

$h\nu_1 =$

$h\nu_2 =$

$h\nu_3 =$

| HOCVIEN | | | | |
|---------|----------------|----------|------------|-------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Linh | Nu | Tay Ninh | K11 |

2.6 Phép chiếu (6)

- ◆ Phép chiếu 1 bộ lên 1 tập thuộc tính
 - tập thuộc tính $K = \{Hoten, Gioitinh\}$
 - $h\nu_2[K] = \{ \text{'Tran Ngoc Han'}, \text{'Nu'} \}$

- $h\nu_1 =$
- $h\nu_2 =$
- $h\nu_3 =$

| HOCVIEN | | | | |
|---------|----------------|----------|------------|-------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Linh | Nu | Tay Ninh | K11 |

2.7 Khóa

2.7.1 Siêu khóa (super key)

2.7.2 Khóa (key)

2.7.3 Khóa chính (primary key)

2.7.5 Khóa ngoại (foreign key)

2.7.1 Siêu khóa (super key) (1)

- ♦ **Siêu khóa** : là một tập con các thuộc tính của Q^+ mà giá trị của chúng có thể phân biệt 2 bộ khác nhau trong cùng một thể hiện T_Q bất kỳ.
Nghĩa là: $\forall t_1, t_2 \in T_Q, t_1[K] \neq t_2[K] \Leftrightarrow K$ là siêu khóa của Q .
- ♦ Một quan hệ có ít nhất một siêu khóa (Q^+) và có thể có nhiều siêu khóa.

2.7.1 Siêu khóa (super key) (2)

- ♦ Ví dụ: các siêu khóa của quan hệ HOCVIEN là:
 $\{Mahv\}; \{Mahv, Hoten\}; \{Hoten\}; \{Noisinh, Hoten\}$

...

| HOCVIEN | | | | |
|----------------|----------------|-----------------|----------------|--------------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Linh | Nu | Tay Ninh | K11 |
| K1105 | Tran Minh Long | Nam | TpHCM | K11 |
| K1106 | Le Nhat Minh | Nam | TpHCM | K11 |

2.7.2 Khóa (key) (1)

Khóa : K là khóa của quan hệ R, thỏa mãn 2 điều kiện:

- ♦ K là một siêu khóa.
- ♦ K là siêu khóa “*nhỏ nhất*” (chứa ít thuộc tính nhất và khác rỗng) nghĩa là

$\neg \exists K_1 \subset K, K_1 \neq \emptyset, K_1 \text{ là siêu khóa.}$

- ♦ Thuộc tính tham gia vào một khóa gọi là *thuộc tính khóa*, ngược lại là *thuộc tính không khóa*.

2.7.2 Khóa (key) (2)

- ♦ **Ví dụ:** các siêu khóa của quan hệ HOCVIEN là:
 $\{Mahv\}; \{Mahv, Hoten\}; \{Hoten\}; \{Hoten, Gioitinh\};$
 $\{Noisinh, Hoten\}; \{Mahv, Hoten, Gioitinh, Noisinh\} \dots$
 \Rightarrow thì khóa của quan hệ HOCVIEN có thể là
 $\{Mahv\}; \{Hoten\}$
- ♦ **Ví dụ:** khóa của quan hệ GIANGDAY là
 $K = \{Magv, Mamh, Malop\}$. Như vậy thuộc tính khóa sẽ là: $Magv, Mamh, Malop$.

2.7.3 Khóa chính (primary key)

- ◆ Khi cài đặt trên một DBMS cụ thể, nếu quan hệ có nhiều hơn một **khóa**, ta chỉ được chọn một và gọi là **khóa chính**
- ◆ **Ký hiệu:** các *thuộc tính nằm trong khóa chính* khi liệt kê trong quan hệ phải được gạch dưới.
- ◆ **Ví dụ:**
 - HOCVIEN (Mahv,Hoten,Gioitinh,Noisinh,Malop)
 - GIANGDAY(Magv,Mamh,Malop,Hocky,Nam)

2.7.4 Khóa ngoại (1)

- ♦ Cho $R(U)$, $S(V)$. $K_1 \subseteq U$ là khóa chính của R , $K_2 \subseteq V$
- ♦ Ta nói **K_2 là khóa ngoại** của S *tham chiếu đến* khóa chính K_1 của R nếu thỏa các điều kiện sau:
 - K_1 và K_2 có cùng số lượng thuộc tính và ngữ nghĩa của các thuộc tính trong K_1 và K_2 cũng giống nhau.
 - Giữa R và S tồn tại mối quan hệ 1-n trên K_1 và K_2 ,
 - $\forall s \in S, \quad \exists r \in R$ sao cho $r.K_1 = s.K_2$

2.7.4 Khóa ngoại (2)

- ♦ Ví dụ, cho 2 quan hệ

LOP (Malop, Tenlop, Siso, Khoahoc)


HOCVIEN (Mahv, Hoten, Gioitinh, Noisinh, Malop)



- ♦ Thuộc tính Malop trong quan hệ LOP là khóa chính của quan hệ LOP. Thuộc tính *Malop* trong quan hệ *HOCVIEN* là khóa ngoại, tham chiếu đến *Malop* trong quan hệ *LOP*

2.7.4 Khóa ngoại (3)

| HOCVIEN | | | | |
|---------|----------------|----------|------------|-------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Linh | Nu | Tay Ninh | K11 |
| K1105 | Tran Minh Long | Nam | TpHCM | K11 |
| K1106 | Le Nhat Minh | Nam | TpHCM | K11 |



| LOP | | | | |
|-------|--------------|--------|------|--------|
| Malop | Tenlop | Trglop | Siso | Magvcn |
| K11 | Lop 1 khoa 1 | K1106 | 11 | GV07 |
| K12 | Lop 2 khoa 1 | K1205 | 12 | GV09 |
| K13 | Lop 3 khoa 1 | K1305 | 12 | GV14 |

2.8 Lược đồ quan hệ (1)

- ♦ **Lược đồ quan hệ** nhằm mục đích **mô tả cấu trúc** của một quan hệ và **các mối liên hệ** giữa các thuộc tính trong quan hệ đó.
- ♦ **Cấu trúc của một quan hệ** là tập thuộc tính hình thành nên quan hệ đó.
- ♦ Một lược đồ quan hệ gồm một tập thuộc tính của quan hệ kèm theo một mô tả để xác định ý nghĩa và mối liên hệ giữa các thuộc tính

2.8 Lược đồ quan hệ (2)

- ♦ Lược đồ quan hệ được đặc trưng bởi:
 - Một tên phân biệt
 - Một tập hợp hữu hạn các thuộc tính (A_1, \dots, A_n)
- ♦ Ký hiệu của lược đồ quan hệ Q gồm n thuộc tính (A_1, A_2, \dots, A_n) là :
 - $Q(A_1, A_2, \dots, A_n)$

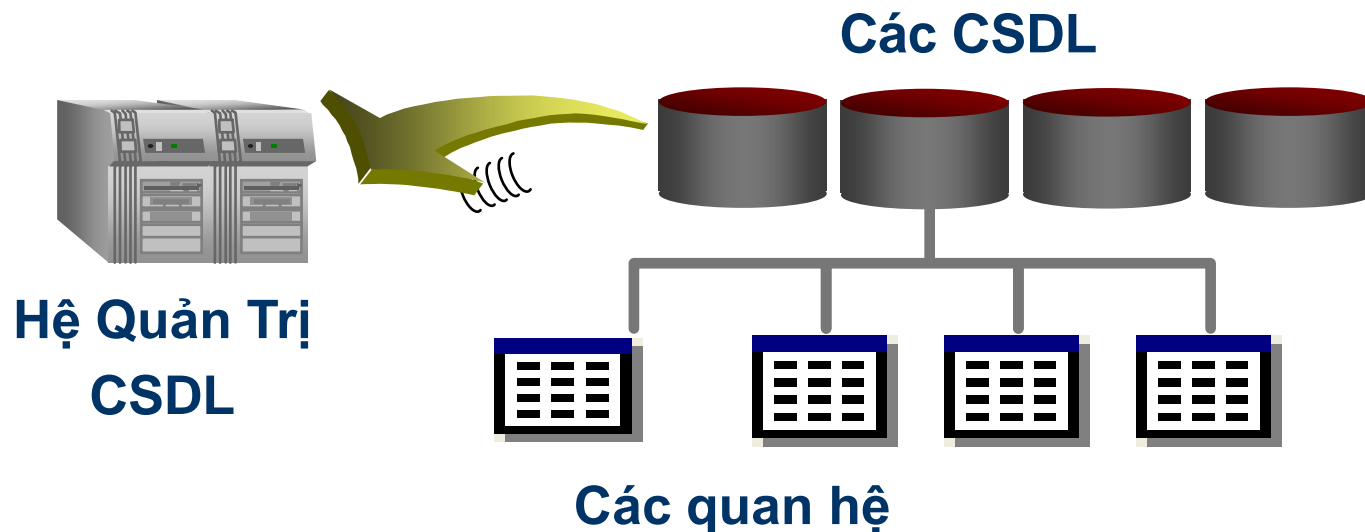
2.8 Lược đồ quan hệ (3)

- ♦ HOCVIEN(Mahv,Hoten,Gioitinh,Noisinh,Malop
- ♦ Tân từ: mỗi học viên có một mã học viên để phân biệt với các học viên khác. Cần lưu trữ họ tên, giới tính, nơi sinh và thuộc lớp nào.

| HOCVIEN | | | | |
|---------|----------------|----------|------------|-------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Linh | Nu | Tay Ninh | K11 |
| K1105 | Tran Minh Long | Nam | TpHCM | K11 |
| K1106 | Le Nhat Minh | Nam | TpHCM | K11 |

2.8 Lược đồ CSDL (1)

- ♦ Là tập hợp gồm các lược đồ quan hệ và các mối liên hệ giữa chúng trong cùng một hệ thống quản lý.



Lược đồ CSDL quản lý bán hàng

HOCVIEN (MAHV, HO, TEN, NGSINH, GIOITINH, NOISINH, MALOP)

Tên từ: mỗi học viên phân biệt với nhau bằng mã học viên, lưu trữ họ tên, ngày sinh, giới tính, nơi sinh, thuộc lớp nào.

LOP (MALOP, TENLOP, TRGLOP, SISO, MAGVCN)

Tên từ: mỗi lớp gồm có mã lớp, tên lớp, học viên làm lớp trưởng của lớp, sĩ số lớp và giáo viên chủ nhiệm.

KHOA (MAKHOA, TENKHOA, NGTLAP, TRGKHOA)

Tên từ: mỗi khoa cần lưu trữ mã khoa, tên khoa, ngày thành lập khoa và trưởng khoa (cũng là một giáo viên thuộc khoa).

MONHOC (MAMH, TENMH, TCLT, TCTH, MAKHOA)

Tên từ: mỗi môn học cần lưu trữ tên môn học, số tín chỉ lý thuyết, số tín chỉ thực hành và khoa nào phụ trách.

DIEUKIEN (MAMH, MAMH_TRUOC)

Tên từ: có những môn học học viên phải có kiến thức từ một số môn học trước.

GIAOVIEN(MAGV,HOTEN,HOCVI,HOCHAM,GIOITINH,NGSINH,NGVL,
HESO, MUCLUONG, MAKHOA)

Tân từ: mã giáo viên để phân biệt giữa các giáo viên, cần lưu trữ họ tên, học vị, học hàm, giới tính, ngày sinh, ngày vào làm, hệ số, mức lương và thuộc một khoa.

GIANGDAY(MALOP,MAMH,MAGV,HOCKY, NAM,TUNGAY,DENNGAY)

Tân từ: mỗi học kỳ của năm học sẽ phân công giảng dạy: lớp nào học môn gì do giáo viên nào phụ trách.

KETQUATHI (MAHV, MAMH, LANTHI, NGTHI, DIEM, KQUA)

Tân từ: lưu trữ kết quả thi của học viên: học viên nào thi môn học gì, lần thi thứ mấy, ngày thi là ngày nào, điểm thi bao nhiêu và kết quả là đạt hay không đạt.

Bài 4: Ngôn ngữ đại số quan hệ

Nội dung

1. Giới thiệu
2. Biểu thức đại số quan hệ
3. Các phép toán
4. Biểu thức đại số quan hệ
5. Ví dụ

1. Giới thiệu

- ◆ Đại số quan hệ (ĐSQH) có nền tảng toán học (cụ thể là lý thuyết tập hợp) để mô hình hóa CSDL quan hệ. Đối tượng xử lý là các quan hệ trong cơ sở dữ liệu quan hệ.
- ◆ **Chức năng:**
 - Cho phép mô tả các phép toán rút trích dữ liệu từ các quan hệ trong cơ sở dữ liệu quan hệ.
 - Cho phép tối ưu quá trình rút trích bằng các phép toán có sẵn của lý thuyết tập hợp.

2. Biểu thức ĐSQH

- ♦ Biểu thức ĐSQH là một biểu thức gồm các phép toán ĐSQH.
- ♦ Biểu thức ĐSQH được xem như một quan hệ (không có tên).
- ♦ Có thể đặt tên cho quan hệ được tạo từ một biểu thức ĐSQH.
- ♦ Có thể đổi tên các thuộc tính của quan hệ được tạo từ một biểu thức ĐSQH.

3. Các phép toán

3.1 Giới thiệu

3.2 Phép chọn

3.3 Phép chiếu

3.4 Phép gán

3.5 Các phép toán trên tập hợp

3.6 Phép kết

3.7 Phép chia

3.8 Hàm tính toán và gom nhóm

3.1 Giới thiệu (1)

- ◆ Có năm phép toán cơ bản:
 - Chọn (σ) hoặc ($:$)
 - Chiếu (π) hoặc ($[]$)
 - Tích (\times)
 - Hiệu ($-$)
 - Hội (\cup)

3.1 Giới thiệu (2)

- ◆ Các phép toán khác không cơ bản nhưng hữu ích:
 - Giao (\cap)
 - Kết (\bowtie)
 - Chia (\div)
 - Phép bù (\neg)
 - Đổi tên (ρ)
 - Phép gán (\leftarrow)
- ◆ Kết quả sau khi thực hiện các phép toán là các quan hệ, do đó có thể kết hợp giữa các phép toán để tạo nên phép toán mới.

3.2 Phép chọn (Selection)

- ♦ Trích chọn các bộ (dòng) từ quan hệ R. Các bộ được trích chọn phải thỏa mãn điều kiện chọn p.
- ♦ **Ký hiệu:** $\sigma_p(R)$
- ♦ **Định nghĩa:** $\sigma_p(R) = \{t / t \in R, p(t)\}$ p(t):thỏa điều kiện p
- ♦ Kết quả trả về là một quan hệ, có cùng danh sách thuộc tính với quan hệ R. Không có kết quả trùng.
- ♦ Phép chọn có tính giao hoán

$$\sigma_{p1}(\sigma_{p2}(R)) = \sigma_{p2}(\sigma_{p1}(R)) = \sigma_{(p1 \wedge p2)}(R)$$

Lược đồ CSDL quản lý giáo vụ

HOCVIEN (MAHV, HO, TEN, NGSINH, GIOITINH, NOISINH, MALOP)

LOP (MALOP, TENLOP, TRGLOP, SISO, MAGVCN)

KHOA (MAKHOA, TENKHOA, NGTLAP, TRGKHOA)

MONHOC (MAMH, TENMH, TCLT, TCTH, MAKHOA)

DIEUKIEN (MAMH, MAMH_TRUOC)

GIAOVIEN (MAGV, HOTEN, HOCVI, HOCHAM, GIOITINH, NGSINH, NGVL, HESO, MUCLUONG, MAKHOA)

GIANGDAY (MALOP, MAMH, MAGV, HOCKY, NAM, TUNGAY, DENNGAY)

KETQUATHI (MAHV, MAMH, LANTHI, NGTHI, DIEM, KQUA)

3.2 Ví dụ phép chọn

- Tìm những học viên ‘Nam’ có nơi sinh ở ‘TpHCM’

$\sigma_{(\text{Gioitinh}=\text{'Nam'}) \wedge (\text{Noisinh}=\text{'TpHCM'})}(\text{HOCVIEN})$

| HOCVIEN | | | | |
|---------|----------------|----------|------------|-------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Linh | Nu | Tay Ninh | K11 |
| K1105 | Tran Minh Long | Nam | TpHCM | K11 |
| K1106 | Le Nhat Minh | Nam | TpHCM | K11 |

3.3 Phép chiếu (Project)

- ♦ Sử dụng để trích chọn giá trị một vài thuộc tính của quan hệ
- ♦ **Ký hiệu:** $\pi_{A_1, A_2, \dots, A_k}(R)$

trong đó A_i là tên các thuộc tính được chiếu.

- ♦ Kết quả trả về một quan hệ có k thuộc tính theo thứ tự như liệt kê. Các dòng trùng nhau chỉ lấy một.
- ♦ Phép chiếu không có tính giao hoán

3.3 Ví dụ

- Tìm mã số, họ tên những học viên “Nam” có nơi sinh ở ‘TpHCM’

$\pi_{Mahv, Hoten} \sigma_{(Gioitinh='Nam') \wedge (Noisinh='TpHCM')}(HOCVIEN)$

| HOCVIEN | | | | |
|---------|----------------|----------|------------|-------|
| Mahv | HoTen | Gioitinh | Noisinh | Malop |
| K1103 | Ha Duy Lap | Nam | Nghe An | K11 |
| K1102 | Tran Ngoc Han | Nu | Kien Giang | K11 |
| K1104 | Tran Ngoc Linh | Nu | Tay Ninh | K11 |
| K1105 | Tran Minh Long | Nam | TpHCM | K11 |
| K1106 | Le Nhat Minh | Nam | TpHCM | K11 |

3.4 Phép gán (Assignment)

- ♦ Dùng để diễn tả câu truy vấn phức tạp.
- ♦ **Ký hiệu:** $A \leftarrow B$
- ♦ **Ví dụ:**
 $R(\text{HO}, \text{TEN}, \text{LUONG}) \leftarrow \pi_{\text{HONV}, \text{TENNV}, \text{LUONG}}(\text{NHANVIEN})$
- ♦ Kết quả bên phải của phép gán được gán cho biến quan hệ nằm bên trái.

3.5 Các phép toán tập hợp

3.5.1 Giới thiệu

3.5.2 Phép hội

3.5.3 Phép trừ

3.5.4 Phép giao

3.5.5 Phép tích

3.5.1 Giới thiệu

- ◆ Các phép toán thực hiện trên 2 quan hệ xuất phát từ lý thuyết tập hợp của toán học: phép hội ($R \cup S$), phép giao ($R \cap S$), phép trừ ($R - S$), phép tích ($R \times S$).
- ◆ Đối với các phép hội, giao, trừ, các quan hệ R và S phải **khả hợp**:
 - Số lượng thuộc tính của R và S phải bằng nhau:
 $R(A_1, A_2, \dots, A_n)$ và $S(B_1, B_2, \dots, B_n)$
 - Miền giá trị của thuộc tính phải tương thích
 $\text{dom}(A_i) = \text{dom}(B_i)$
- ◆ Quan hệ kết quả của phép hội, giao, trừ có cùng tên thuộc tính với quan hệ đầu tiên.

3.5.2 Phép hội (Union)

- ♦ **Ký hiệu:** $R \cup S$
- ♦ **Định nghĩa:** $R \cup S = \{t | t \in R \vee t \in S\}$ trong đó R,S là hai quan hệ khả hợp.
- ♦ **Ví dụ:** Học viên được khen thưởng đợt 1 hoặc đợt 2

| DOT1 | |
|-------|---------------|
| Mahv | Hoten |
| K1103 | Le Van Tam |
| K1114 | Tran Ngoc Han |
| K1203 | Le Thanh Hau |
| K1308 | Nguyen Gia |

| DOT2 | |
|-------|---------------|
| Mahv | Hoten |
| K1101 | Le Kieu My |
| K1114 | Tran Ngoc Han |

| Mahv | Hoten |
|-------|---------------|
| K1101 | Le Kieu My |
| K1103 | Le Van Tam |
| K1114 | Tran Ngoc Han |
| K1203 | Le Thanh Hau |
| K1308 | Nguyen Gia |

DOT1 ∪ DOT2

3.5.3 Phép trừ (Set Difference)

- ♦ **Ký hiệu:** $R - S$
- ♦ **Định nghĩa:** $R - S = \{t | t \in R \vee t \notin S\}$ trong đó R, S là hai quan hệ khả hợp.
- ♦ **Ví dụ:** Học viên được khen thưởng đợt 1 nhưng không được khen thưởng đợt 2

| DOT1 | |
|-------|---------------|
| Mahv | Hoten |
| K1103 | Le Van Tam |
| K1114 | Tran Ngoc Han |
| K1203 | Le Thanh Hau |
| K1308 | Nguyen Gia |

| DOT2 | |
|-------|---------------|
| Mahv | Hoten |
| K1101 | Le Kieu My |
| K1114 | Tran Ngoc Han |

| Mahv | Hoten |
|-------|--------------|
| K1103 | Le Van Tam |
| K1203 | Le Thanh Hau |

DOT1- DOT2

3.5.4 Phép giao (Set-Intersection)

- ♦ **Ký hiệu:** $R \cap S$
- ♦ **Định nghĩa:** $R \cap S = \{t | t \in R \wedge t \in S\}$ trong đó R,S là hai quan hệ khả hợp. Hoặc $R \cap S = R - (R - S)$
- ♦ **Ví dụ:** Học viên được khen thưởng cả hai đợt 1 và 2

| KT_D1 | |
|-------|---------------|
| Mahv | Hoten |
| K1103 | Le Van Tam |
| K1114 | Tran Ngoc Han |
| K1203 | Le Thanh Hau |
| K1308 | Nguyen Gia |

| KT_D2 | |
|-------|---------------|
| Mahv | Hoten |
| K1101 | Le Kieu My |
| K1114 | Tran Ngoc Han |

| Mahv | Hoten |
|-------|---------------|
| K1114 | Tran Ngoc Han |

DOT1 \cap DOT2

3.5.5 Phép tích (1)

- ♦ **Ký hiệu:** $R \times S$
- ♦ **Định nghĩa:** $R \times S = \{ t_r t_s / t_r \in R \wedge t_s \in S \}$
- ♦ Nếu R có n bộ và S có m bộ thì kết quả là $n \times m$ bộ
 $KQ(A_1, A_2, \dots, A_m, B_1, B_2, \dots, B_n) \leftarrow R(A_1, A_2, \dots, A_m) \times S(B_1, B_2, \dots, B_n)$
- ♦ Phép tích thường dùng kết hợp với các phép chọn để kết hợp các bộ có liên quan từ hai quan hệ.
- ♦ **Ví dụ:** từ hai quan hệ HOCVIEN và MONHOC, có tất cả những trường hợp nào “*học viên đăng ký học môn học*”, giả sử không có bất kỳ điều kiện nào

3.5.5 Phép tích (2)

| HOCVIEN | |
|---------|---------------|
| Mahv | Hoten |
| K1103 | Le Van Tam |
| K1114 | Tran Ngoc Han |
| K1203 | Le Thanh Hau |

| MONHOC |
|--------|
| Mamh |
| CTRR |
| THDC |
| CTDL |

| Mahv | Hoten | Mamh |
|-------|---------------|------|
| K1103 | Le Van Tam | CTRR |
| K1114 | Tran Ngoc Han | CTRR |
| K1203 | Le Thanh Hau | CTRR |
| K1103 | Le Van Tam | THDC |
| K1114 | Tran Ngoc Han | THDC |
| K1203 | Le Thanh Hau | THDC |
| K1103 | Le Van Tam | CTDL |
| K1114 | Tran Ngoc Han | CTDL |
| K1203 | Le Thanh Hau | CTDL |

HOCVIEN × MONHOC

3.6 Phép kết

3.6.1 Phép kết

3.6.2 Phép kết bằng, phép kết tự nhiên

3.6.3 Phép kết ngoài

3.6.1 Phép kết (Theta-Join) (1)

- ♦ Theta-join (θ): Tương tự như phép tích kết hợp với phép chọn. Điều kiện chọn gọi là *điều kiện kết*.
- ♦ **Ký hiệu:** $R \bowtie_p S$
trong đó R,S là các quan hệ, p là điều kiện kết
- ♦ Các bộ có giá trị NULL tại thuộc tính kết nối không xuất hiện trong kết quả của phép kết.
- ♦ Phép kết với điều kiện tổng quát gọi là θ -kết với θ là một trong những phép so sánh ($\neq, =, >, \geq, <, \leq$)

3.6.1 Phép kết (2)

$$A_1 > B_2$$

$$R \triangleright \triangleleft S$$

| R | |
|----------------|----------------|
| A ₁ | A ₂ |
| 1 | 2 |
| 1 | 8 |
| 0 | 0 |
| 8 | 4 |
| 0 | 3 |

| S | | |
|----------------|----------------|----------------|
| B ₁ | B ₂ | B ₃ |
| 0 | 2 | 8 |
| 7 | 8 | 7 |
| 8 | 0 | 4 |
| 1 | 0 | 7 |
| 2 | 1 | 5 |

| A ₁ | A ₂ | B ₁ | B ₂ | B ₃ |
|----------------|----------------|----------------|----------------|----------------|
| 1 | 2 | 8 | 0 | 4 |
| 1 | 2 | 1 | 0 | 7 |
| 1 | 8 | 8 | 0 | 4 |
| 1 | 8 | 1 | 0 | 7 |
| 8 | 4 | 0 | 2 | 8 |
| 8 | 4 | 8 | 0 | 4 |
| 8 | 4 | 1 | 0 | 7 |
| 8 | 4 | 2 | 1 | 5 |

3.6.2 Phép kết bằng, kết tự nhiên

- ◆ Nếu θ là phép so sánh bằng ($=$), phép kết gọi là phép kết bằng (equi-join).

Ký hiệu: $\overset{Mahv = Trglop}{HOCVIEN} \triangleright \triangleleft LOP$

- ◆ Nếu điều kiện của equi-join là các thuộc tính giống nhau thì gọi là phép kết tự nhiên (natural-join). Khi đó kết quả của phép kết loại bỏ bớt 1 cột (bỏ 1 trong 2 cột giống nhau)

Ký hiệu: $\overset{Mahv}{HOCVIEN} \triangleright \triangleleft \overset{Mahv}{KETQUATHI}$ hoặc
 $HOCVIEN * KETQUATHI$

3.6.3 Phép kết ngoài (outer join)

- ◆ Mở rộng phép kết để tránh mất thông tin
- ◆ Thực hiện phép kết và sau đó thêm vào kết quả của phép kết các bộ của quan hệ mà không phù hợp với các bộ trong quan hệ kia.
- ◆ Có 3 loại:
 - Left outer join $R \bowtie S$
 - Right outer join $R \bowtie S$
 - Full outer join $R \bowtie S$
- ◆ **Ví dụ:** In ra danh sách tất cả các học viên và điểm số của các môn học mà học viên đó thi (nếu có)

3.6.3 Phép kết ngoài (2)

- ♦ HOCVIEN $\overset{\text{mahv}}{\bowtie}$ KETQUATHI

| Mahv | Hoten | Mahv | Mam h | Diem |
|------|----------------|------|----------|------|
| HV01 | Nguyen Van Lan | HV01 | CSDL | 7.0 |
| HV01 | Nguyen Van Lan | HV01 | CTRR | 8.5 |
| HV02 | Tran Hong Son | HV02 | CSDL | 8.5 |
| HV03 | Nguyen Le | HV03 | CTRR | 9.0 |
| HV04 | Le Minh | Null | Null | Null |

| HOCVIEN | |
|---------|----------------|
| Mahv | Hoten |
| HV01 | Nguyen Van Lan |
| HV02 | Tran Hong Son |
| HV03 | Nguyen Le |
| HV04 | Le Minh |

| KETQUATHI | | |
|-----------|------|------|
| Mahv | Mamh | Diem |
| HV01 | CSDL | 7.0 |
| HV02 | CSDL | 8.5 |
| HV01 | CTRR | 8.5 |
| HV03 | CTRR | 9.0 |

3.7 Phép chia (Division)

- ♦ **Định nghĩa:** $Q = R \div S = \{t / \forall s \in S, (t, s) \in R\}$
- ♦ R và S là hai quan hệ, R^+ và S^+ lần lượt là tập thuộc tính của R và S. Điều kiện $S^+ \neq \emptyset$ là **tập con không bằng** của R^+ . Q là kết quả phép chia giữa R và S, $Q^+ = R^+ - S^+$
- ♦ Có thể diễn đạt bằng phép toán đại số như sau:

$$T_1 \leftarrow \pi_{R^+ - S^+}(R)$$

$$T_2 \leftarrow \pi_{R^+ - S^+}((S \times T_1) - R)$$

$$T \leftarrow T_1 - T_2$$

3.7 Phép chia (2)

| KETQUATHI | | |
|-----------|------|------|
| Mahv | Mamh | Diem |
| HV01 | CSDL | 7.0 |
| HV02 | CSDL | 8.5 |
| HV01 | CTRR | 8.5 |
| HV03 | CTRR | 9.0 |
| HV01 | THDC | 7.0 |
| HV02 | THDC | 5.0 |
| HV03 | THDC | 7.5 |
| HV03 | CSDL | 6.0 |

| MONHOC | |
|--------|-------------------|
| Mamh | Tenmh |
| CSDL | Có sở dữ liệu |
| CTRR | Cấu trúc rời rạc |
| THDC | Tin học đại cương |



| Mahv |
|------|
| HV01 |
| HV03 |

KETQUA = MONHOC

$KETQUA \leftarrow KETQUATHI [Mahv, Mamh]$

$MONHOC \leftarrow MONHOC [Mamh]$

3.8 Hàm tính toán và gom nhóm (1)

- ◆ Hàm tính toán gồm các hàm: avg(giatri), min(giatri), max(giatri), sum(giatri), count(giatri).
- ◆ Phép toán gom nhóm:

$$\mathfrak{S}_{G_1, G_2, \dots, G_n}^{F_1(A_1), F_2(A_2), \dots, F_n(A_n)}(E)$$

- E là biểu thức đại số quan hệ
- G_i là thuộc tính gom nhóm (rỗng, nếu không gom nhóm)
- F_i là hàm tính toán
- A_i là tên thuộc tính

3.8 Hàm tính toán và gom nhóm (2)

- ♦ Điểm thi cao nhất, thấp nhất, trung bình của môn CSDL ?

$$\mathfrak{I}_{\max(Diem), \min(Diem), avg(Diem)}^{\sigma_{Mamh = 'CSDL'}} (KETQUATHI)$$

- ♦ Điểm thi cao nhất, thấp nhất, trung bình của từng môn ?

$$Mamh \mathfrak{I}_{\max(Diem), \min(Diem), avg(Diem)} (KETQUATHI)$$

Bài tập

Lược đồ CSDL quản lý bán hàng gồm có các quan hệ sau:

KHACHHANG (MAKH, HOTEN, DCHI, SODT, NGSINH,
DOANHISO, NGDK)

NHANVIEN (MANV,HOTEN, NGVL, SODT)

SANPHAM (MASP,TENSP, DVT, NUOCSX, GIA)

HOADON (SOHD, NGHD, MAKH, MANV, TRIGIA)

CTHD (SOHD,MASP,SL)

Mô tả các câu truy vấn sau bằng ĐSQH

1. In ra danh sách các sản phẩm (MASP,TENSP) do “Trung Quốc” sản xuất có giá từ 30.000 đến 40.000
2. In ra danh sách các khách hàng (MAKH, HOTEN) đã mua hàng trong ngày 1/1/2007.
3. In ra danh sách các sản phẩm (MASP,TENSP) do “Trung Quoc” sản xuất hoặc các sản phẩm được bán ra trong ngày 1/1/2007.
4. Tìm các số hóa đơn mua cùng lúc 2 sản phẩm có mã số “BB01” và “BB02”.
5. In ra danh sách các sản phẩm (MASP,TENSP) do “Trung Quoc” sản xuất không bán được trong năm 2006.
6. Tìm số hóa đơn đã mua tất cả các sản phẩm do Singapore sản xuất

Câu 1

- ◆ In ra danh sách các sản phẩm (MASP, TENS²P) do “Trung Quốc” sản xuất có giá từ 30.000 đến 40.000.

SANPHAM : ((*nuocsx* = 'TrungQuoc') ^ (30 .000 ≤ *gia* ≤ 40 .000)) [*masp* , *tensp*]

$\pi_{masp, tensp} \sigma_{(nuocsx = 'TrungQuoc') \wedge (30.000 \leq gia \leq 40.000)} SANPHAM$

Câu 2

- ♦ In ra danh sách các khách hàng (MAKH, HOTEN) đã mua hàng trong ngày 1/1/2007.

$$\left(KHACHHANG \xrightarrow{MAKH} HOADON : (nght = \#1 / 1 / 2007 \#) [makh, hoten] \right)$$

$$\pi_{masp, hoten} \sigma_{(nght = \#1 / 1 / 2007 \#)} (HOADON \xrightarrow{MAKH} KHACHHANG)$$

Câu 3

- ◆ In ra danh sách các sản phẩm do “Trung Quoc” sản xuất hoặc các sản phẩm được bán ra trong ngày 1/1/2007.

$$A \leftarrow SANPHAM : (nuocsx = 'TrungQuoc')[masp, tensp]$$

$$B \leftarrow (SANPHAM \bowtie_{MASP} CTHD \bowtie_{SOHD} HOADON : (nghd = \#1 / 1 / 2007 \#))[masp, tensp]$$

$$C \leftarrow A \cup B$$

Hoặc

$$A \leftarrow \pi_{masp, tensp} \sigma_{nuocsx = 'TrungQuoc'}(SANPHAM)$$

$$B \leftarrow \pi_{masp, tensp} ((\sigma_{nghd = \#1 / 1 / 2007 \#} (HOADON \bowtie_{SOHD} CTHD \bowtie_{MASP} SANPHAM))$$

$$C \leftarrow A \cup B$$

Câu 4

- ♦ Tìm các số hóa đơn đã mua cùng lúc các sản phẩm có mã số “BB01” và “BB02”.

$$A \leftarrow CTHD : (masp = 'BB\ 01')[sohd]$$

$$B \leftarrow CTHD : (masp = 'BB\ 02')[sohd]$$

$$C \leftarrow A \cap B$$

Hoặc

$$A \leftarrow \pi_{sohd} \sigma_{masp = 'BB\ 01'}(CTHD)$$

$$B \leftarrow \pi_{sohd} \sigma_{masp = 'BB\ 02'}(CTHD)$$

$$C \leftarrow A \cap B$$

Câu 5

- ◆ In ra danh sách các sản phẩm do “TrungQuoc” sản xuất không bán được trong năm 2006.

$$A \leftarrow \pi_{\text{masp}, \text{tensp}} \sigma_{\text{nuocsx} = \text{'TrungQuoc'}} (\text{SANPHAM})$$

$$B \leftarrow ((\text{SANPHAM} \bowtie_{\text{MASP}} \text{CTHD}) \bowtie_{\text{SOHD}} \text{HOADON})$$

$$C \leftarrow \pi_{\text{masp}, \text{tensp}} \sigma_{(\text{nuocsx} = \text{'TrungQuoc'}) \wedge (\text{year}(\text{nghd}) = 2006)} (B)$$

$$D \leftarrow (A - C)$$

Câu 6

- ♦ Tìm số hóa đơn đã mua tất cả các sản phẩm do Singapore sản xuất

$$A \leftarrow \pi_{\text{masp}} \sigma_{\text{nuocsx} = \text{'Singapore'}, (\text{SANPHAM} \quad)}$$

$$B \leftarrow \pi_{\text{masp}, \text{sohd}} \sigma_{\text{nuocsx} = \text{'Singapore'}, (\text{SANPHAM} \quad \triangleright \triangleleft \text{CTHD} \quad)}$$

$$C \leftarrow B \div A$$

Bài 5: Ngôn ngữ SQL

Nội dung

1. Giới thiệu
2. Các ngôn ngữ giao tiếp
3. Ngôn ngữ định nghĩa dữ liệu
4. Ngôn ngữ thao tác dữ liệu
5. Ngôn ngữ truy vấn dữ liệu có cấu trúc
6. Ngôn ngữ điều khiển dữ liệu

1. Giới thiệu

- ♦ Là ngôn ngữ chuẩn để truy vấn và thao tác trên CSDL quan hệ
- ♦ Là ngôn ngữ phi thủ tục
- ♦ Khởi nguồn của SQL là SEQUEL - *Structured English Query Language*, năm 1974)
- ♦ *Các chuẩn SQL*
 - SQL89
 - SQL92 (SQL2)
 - SQL99 (SQL3)

2. Các ngôn ngữ giao tiếp

- ♦ **Ngôn ngữ định nghĩa dữ liệu** (*Data Definition Language - DDL*): cho phép khai báo cấu trúc bảng, các mối quan hệ và các ràng buộc.
- ♦ **Ngôn ngữ thao tác dữ liệu** (*Data Manipulation Language - DML*)
liệu.
- ♦ **Ngôn ngữ truy vấn dữ liệu** (*Structured Query Language – SQL*): cho phép truy vấn dữ liệu.
- ♦ **Ngôn ngữ điều khiển dữ liệu** (*Data Control Language – DCL*)
u.

3. Ngôn ngữ định nghĩa dữ liệu

3.1 Tạo bảng (CREATE)

- 3.1.1 Cú pháp

- 3.1.2 Một số kiểu dữ liệu

3.2 Sửa cấu trúc bảng (ALTER)

- 3.2.1 Thêm thuộc tính

- 3.2.2 Sửa kiểu dữ liệu của thuộc tính

- 3.2.3 Xóa thuộc tính

- 3.2.4 Thêm ràng buộc toàn vẹn

- 3.2.5 Xóa ràng buộc toàn vẹn

3.3 Lệnh xóa bảng (DROP)

3.1 Lệnh tạo bảng

3.1.1 Cú pháp

CREATE TABLE <tên_bảng>

(

<tên_cột1> <kiểu_dữ_liệu> [not null],

<tên_cột2> <kiểu_dữ_liệu> [not null],

...

<tên_cột1> <kiểu_dữ_liệu> [not null],

khai báo khóa chính, khóa ngoại, ràng buộc

)

3.1 Lệnh tạo bảng (2)

3.1.2 Một số kiểu dữ liệu

| Kiểu dữ liệu | SQL Server |
|--------------|--|
| Chuỗi ký tự | varchar(n), char(n), nvarchar(n), nchar(n) |
| Số | tinyint, smallint, int, numeric(m,n), decimal(m,n), float, real, smallmoney, money |
| Ngày tháng | smalldatetime, datetime |
| Luận lý | bit |

3.1 Lệnh tạo bảng (3)

Lược đồ CSDL quản lý bán hàng gồm có các quan hệ sau:

KHACHHANG (MAKH, HOTEN, DCHI, SODT, NGSINH,
DOANHISO, NGDK, CMND)

NHANVIEN (MANV,HOTEN, NGVL, SODT)

SANPHAM (MASP,TENSP, DVT, NUOCSX, GIA)

HOADON (SOHD, NGHD, MAKH, MANV, TRIGIA)

CTHD (SOHD,MASP,SL)

3.1 Lệnh tạo bảng (4)

Create table KHACHHANG

```
(  
    MAKH                char(4) primary key,  
    HOTEN                varchar(40),  
    DCHI                 varchar(50),  
    SODT                 varchar(20),  
    NGSINH               smalldatetime,  
    DOANHISO             money,  
    NGDK                 smalldatetime,  
    CMND                 varchar(10)  
)
```

3.1 Lệnh tạo bảng (5)

Create table CTHD

```
(  
    SOHD      int foreign key  
                references HOADON(SOHD),  
    MASP      char(4) foreign key  
                references SANPHAM(MASP),  
    SL        int,  
    constraint PK_CTHD primary key (SOHD, MASP)  
)
```

3.2 Sửa cấu trúc bảng(1)

3.2.1 Thêm thuộc tính

ALTER TABLE tên_bảng ADD tên_cột kiểu_dữ_liệu

- Ví dụ thêm Ghi_chu vào bảng khách hàng
ALTER TABLE KHACHHANG ADD GHI_CHU varchar(20)

3.2.2 Sửa kiểu dữ liệu thuộc tính

***ALTER TABLE tên_bảng ALTER COLUMN tên_cột
kiểu_dữ_liệu_mới***

- ♦ Lưu ý:

Không phải sửa bất kỳ kiểu dữ liệu nào cũng được

3.2 Sửa cấu trúc bảng(2)

- **Ví dụ:** Sửa Cột Ghi_chu thành kiểu dữ liệu varchar(50)
ALTER TABLE KHACHHANG ALTER COLUMN GHI_CHU varchar(50)
- Nếu sửa kiểu dữ liệu của cột Ghi_chu thành varchar(5), mà trước đó đã nhập giá trị cho cột Ghi_chu có độ dài hơn 5 ký tự thì không được phép.
- Hoặc sửa từ kiểu chuỗi ký tự sang kiểu số, ...

3.2.3 Xóa thuộc tính

ALTER TABLE tên_bảng DROP COLUMN tên_cột

- Ví dụ: xóa cột Ghi_chu trong bảng KHACHHANG
ALTER TABLE NHANVIEN DROP COLUMN Ghi_chu

3.2 Sửa cấu trúc bảng(3)

3.2.4 Thêm ràng buộc toàn vẹn

**ALTER TABLE <tên_bảng>
ADD CONSTRAINT
<tên_ràng_buộc>**

UNIQUE tên_cột

PRIMARY KEY (tên_cột)

FOREIGN KEY (tên_cột)
REFERENCES tên_bảng
(cột_là_khóa_chính) [**ON
DELETE CASCADE**] [**ON
UPDATE CASCADE**]

CHECK (tên_cột điều_kiện)

3.2 Sửa cấu trúc bảng(4)

♦ Ví dụ

- ALTER TABLE NHANVIEN ADD CONSTRAINT PK_NV PRIMARY KEY (MANV)
- ALTER TABLE CTHD ADD CONSTRAINT FK_CT_SP FOREIGN KEY (MASP) REFERENCES SANPHAM(MASP)
- ALTER TABLE SANPHAM ADD CONSTRAINT CK_GIA CHECK (GIA >=500)
- ALTER TABLE KHACHHANG ADD CONSTRAINT UQ_KH UNIQUE (CMND)

3.2 Sửa cấu trúc bảng(5)

3.2.5 Xóa ràng buộc toàn vẹn

*ALTER TABLE tên_bảng DROP CONSTRAINT
tên_ràng_buộc*

■ Ví dụ:

- Alter table CTHD drop constraint FK_CT_SP
 - Alter table SANPHAM drop constraint ck_gia
- ♦ Lưu ý: đối với ràng buộc khóa chính, muốn xóa ràng buộc này phải xóa hết các ràng buộc khóa ngoại tham chiếu tới nó

3.3 Lệnh xóa bảng

- ♦ **Cú pháp**

DROP TABLE tên_bảng

- ♦ **Ví dụ:** xóa bảng KHACHHANG.

DROP TABLE KHACHHANG

- ♦ **Lưu ý:** khi muốn xóa một bảng phải xóa tất cả những khóa ngoại tham chiếu tới bảng đó trước.

4. Ngôn ngữ thao tác dữ liệu

- ♦ Nhóm các lệnh:

- 4.1 Lệnh thêm dữ liệu (INSERT)

- 4.2 Lệnh sửa dữ liệu (UPDATE)

- 4.3 Lệnh xóa dữ liệu (DELETE)

4.1 Thêm dữ liệu

- ◆ **Cú pháp**

INSERT INTO tên_bảng (cột1,...,cộtn) VALUES
(giá_trị_1,..., giá_trị_n)

INSERT INTO tên_bảng VALUES (giá_trị_1,
giá_trị_2,..., giá_trị_n)

- ◆ **Ví dụ:**

- insert into SANPHAM values('BC01','But chi', 'cay', 'Singapore', 3000)
- insert into SANPHAM(masp,tensp,dvt,nuocsx,gia) values ('BC01','But chi','cay','Singapore',3000)

4.2 Sửa dữ liệu

- ◆ **Cú pháp**

UPDATE tên_bảng
SET cột_1 = giá_trị_1, cột_2 = giá_trị_2
[WHERE điều_kiện]

- ◆ **Lưu ý:** cẩn thận với các lệnh xóa và sửa, nếu không có điều kiện ở WHERE nghĩa là xóa hoặc sửa tất cả.
- ◆ **Ví dụ:** Tăng giá 10% đối với những sản phẩm do “Trung Quoc” sản xuất

```
UPDATE SANPHAM  
SET Gia = Gia*1.1  
WHERE Nuocsx='Trung Quoc'
```

4.3 Xóa dữ liệu

- ♦ **Cú pháp**

DELETE FROM tên_bảng [WHERE điều_kiện]

- ♦ **Ví dụ:**

- Xóa toàn bộ nhân viên

DELETE FROM NHANVIEN

- Xóa những sản phẩm do Trung Quốc sản xuất có giá thấp hơn 10000

DELETE FROM SANPHAM

WHERE (Gia < 10000) and (Nuocsx='Trung Quoc')

5. Ngôn ngữ truy vấn dữ liệu có cấu trúc

- 5.1 Câu truy vấn tổng quát
- 5.2 Truy vấn đơn giản
- 5.3 Phép kết
- 5.4 Đặt bí danh, sử dụng *, distinct
- 5.5 Các toán tử
- 5.6 Câu truy vấn con (subquery)
- 5.7 Phép chia
- 5.8 Hàm tính toán, gom nhóm

5.1 Câu truy vấn tổng quát

SELECT [DISTINCT] *|tên_cột | hàm
FROM bảng
[WHERE điều_kiện]
[GROUP BY tên_cột]
[HAVING điều_kiện]
[ORDER BY tên_cột ASC | DESC]

5.2 Truy vấn đơn giản(1)

◆ SELECT

- Tương đương phép chiếu của ĐSQH
- Liệt kê các thuộc tính cần hiển thị trong kết quả

◆ WHERE

- Tương ứng với điều kiện chọn trong ĐSQH
- Điều kiện liên quan tới thuộc tính, sử dụng các phép nối luận lý AND, OR, NOT, các phép toán so sánh, BETWEEN

◆ FROM

- Liệt kê các quan hệ cần thiết, các phép kết

5.2 Truy vấn đơn giản(2)

- ♦ Tìm masp, tensp do “Trung Quoc” sản xuất có giá từ 20000 đến 30000

```
Select      masp,tensp
From        SANPHAM
Where       nuocsx='Trung Quoc'
           and gia between 20000 and 30000
```

5.3 Phép kết(1)

- ◆ Inner Join, Left Join, Right Join, Full Join

- ◆ Ví dụ:

- In ra danh sách các khách hàng (MAKH, HOTEN) đã mua hàng trong ngày 1/1/2007.

```
select  KHACHHANG.makh,hoten
from    KHACHHANG inner join HOADON on
        KHACHHANG.makh=HOADON.makh
where   nghd='1/1/2007'
```

5.3 Phép kết (2)

- ◆ **Ví dụ:** In ra danh sách tất cả các hóa đơn và họ tên của khách hàng mua hóa đơn đó (nếu có)
 - Select sohd, hoten
From HOADON left join KHACHHANG on
 HOADON.makh=KHACHHANG.makh
 - Select sohd, hoten
From HOADON ,KHACHHANG
where HOADON.makh*=KHACHHANG.makh

5.4 Đặt bí danh, sử dụng *, distinct

- ◆ Đặt bí danh – Alias: cho thuộc tính và quan hệ:
tên_cũ AS tên_mới
 - Select manv,hoten as [ho va ten] From NHANVIEN
- ◆ Liệt kê tất cả các thuộc tính của quan hệ:
 - Select * from Nhanvien
 - Select NHANVIEN.* from NHANVIEN
- ◆ Distinct: trùng chỉ lấy một lần
 - Select distinct nuocsx from SANPHAM
- ◆ Sắp xếp kết quả hiển thị: Order by
 - Select * from SANPHAM order by nuocsx, gia DESC

5.5 Toán tử truy vấn(1)

- ♦ Toán tử so sánh: =, >, <, >=, <=, <>
- ♦ Toán tử logic: AND, OR, NOT
- ♦ Phép toán: +, -, *, /
- ♦ BETWEEN AND
- ♦ IS NULL, IS NOT NULL
- ♦ LIKE (_ %)
- ♦ IN, NOT IN
- ♦ EXISTS , NOT EXISTS
- ♦ SOME, ALL

5.5 Toán tử truy vấn(2)

♦ IS NULL, IS NOT NULL

- Select sohd from HOADON where makh is Null
- Select * from HOADON where makh is Not Null

♦ Toán tử so sánh, phép toán

- Select gia*1.1 as [gia ban] from SANPHAM where nuocsx<>'Viet Nam'
- Select * from SANPHAM where (gia between 20000 and 30000) OR (nuocsx='Viet Nam')

♦ Toán tử IN, NOT IN

- Select * from SANPHAM where masp NOT IN ('BB01','BB02','BB03')

5.5 Toán tử so sánh(3)

Toán tử LIKE

- So sánh chuỗi tương đối
- Cú pháp: s LIKE p, p có thể chứa % hoặc _
- % : thay thế một chuỗi ký tự bất kỳ
- _ : thay thế một ký tự bất kỳ
- Ví dụ: `Select masp,tensp from SANPHAM
where masp like 'B%01'`

5.6 Câu truy vấn con (1)

In hoặc Exists

- ♦ **Ví dụ:** Tìm các số hóa đơn mua cùng lúc 2 sản phẩm có mã số “BB01” và “BB02”.
 - select distinct sohd
from CTHD where masp='BB01' and sohd **IN**
(select distinct sohd from CTHD where masp='BB02')
 - select distinct A.sohd
from CTHD A where A.masp='BB01' and
EXISTS (select * from CTHD B
where B.masp='BB02' and A.sohd=B.sohd)

5.6 Câu truy vấn con (2)

Not In hoặc Not Exists

- ♦ **Ví dụ:** Tìm các số hóa đơn có mua sản phẩm mã số 'BB01' nhưng không mua sản phẩm mã số 'BB02'.
 - select distinct sohd
from CTHD where masp='BB01' and sohd **NOT IN**
(select distinct sohd from CTHD where masp='BB02')
 - select distinct A.sohd
from CTHD A where A.masp='BB01' and
NOT EXISTS (select * from CTHD B
where B.masp='BB02' and A.sohd=B.sohd)

5.7 Phép chia

Sử dụng NOT EXISTS

- ♦ **Ví dụ:** Tìm số hóa đơn đã mua tất cả những sản phẩm do “Trung Quoc” sản xuất.
- ♦ Select sohd from HOADON where not exists
(select * from SANPHAM
where nuocsx=‘Trung Quoc’ and not exists
(select * from CTHD where
HOADON.sohd=CTHD.sohd and
CTHD.masp=SANPHAM.masp))

5.8 Các hàm tính toán và gom nhóm (1)

5.8.1 Các hàm tính toán cơ bản

- COUNT: Đếm số bộ dữ liệu của thuộc tính
- MIN: Tính giá trị nhỏ nhất
- MAX: Tính giá trị lớn nhất
- AVG: Tính giá trị trung bình
- SUM: Tính tổng giá trị các bộ dữ liệu

NHANVIEN

| MANV | HOTEN | PHAI | MANQL | PHONG | LUONG |
|-------|---------------------|------|-------|-------|-----------|
| NV001 | Nguyễn Ngọc Linh | Nữ | Null | NC | 2.800.000 |
| NV002 | Đinh Bá Tiến | Nam | NV002 | DH | 2.000.000 |
| NV003 | Nguyễn Văn Mạnh | Nam | NV001 | NC | 2.300.000 |
| NV004 | Trần Thanh Long | Nam | NV002 | DH | 1.800.000 |
| NV005 | Nguyễn Thị Hồng Vân | Nữ | NV001 | NC | 2.500.000 |
| NV006 | Nguyễn Minh | Nam | NV002 | DH | 2.000.000 |
| NV007 | Hà Duy Lập | Nam | NV003 | NC | 1.800.000 |
| NV008 | Trần Kim Duyên | Nữ | NV003 | NC | 1.800.000 |

Ví dụ

1. Tính lương thấp nhất, cao nhất, trung bình và tổng lương của tất cả các nhân viên.
2. Có tất cả bao nhiêu nhân viên
3. Bao nhiêu nhân viên có người quản lý
4. Bao nhiêu phòng ban có nhân viên trực thuộc
5. Tính lương trung bình của các nhân viên
6. Tính lương trung bình của các nhân viên theo từng phòng ban

1. Tính lương thấp nhất, cao nhất, trung bình và tổng lương của tất cả các nhân viên.

```
SELECT  min(luong) as thapnhat,  
        max(luong) as caonhat,  
        avg(luong) as trungbinh,  
        sum(luong) as tongluong  
  
FROM    NhanVien
```

2. Có tất cả bao nhiêu nhân viên

SELECT count(*) FROM NhanVien

3. Bao nhiêu nhân viên có người quản lý

- **Select count(*) FROM NhanVien WHERE manql is not null**
- **SELECT count(Manql) FROM NhanVien**

4. Bao nhiêu phòng ban có nhân viên trực thuộc

SELECT count(distinct phong) FROM NhanVien

5.8 Các hàm tính toán và gom nhóm (2)

5.8.2 Gom nhóm: mệnh đề GROUP BY

- ◆ Sử dụng hàm gom nhóm trên các bộ trong quan hệ.
- ◆ Mỗi nhóm bộ bao gồm tập hợp các bộ có cùng giá trị trên các thuộc tính gom nhóm
- ◆ Hàm gom nhóm áp dụng trên mỗi bộ độc lập nhau.
- ◆ SQL có mệnh đề GROUP BY để chỉ ra các thuộc tính gom nhóm, các thuộc tính này phải xuất hiện trong mệnh đề SELECT

5. Tính lương trung bình của các nhân viên

```
SELECT          avg(LUONG) as LUONGTB
FROM            NhanVien
```

6. Tính lương trung bình của các nhân viên theo từng phòng ban.

```
SELECT          phong, avg(LUONG) as LUONGTB
FROM            NhanVien
GROUP BY        phong
```

5.8 Các hàm tính toán và gom nhóm (3)

5.8.3 Điều kiện sau gom nhóm: mệnh đề HAVING

- Lọc kết quả theo điều kiện, sau khi đã gom nhóm
 - Điều kiện ở HAVING được thực hiện sau khi gom nhóm, các điều kiện có liên quan đến thuộc tính Group By
- ♦ **Ví dụ:** tìm phòng có số lượng nhân viên “Nữ” trên 5 người

```
SELECT      phong
FROM        NhanVien
WHERE       phai = 'Nữ'
GROUP BY    phong
HAVING      count(manv) > 5
```

Bài 6: Ngôn ngữ tân từ

Nội dung

1. Giới thiệu
2. Cú pháp
3. Các định nghĩa
4. Diễn giải của một công thức
5. Quy tắc lượng giá công thức
6. Ngôn ngữ tân từ có biến là n bộ
7. Ngôn ngữ tân từ có biến là miền giá trị

1. Giới thiệu

- ♦ Ngôn ngữ tân từ là ngôn ngữ truy vấn hình thức do Codd đề nghị (1972-1973) được Lacroit, Proix và Ullman phát triển, cài đặt trong một số ngôn ngữ như QBE, ALPHA..
- ♦ **Đặc điểm:**
 - Ngôn ngữ phi thủ tục
 - Rút trích cái gì chứ không phải rút trích như thế nào
 - Khả năng diễn đạt tương đương với đại số quan hệ
- ♦ **Có hai loại:**
 - Có biến là n bộ
 - Có biến là miền giá trị

2. Cú pháp

- ♦ $()$: biểu thức trong ngoặc
- ♦ **Biến**: dùng chữ thường ở cuối bộ ký tự: x, y, z, t, s, \dots
- ♦ **Hằng**: dùng chữ thường ở đầu bộ ký tự: a, b, c, \dots
- ♦ **Hàm**: là một ánh xạ từ một miền giá trị vào tập hợp gồm 2 giá trị: đúng hoặc sai. Thường dùng chữ thường ở giữa bộ ký tự: h, g, f, \dots
- ♦ **Tân từ**: là một biểu thức được xây dựng dựa trên biểu thức logic. Dùng chữ in hoa ở giữa bộ ký tự P, Q, R, \dots
- ♦ **Các phép toán logic**: phủ định (\neg) , kéo theo (\Rightarrow) , và (\wedge) , hoặc (\vee) .
- ♦ **Các lượng từ**: với mọi (\forall) , tồn tại (\exists)

3. Các định nghĩa (1)

♦ Định nghĩa 1: Tân từ 1 ngôi

- Tân từ 1 ngôi được định nghĩa trên tập X và biến x có giá trị chạy trên các phần tử của X .
- Với mỗi giá trị của x , tân từ $P(x)$ là một mệnh đề logic, tức là nó có giá trị đúng (Đ) hoặc sai (S)
- Ví dụ
 - $P(x)$, x là biến chạy trên X , là một tân từ
 - $P(gt)$, $gt \in X$ là một mệnh đề, $X = \{\text{Nguyen Van A, Tran Thi B}\}$
 - Với tân từ $NỮ(x)$ được xác định: “ x là người nữ”. Khi đó
 - Mệnh đề $NỮ(\text{Nguyen Van A})$: cho kết quả Sai
 - $NỮ(\text{Tran Thi B})$: cho kết quả Đúng

3. Các định nghĩa (2)

♦ Định nghĩa 2: Tân từ n ngôi

- Tân từ n ngôi được định nghĩa trên các tập X_1, X_2, \dots, X_n và n biến x_1, x_2, \dots, x_n lấy giá trị trên các tập X_i tương ứng.
- Với mỗi giá trị $a_i \in X_i$, $x_i = a_i$. Tân từ n ngôi là một mệnh đề.
- Ký hiệu: $P(x_1, x_2, \dots, x_n)$
- Ví dụ: $CHA(x_1, x_2)$: “ x_1 là CHA của x_2 ”
- Chú ý:
 - Các X_i không nhất thiết phải là rời nhau
 - Với $x_i = a_i$, $P(x_1, x_2, \dots, a_i, \dots, x_n)$ là tân từ n-1 ngôi

3. Các định nghĩa (3)

♦ Định nghĩa 3: Từ

- Từ là một hằng hay là một biến
- Nếu $f(t_1, t_2, \dots, t_n)$ là hàm n ngôi thì f là một từ

♦ Định nghĩa 4: Công thức

- Công thức nguyên tố: $P(t_1, t_2, \dots, t_n)$, t_i là các từ
- Nếu F_1, F_2 là các công thức thì các biểu thức sau cũng là các công thức: $F_1 \vee F_2$, $F_1 \wedge F_2$, $F_1 \Rightarrow F_2$, $\neg F_1$
- Nếu F_1 là một công thức thì $\forall x:F_1$, $\exists x:F_1$ cũng là các công thức
- Nếu F_1 là công thức thì (F_1) cũng là một công thức

3. Các định nghĩa (4)

♦ Định nghĩa 4:

- Công thức *đóng* là công thức nếu mọi biến đều có kèm với lượng từ. (khẳng định Đ, S)
- Công thức *mở* là công thức tồn tại 1 biến không kèm lượng từ. (tìm kiếm thông tin)

♦ Ví dụ:

- $C_1: \forall x \exists t \forall y (P(x, y, a) \Rightarrow \exists z (Q(y, z, t) \wedge R(x, t)))$ là công thức đóng vì các biến x, y, z, t đều có kèm lượng từ \forall, \exists
- $C_2: \forall x \exists t (P(x, y, a) \Rightarrow \exists z (Q(y, z, t) \wedge R(x, t)))$ là công thức mở vì biến y không có lượng từ \forall, \exists

4. Diễn giải của một công thức

Gồm 4 phần:

- ♦ Miền giá trị của các biến của công thức (ký hiệu là tập M)
- ♦ Sử dụng các hằng, các tân từ (ý nghĩa tân từ, xác định được quan hệ n ngôi)
- ♦ Ý nghĩa của công thức
- ♦ Xác định 1 quan hệ n ngôi trên tập M^n

5. Quy tắc lượng giá công thức

- ♦ Lượng giá tân từ: xét tân từ bậc n : $P(x_1, x_2, \dots, x_n)$ và liên kết với quan hệ R , n ngôi.

$$P(a_1, a_2, \dots, a_n): \text{Đ} \Leftrightarrow (a_1, a_2, \dots, a_n) \in R$$

$$P(a_1, a_2, \dots, a_n): \text{S} \Leftrightarrow (a_1, a_2, \dots, a_n) \notin R$$

- ♦ Các phép toán $\wedge, \vee, \neg, \Rightarrow$ dùng bảng chân trị
- ♦ Lượng từ \exists : gọi x là biến. Công thức $\exists x F(x)$ là đúng khi có ít nhất một $a_i \in M / F(a_i): \text{Đ}$

$$M = \{a_1, a_2, \dots, a_n\} \equiv \vee F(a_i), a_i \in M$$

- ♦ Lượng từ \forall : x là biến, $\forall x F(x): \text{Đ}$ với $\forall a_i \in M / F(a_i): \text{Đ}$

$$M = \{a_1, a_2, \dots, a_n\} \equiv \wedge F(a_i), a_i \in M$$

6. Ngôn ngữ tân từ có biến là n bộ

6.1 Qui tắc

6.2 Định nghĩa

6.3 Công thức an toàn

6.4 Biểu diễn các phép toán

6.1 Quy tắc (1)

1. Biến là 1 bộ của quan hệ
2. Từ: hằng, biến hoặc biểu thức có dạng $s[C]$, s : biến, C : tập các thuộc tính của quan hệ được gọi là từ chiếu.
3. Công thức:
 - R_s (R là quan hệ, s là biến) được gọi là từ. Miền giá trị sẽ định nghĩa miền biến thiên của s .
 - $t_1 \theta a$, $t_1 \theta t_2$ ở đây t_1, t_2 là các từ chiếu, còn a là một hằng, θ là toán tử so sánh được gọi là công thức nguyên tố

6.1 Quy tắc (2)

4. Một *công thức nguyên tố* là một công thức
5. F_1 và F_2 là công thức: $F_1 \vee F_2$, $F_1 \wedge F_2$, $F_1 \Rightarrow F_2$, $\neg F_1$ là công thức
6. F là công thức, s là biến $\exists sF$, $\forall sF$ là công thức
7. F là công thức, (F) là công thức

6.2 Định nghĩa

- ♦ Một câu hỏi trong ngôn ngữ tân từ có biến là n bộ được biểu diễn như sau: $\{s \mid F\}$. Trong đó s là biến n bộ, F là một công thức chỉ có một biến tự do là s .
- ♦ **Ví dụ:** BIENGIOI(nuoc,tinh_{tp}). Phép toán quan hệ BIENGIOI[nuoc] được chuyển thành câu hỏi trong ngôn ngữ tân từ có biến là bộ: $\{s[\text{nuoc}] \text{ BIENGIOI } s\}$

6.3 Công thức an toàn

F là công thức an toàn: nếu nó thoả mãn 3 điều kiện sau:

- i) Nếu s là bộ n thỏa: $F(s)$ là đúng thì mọi thành phần của s là phần tử của $DOM(F)$:

$$(F_s : \text{Đúng}) \rightarrow s \in DOM(F)$$

- ii) F' là công thức con của F :

$$\exists s F'_s, F'_s : \text{Đúng} \rightarrow s \in DOM(F')$$

- iii) $\forall s F'_s, F'_s : \text{Đúng} \rightarrow s \notin DOM(F')$

6.4 Biểu diễn các phép toán (1)

♦ 1. Phép hội

- Q_1, Q_2 là các quan hệ n chiều
- F_1, F_2 là các công thức ứng với Q_1, Q_2
- Công thức của $Q = Q_1 \cup Q_2$
- $F_s = F_{1s} \vee F_{2s}$

♦ 2. Phép trừ

- Q_1, Q_2 là các quan hệ n chiều
- F_1, F_2 là các công thức ứng với Q_1, Q_2
- Công thức của $Q = Q_1 - Q_2$
- $F_s = F_1 \wedge \neg F_{2s}$

6.4 Biểu diễn các phép toán (2)

♦ 3. Phép tích

- $Q_1(x_1, \dots, x_m), Q_2(y_1, \dots, y_n)$
- F_1, F_2 là các công thức ứng với Q_1, Q_2
- Công thức của $Q = Q_1 \times Q_2$

$$F_s: s(x_1, \dots, x_m, y_1, \dots, y_n)$$

$$F_s = (\exists v) (\exists p) (F_{1v} \wedge F_{2p} \wedge \\ s_1 = v_1 \wedge \dots \wedge s_m = v_m \wedge s_{m+1} = p_1 \wedge \dots \wedge s_{m+n} = p_n)$$

6.4 Biểu diễn các phép toán (3)

♦ 4. Phép chiếu

- $Q_1(x_1, \dots, x_n)$, F_1 là các công thức ứng với Q_1
- Công thức của $Q = Q_1[x_{i1}, x_{i2}, \dots, x_{ik}]$

$$F_s = (\exists v) (F_{1v} \wedge s_1 = v_{i1} \wedge s_2 = v_{i2} \wedge \dots \wedge s_k = v_{ik})$$

♦ 5. Phép chọn

- Q_1 là quan hệ n chiều, F_1 là công thức ứng với Q_1
- Công thức $Q = Q_1$: điều kiện ĐK ($\text{ĐK}: x_i \theta x_j$ hoặc $x_i \theta a$)

$$F_s = F_{1s} \wedge s_i \theta s_j \text{ hoặc } F_{1s} \wedge s_i \theta a \quad (1 \leq i, j \leq n, i \neq j)$$

7. Ngôn ngữ tân từ có biến là miền giá trị

7.1 Quy tắc

7.2 Biểu diễn câu hỏi

7.3 Công thức an toàn

7.4 Biểu diễn các phép toán

7.1 Quy tắc

1. Từ: là hằng hoặc biến
2. Công thức nguyên tố
 - $Q(t_1, t_2, \dots, t_n)$: t_i là các từ, Q là quan hệ
 - $t_i \theta t_j, t_i \theta a$ với t_i là từ, a là một hằng, θ là phép toán
3. Một *công thức nguyên tố* là một công thức
4. F_1 và F_2 là công thức: $F_1 \vee F_2, F_1 \wedge F_2, F_1 \Rightarrow F_2, \neg F_1$ là công thức
5. F là công thức, t : biến tự do, $\exists sF, \forall sF$ cũng công thức
6. F là công thức, (F) là công thức

7.2 Biểu diễn câu hỏi

$$\{(x_1, x_2, \dots, x_n) \mid F(x_1, x_2, \dots, x_n)\}$$

- ♦ x_i là các biến tự do của F
- ♦ $Q = \{(x_1, x_2, \dots, x_n) \mid F(x_1, x_2, \dots, x_n)\}$ nên
 $(x_1, x_2, \dots, x_n) \in Q \Rightarrow F(x_1, x_2, \dots, x_n)$:Đúng

7.3 Công thức an toàn

F là công thức an toàn: nếu nó thoả mãn 3 điều kiện sau:

i) Nếu s là bộ n thỏa: F(s) là đúng thì mọi thành phần của s là phần tử của DOM(F):

$$(F(x_1, \dots, x_n) : \text{Đúng}) \rightarrow x_i \in \text{DOM}(F), i = 1, \dots, n$$

ii) F' là công thức con của F:

$$\exists x F' : \text{Đúng} \rightarrow x \in \text{DOM}(F')$$

iii) $\forall x F' : \text{Đúng} \rightarrow \exists x \notin \text{DOM}(F')$

$$(F(x_1, \dots, x_n) : \text{Đúng}) \rightarrow \exists x_i \notin \text{DOM}(F), i = 1, \dots, n$$

7.4 Biểu diễn các phép toán (1)

♦ 1. Phép hội

- Q_1, Q_2 là các quan hệ n chiều
- F_1, F_2 là các công thức ứng với Q_1, Q_2
- Công thức của $Q = Q_1 \cup Q_2$
- $F = F_1 \vee F_2$

♦ 2. Phép trừ

- Q_1, Q_2 là các quan hệ n chiều
- F_1, F_2 là các công thức ứng với Q_1, Q_2
- Công thức của $Q = Q_1 - Q_2$
- $F = F_1 \wedge \neg F_2$

7.4 Biểu diễn các phép toán (2)

♦ 3. Phép tích

- $Q_1(x_1, \dots, x_m), Q_2(y_1, \dots, y_n)$
- F_1, F_2 là các công thức ứng với Q_1, Q_2
- Công thức của $Q = Q_1 \times Q_2$

$$F(x_1, \dots, x_m, y_1, \dots, y_n) = F_1(x_1, \dots, x_m) \wedge F_2(y_1, \dots, y_n)$$

7.4 Biểu diễn các phép toán (3)

♦ 4. Phép chiếu

- $Q_1(x_1, \dots, x_n), F_1(x_1, \dots, x_n)$ là các công thức ứng với Q_1
- Công thức của $Q = Q_1[x_{i1}, x_{i2}, \dots, x_{ik}]$
 $F_s(x_{i1}, x_{i2}, \dots, x_{ik}) = (\exists x_{j1})(\exists x_{j2}) \dots (\exists x_{jn-k})(F_1(x_1, \dots, x_n))$
trong đó $(x_{i1}, x_{i2}, \dots, x_{ik}) \cup (x_{j1}, x_{j2}, \dots, x_{jn-k}) = (x_1, x_2, \dots, x_n)$

♦ 5. Phép chọn

- $Q_1(x_1, \dots, x_n), F_1(x_1, \dots, x_n)$ là các công thức ứng với Q_1
- Công thức $Q = Q_1$: điều kiện ĐK $(\text{ĐK}: x_i \theta x_j \text{ hoặc } x_i \theta a)$
 $F_1(x_1, \dots, x_n) = F_1(x_1, \dots, x_n) \wedge x_i \theta x_j \text{ hoặc}$
 $= F_1(x_1, \dots, x_n) \wedge x_i \theta a$

Bài 7: Ràng buộc toàn vẹn

Nội dung chính

1. Giới thiệu ràng buộc toàn vẹn (RBTV)
2. Các đặc trưng của một RBTV
3. Phân loại RBTV
4. Bảng tầm ảnh hưởng tổng hợp

1. Giới thiệu

- ♦ Ràng buộc toàn vẹn là các quy định, điều kiện từ ứng dụng thực tế, các điều kiện này là bất biến.

⇒ Vì thế phải luôn đảm bảo cơ sở dữ liệu thoả ràng buộc toàn vẹn sau mỗi thao tác làm thay đổi tình trạng của cơ sở dữ liệu.

2. Các đặc trưng của một RBTV

2.1 Nội dung

2.2 Bối cảnh

2.3 Bảng tầm ảnh hưởng

2.1 Nội dung

- ◆ Mô tả chặt chẽ ý nghĩa của ràng buộc toàn vẹn.
- ◆ Nội dung được phát biểu bằng ngôn ngữ tự nhiên hoặc bằng ngôn ngữ hình thức (ngôn ngữ tân từ, đại số quan hệ, mã giả,...)
 - Ngôn ngữ tự nhiên: dễ hiểu nhưng không chặt chẽ, logic.
 - Ngôn ngữ hình thức: chặt chẽ, cô đọng

2.2 Bối cảnh

- ♦ Là tập các quan hệ khi thao tác trên những quan hệ đó có khả năng làm cho ràng buộc bị vi phạm.
- ♦ Đó là những quan hệ có thể vi phạm ràng buộc toàn vẹn khi thực hiện các thao tác thêm, xoá, sửa.

2.3 Bảng tầm ảnh hưởng (1)

- ◆ Nhằm xác định khi nào tiến hành kiểm tra ràng buộc toàn vẹn. Thao tác nào thực hiện có thể làm vi phạm ràng buộc toàn vẹn.
- ◆ Phạm vi ảnh hưởng của một ràng buộc toàn vẹn được biểu diễn bằng một bảng 2 chiều gọi là *bảng tầm ảnh hưởng*.

2.3 Bảng tầm ảnh hưởng (2)

Một số quy định

- ◆ Những thuộc tính khoá (những thuộc tính nằm trong khoá chính của quan hệ) không được phép sửa giá trị
- ◆ Thao tác thêm và xoá xét trên một bộ của quan hệ. Thao tác sửa xét sửa từng thuộc tính trên bộ của quan hệ
- ◆ Trước khi xét thao tác thực hiện có thể làm vi phạm ràng buộc hay không thì CSDL phải thoả ràng buộc toàn vẹn trước.

2.3 Bảng tầm ảnh hưởng (3)

♦ Bảng tầm ảnh hưởng của một ràng buộc

| Ràng buộc Ri | Thêm | Xóa | Sửa |
|-----------------|------|-----|-----|
| Quan hệ 1 | | | |
| | | | |
| Quan hệ n | | | |

- + : thực hiện thao tác có thể làm vi phạm RBTV
- : thực hiện thao tác không thể làm vi phạm RBTV
- +(A) : có thể làm vi phạm RBTV khi sửa trên thuộc tính A
- (*) : không vi phạm RBTV do thao tác không thực hiện được

3. Phân loại

- 3.1 RBTV có bối cảnh trên 1 quan hệ
- 3.2 RBTV có bối cảnh trên nhiều quan hệ
- 3.3 Phụ thuộc hàm (functional dependency)

3.1 RBTV có bối cảnh 1 quan hệ

3.1.1 RBTV miền giá trị.

3.1.2 RBTV liên thuộc tính

3.1.3 RBTV liên bộ

Lược đồ CSDL quản lý giáo vụ

HOCVIEN (MAHV, HO, TEN, NGSINH, GIOITINH, NOISINH, MALOP)

LOP (MALOP, TENLOP, TRGLOP, SISO, MAGVCN)

KHOA (MAKHOA, TENKHOA, NGTLAP, TRGKHOA)

MONHOC (MAMH, TENMH, TCLT, TCTH, MAKHOA)

DIEUKIEN (MAMH, MAMH_TRUOC)

GIAOVIEN(MAGV,HOTEN,HOCVI,HOCHAM,GIOITINH,NGSINH,NGVL,
HESO, MUCLUONG, MAKHOA)

GIANGDAY(MALOP,MAMH,MAGV,HOCKY, NAM,TUNGAY,DENNGAY)

KETQUATHI (MAHV, MAMH, LANTHI, NGTHI, DIEM, KQUA)

3.1.1 Ràng buộc miền giá trị

- ♦ Là tập giá trị mà một thuộc tính có thể nhận.
- ♦ **R1:** Giới tính của học viên chỉ là Nam hoặc Nữ
 - Nội dung:
 $\forall hv \in HOCVIEN: hv.Gioitinh \in \{ 'Nam', 'Nữ' \}$
 - Bối cảnh: quan hệ HOCVIEN
 - Bảng tầm ảnh hưởng:

| R1 | Thêm | Xóa | Sửa |
|---------|------|-----|-------------|
| HOCVIEN | + | - | +(Gioitinh) |

3.1.2 Ràng buộc liên thuộc tính

- ♦ Là ràng buộc giữa các thuộc tính với nhau trên 1 bộ của quan hệ
- ♦ **R2:** Ngày bắt đầu (TUNGAY) giảng dạy một môn học cho một lớp luôn nhỏ hơn ngày kết thúc (DENNGAY)
 - Nội dung:
 $\forall gd \in GIANGDAY: gd.TUNGAY < gd.DENNGAY$
 - Bối cảnh : GIANGDAY
 - Bảng tầm ảnh hưởng:

| R2 | Thêm | Xóa | Sửa |
|----------|------|-----|--------------------|
| GIANGDAY | + | - | +(Tungay, Denngay) |

3.1.3 Ràng buộc liên bộ (1)

- ♦ Là ràng buộc giữa các bộ trên cùng một quan hệ (có thể liên quan đến nhiều thuộc tính).
- ♦ **R3:** Tất cả các học viên phải có mã số phân biệt với nhau
 - Nội dung:
 $\forall h_1, h_2 \in \text{HOCVIEN}$: Nếu $h_1 \neq h_2$ thì $h_1.\text{Mahv} \neq h_2.\text{Mahv}$
 - Bối cảnh: quan hệ HOCVIEN
 - Bảng tầm ảnh hưởng:

| R3 | Thêm | Xóa | Sửa |
|---------|------|-----|------|
| HOCVIEN | + | - | -(*) |

3.1.3 Ràng buộc liên bộ (2)

- ♦ **R4:** Các giáo viên có cùng học vị, cùng hệ số lương thì mức lương sẽ bằng nhau
 - Nội dung:
 $\forall gv_1, gv_2 \in GIAOVIEN:$
Nếu $(gv_1.Hocvi = gv_2.Hocvi) \wedge (gv_1.Heso = gv_2.Heso)$ thì
 $gv.Mucluong = gv.Mucluong$
 - Bối cảnh: quan hệ GIAOVIEN
 - Bảng tầm ảnh hưởng:

| R4 | Thêm | Xóa | Sửa |
|----------|------|-----|--------------------------|
| GIAOVIEN | + | - | +(Hocvi, Heso, Mucluong) |

3.2 RBTV có bối cảnh nhiều quan hệ

3.2.1 RBTV tham chiếu (khoá ngoại, phụ thuộc tồn tại)

3.2.2 RBTV liên thuộc tính

3.2.3 RBTV do thuộc tính tổng hợp

3.2.4 RBTV do chu trình trong lược đồ biểu diễn quan hệ

3.2.1 Ràng buộc tham chiếu (1)

- ♦ Là ràng buộc quy định giá trị thuộc tính trong một bộ của quan hệ R (tập thuộc tính này gọi là khoá ngoại), phải phụ thuộc vào sự tồn tại của một bộ trong quan hệ S (tập thuộc tính này là khoá chính trong quan hệ S).
- ♦ RBTV tham chiếu còn gọi là ràng buộc phụ thuộc tồn tại hay ràng buộc khóa ngoại

3.2.1 Ràng buộc tham chiếu (2)

- ♦ **R5:** Học viên thi một môn học nào đó thì môn học đó phải có trong danh sách các môn học

- Nội dung:

- $\forall k \in \text{KETQUATHI}, \exists m \in \text{MONHOC}: k.\text{Mamh} = m.\text{Mamh}$
- Hoặc: $\text{KETQUATHI}[\text{Mamh}] \subseteq \text{MONHOC}[\text{Mamh}]$

- Bối cảnh: quan hệ KETQUATHI, MONHOC

- Bảng tầm ảnh hưởng:

| R5 | Thêm | Xóa | Sửa |
|-----------|------|-----|------|
| KETQUATHI | + | - | -(*) |
| MONHOC | - | + | -(*) |

3.2.2 Ràng buộc liên thuộc tính (1)

- ♦ Là ràng buộc giữa các thuộc tính trên những quan hệ khác nhau
- ♦ **R6:** Ngày giáo viên giảng dạy một môn học phải lớn hơn hoặc bằng ngày giáo viên đó vào làm.
 - Nội dung: $\forall gd \in GIANGDAY$
Nếu $\exists gv \in GIAOVIEN: gd.Magv = gv.Magv$ thì
 $gv.NGVl \leq gd.TUNGAY$
 - Bối cảnh: GIANGDAY, GIAOVIEN
 - Bảng tầm ảnh hưởng:

| R6 | Thêm | Xóa | Sửa |
|----------|------|-----|-----------|
| GIANGDAY | + | - | +(Tungay) |
| GIAOVIEN | - | - | +(Ngvl) |

3.2.2 Ràng buộc liên thuộc tính (2)

- ♦ **R7:** Ngày thi một môn học phải lớn hơn ngày kết thúc học môn học đó.
 - Nội dung:
 $\forall kq \in KETQUATHI$
Nếu $\exists gd \in GIANGDAY, \exists hv \in HOCVIEN$:
 $(gd.Malop = hv.Malop) \wedge (kq.Mamh = gd.Mamh)$ thì
 $gd.Denngay < kq.Ngthi$
 - Bối cảnh: GIANGDAY, HOCVIEN, KETQUATHI

3.2.2 Ràng buộc liên thuộc tính (3)

- Bảng tầm ảnh hưởng:

| R7 | Thêm | Xóa | Sửa |
|-----------|------|-----|------------|
| HOCVIEN | - | - | +(Malop) |
| GIANGDAY | - | - | +(Denngay) |
| KETQUATHI | + | - | +(Ngthi) |

3.2.3 RBTV do thuộc tính tổng hợp (1)

- ♦ Là ràng buộc giữa các thuộc tính, các bộ trên những quan hệ khác nhau.
- ♦ Thuộc tính tổng hợp là thuộc tính được tính toán từ giá trị của các thuộc tính khác, các bộ khác.
- ♦ **Ví dụ :** SANPHAM(Masp, Tensp, Nuocsx, Gia)
 KHACHHANG(Makh, Hoten, Doanhso)
 HOADON(Sohd, Nghd, Makh, Trigia)
 CTHD(Sohd, Masp, Soluong, Gia)
 - Trị giá của một hoá đơn bằng tổng thành tiền của các chi tiết thuộc hoá đơn đó

3.2.3 RBTV do thuộc tính tổng hợp (2)

- ◆ Doanh số của một khách hàng bằng tổng trị giá các hoá đơn mà khách hàng đó đã mua

- Nội dung:

$\forall kh \in KHACHHANG,$

$$kh.Doanhso = \sum_{(hd \in HOADON: hd.Makh=kh.Makh)} (hd.Trigia)$$

- Bối cảnh: KHACHHANG, HOADON

- Bảng tầm ảnh hưởng:

| | Thêm | Xóa | Sửa |
|-----------|------|-----|------------|
| KHACHHANG | + | - | +(Doanhso) |
| HOADON | + | + | +(Makh) |

3.2.3 RBTV do thuộc tính tổng hợp (3)

- ♦ **R8:** Sĩ số của một lớp là số lượng học viên thuộc lớp đó

- Nội dung:

$\forall l \in \text{LOP},$

$$l.\text{Siso} = \text{Count}_{(hv \in \text{HOCVIEN}: hv.\text{Malop} = l.\text{Malop})}^{(*)}$$

- Bối cảnh: quan hệ LOP, HOCVIEN

- Bảng tầm ảnh hưởng:

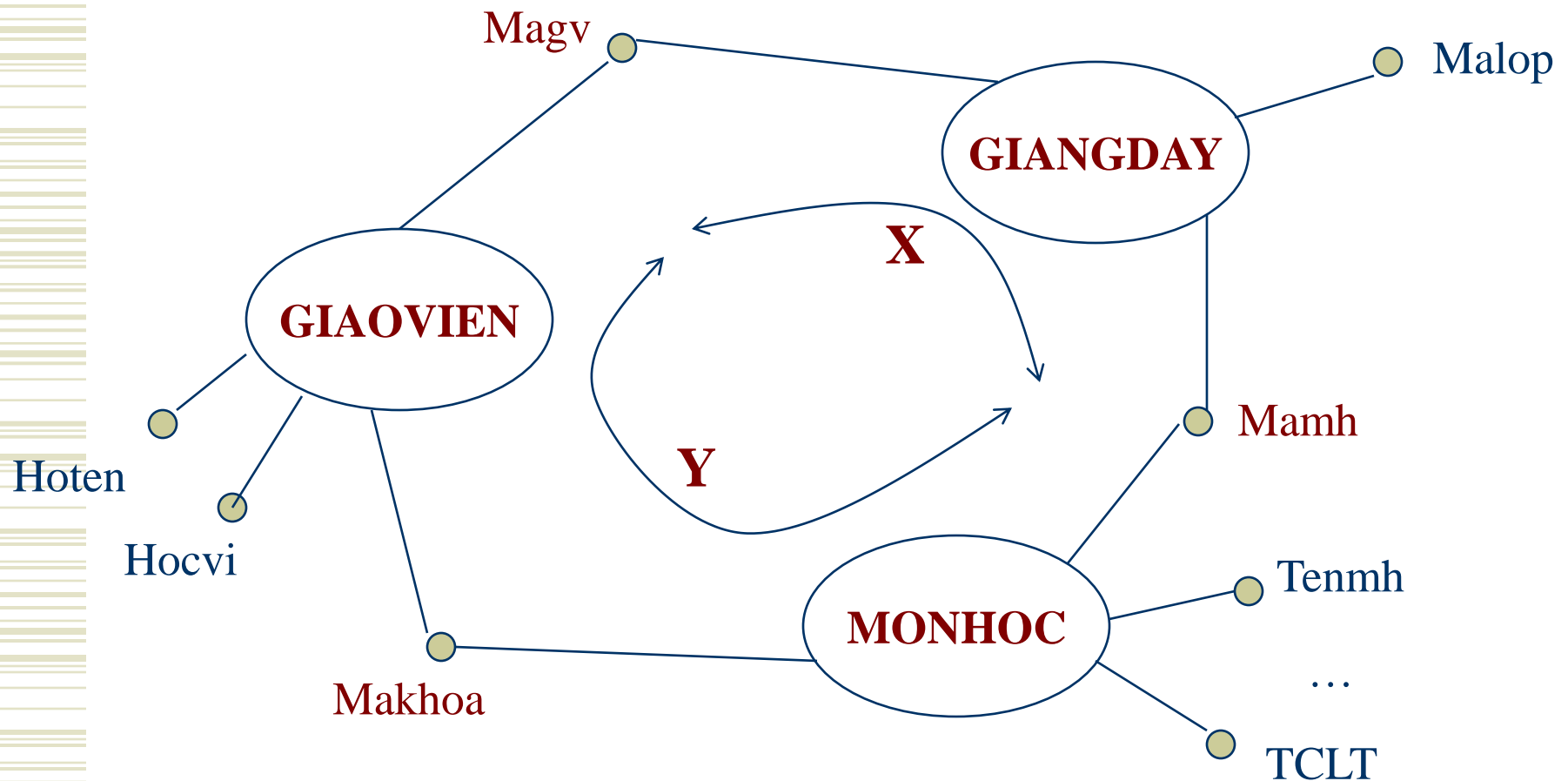
| R8 | Thêm | Xóa | Sửa |
|---------|------|-----|----------|
| LOP | + | - | +(Siso) |
| HOCVIEN | + | + | +(Malop) |

3.2.4 Do hiện diện của chu trình (1)

Biểu diễn lược đồ quan hệ dưới dạng đồ thị:

- Quan hệ được biểu diễn bằng nút tròn rỗng to
- Thuộc tính được biểu diễn bằng nút tròn đặc nhỏ
- Tất cả các nút đều được chỉ rõ bằng tên của quan hệ hoặc thuộc tính. Thuộc tính thuộc một quan hệ được biểu diễn bởi một cung nối giữa nút tròn to và nút tròn nhỏ
- Nếu đồ thị biểu diễn xuất hiện một đường khép kín
=> lược đồ CSDL có sự hiện diện của chu trình.

3.2.4 Do hiện diện của chu trình (2)



3.2.4 Do hiện diện của chu trình (3)

- ◆ $X = \text{GIANGDAY} \underset{\text{Makhoa}}{Y} [\text{Magv}, \text{Mamh}]$
- ◆ $Y = (\text{GIAOVIEN} \bowtie \text{MONHOC}) [\text{Magv}, \text{Mamh}]$
- ◆ **Ý nghĩa:**
 - X: giáo viên và những môn học đã được phân công cho giáo viên đó giảng dạy
 - Y: giáo viên và những môn học thuộc khoa giáo viên đó phụ trách
- ◆ Mỗi quan hệ giữa X và Y trong các ràng buộc sau:

3.2.4 Do hiện diện của chu trình (4)

- ♦ **Ràng buộc 1:** giáo viên chỉ được phân công giảng dạy những môn thuộc khoa giáo viên đó phụ trách $X \subseteq Y$
- ♦ **Ràng buộc 2:** giáo viên phải được phân công giảng dạy tất cả những môn thuộc khoa giáo viên đó phụ trách $X = Y$
- ♦ **Ràng buộc 3:** có thể phân công giáo viên giảng dạy bất kỳ môn học nào $X \neq Y$

3.2.4 Do hiện diện của chu trình (4)

- ♦ **R9**: giáo viên chỉ được phân công giảng dạy những môn thuộc khoa giáo viên đó phụ trách $X \subseteq Y$

| R9 | Thêm | Xóa | Sửa |
|----------|------|-----|-----------|
| MONHOC | - | - | +(Makhoa) |
| GIAOVIEN | - | - | +(Makhoa) |
| GIANGDAY | + | - | +(Magv) |

3.3 Phụ thuộc hàm (1)

- ♦ Cho quan hệ $Q(A, B, C)$. Phụ thuộc hàm A xác định B . Ký hiệu $A \rightarrow B$ nếu:

$$\forall q_1, q_2 \in Q: \text{Nếu } q_1.A = q_2.A \text{ thì } q_1.B = q_2.B$$

- ♦ $A \rightarrow B$ được gọi là *phụ thuộc hàm hiển nhiên* nếu $B \subseteq A$
- ♦ $A \rightarrow B$ được gọi là *phụ thuộc hàm nguyên tố* nếu $\neg \exists A' \subset A, A' \neq \emptyset$ sao cho $A' \rightarrow B$

3.3 Phụ thuộc hàm (2)

- ♦ Mỗi quan hệ đều có ít nhất một phụ thuộc hàm
- ♦ Ràng buộc khoá cũng là một phụ thuộc hàm
 $Mamh \rightarrow Tenmh, Tcvt, Tcth, Makhoa$
- ♦ **R4:** Các giáo viên có cùng học vị, cùng hệ số lương thì mức lương sẽ bằng nhau. Ràng buộc này có thể biểu diễn bằng phụ thuộc hàm như sau:

$Hocvi, Heso \rightarrow Mucluong$

4. Bảng tầm ảnh hưởng tổng hợp (1)

- ♦ **Bảng tầm ảnh hưởng tổng hợp** của m ràng buộc trên n quan hệ bối cảnh

| | QH₁ | | | QH₂ | | | ... | QH_n | | |
|------------|-----------------------|---|---|-----------------------|---|---|------------|-----------------------|---|---|
| | T | X | S | T | X | S | ... | T | X | S |
| R1 | | | | | | | | | | |
| R2 | | | | | | | | | | |
| ... | | | | | | | | | | |
| Rm | | | | | | | | | | |

4. Bảng tầm ảnh hưởng tổng hợp (2)

| | HOCVIEN | | | GIAOVIEN | | | LOP | | | MONHOC | | | GIANGDAY | | | KETQUA THI | | |
|----|---------|---|----|----------|---|---|-----|---|---|--------|---|----|----------|---|---|------------|---|----|
| | T | X | S | T | X | S | T | X | S | T | X | S | T | X | S | T | X | S |
| R1 | + | - | + | | | | | | | | | | | | | | | |
| R2 | | | | | | | | | | | | | + | - | + | | | |
| R3 | + | - | _* | | | | | | | | | | | | | | | |
| R4 | | | | + | - | + | | | | | | | | | | | | |
| R5 | | | | | | | | | | - | + | _* | | | | + | - | _* |
| R6 | | | | - | - | + | | | | | | | + | - | + | | | |
| R7 | - | - | + | | | | | | | | | | - | - | + | + | - | + |
| R8 | + | + | + | | | | + | - | + | | | | | | | | | |
| R9 | | | | - | - | + | | | | - | - | + | + | - | + | | | |

Bài 8: Tối ưu hóa câu hỏi

Nội dung

1. Giới thiệu
2. Các nguyên tắc tổng quát để tối ưu hóa câu hỏi
 - 2.1 Biểu thức tương đương
 - 2.1.1 Định nghĩa
 - 2.1.2 Tính chất của phép kết và phép tích
 - 2.2 Nguyên tắc tổng quát
 - 2.3 Các phép biến đổi tương đương
3. Một số kỹ thuật tối ưu hóa câu hỏi bằng ĐSQH
 - 3.1 Kỹ thuật (dãy phép chọn, phép chiếu, hoán vị ...)
 - 3.2 Thuật giải tối ưu hoá câu hỏi trong .

1. Giới thiệu (1)

◆ Mục đích:

- Giảm thời gian xử lý câu hỏi, giảm khối lượng dữ liệu trung gian.
- Kết hợp giữa các phép tích, phép kết với phép chọn với phép chiếu.

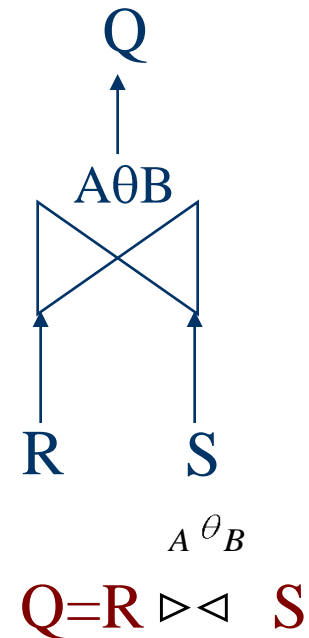
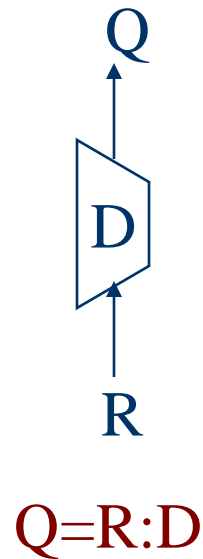
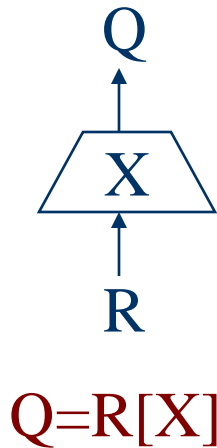
◆ Ví dụ:

$$+ ((Q_1 \triangleright \triangleleft Q_2) : A = a_0)[C]$$

$$+ ((Q_1 : A = a_0) \triangleright \triangleleft Q_2)[C]$$

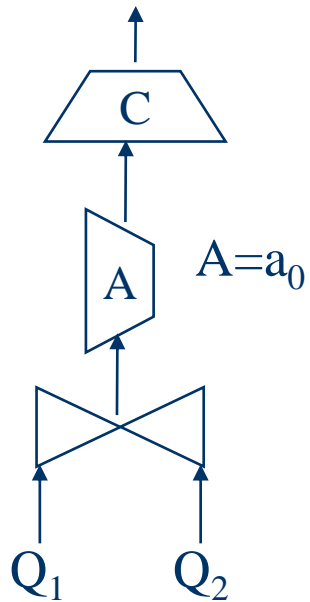
1. Giới thiệu (2)

◆ Ký hiệu:

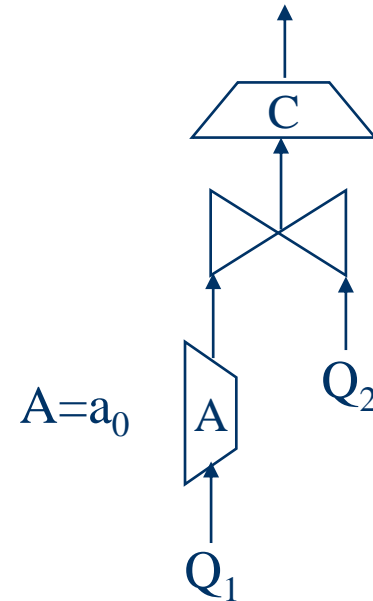


1. Giới thiệu (3)

♦ Ví dụ



$$((Q_1 \triangleright\triangleleft Q_2) : A = a_0)[C]$$



$$((Q_1 : A = a_0) \triangleright\triangleleft Q_2)[C]$$

2.1 Tính tương đương (1)

- ♦ **2.1.1 Định nghĩa:** hai biểu thức A, B là tương đương nếu có cùng một tình trạng CSDL thì đều cho một kết quả.
- ♦ **2.1.2 Tính chất của phép kết và phép tích**
 - **Phép kết**
 - Giao hoán $Q_1 \times Q_2 = Q_2 \times Q_1$
 - Kết hợp $Q_1 \times (Q_2 \times Q_3) = (Q_1 \times Q_2) \times Q_3$
 - **Phép tích**
 - Giao hoán: $Q_1 \begin{smallmatrix} dk \\ \triangleright \triangleleft \end{smallmatrix} Q_2 = Q_2 \begin{smallmatrix} dk \\ \triangleright \triangleleft \end{smallmatrix} Q_1$
 - Kết hợp: $Q_1 \begin{smallmatrix} dk\ 1 \\ \triangleright \triangleleft \end{smallmatrix} (Q_2 \begin{smallmatrix} dk\ 2 \\ \triangleright \triangleleft \end{smallmatrix} Q_3) = (Q_1 \begin{smallmatrix} dk\ 1 \\ \triangleright \triangleleft \end{smallmatrix} Q_2) \begin{smallmatrix} dk\ 2 \\ \triangleright \triangleleft \end{smallmatrix} Q_3$

2.1 Tính tương đương (2)

2.1.3 Các phép biến đổi tương đương

$$1. \quad Q_1(A, B) \triangleright \triangleleft^B Q_2(B, C) \equiv (Q_1 \times Q_2 : Q_1[B] = Q_2[B])$$

$$2. \quad Q_1(A, B) \triangleright \triangleleft^{B \theta_D} Q_2(C, D) \equiv (Q_1 \times Q_2 : B \theta_D)$$

$$3. \quad Q_1 \cap Q_2 \equiv \neg ((\neg Q_1) \cup (\neg Q_2))$$

$$4. \quad \neg Q(X_1, \dots, X_n) \equiv (Q[X_1] \times Q[X_2] \times \dots \times Q[X_n]) \neg Q(X_1, \dots, X_n)$$

$$5. \quad Q_1(A, B) \cap Q_2(A, B) \equiv Q_1[B] \neg ((Q_1[B] \times Q_2[A] \neg Q_1(A, B))[B])$$

2.2 Nguyên tắc tổng quát

1. Thực hiện phép chiếu, phép chọn càng sớm càng tốt
2. Gom các phép chọn và chiếu cùng quan hệ để thực hiện cùng lúc
3. Biến phép tích thành phép kết tự nhiên hay theta kết
4. Tìm các biểu thức con chung trong một biểu thức
5. Tiền xử lý các quan hệ: lập chỉ mục
6. Đánh giá trước khi thực hiện tính toán

3.1 Các kỹ thuật tối ưu (1)

1. Dãy các phép chọn
2. Dãy các phép chiếu
3. Hoán vị giữa phép chiếu và phép chọn
4. Hoán vị giữa phép chọn và phép tích
5. Hoán vị giữa phép hợp và phép chọn
6. Hoán vị giữa phép chọn và phép trừ
7. Hoán vị giữa phép chiếu và phép hội
8. Hoán vị giữa phép chiếu và phép tích

3.1 Các kỹ thuật tối ưu (2)

1. Dãy các phép chọn

$$(((Q : dk\ 1) : dk\ 2) \dots : dkn) \equiv Q : dk\ 1 \wedge dk\ 2 \wedge \dots dkn$$

2. Dãy phép chiếu

$$(Q[Y])[Z] \equiv Q[Z], \quad Z \subseteq Y$$

Ví dụ: Cho $Q(A, B, C, D)$

$$(Q[A, C, D])[AD] \equiv Q[AD]$$

3.1 Các kỹ thuật tối ưu (3)

3. Hoán vị giữa phép chiếu và phép chọn

- Nếu $X \subseteq Y$

$$(Q : dk(X))[Y] \equiv (Q[Y] : dk(X))$$

- Nếu $X \not\subseteq Y$

$$(Q : dk(X))[Y] \equiv (Q[X \cup Y] : dk(X))$$

3.1 Các kỹ thuật tối ưu (4)

4. Hoán vị giữa phép chọn và phép tích:

- Điều kiện dk xác lập trên các thuộc tính của X

$$(Q_1(X) : dk(X) \times Q_2(Y) \equiv (Q_1(X) \times Q_2(Y)) : dk$$

- Nếu $dk = dk_1 \wedge dk_2$, dk1 xác lập trên các thuộc tính của X, dk2 xác lập trên các thuộc tính của Y.

$$((Q_1(X) \times Q_2(Y)) : dk_1(X) \wedge dk_2(Y) \equiv ((Q_1(X) : dk_1) \times (Q_2(Y) : dk_2))$$

- Nếu dk1 xác lập trên các thuộc tính của X và dk2 xác lập trên các thuộc tính của $X \cup Y$

$$((Q_1(X) \times Q_2(Y)) : dk_1(X) \wedge dk_2(X \cup Y) \equiv$$

$$((Q_1(X) : dk_1) \times (Q_2(Y) : dk_2(X \cup Y))$$

3.1 Các kỹ thuật tối ưu (5)

5. Hoán vị giữa phép hội và phép chọn

$$(Q_1 \cup Q_2) : dk \equiv (Q_1 : dk) \cup (Q_2 : dk)$$

6. Hoán vị giữa phép chọn và phép trừ

$$(Q_1 - Q_2) : dk \equiv (Q_1 : dk) - (Q_2 : dk)$$

7. Hoán vị giữa phép chiếu và phép hội

$$(Q_1 \cup Q_2)[Z] \equiv (Q_1[Z]) \cup (Q_2[Z])$$

8. Hoán vị giữa phép chiếu và phép tích

$$(Q_1(X) \times Q_2(Y))[Z] \equiv (Q_1[Y \cap Z]) \times (Q_2[Y \cap Z]), \quad Z \in X \cup Y$$

3.2 Thuật toán

- ◆ **Bước 1:** Áp dụng các phép biến đổi tương đương
- ◆ **Bước 2:** Áp dụng (1)
- ◆ **Bước 3:** Đối với các phép chọn áp dụng (3), (4), (5), (6) nhằm đưa phép chọn càng sâu càng tốt
- ◆ **Bước 4:** Đối với các phép chiếu áp dụng (2), (3), (7), (8) nhằm đưa phép chiếu càng sâu càng tốt
- ◆ **Bước 5:**
 - Tập trung các phép chọn để áp dụng (1)
 - Kết hợp phép tích và phép chọn để chuyển thành phép kết