

MỤC LỤC

<u>MUC LUC.....</u>	<u>1</u>
<u>CHƯƠNG TRÌNH MÔN HỌC CƠ SỞ DỮ LIỆU.....</u>	<u>2</u>
<u>Chương 1. TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU.....</u>	<u>4</u>
<u>Chương 2. MÔ HÌNH DỮ LIỆU QUAN HỆ.....</u>	<u>20</u>
<u>Chương 3. NGÔN NGỮ TRUY VẤN DỮ LIỆU.....</u>	<u>35</u>
<u>Chương 4. RÀNG BUỘC TOÀN Vẹn.....</u>	<u>65</u>
<u>Chương 5. LÝ THUYẾT THIẾT KẾ CƠ SỞ DỮ LIỆU.....</u>	<u>78</u>

CHƯƠNG TRÌNH MÔN HỌC CƠ SỞ DỮ LIỆU

Mã môn học: MH16

I. Vị trí, ý nghĩa, vai trò môn học:

Môn học Cơ sở dữ liệu được bố trí sau khi học xong các môn Tin học đại cương, lập trình căn bản, toán ứng dụng.

Là môn học kỹ thuật cơ sở, thuộc các môn học, mô đun đào tạo nghề.

II. Mục tiêu của môn học:

- Trình bày được các khái niệm, các thuật ngữ trong cơ sở dữ liệu;
- Vận dụng được các mô hình dữ liệu của cơ sở dữ liệu quan hệ vào việc thiết kế cơ sở dữ liệu cho bài toán cụ thể;
- Sử dụng ngôn ngữ truy vấn dữ liệu quan hệ thành thạo;
- Mô tả được các dạng chuẩn và chuẩn hóa được bài toán cơ sở dữ liệu trước khi cài đặt;
- Rèn luyện cho sinh viên khả năng tự nghiên cứu tài liệu và tự giác trong làm việc nhóm.

III. Nội dung môn học:

Mã chương	Tên chương	Loại bài dạy	Địa điểm	Thời lượng			
				Tổng số	Lý thuyết	Thực hành	Kiểm tra
MH16-01	Tổng quan về cơ sở dữ liệu	Tích hợp	Lớp học	10	5	5	0
MH16-02	Mô hình dữ liệu quan hệ	Tích hợp	Lớp học	15	8	6	1
MH16-03	Ngôn ngữ truy vấn dữ liệu	Tích hợp	Lớp học	20	9	10	1
MH16-04	Ràng buộc toàn vẹn	Tích hợp	Lớp học	7	4	3	0
MH16-05	Lý thuyết thiết kế cơ sở dữ liệu	Tích hợp	Lớp học	38	19	17	2

IV. YÊU CẦU VỀ ĐÁNH GIÁ HOÀN THÀNH MÔN HỌC

Về kiến thức: Đánh giá thông qua bài kiểm tra viết, trắc nghiệm đạt được các yêu cầu sau:

- + Trình bày được các khái niệm về các thực thể, bộ, quan hệ, khóa, phụ thuộc hàm,..
- + Trình bày được các mô hình cơ sở dữ liệu quan hệ.
- + Trình bày được cú pháp của ngôn ngữ SQL
- + Trình bày được các dạng chuẩn của lược đồ quan hệ.

Về kỹ năng: Sau khi học xong môn học này học viên có khả năng:

- + Phân tích dữ liệu và vẽ được các mô hình dữ liệu thực thể - kết hợp (mô hình E-R); chuyển đổi E-R sang lược đồ quan hệ. Xác định được các khóa, chuẩn hóa được lược đồ ở mức tốt nhất có thể.
- + Sử dụng thành thạo ngôn ngữ truy vấn dữ liệu SQL chuẩn cho việc truy vấn dữ liệu đã cài đặt.

Về thái độ: Cẩn thận, tự tìm thêm tài liệu tham khảo, sưu tầm các mô hình cơ sở dữ liệu của các doanh nghiệp để tự học hỏi thêm.

Chương 1. TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU

Mã chương MH16-01

Giới thiệu:

.....

Mục tiêu:

Trình bày sơ lược các khái niệm về cơ sở dữ liệu, các mô hình dữ liệu;
Trình bày chi tiết mô hình thực thể kết hợp (ERD), có thể phân tích dữ liệu và thiết kế được mô hình thực thể kết hợp.

Thực hiện các thao tác an toàn với máy tính.

Nội dung:

1. Một số khái niệm cơ bản.

1.1. Định nghĩa cơ sở dữ liệu

Dữ liệu được lưu trữ trên các thiết bị lưu trữ theo một cấu trúc nào đó để phục vụ cho nhiều người dùng với nhiều mục đích khác nhau gọi là **cơ sở dữ liệu**.

Hình dung: Cơ sở dữ liệu như một bảng hai chiều

Chiều ngang: tập hợp các đặc điểm của một đối tượng cần quản lý gọi là bản ghi hay bộ.

Chiều dọc: gồm các điểm của một đối tượng quản lý gọi là trường.

1.2. Ưu điểm của cơ sở dữ liệu

- Giảm sự trùng lặp thông tin xuống mức thấp nhất và do đó bảo đảm được tính nhất quán và toàn vẹn dữ liệu.

- Đảm bảo dữ liệu có thể truy xuất theo nhiều cách khác nhau.

- Khả năng chia sẻ thông tin cho nhiều người sử dụng.

1.3. Các đặc trưng của phương pháp cơ sở dữ liệu

- Tính chia sẻ dữ liệu: Dữ liệu được chia sẻ bởi nhiều người dùng hợp pháp
- Tính giảm thiểu dư thừa dữ liệu: Dữ liệu dùng chung cho nhiều bộ phận được lưu một chỗ theo cấu trúc thống nhất.
- Tính tương thích: Việc loại bỏ dư thừa kéo theo hệ quả là sự tương thích.
- Tính toàn vẹn dữ liệu: Đảm bảo một số ràng buộc toàn vẹn. Khi người dùng chèn, xoá, sửa thì ràng buộc phải được kiểm tra chặt chẽ.
- Tính bảo mật dữ liệu: Đảm bảo an toàn dữ liệu và bảo mật thông tin là quan trọng.
- Tính đồng bộ dữ liệu: Thông thường cơ sở dữ liệu được nhiều người dùng truy cập đồng thời. Cần có cơ chế bảo vệ chống sự không tương thích (cả 2 cùng đặt chỗ ghế không trùng nhau)
- Tính độc lập dữ liệu: Sự tách biệt cấu trúc mô tả dữ liệu khỏi chương trình ứng dụng sử dụng dữ liệu gọi là độc lập dữ liệu. Điều này cho phép phát triển tổ chức dữ liệu mà không sửa đổi chương trình ứng dụng.

1.4. Các đối tượng sử dụng CSDL

Đối tượng sử dụng là người khai thác cơ sở dữ liệu thông qua hệ quản trị CSDL. Có thể phân làm 3 loại đối tượng: Người quản trị CSDL, người phát triển và lập trình ứng dụng, người dùng cuối cùng.

- Người quản trị CSDL: Là người hàng ngày chịu trách nhiệm quản lí và bảo trì CSDL như:
 - + Sự chính xác, toàn vẹn và bảo mật của dữ liệu và ứng dụng trong CSDL.
 - + Lưu trữ dự phòng và phục hồi CSDL.
 - + Giữ liên lạc với người phát triển và lập trình ứng dụng, người dùng cuối.
 - + Bảo đảm sự hoạt động hiệu quả của CSDL và hệ quản trị CSDL

- Người phát triển và lập trình ứng dụng: là những người chuyên nghiệp về lĩnh vực tin học có trách nhiệm thiết kế, tạo dựng và bảo trì thông tin cuối cùng cho người dùng.
- Người dùng cuối là những người không chuyên trong lĩnh vực tin học, họ là các chuyên gia trong các lĩnh vực khác có trách nhiệm cụ thể trong công việc. Họ khai thác CSDL thông qua chương trình (phần mềm ứng dụng) được xây dựng bởi người phát triển ứng dụng hay công cụ truy vấn của hệ quản trị CSDL.

1.5. Hệ Quản Trị Cơ Sở Dữ Liệu (Data Base Management System)

Để giải quyết tốt những vấn đề mà cách tổ chức CSDL đặt ra như đã nói ở trên, cần thiết phải có những phần mềm chuyên dùng để khai thác chúng. Những phần mềm này được gọi là các hệ quản trị CSDL. Các hệ quản trị CSDL có nhiệm vụ hỗ trợ cho các nhà phân tích thiết kế CSDL cũng như những người khai thác CSDL. Hiện nay trên thị trường phần mềm đã có những hệ quản trị CSDL hỗ trợ được nhiều tiện ích như: MS Access, Visual Foxpro, SQL Server Oracle, ...

Mỗi hệ quản trị CSDL đều được cài đặt dựa trên một mô hình dữ liệu cụ thể. Dù là dựa trên mô hình dữ liệu nào, một hệ quản trị CSDL cũng phải hội đủ các yếu tố sau:

- **Ngôn ngữ giao tiếp giữa người sử dụng và CSDL**, bao gồm: *Ngôn ngữ mô tả dữ liệu*: Để cho phép khai báo cấu trúc của CSDL, khai báo các mối liên hệ của dữ liệu và các quy tắc quản lý áp đặt lên các dữ liệu đó. *Ngôn ngữ thao tác dữ liệu*: Cho phép người sử dụng có thể cập nhật dữ liệu (thêm/sửa/xoá). *Ngôn ngữ truy vấn dữ liệu*: Cho phép người khai thác sử dụng để truy vấn các thông tin cần thiết trong CƠ SỞ DỮ LIỆU. *Ngôn ngữ quản lý dữ liệu*: Cho phép những người quản trị hệ thống thay đổi cấu trúc của các bảng dữ liệu, khai báo bảo mật thông tin và cấp quyền hạn

khai thác CSDL cho người sử dụng.....

-Từ điển dữ liệu: Dùng để mô tả các ánh xạ liên kết, ghi nhận các thành phần cấu trúc của CSDL, các chương trình ứng dụng, mật mã, quyền hạn sử dụng,...

-Cơ chế giải quyết vấn đề tranh chấp dữ liệu: Mỗi hệ quản trị CSDL cũng có thể cài đặt một cơ chế riêng để giải quyết các vấn đề này. Một số biện pháp sau đây thường được sử dụng: thứ nhất: cấp quyền ưu tiên cho từng người sử dụng; thứ hai: Đánh dấu yêu cầu truy xuất dữ liệu, phân chia thời gian, người nào có yêu cầu trước thì có quyền truy xuất dữ liệu trước,...

-Hệ quản trị CSDL cũng phải có cơ chế sao lưu (backup) và phục hồi (restore) dữ liệu khi có sự cố xảy ra. Điều này có thể thực hiện sau một thời gian nhất định hệ quản trị CSDL sẽ tự động tạo ra một bản sao CSDL, cách này hơi tốn kém, nhất là đối với CSDL lớn.

-Hệ quản trị CSDL phải cung cấp một giao diện thân thiện, dễ sử dụng.

1.6. Các Ứng Dụng Của Cơ Sở Dữ Liệu

Hiện nay, hầu như CSDL gắn liền với mọi ứng dụng của tin học; chẳng hạn như việc quản lý hệ thống thông tin trong các cơ quan nhà nước, việc lưu trữ và xử lý thông tin trong các doanh nghiệp, trong các lĩnh vực nghiên cứu khoa học, trong công tác giảng dạy, cũng như trong việc tổ chức thông tin đa phương tiện,...

2. Các mô hình dữ liệu

Mô hình dữ liệu là sự trừu tượng hoá môi trường thực. Mỗi loại mô hình dữ liệu đặc trưng cho một cách tiếp cận dữ liệu khác nhau của những nhà phân tích thiết kế CSDL. Mỗi loại mô hình dữ liệu đều có những ưu

điểm và những mặt hạn chế của nó, nhưng vẫn có những mô hình dữ liệu nổi trội và được nhiều người quan tâm nghiên cứu.

Sau đây chúng ta sẽ điếm qua lịch sử phát triển của các mô hình dữ liệu.

Vào những năm sáu mươi, thế hệ đầu tiên của CSDL ra đời dưới dạng mô hình thực thể kết hợp, mô hình mạng và mô hình phân cấp.

Vào những năm bảy mươi, thế hệ thứ hai của CSDL ra đời. Đó là mô hình dữ liệu quan hệ do EF. Codd phát minh. Mô hình này có cấu trúc logic chặt chẽ. Đây là mô hình đã và đang được sử dụng rộng khắp trong công tác quản lý trên phạm vi toàn cầu. Việc nghiên cứu mô hình dữ liệu quan hệ nhằm vào lý thuyết chuẩn hoá các quan hệ và là một công cụ quan trọng trong việc phân tích thiết kế các hệ CSDL hiện nay. Mục đích của nghiên cứu này nhằm bỏ đi các phần tử không bình thường của quan hệ khi thực hiện các phép cập nhật, loại bỏ các phần tử dư thừa.

Sang thập kỷ tám mươi, mô hình CSDL thứ ba ra đời, đó là mô hình cơ sở dữ liệu hướng đối tượng, mô hình cơ sở dữ liệu phân tán, mô hình cơ sở dữ liệu suy diễn,...

Trong phần tiếp theo sau đây, chúng tôi sẽ trình bày về mô hình dữ liệu tiêu biểu nhất để thiết kế (bước đầu) một ứng dụng tin học đó là mô hình thực thể kết hợp. Trong các chương còn lại của giáo trình này chúng tôi sẽ trình bày về mô hình dữ liệu quan hệ.

3. Mô hình thực thể kết hợp

Hiện nay mô hình dữ liệu quan hệ thường được dùng trong các hệ quản trị CSDL, đây là mô hình dữ liệu ở mức vật lý. Để thành lập được mô hình này, thường là phải dùng mô hình dữ liệu ở mức quan niệm để đặc tả, một trong những mô hình ở dạng đó là mô hình thực thể kết hợp (sau

đó mới dùng một số quy tắc để chuyển hệ thống từ mô hình này về mô hình dữ liệu quan hệ – các quy tắc này sẽ được nói đến trong mục 2.2).

Sau đây là các khái niệm của mô hình thực thể kết hợp.

3.1. Thực Thể (entity)

Thực thể là một sự vật tồn tại và phân biệt được, chẳng hạn sinh viên Nguyễn Văn Thành, lớp Cao Đẳng Tin Học 2A, môn học Cơ Sở Dữ Liệu, xe máy có biển số đăng ký 52-0549,... là các ví dụ về thực thể.

3.2. Thuộc tính (attribute)

Các đặc điểm riêng của thực thể gọi là các *thuộc tính*.

Chẳng hạn các thuộc tính của sinh viên Nguyễn Văn Thành là: mã số, sinh viên, giới tính, ngày sinh, hộ khẩu thường trú, lớp đang theo học, ... (Trong giáo trình này, tên thuộc tính được viết bằng chữ in hoa)

3.3. Loại thực thể (entity type)

Là tập hợp các thực thể có cùng thuộc tính. Mỗi loại thực thể đều phải được đặt tên sao cho có ý nghĩa. Một loại thực thể được biểu diễn bằng một hình chữ nhật.

Ví dụ các sinh viên có mã sinh viên là “02CDTH019”, “02CDTH519”, “02TCTH465”,... nhóm lại thành một loại thực thể, được đặt tên là Sinhvien chẳng hạn.

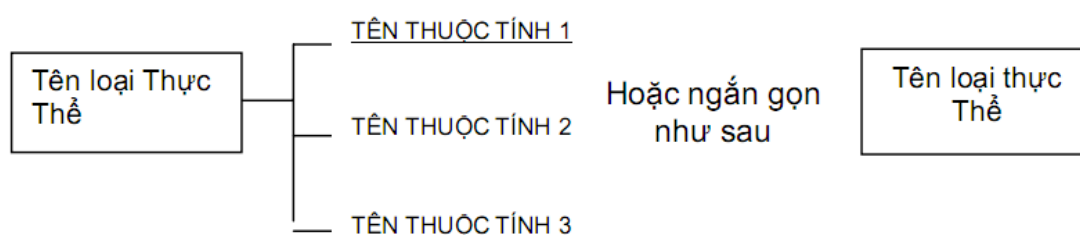
Tương tự trong ứng dụng quản lý điểm của sinh viên (sẽ được trình bày ngay sau đây) ta có các loại thực thể như Monhoc, Lop, Khoa,...(Trong giáo trình này, tên của loại thực thể được in hoa ký tự đầu tiên, các ký tự còn lại viết thường).

3.4. Khoá (key)

Khoá của loại thực thể E là một hay một tập các thuộc tính của E có thể dùng để phân biệt hai thực thể bất kỳ của E.

Ví dụ khoá của loại thực thể Sinhvien là MASV, của Lớp là MALOP, của Khoa là MAKHOA, của Monhoc là MAMH,...

Cần chú ý rằng khi biểu diễn một hệ thống bằng mô hình thực thể kết hợp thì tên của các loại thực thể phải khác nhau. Trong danh sách các thuộc tính của một loại thực thể thì tập thuộc tính khoá thường được gạch dưới liền nét. Nếu một hệ thống có nhiều loại thực thể, để đơn giản hoá mô hình, người ta có thể chỉ nêu tên các loại thực thể; còn các thuộc tính của loại thực thể được liệt kê riêng.



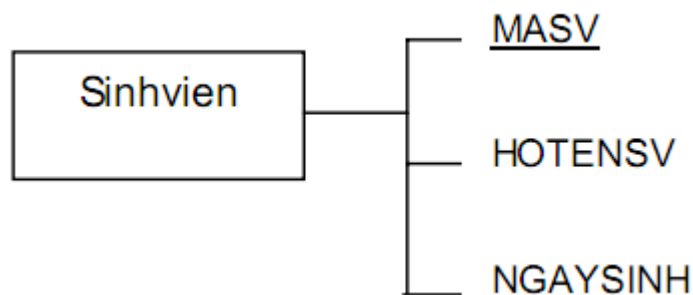
Ví dụ 1.1:

Bài toán quản lý điểm của sinh viên được phát biểu sơ bộ như sau: Mỗi sinh viên cần quản lý các thông tin như: họ và tên (HOTENSV), ngày tháng năm sinh(NGAYSINH), giới tính (NU), nơi sinh(NƠISINH), hộ khẩu thường trú (TINH). Mỗi sinh viên được cấp một mã số sinh viên duy nhất (MASV) để phân biệt với mọi sinh viên khác của trường, mỗi sinh viên chỉ thuộc về một lớp nào đó. Mỗi lớp học có một mã số lớp (MALOP) duy nhất để phân biệt với tất cả các lớp học khác trong trường: có một tên gọi (TENLOP) của lớp, mỗi lớp chỉ thuộc về một khoa. Mỗi khoa có một tên gọi (TENKHOA) và một mã số duy nhất (MAKHOA) phân biệt với các khoa khác. Mỗi môn học có một tên gọi (TENMH) cụ thể, được học trong một số đơn vị học trình (DONVIHT) và ứng với môn học là một mã số duy nhất (MAMH) để phân biệt với các môn học khác. Mỗi giảng viên cần quản lý các thông tin: họ và tên(HOTENGV), cấp học vị (HOCVI), thuộc một chuyên ngành (CHUYENNGANH) và được gán cho một mã số

duy nhất gọi là mã giảng viên(MAGV) để phân biệt với các giảng viên khác. Mỗi giảng viên có thể dạy nhiều môn ở nhiều khoa, nhưng chỉ thuộc về sự quản lý hành chính của một khoa. Mỗi sinh viên với một môn học được phép thi tối đa 3 lần, mỗi lần thi (LANTHI), điểm thi (DIEMTHI). Mỗi môn học ở mỗi lớp học chỉ phân công cho một giảng viên dạy (tất nhiên là một giảng viên thì có thể dạy nhiều môn ở một lớp).

Với bài toán trên thì các loại thực thể cần quản lý như: Sinhviên, Môn học, Khoa, Lớp, Giảngviên.

Ví dụ với loại thực thể Sinhviên thì cần quản lý các thuộc tính như: MASV, HOTENSV, NGAYSINH,... và ta có thể biểu diễn như sau:



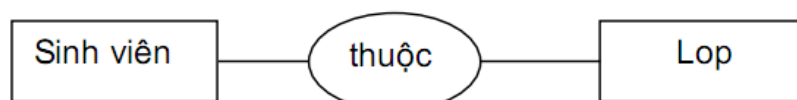
3.5. **Mối kết hợp (relationship)**

Mối kết hợp diễn tả sự liên hệ giữa các loại thực thể trong một Ứng dụng tin học.

Ví dụ mối kết hợp giữa hai loại thực thể Sinhviên và Lop, mối kết hợp giữa Sinhviên với Môn học,...

Mối kết hợp được biểu diễn bằng một hình elip và hai bên là hai nhánh gắn kết với các loại thực thể (hoặc mối kết hợp) liên quan, tên mối kết hợp thường là: *thuộc, gồm, chứa*,...

Chẳng hạn giữa hai loại thực thể Lớp và Khoa có mối kết hợp “thuộc” như sau:

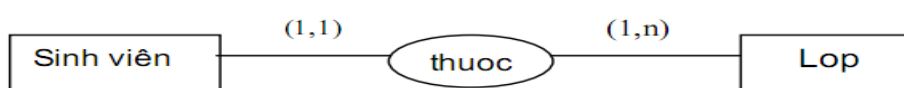


Bản số của mỗi kết hợp:

Bản số của một nhánh R trong mỗi kết hợp thể hiện số lượng các thực thể thuộc thực thể ở nhánh “bên kia” có liên hệ với một thực thể của nhánh R.

Mỗi bản số là một cặp số (min,max), chỉ số lượng tối thiểu và số lượng tối đa của thực thể khi tham gia vào mỗi kết hợp đó.

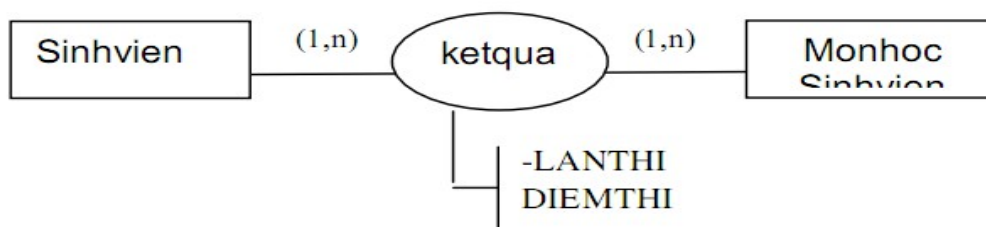
Ví dụ:



Có nghĩa là: “mỗi sinh viên thuộc một và chỉ một lớp nên bản số bên nhánh

Sinhviên là (1,1), mỗi lớp có 1 đến n sinh viên nên bản số bên nhánh Lop là (1,n)”. Trong một số trường hợp đặc biệt, mỗi kết hợp có thể có các thuộc tính đi kèm và do đó chúng thường được đặt tên ý với nghĩa đầy đủ hơn.

Ví dụ giữa hai loại thực thể Monhoc và Sinhvien có mỗi kết hợp ketqua với ý nghĩa: “mỗi sinh viên ứng với mỗi lần thi của mỗi môn học có một kết quả điểm thi duy nhất”.



Khoá

của mỗi kết hợp: là hợp của các khoá của các loại thực thể liên quan. Chẳng hạn như thuộc tính MAGV là khoá của loại thực thể Giangvien, MALOP là thuộc tính khoá của loại thực thể Lop, MAMH là thuộc tính khoá của loại

thực thể Monhoc, do đó mỗi kết hợp phancong (giữa các loại thực thể Giangvien,Lop,Monhoc) có khoá là {MAGV,MAMH,MALOP} - phancong là mỗi kết hợp 3 ngôi.

(Trong giáo trình này, tên của mỗi kết hợp được viết toàn bằng chữ thường). Việc thành lập mô hình thực thể kết hợp cho một ứng dụng tin học có thể tiến hành theo các bước sau:

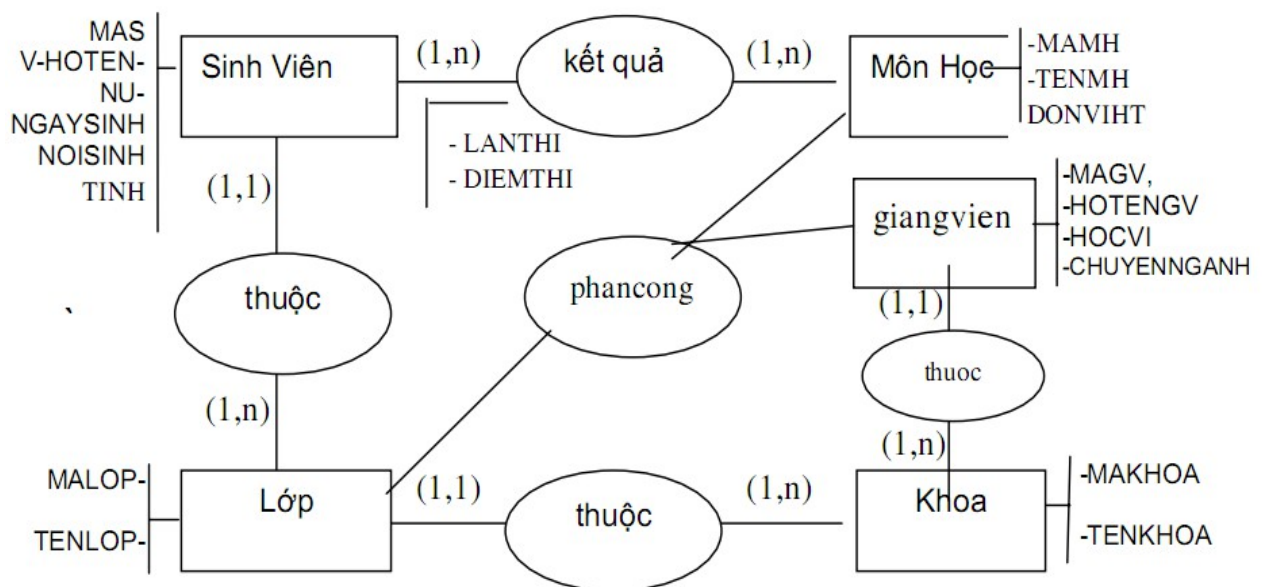
b1.Xác định danh sách các loại thực thể

b2.Xác định các mối kết hợp giữa các loại thực thể để phác thảo mô hình.

b3.Lập bản số của các mối kết hợp.

Để kết thúc chương này, chúng tôi sẽ lập mô hình thực thể kết hợp cho bài toán quản lý điểm của sinh viên đã được nêu trong ví dụ 1.1

Ví dụ 1.2:



BÀI TẬP

Dựa vào các phân tích sơ bộ dưới đây, hãy lập mô hình thực thể kết hợp

(gồm loại thực thể, mối kết hợp, bản số, thuộc tính của loại thực thể, khoá của loại thực thể) cho mỗi bài toán quản lý sau:

Bài 1. QUẢN LÝ SỐ LƯỢNG NGÀY CÔNG CỦA CÁC NHÂN VIÊN

Để quản lý việc phân công các nhân viên tham gia vào xây dựng các công trình. Công ty xây dựng ABC tổ chức quản lý như sau: Cùng lúc công ty có thể tham gia xây dựng nhiều công trình, mỗi công trình có một mã số công trình duy nhất (MACT), mỗi mã số công trình xác định các thông tin như: Tên gọi công trình (TENCT), địa điểm(ĐIADIEM), ngày công trình được cấp giấy phép xây dựng (NGAYCAPGP), ngày khởi công (NGAYKCC), ngày hoàn thành (NGAYHT). Mỗi nhân viên của công ty ABC có một mã số nhân viên(MANV) duy nhất, một mã số nhân viên xác định các thông tin như: Họ tên (HOTEN), ngày sinh(NGSINH), phái (PHAI), địa chỉ (ĐIACHI), phòng ban, ... Công ty phân công các nhân viên tham gia vào các công trình, mỗi công trình có thể được phân cho nhiều nhân viên và mỗi nhân viên cùng lúc cũng có thể tham gia vào nhiều công trình. Với mỗi công trình một nhân viên có một số lượng ngày công (SLNGAYCONG) đã tham gia vào công trình đó.

Công ty có nhiều phòng ban(Phòng kế toán, phòng kinh doanh, phòng kỹ thuật, phòng tổ chức, phòng chuyên môn, Phòng phục vụ,...). Mỗi phòng ban có một mã số phòng ban(MAPB) duy nhất, một phòng ban ứng với một tên phòng ban(TENPB).

Bài 2. QUẢN LÝ VIỆC MƯỢN/TRẢ SÁCH Ở MỘT THƯ VIỆN

Một thư viện tổ chức việc cho mượn sách như sau: Mỗi quyển sách được đánh một mã sách (MASH) dùng để phân biệt với các quyển sách khác (giả sử nếu một tác phẩm có nhiều bản giống nhau hoặc có nhiều tập thì cũng xem là có mã sách khác nhau), mỗi mã sách xác định các thông tin khác như : tên sách (TENSACH), tên tác giả (TACGIA), nhà xuất bản

(NHAXB), năm xuất bản (NAMXB). Mỗi độc giả được thư viện cấp cho một thẻ thư viện, trong đó có ghi rõ mã độc giả (MADG), cùng với các thông tin khác như: họ tên (HOTEN), ngày sinh (NGAYSINH), địa chỉ (DIACHI), nghề nghiệp(NGHENGHIEP).

Cứ mỗi lượt mượn sách, độc giả phải đăng ký các quyển sách cần mượn vào một phiếu mượn, mỗi phiếu mượn có một số phiếu mượn (SOPM) khác nhau, mỗi phiếu mượn xác định các thông tin như: ngày mượn sách (NGAYMUON), mã độc giả. Các các quyển sách trong cùng một phiếu mượn không nhất thiết phải trả trong một lần. Mỗi quyển sách có thể thuộc nhiều phiếu mượn khác nhau (tất nhiên là tại các thời điểm khác nhau).

Bài 3. QUẢN LÝ LỊCH DẠY CỦA GIÁO VIÊN

Để quản lý lịch dạy của các giáo viên và lịch học của các lớp, một trường tổ chức như sau: Mỗi giáo viên có một mã số giáo viên (MAGV) duy nhất, mỗi MAGV xác định các thông tin như: họ và tên giáo viên (HOTEN), số điện thoại (DTGV). Mỗi giáo viên có thể dạy nhiều môn cho nhiều khoa nhưng chỉ thuộc sự quản lý hành chính của một khoa nào đó. Mỗi môn học có một mã số môn học (MAMH) duy nhất, mỗi môn học xác định tên môn học(TENMH). Ứng với mỗi lớp thì mỗi môn học chỉ được phân cho một giáo viên. Mỗi phòng học có một số phòng học (PHONG) duy nhất, mỗi phòng có một chức năng (CHUCNANG); chẳng hạn như phòng lý thuyết, phòng thực hành máy tính, phòng nghe nhìn, xưởng thực tập cơ khí,... Mỗi khoa có một mã khoa (MAKHOA) duy nhất, mỗi khoa xác định các thông tin như: tên khoa (TENKHOA), điện thoại khoa(DTKHOA). Mỗi lớp có một mã lớp (MALOP) duy nhất, mỗi lớp có một tên lớp (TENLOP), sĩ số lớp (SISO). Mỗi lớp có thể học nhiều môn của nhiều khoa nhưng chỉ thuộc sự quản lý hành chính của một khoa nào đó. Hàng tuần, mỗi giáo viên phải lập lịch báo giảng cho biết giáo viên đó sẽ dạy những lớp nào, ngày nào (NGAYDAY), môn gì?, tại

phòng nào, từ tiết nào (TUTIET) đến tiết nào (ĐENTIET), tựa đề bài dạy (BAIDAY), những ghi chú (GHICHU) về các tiết dạy này, đây là giờ dạy lý thuyết (LYTHUYET) hay thực hành - giả sử nếu LYTHUYET=1 thì đó là giờ dạy thực hành và nếu LYTHUYET=2 thì đó là giờ lý thuyết, một ngày có 16 tiết, sáng từ tiết 1 đến tiết 6, chiều từ tiết 7 đến tiết 12, tối từ tiết 13 đến 16. Một số yêu cầu của hệ thống này như:: Lập lịch dạy trong tuần của các giáo viên. Tổng số dạy của các giáo viên theo từng môn cho từng lớp,

Bài 4. QUẢN LÝ HỌC VIÊN Ở MỘT TRUNG TÂM TIN HỌC

Trung tâm tin học KTCT thường xuyên mở các lớp tin học ngắn hạn và dài hạn. Mỗi lớp ngắn hạn có một hoặc nhiều môn học (chẳng hạn như lớp Tin học văn phòng thì có các môn : Word, Power Point, Excel, còn lớp lập trình Pascal thì chỉ học một môn Pascal). Các lớp dài hạn (chẳng hạn như lớp kỹ thuật viên đồ họa đa truyền thông, lớp kỹ thuật viên lập trình, lớp kỹ thuật viên phần cứng và mạng,...) thì có thể học nhiều học phần và mỗi học phần có thể có nhiều môn học. Mỗi học viên có một mã học viên(MAHV) duy nhất và chỉ thuộc về một lớp duy nhất (nếu học viên cùng lúc học nhiều lớp thì ứng với mỗi lớp, học viên đó có một MAHV khác nhau). Mỗi học viên xác định họ tên (HOTEN), ngày sinh (NGAYSINH), nơi sinh (NOISINH), phái nam hay nữ (PHAI), nghề nghiệp (NGHENGHIEP) - nghề nghiệp là SINH VIÊN, GIÁO VIÊN, KỸ SƯ, HỌC SINH, BUÔN BÁN,... Trung tâm KTCT có nhiều lớp, mỗi lớp có một mã lớp duy nhất (MALOP), mỗi lớp xác định các thông tin: tên lớp (TENLOP), thời khoá biểu, ngày khai giảng (NGAYKG), học phí (HOCPHI).

Chú ý rằng tại một thời điểm, trung tâm có thể mở nhiều lớp cho cùng một chương trình học. Với các lớp dài hạn thì ngày khai giảng được xem là ngày bắt đầu của mỗi học phần và HỌC PHÍ là học phí của mỗi học phần, với lớp ngắn hạn thì HỌC PHÍ là học phí của toàn khoá học đó.

Trung tâm có nhiều môn học, mỗi môn học có mã môn học (MAMH) duy nhất, mỗi môn học xác định tên môn học (TENMH), số tiết lý thuyết (SOTIETLT), số tiết thực hành (SOTIETTH). Mỗi học viên ứng với mỗi môn học có một điểm thi (DIEMTHI) duy nhất. Mỗi lần đóng học phí, học viên sẽ được trung tâm giao cho một phiếu biên lai thu tiền, mỗi biên lai có một số biên lai duy nhất để quản lý.

Một số yêu cầu của hệ thống này như: Lập danh sách những học viên khai giảng khoá ngày nào đó. Lập danh sách các học viên của một lớp ? Cho biết số lượng học viên của mỗi lớp khai giảng khoá ngày nào đó ?

Bài 5. QUẢN LÝ COI THI TUYỂN SINH

Một hội đồng coi thi tuyển sinh có nhiều điểm thi, mỗi điểm thi được đặt tại một trường nào đó. Các điểm thi (DIEMTHISO) được đánh số là điểm thi số 1, điểm thi số 2, điểm thi số 3,... Mỗi điểm thi xác định địa chỉ (DIACHIDIEMTHI). Ví dụ: điểm thi số 1, đặt tại trường PTTT Nguyễn Thị Minh Khai, điểm thi số 2 đặt tại trường PTTT Bùi Thị Xuân,... Mỗi thí sinh có một số báo danh (SOBD) duy nhất, mỗi số báo danh xác định các thông tin: họ và tên (HOTEN), ngày sinh (NGAYSINH), phái (PHAI), hộ khẩu thường trú (TINH), đối tượng dự thi (DOITUONG), ngành đăng ký thi, khu vực của thí sinh (KHUVUC), số hiệu phòng thi. Ví dụ: thí sinh Vũ Mạnh Cường, có số báo danh là 02978, sinh ngày 12/12/1984, phái nam, hộ khẩu thường trú tại Chợ Gạo - Tiền Giang, thuộc khu vực 1, đối tượng là 5B, đăng ký dự thi vào ngành có mã ngành là 01, thi tại phòng thi 0178, điểm thi số 1. Mỗi ngành có một mã ngành (MANGANH) duy nhất, mỗi mã ngành xác định tên ngành (TENNGANH). Mỗi điểm thi có nhiều phòng thi – mỗi phòng thi (PHONGTHI) được đánh số khác nhau ở tất cả các điểm thi. Trong một phòng thi, danh sách các thí sinh được sắp xếp theo thứ tự alphabet (do đó trong một phòng thi có thể có thí sinh của nhiều ngành khác nhau). Mỗi phòng thi có thêm

cột ghi chú (GHICHU) - ghi thêm các thông tin cần thiết như phòng thi đó nằm tại dãy nhà nào. Ví dụ phòng thi 0060 nằm ở dãy nhà H lầu 2 - điểm thi số 1 - trường PTTH Bùi Thị Xuân. Mỗi môn thi có một mã môn thi duy nhất (MAMT), mỗi mã môn thi biết các thông tin như : tên môn thi (TENMT), ngày thi (NGAYTHI), buổi thi (BUOITHI), thời gian làm bài thi được tính bằng phút (PHUT). Thời gian làm bài thi của các môn tối thiểu là 90 phút và tối đa là 180 phút (tuỳ theo kỳ tuyển sinh công nhân, trung cấp, cao đẳng hay đại học). Mỗi ngành có một mã ngành, chẳng hạn ngành Công Nghệ Thông Tin có mã ngành là 1, ngành Công Nghệ Hoá Thực Phẩm có mã ngành là 10,...

Mỗi đơn vị có cán bộ tham gia vào kỳ thi có một mã đơn vị duy nhất (MADONVI), mã đơn vị xác định tên đơn vị (TENDONVI). Nếu là cán bộ, công nhân viên của trường thì đơn vị là khoa/phòng quản lý cán bộ đó, nếu là giáo viên từ các trường khác thì ghi rõ tên đơn vị đó. Chẳng hạn cán bộ Nguyễn Thanh Liêm đơn vị Khoa Công Nghệ Thông Tin, cán bộ coi thi Nguyễn Thị Tuyết Mai, đơn vị trường PTTH Ngôi Sao - Quận 1,...

Mỗi cán bộ coi thi chỉ làm việc tại một điểm thi nào đó. Mỗi cán bộ có một mã số duy nhất (MACANBO), mỗi MACANBO xác định các thông tin khác như : họ và tên (HOTENCB), đơn vị công tác, chức vụ (CHUCVU) được phân công tại điểm thi, chẳng hạn chức vụ là điểm trưởng, điểm phó, giám sát, thư ký, cán bộ coi thi, phục vụ,... Ví dụ cán bộ Nguyen Van Thanh đơn vị Khoa Công Nghệ Thông Tin, làm nhiệm vụ thi tại điểm thi số 1, chức vụ là giám sát phòng thi.

ĐÁNH GIÁ KẾT QUẢ HỌC TẬP

TT	Tiêu chí đánh giá	Cách thức và phương pháp đánh giá	Điểm tối đa	Kết quả thực hiện của người học
I	Kiến thức			
II	Kỹ năng		0,5	
III	Thái độ			
<i>Cộng:</i>			<i>10 đ</i>	

KẾT QUẢ HỌC TẬP

Tiêu chí đánh giá	Kết quả thực hiện	Hệ số	Kết quả học tập
<i>Kiến thức</i>		0,3	
<i>Kỹ năng</i>		0,5	
<i>Thái độ</i>		0,2	
<i>g:</i>		<i>Cộng</i>	

Chương 2. MÔ HÌNH DỮ LIỆU QUAN HỆ

Mã chương MH16-02

Giới thiệu:

.....

Mục tiêu:

Trình bày cụ thể về mô hình dữ liệu quan hệ, các khái niệm, cơ sở lý thuyết của mô hình dữ liệu quan hệ (đại số quan hệ);

Áp dụng được các quy tắc chuyển đổi từ mô hình ERD sang mô hình dữ liệu quan hệ (ở dạng lược đồ).

Thực hiện các thao tác an toàn với máy tính.

Nội dung:

1. Một số khái niệm cơ bản.

1.1. Thuộc tính

Thuộc tính là các đặc điểm riêng của một đối tượng (*đối tượng* được hiểu như là một loại thực thể ở mô hình thực thể kết hợp), mỗi thuộc tính có một *tên gọi* và phải thuộc về một *kiểu dữ liệu* nhất định.

Kiểu dữ liệu (data type): Các thuộc tính được phân biệt qua tên gọi và phải thuộc một kiểu dữ liệu nhất định (số, chuỗi, ngày tháng, logic, hình ảnh,...). Kiểu dữ liệu ở đây có thể là kiểu vô hướng hoặc là kiểu có cấu trúc. Nếu thuộc tính có kiểu dữ liệu là vô hướng thì nó được gọi là thuộc tính đơn hay thuộc tính nguyên tố, nếu thuộc tính có kiểu dữ liệu có cấu trúc thì ta nói rằng nó không phải là thuộc tính nguyên tố.

Chẳng hạn với sinh viên Nguyễn Văn Thành thì các thuộc tính họ và tên, mã số sinh viên thuộc kiểu chuỗi, thuộc tính ngày sinh thuộc kiểu ngày tháng, hộ khẩu thường trú kiểu chuỗi, thuộc tính hình ảnh kiểu hình ảnh,...

Miền giá trị (domain of values): Thông thường mỗi thuộc tính chỉ chọn lấy giá trị trong một tập con của kiểu dữ liệu và tập hợp con đó gọi là miền giá trị của thuộc tính đó. Chẳng hạn thuộc tính NỮ có miền giá trị là {nam,nữ}, thuộc tính màu da có miền giá trị là {da trắng, da vàng, da đen, da đỏ}, thuộc tính điểm thi là các số thuộc tập {0; 1; 2;...,10}.

Lưu ý rằng nếu không lưu ý đến ngữ nghĩa thì tên của các thuộc tính thường được ký hiệu bằng các chữ cái in hoa đầu tiên trong bảng chữ cái la tinh: A,B,C,D,... Những chữ cái in hoa X,Y,Z,W,... thường dùng thay cho một nhóm nhiều thuộc tính. Đôi khi còn dùng các ký hiệu chữ cái với các chỉ số A_1, A_2, \dots, A_n để chỉ các thuộc tính trong trường hợp tổng quát hay muốn đề cập đến số lượng các thuộc tính. Tên thuộc tính phải được đặt một cách gợi nhớ, không nên đặt tên thuộc tính quá dài (vì như thế sẽ làm cho việc viết các câu lệnh truy vấn trở nên vất vả hơn), nhưng cũng không nên đặt tên thuộc tính quá ngắn (vì nó sẽ không cho thấy ngữ nghĩa của thuộc tính), đặc biệt không đặt trùng tên hai thuộc tính mang ngữ nghĩa khác nhau thuộc hai đối tượng khác nhau.

Trong nhiều hệ quản trị cơ sở dữ liệu, người ta thường đưa thêm vào miền giá trị của các thuộc tính một giá trị đặc biệt gọi là giá trị rỗng (NULL). Tùy theo ngữ cảnh mà giá trị này có thể đặc trưng cho một giá trị không thể xác định được hoặc một giá trị chưa được xác định ở vào thời điểm nhập tin nhưng có thể được xác định vào một thời điểm khác.

1.2 Lược Đồ Quan Hệ (relation schema)

Tập tất cả các thuộc tính cần quản lý của một đối tượng cùng với các mối liên hệ giữa chúng được gọi là lược đồ quan hệ. Lược đồ quan hệ Q với tập thuộc tính $\{A_1, A_2, \dots, A_n\}$ được viết là $Q(A_1, A_2, \dots, A_n)$,

ký hiệu: $Q^+ = \{A_1, A_2, \dots, A_n\}$.

Chẳng hạn lược đồ quan hệ Sinhvien với các thuộc tính như đã được liệt kê trong ví dụ 1.1 được viết như sau:

Sinhvien(MASV, HOTENSV, NU, NGAYSINH, NOISINH,TINH, MALOP) Thường thì khi thành lập một lược đồ quan hệ, người thiết kế gán cho nó một ý nghĩa nhất định, gọi là *tên từ của lược đồ quan hệ*. chẳng hạn tên từ của lược đồ quan hệ Sinhvien là: "Mỗi sinh viên có mỗi MASV duy nhất. Mỗi MASV xác định các thuộc tính còn lại của sinh viên đó như HOTENSV,NU, NGAYSINH, NOISINH,TINH,MALOP".

Khi phát biểu tên từ cho một lược đồ quan hệ, người thiết kế cần phải mô tả đầy đủ ý nghĩa để người khác tránh hiểu nhầm. Dựa vào tên từ này, người ta xác định được tập khoá, siêu khoá của lược đồ quan hệ (sẽ được trình bày trong những mục kế tiếp).

Nhiều lược đồ quan hệ cùng nằm trong một hệ thống thông tin được gọi là một *lược đồ cơ sở dữ liệu*.

Khái niệm lược đồ quan hệ ứng với khái niệm loại thực thể ở mô hình thực thể kết hợp.

1.3. Quan Hệ (relation)

Sự thể hiện của lược đồ quan hệ ở một thời điểm nào đó được gọi là *quan hệ*, rõ ràng là trên một lược đồ quan hệ có thể xác định nhiều quan hệ. Thường ta dùng các ký hiệu như R,S,Q để chỉ các lược đồ quan hệ, còn quan hệ thường được dùng bởi các ký hiệu là r, s,q,...

Về trực quan thì quan hệ (hay bảng quan hệ) như là một bảng hai chiều gồm các dòng và các cột.

Một quan hệ có n thuộc tính được gọi là quan hệ n ngôi.

Để chỉ quan hệ r xác định trên lược đồ quan hệ Q ta có thể viết r(Q).

1.4 Bộ (Tuple)

Mỗi *bộ* là những thông tin về một đối tượng thuộc một quan hệ, bộ cũng còn được gọi là mẫu tin.

Thường người ta dùng các chữ cái thường (như t,...) để biểu diễn bộ trong quan hệ.

2. Chuyển mô hình thực thể kết hợp sang mô hình dữ liệu quan hệ

Sau đây là một số quy tắc được sử dụng trong việc chuyển đổi mô hình thực thể kết hợp sang mô hình dữ liệu quan hệ.

Quy tắc 1:

Chuyển đổi mỗi loại thực thể thành một lược đồ quan hệ, các thuộc tính của loại thực thể thành các thuộc tính của lược đồ quan hệ, thuộc tính khoá của loại thực thể là thuộc tính khoá của lược đồ quan hệ.

Chẳng hạn loại thực thể Sinhvien ở ví dụ 1.2 khi áp dụng quy tắc 1 thì sẽ được chuyển thành lược đồ quan hệ Sinhvien như sau:

Sinhvien(MASV, HOTENSV, NU, NGAYSINH, TINH,...)

Quy tắc 2:

Nếu mỗi kết hợp mà cả hai nhánh của nó đều có bản số max là n thì mỗi kết hợp này sẽ được chuyển thành một lược đồ quan hệ K' gồm các thuộc tính của mỗi kết hợp K, cộng thêm các thuộc tính khoá của hai lược đồ quan hệ A, B tương ứng với hai thực thể tham gia vào mỗi kết hợp. Khoá của lược đồ quan hệ K' gồm cả hai khoá của hai lược đồ quan hệ A và B.

Chẳng hạn mỗi kết hợp Phancong giữa ba loại thực thể Giangvien, Monhoc và Lop được chuyển thành lược đồ quan hệ Phancong và có tập khoá là {MAGV,MAMH,MALOP} như sau:

Phancong(MAGV,MAMH,MALOP)

Quy tắc 3:

Mỗi kết hợp mà một nhánh có bản số là n (nhánh B) và nhánh còn lại có bản số max là 1 (nhánh A) thì loại bỏ mỗi kết hợp này khỏi mô hình thực thể kết hợp và thêm các thuộc tính khoá của lược đồ tương ứng với loại thực thể ở nhánh B vào lược đồ tương ứng với loại thực thể ở nhánh A (khoá của B sẽ thành khoá ngoại của A). Nếu mỗi kết hợp có các thuộc tính thì những thuộc tính này cũng được thêm vào lược đồ quan hệ tương ứng với loại thực thể ở nhánh A.

Chẳng hạn mỗi kết hợp *thuộc* giữa hai loại thực thể Sinhvien và Lop nên lược đồ quan hệ Sinhvien được sửa thành như sau:

Sinhvien(MASV,HOTENSV,NU,NGAYSINH, TINH,MALOP)

Quy tắc 4:

Nếu mỗi kết hợp mà cả hai nhánh đều có bản số max là 1 thì áp dụng quy tắc 3 cho một trong hai nhánh tuỳ chọn.

Ví dụ 2.1:

Sau đây là mô hình dữ liệu quan hệ được chuyển từ mô hình thực thể kết hợp ở ví dụ 1.2.

Sinhvien(MASV, HOTENSV, NU, NGAYSINH, NOISINH,TINH, MALOP)

Lop(MALOP,TENLOP,MAKHOA)

Khoa(MAKHOA,TENKHOA)

Monhoc(MAMH,TENMH,DONVIHT)

Giangvien(MAGV,HOTENGV,HOCVI,CHUYENNGANH,MAKHOA)

Ketqua(MASV, MAMH, LANTHI,DIEMTHI)

3. Ngôn ngữ đại số quan hệ

3.1. Phép Hợp 2 quan hệ(Union)

Ta nói hai quan hệ r_1 và r_2 là tương thích nếu chúng được định nghĩa trên cùng một lược đồ quan hệ.

Cho hai quan hệ tương thích r_1 và r_2 . Hợp của hai quan hệ r_1 và r_2 ký hiệu là $r_1 + r_2$ là một quan hệ trên lược đồ quan hệ Q gồm các phần tử thuộc r_1 hoặc thuộc r_2 , tức là: $r_1 + r_2 = \{t / t \in r_1 \text{ hoặc } t \in r_2\}$

 r_1

A	B	C	D
a1	b1	c1	d1
a2	b2	c2	d2
a3	b3	c3	d3
a4	b4	c4	d4

 r_2

A	B	C	D
x1	y1	z1	v1
a2	b2	c2	d2
x3	y3	z3	v3

Ví dụ 2.2

Khi đó, nội dung của quan hệ $r_1 + r_2$ là:

A	B	C	D
a1	b1	c1	d1
a2	b2	c2	d2
a3	b3	c3	d3
a4	b4	c4	d4
x1	y1	z1	v1
x3	y3	z3	v3

Do thứ tự trước/sau của các bộ trong các quan hệ là không quan trọng nên ta có: - r_1, r_2 thì $r_1 + r_2 = r_2 + r_1$

- r thì $r + r = r$

Một cách tổng quát có thể lấy hợp của n quan hệ tương thích: cho n quan hệ tương thích r_1, r_2, \dots, r_n

Hợp của n quan hệ r_1, r_2, \dots, r_n là một quan hệ $r_1 + r_2 + \dots + r_n$ gồm các phần tử thuộc r_1 hoặc thuộc $r_2 \dots$ hoặc thuộc r_n

3.2. Phép Giao 2 quan hệ (Intersection)

Cho lược đồ quan hệ $Q(A_1, A_2, \dots, A_n)$. r_1 và r_2 là hai quan hệ tương thích trên Q . Giao của hai quan hệ r_1 và r_2 ký hiệu là $r_1 * r_2$ là một quan hệ trên Q gồm các phần tử vừa thuộc r_1 vừa thuộc r_2 .

Vậy: $r_1 * r_2 = \{ t / t \in r_1 \text{ và } t \in r_2 \}$

Chẳng hạn với ví dụ 2.2 ở trên thì $r_1 * r_2$ là:

A	B	C	D
a2	b2	c2	d2

3.3. Phép Trừ 2 quan hệ (Minus)

Cho hai quan hệ tương thích r_1 và r_2 có tập thuộc tính $Q(A_1, A_2, \dots, A_n)$. Hiệu của r_1 cho r_2 ký hiệu là $r_1 - r_2$ là một quan hệ trên Q gồm các phần tử chỉ thuộc r_1 mà không thuộc r_2 , nghĩa là $r_1 - r_2 = \{ t \in r_1 \text{ và } t \notin r_2 \}$

Chẳng hạn với ví dụ 2.2. thì $r_1 - r_2$ là:

A	B	C	D
a1	b1	c1	d1
a3	b3	c3	d3
a4	b4	c4	d4

3.4. Tích Descartes của 2 quan hệ (Cartesian Product)

Cho hai lược đồ quan hệ $Q_1(A_1, A_2, \dots, A_n)$, $Q_2(B_1, B_2, \dots, B_m)$. Giả sử r_1, r_2 là hai quan hệ trên Q_1, Q_2 tương ứng.

Tích Descartes (decac) của r_1 và r_2 ký hiệu là $r_1 \times r_2$ là quan hệ trên lược đồ quan hệ có tập thuộc tính $Q = Q_1 \cup Q_2$.

Vậy quan hệ $r_1 \times r_2$ là quan hệ trên lược đồ: $Q = Q_1 \cup Q_2 = \{A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m\}$ với $r_1 \times r_2 = \{(t_1, t_2) : t_1 \in r_1, t_2 \in r_2\}$

Ví dụ 2.3. cho r_1 và r_2 là

r_1		
A	B	C
6	5	4
7	5	5

r_2		
E	F	H
1	5	9
4	6	8
7	5	3

Thì kết quả $r_1 \times r_2$ như sau:

A	B	C	E	F	H
6	5	4	1	5	9
6	5	4	4	6	8
6	5	4	7	5	3
7	5	5	1	5	9
7	5	5	4	6	8
7	5	5	7	5	3

3.5. phép chia 2 quan hệ:

Cho 2 lược đồ quan hệ $Q_1(A_1, A_2, \dots, A_n)$, $Q_2(B_1, B_2, \dots, B_m)$, r là quan hệ xác định trên Q_1 ; s là quan hệ xác định trên Q_2 ($n > m$ và s khác rỗng), có m thuộc tính chung (giống nhau về mặt ngữ nghĩa, hoặc các thuộc tính có thể so sánh được) giữa r và s . phép chia 2 quan hệ r và s ký hiệu $r \div s$, là một quan hệ q có $n - m$ thuộc tính được định nghĩa như sau:

$$q = r \div s = \{t / \exists u \in s, (t, u) \in r\}$$

Ví dụ 2.4:

r						
A	B	C	D			
a	b	c	d			
a	b	e	f			
b	c	e	f			
c	d	c	d			
c	d	e	f			
a	b	d	e			
s					r ÷ s	
C	D			A	B	
c	d			a	b	
e	f			c	d	

3.6. Phép Chiếu (projection)

Cho lược đồ quan hệ $Q(A_1, A_2, \dots, A_n)$, r là quan hệ trên Q và $X \subseteq Q^+$.

Phép chiếu của r lên tập thuộc tính X , ký hiệu là $r[X]$ (hoặc $r.X$) sẽ tạo thành lược đồ quan hệ r' , trong đó tập thuộc tính của r' chính là X và quan hệ r' được trích từ r bằng cách chỉ lấy các thuộc tính có trong X .

Phép chiếu chính là phép rút trích dữ liệu theo cột. Chẳng hạn với r_1 ở ví dụ 2.2 thì khi đó ta có quan hệ con của r_1 chiếu lên $X = \{A, C\}$ là:

$r_1[X]$

A	C
a1	c1
a2	c2
a3	c3
a4	c4

3.7. Phép Chọn (Selection)

Cho lược đồ quan hệ $Q(A_1, A_2, \dots, A_n)$, r là một quan hệ trên lược đồ quan hệ Q . X là một tập con của Q^+ và E là một mệnh đề logic được phát biểu trên tập X . Phần tử $t \in r$ thoả mãn điều kiện E ký hiệu là $t(E)$. Phép chọn từ quan hệ r theo điều kiện E (ký hiệu là $r : E$) sẽ tạo thành một quan hệ mới ký hiệu là $r(E)$, trong đó $r(E) = \{t: t \in r \text{ và } t(E)\}$

Phép chọn chính là phép rút trích dữ liệu theo dòng. Chẳng hạn với r_2 ở ví dụ 2.3 và điều kiện E là: “ $F \geq 6$ ” thì kết quả $r_2(E)$ hay $r_2: “F \geq 6”$ có nội dung là

E	F	H
4	6	8

3.8. Phép kết

Cho hai lược đồ quan hệ Q_1 và Q_2 như sau: $Q_1(A_1, A_2, \dots, A_n)$ $Q_2(B_1, B_2, \dots, B_m)$; r và s lần lượt là hai quan hệ trên Q_1 và Q_2 .

A_i và B_j lần lượt là thuộc tính của Q_1 , Q_2 sao cho $MGT(A_i) = MGT(B_j)$.

là một trong các phép so sánh ($=, <, >, \leq, \geq, \neq$) trên $MGT(A_i)$.

A_i

B_j

Phép kết giữa r và s theo điều kiện $A_i \text{ } B_j$ ký hiệu là $r \bowtie s$ là một quan hệ trên lược đồ quan hệ có tập thuộc tính là $Q_1 \cup Q_2$. gồm những bộ thuộc tích Descartes của r và s sao $A_i B_j$.

$A_i \theta B_j$

$r |><| s = \{t_{12} / \exists t_1 \in r_1, \exists t_2 \in r_2 \text{ sao cho } t_{12}.Q_1^+ = t_1; t_{12}.Q_2^+ = t_2; t_{12}A_i \theta t_{12}.B_j\}$

Ví dụ 2.5 Cho hai quan hệ r_1 và r_2 như sau:

r_1

A	B	C
6	5	4
7	5	5
4	2	6

r_2

E	F	H
1	5	9
4	6	8
7	5	3

A_i là thuộc tính B, B_j là thuộc tính F và θ là phép so sánh “>=”. Ta được kết quả là quan hệ sau:

A	B	C	E	F	H
6	5	4	1	5	9
6	5	4	7	5	3
7	5	5	1	5	9
7	5	5	7	5	3

3.9. Phép Kết Tự Nhiên (natural join)

Nếu θ được sử dụng trong phép kết trên là phép so sánh bằng (=) thì gọi là phép kết bằng. Hơn nữa nếu $A_i \equiv B_j$ thì phép kết bằng này được gọi là phép kết tự nhiên. Phép kết tự nhiên là phép kết thường dùng nhất trong thực tế.

Ngôn ngữ với các phép toán trên gọi là ngôn ngữ đại số quan hệ. Sau đây là một ví dụ về ngôn ngữ đại số quan hệ.

Ví dụ 2.6

Cho lược đồ CSDL dùng để quản lý điểm sinh viên được mô tả như ở ví dụ 2.1. Hãy thực hiện các yêu cầu sau bằng ngôn ngữ đại số quan hệ:

1. Lập danh sách các sinh viên lớp có mã lớp là CDTH2A, danh sách cần MASV, HOTENSV

2.Lập danh sách sinh viên nữ và có mã khoa là “CNTT”, danh sách cần MASV, HOTENSV.

3.Lập bảng điểm thi lần 1 của tất cả các môn cho sinh viên lớp CDTH2A, danh sách cần MASV, HOTENSV, TENMH, DIEMTHI.

4.Lập phiếu điểm thi lần 1 các môn cho sinh viên có MASV=”00CDTH189”. danh sách cần MAMH,TENMH, DONVIHT, DIEMTHI.

Giải:

1. Sinhvien: MALOP="CDTH2A" [MASV,HOTENSV]

MALOP

2. (Sinhvien|><| Lop: NU and MAKHOA="CNTT")

[MASV,HOTENSV]

MASV

MAMH

3. (((Sinhvien |><| Ketqua) |><| Monhoc): MALOP = "CDTH2A" and LANTHI=1) [MASV,HOTENSV,TENMH, DIEMTHI]

MAMH

4. (Ketqua |><| Monhoc : MASV='00CDTH189' and LANTHI=1)

[MAMH,TENMH,DONVIHT,DIEMTHI]

BÀI TẬP

2.1. Hãy lập mô hình dữ liệu quan hệ cho các bài toán quản lý 1.1, 1.2, 1.3, 1.4, 1.5. Hãy xác định khoá cho từng lược đồ cho mỗi bài toán trên.

2.2. Cho lược đồ cơ sở dữ liệu

Sinhvien(MASV, HOTENSV, NU, NGAYSINH,
NOISINH, TINH, MALOP)

Lop(MALOP, TENLOP, MAKHOA) Khoa(MAKHOA, TENKHOA)
Monhoc(MAMH, TENMH, DONVIHT)

Giangvien(MAGV, HOTENGV, HOCVI, CHUYENNGANH, MAKHOA)
Ketqua(MASV, MAMH, LANTHI, DIEMTHI)
Phancong(MALOP, MAMH, MAGV)

Thực hiện các yêu cầu sau bằng ngôn ngữ đại số quan hệ:

a. Lập danh sách những sinh viên có hộ khẩu thường trú ở tỉnh “LONG AN”, danh sách cần các thông tin: MASV, HOTENSV, NGAYSINH, TENLOP

b. Lập danh sách các sinh viên của lớp có MALOP là CDTH2A, danh sách cần các thông tin: MASV, HOTENSV, NGAYSINH, TINH.

c. Lập danh sách các giảng viên có cấp học vị là THAC SY của khoa có MAKHOA là “CNTT”, danh sách cần: MAGV, HOTENGV, CHUYENNGANH.

d. Lập bảng điểm thi lần 1 môn học “869” cho tất cả sinh viên thuộc hai lớp có MALOP là “CDTH2A” và “CDTH2B”, danh sách cần: MASV, HOTENSV, DIEMTHI.

e. Lập danh sách các giảng viên đã dạy lớp CDTH2A, danh sách cần các thông tin: MAGV, HOTENGV, TENKHOA, HOCVI, TENMH.

f.Lập danh sách các môn mà lớp CDTH2A đã học, danh sách cần các thông tin: MAMH,TENMH,DONVIHT,HOTENGV.

g.Lập danh sách những giảng viên đã dạy sinh viên có MASV là “00CDTH189“, danh sách cần MAGV,HOTENGV,HOCVI,CHUYENNGANH,TENKHOA,TENMH

h.Lập danh sách các sinh viên có mã khoa “CNTT” có điểm thi lần 1 môn học “869” lớn hơn hoặc bằng 8, danh sách cần MASV, HOTENSV, DIEMTHI, TENLOP.

ĐÁNH GIÁ KẾT QUẢ HỌC TẬP

TT	Tiêu chí đánh giá	Cách thức và phương pháp đánh giá	Điểm tối đa	Kết quả thực hiện của người học
I	Kiến thức			
II	Kỹ năng		0,5	
III	Thái độ			
<i>Cộng:</i>			<i>10 đ</i>	

KẾT QUẢ HỌC TẬP

Tiêu chí đánh giá	Kết quả thực hiện	Hệ số	Kết quả học tập
<i>Kiến thức</i>		0,3	
<i>Kỹ năng</i>		0,5	
<i>Thái độ</i>		0,2	
<i>g:</i>		Cộng	

Chương 3. NGÔN NGỮ TRUY VẤN DỮ LIỆU

Mã chương MH16-03

Giới thiệu:

SQL(STRUCTURE QUERY LANGUAGE) là ngôn ngữ truy vấn dựa trên đại số quan hệ. Ngôn ngữ truy vấn SQL có tập lệnh khá phong phú dùng để thao tác trên cơ sở dữ liệu. Chẳng hạn lệnh create để tạo các bảng quan hệ, lệnh update để cập nhật dữ liệu, lệnh delete để xoá dữ liệu, lệnh insert để thêm dữ liệu,...

Trong chương này, chúng ta chỉ nghiên cứu về câu lệnh quan trọng nhất của SQL đó là câu lệnh truy vấn dữ liệu SELECT. Kết quả của lệnh select là một quan hệ, quan hệ kết quả này có thể kết xuất ra màn hình, máy in, hoặc là trên các thiết bị lưu trữ thông tin khác.

Mục tiêu:

- Bài này giúp sinh viên hiểu cách thức truy vấn của dữ liệu quan hệ, điển hình là ngôn ngữ truy vấn SQL chuẩn, thực hiện truy vấn được trên các dữ liệu đã cài đặt.
- Thực hiện các thao tác an toàn với máy tính..

Nội dung:

1. Mở đầu

Vào những năm 1970, SQL(STRUCTURE QUERY LANGUAGE) lần đầu được hãng IBM phát triển như một bộ phận của hệ quản trị CSDL mô hình quan hệ có tên là SYSTEM R. Sau đó vào các năm 1980 IBM tiếp tục phát triển SQL cho các hệ quản trị cơ sở dữ liệu nổi tiếng là SQL/DS trên nền hệ điều hành VM, DB2 trên nền hệ điều hành MVS, Hệ quản trị cơ sở dữ liệu mở rộng trên nền hệ điều hành IBM OS/2, Hệ quản trị cơ sở dữ liệu cho hệ thống IBM AS/400. Năm 1986 , Viện tiêu chuẩn quốc gia Hoa kỳ (ANSI – American National Standards Institute) và Tổ chức Tiêu chuẩn Quốc

tế (ISO – International Standards Organization) đã thừa nhận SQL như là *ngôn ngữ chuẩn xử lý dữ liệu*. Ngôn ngữ chuẩn ANSI SQL tiếp tục được cập nhật vào những năm 1989 và 1992 sau đó. Hiện nay một phiên bản mới SQL3 đang được phát triển.

SQL được cài đặt cho hệ thống máy tính lớn (mainframe) cũng như máy tính cá nhân. Bên cạnh các sản phẩm của hãng IBM, cũng cần phải kể đến các hệ quản trị cơ sở dữ liệu nổi tiếng khác như ORACLE của Công ty Oracle, SQL Server của hãng Microsoft, SQLBase của hãng Sybase, Ingres của hãng Relational Technologies, ...

Mục đích của chuẩn SQL là

- Xác định cú pháp và ngữ nghĩa của ngôn ngữ SQL định nghĩa và thao tác dữ liệu
- Định nghĩa các cấu trúc dữ liệu và phép toán cơ bản để thiết kế, truy cập, lưu trữ, kiểm soát và bảo vệ cơ sở dữ liệu SQL
- Cung cấp công cụ đảm bảo tính tương thích của cấu trúc dữ liệu và các modul ứng dụng giữa các hệ quản trị cơ sở dữ liệu
- Xác định chuẩn tối thiểu (Mức 1) và chuẩn hoàn chỉnh (Mức 2), cho phép các cấp độ sử dụng SQL khác nhau trong các sản phẩm
- Cung cấp chuẩn ban đầu, có thể chưa hoàn chỉnh, cho phép mở rộng các chức năng xử lý những vấn đề như sự toàn vẹn tham chiếu, giao thức chuyển đổi, các hàm người dùng, các toán tử nối ngoài phép đẳng nối, và các hệ thống ký tự quốc gia,...

Một ngôn ngữ quan hệ chuẩn như SQL sẽ mang lại các lợi ích sau đây:

- Giảm thiểu chi phí đào tạo. Các hệ quản trị cơ sở dữ liệu có chung ngôn ngữ chuẩn SQL sẽ làm giảm chi phí chuyển đổi từ hệ này sang hệ khác.

- Nâng cao hiệu năng công việc. Các chuyên gia hệ thống thông tin với kiến thức sâu sắc về SQL sẽ nhanh chóng nắm bắt các chương trình ứng dụng của các hệ quản trị cơ sở dữ liệu, vì họ đã quen thuộc với ngôn ngữ của các chương trình này.
- Tính khả chuyển của các ứng dụng. Các ứng dụng có thể dễ dàng sử dụng trên các hệ thống khác nhau, nhưng cùng sử dụng SQL.
- Tăng tuổi thọ của các ứng dụng. Một ngôn ngữ chuẩn có xu hướng tồn tại thời gian dài, điều đó làm giảm áp lực viết lại chương trình.
- Làm giảm sự phụ thuộc vào nhà cung cấp. Vì SQL là ngôn ngữ chung nên người dùng dễ dàng sử dụng nhiều sản phẩm của các nhà cung ứng khác nhau, với giá cả cạnh tranh.
- Khả năng giao tiếp giữa các hệ thống chéo. các hệ quản trị cơ sở dữ liệu quan hệ và các chương trình ứng dụng khác nhau có thể dễ dàng giao tiếp và hợp tác để xử lý dữ liệu và thực hiện chương trình người dùng.

SQL là ngôn ngữ có cấu trúc

Trong câu lệnh của SQL có một số mệnh đề tuân theo những cú pháp riêng của nó. Có 4 loại lệnh trong SQL :

- Các lệnh truy vấn dữ liệu.
- Các lệnh định nghĩa dữ liệu (DDL) .
- Các lệnh xử lý cập nhật dữ liệu (DML) .
- Các lệnh kiểm soát dữ liệu .

Kiểu dữ liệu

SQL có các kiểu dữ liệu sau

INTEGER : Kiểu số nguyên 2 byte -2 147 483 648 2 147 483 647 .

SMALLINT : Kiểu số nguyên 1 byte - 32 768 32 767.

DECIMAL(n,d) : Kiểu số thực độ dài n (kể cả dấu), số chữ số thập phân d

FLOAT(n,d) : Kiểu số thực khoa học độ dài n , số chữ số thập phân d

CHAR(n) : Kiểu chuỗi ký tự độ dài n

DATE : Kiểu ngày tháng (ngày/tháng/năm)

LOGICAL : Kiểu logic, nhận giá trị *true* hoặc *false*

VARCHAR(n) : Kiểu chuỗi ký tự độ dài thay đổi, tối đa n ký tự

LONGVARCHAR : Kiểu chuỗi ký tự độ dài thay đổi (chứa dữ liệu như ghi chú,...)

Kiểu tiền tệ

Kiểu hình ảnh

Kiểu âm thanh...

Cú pháp tổng quát:

SELECT [DISTINCT] <biểu thức 1> [AS <tên 1>] [...] | *

FROM <bảng 1> [<bí danh 1>] [...]

[INTO <dbf đích>]

[**WHERE** <điều kiện nối > [AND | OR <điều kiện lọc>]]

[GROUP BY <cột nhóm 1> [...]

[HAVING <điều kiện nhóm>]]

[ORDER BY <biểu thức sắp xếp 1> [ASC | DESC] [...]]

[UNION | INTERSECT | MINUS <câu truy vấn khác>]

Các thành phần cơ bản của câu lệnh:

+ Biểu thức sau SELECT có thể bao gồm :

- Danh sách các cột, kể cả các biểu thức chứa các cột, của các bảng hoặc khung nhìn khai báo sau FROM. Các biểu thức ngăn cách nhau bằng dấu phẩy (,) và có thể đổi tên bằng tùy chọn AS <tên >.

- * là ký tự đại diện tất cả các thuộc tính của bảng sau FROM

- Các hàm tính toán : COUNT, SUM, AVG, MIN, MAX

+ Biểu thức sau FROM gồm một hoặc danh sách các bảng quan hệ (có thể đặt bí danh).

+ INTO <dbf đích>: Lưu bảng kết quả vào đĩa.

+ Biểu thức sau WHERE bao gồm :

- Các thuộc tính của các bảng quan hệ sau FROM

- Các toán tử số học : +, -, *, /

- Các toán tử so sánh, có thể so sánh với ANY (giá trị nào đó trong 1 tập hợp) hoặc ALL (tất cả giá trị trong 1 tập hợp)

- Các toán tử Boolean : AND, OR, NOT

- Các toán tử tập hợp : UNION (hợp), INTERSECT (giao), MINUS (hiệu)

- Các quan hệ bao hàm : IN (), NOT IN (), CONTAINS (chứa), DOES NOT CONTAIN (không chứa)

- Điều kiện tồn tại : EXISTS (tồn tại), NOT EXIST (không tồn tại)

- Câu văn tin con kiểu SELECT ... FROM ... WHERE ...

2. Tìm thông tin từ các cột của bảng – Mệnh đề Select

Cú pháp:

SELECT [DISTINCT] <biểu thức 1> [AS <tên 1>] [...] | *

FROM <bảng >

Công dụng:

Xác định nội dung kết quả gồm các cột liệt kê sau SELECT hoặc tất cả các cột <*> lấy từ <Bảng>.

Từ khoá DISTINCT dùng liệt kê những giá trị duy nhất tránh trùng lặp.

Từ khoá AS dùng đặt bí danh

+ Ví dụ

Cho lược đồ cơ sở dữ liệu khách hàng

KHACH(TenKH, DiaChi, SoDu)

HOPDONG(SoHD, TenKH, Hang, SoLuong)

CUNGUNG(Ten, DiaChi, Hang, DonGia)

- Trả về các bản ghi gồm tất cả các cột của bảng CUNGUNG.

```
SELECT *  
FROM CUNGUNG
```

- Trả về các bản ghi gồm 2 cột Hang, DonGia của bảng CUNGUNG.

```
SELECT Hang, DonGia  
FROM CUNGUNG
```

hoặc

```
SELECT DISTINCT Hang, DonGia  
FROM CUNGUNG
```

- Trả về các bản ghi gồm 1 cột Hang với giá trị duy nhất của bảng CUNGUNG.

```
SELECT DISTINCT Hang  
FROM CUNGUNG
```

- Có thể cho hiển thị giá trị *hàng*, chẳng hạn

```
SELECT “Mặt hàng”, Hang, “có đơn giá là”, DonGia  
FROM CUNGUNG
```

3. Chọn các dòng của bảng – Mệnh đề Where

Cú pháp:

```
SELECT [DISTINCT] <biểu thức 1> [AS <tên 1>] [... ] | *  
FROM <bảng >
```


WHERE <điều kiện lọc>

Công dụng:

Trả về các bản ghi gồm các cột liệt kê sau SELECT hoặc tất cả các cột <*> lấy từ <Bảng> gồm các bộ thỏa mãn <điều kiện lọc>.

Ghi chú : Sau WHERE có thể dùng các mệnh đề :

- <biểu thức> BETWEEN <giá trị 1> AND <giá trị 2> : biểu diễn điều kiện

<giá trị 1> <biểu thức> <giá trị 2>

- <biểu thức> NOT BETWEEN <giá trị 1> AND <giá trị 2> : biểu diễn điều kiện

<biểu thức> < <giá trị 1> hoặc <biểu thức> > <giá trị 2>

- <biểu thức> IN (<giá trị 1>, <giá trị 2>, ...) : biểu diễn điều kiện

<biểu thức> bằng một trong các giá trị <giá trị 1>, <giá trị 2>, ...

- <biểu thức> NOT IN (<giá trị 1>, <giá trị 2>, ...) : biểu diễn điều kiện <biểu thức> khác các giá trị <giá trị 1>, <giá trị 2>, ...

- <biểu thức ký tự> [NOT] LIKE <chuỗi ký tự> : biểu diễn điều kiện <biểu thức ký tự> [không] giống <chuỗi ký tự>

Trong <chuỗi ký tự> có thể dùng dấu

% (dấu phần trăm) đại diện cho chuỗi ký tự bất kỳ

_ (dấu nối) đại diện cho một ký tự bất kỳ

- <trường> IS [NOT] NULL: biểu diễn điều kiện giá trị <trường> là [khác] NULL

+ Ví dụ

- Trả về các bản ghi gồm các cột Ten, DiaChi có giá trị duy nhất từ bảng CUNGUNG của các nhà cung ứng hoặc Đường hoặc Sữa.

SELECT DISTINCT Ten, DiaChi

FROM CUNGUNG

WHERE (Hang = 'Đường') OR (Hang = 'Sữa')

hoặc

```
SELECT DISTINCT Ten, DiaChi
FROM CUNGUNG
WHERE Hang IN ('Đường', 'Sữa')
```

- Trả về các bản ghi gồm các Ten, DiaChi của các khách hàng có SoDu không âm từ bảng KHACH

```
SELECT Ten, DiaChi
FROM KHACH
WHERE SoDu >= 0
```

- Trả về các bản ghi gồm các hợp đồng có SoLuong từ 100 đến 200 từ bảng HOPDONG

```
SELECT *
FROM HOPDONG
WHERE SoLuong BETWEEN 100 AND 200
```

- Trả về các bản ghi gồm các nhà cung ứng có tên bắt đầu bằng 'Cty' từ bảng CUNGUNG

```
SELECT *
FROM CUNGUNG
WHERE Ten LIKE 'Cty %'
```

- Trả về các bản ghi gồm các nhà cung ứng có mặt hàng 'Bia' từ bảng CUNGUNG

```
SELECT *
FROM CUNGUNG
WHERE Hang LIKE 'Bia %'
```

- Chọn những khách hàng chưa có địa chỉ

```
SELECT *
FROM KHACH
```

WHERE DiaChi IS NULL

- Chọn những khách hàng có số dư xác định :

```
SELECT *
FROM KHACH
WHERE SoDu IS NOT NULL
```

4. Sắp xếp các dòng của bảng – Mệnh đề Order by

Cú pháp:

```
SELECT [DISTINCT] <biểu thức 1> [AS <tên 1>] [,...] | *
FROM <bảng >
ORDER BY <biểu thức sắp xếp 1> [ASC | DESC] [,...]
```

Công dụng:

Trả về các bản ghi gồm các cột liệt kê sau SELECT hoặc tất cả các cột <*> lấy từ <Bảng> gồm các bộ được sắp xếp theo các biểu thức sau ORDER BY.

Ý nghĩa các tham số :

ASC : tăng dần (ngầm định)

DESC: giảm dần .

+ Ví dụ :

- Cho biết danh sách khách hàng sắp xếp theo TÊN tăng dần :

```
SELECT *
FROM KHACH
ORDER BY Ten
```

- Cho biết danh sách các nhà cung ứng ‘Sữa’ sắp xếp theo thứ tự DonGia giảm dần :

```
SELECT *
FROM CUNGUNG
WHERE Hang = ‘Sữa’
```

ORDER BY DonGia DESC

- Cho biết danh sách các nhà cung ứng sắp xếp theo thứ tự Ten và Hang :

```
SELECT *  
FROM CUNGUNG  
ORDER BY Ten, Hang
```

5. Câu lệnh truy vấn lồng nhau

Cú pháp:

```
SELECT [DISTINCT] <biểu thức 1> [AS <tên 1>] [...] / *  
FROM <bảng 1> [<bí danh 1>] [...]  
[WHERE <điều kiện nối , lọc có chứa câu truy vấn con >]  
[ORDER BY <biểu thức sắp xếp 1> [ASC / DESC] [...]]
```

Công dụng:

Cho phép tạo ra các điều kiện chứa quan hệ tập hợp và số liệu tổng hợp.

Trong quan hệ tập hợp, câu truy vấn con đặt trong ngoặc đơn .

❖ **Sử dụng các toán tử so sánh : <, <=, =, >, >=, <>**

- Liệt kê các khách hàng có số dư lớn nhất:

```
SELECT Ten  
FROM KHACH  
WHERE Sodu =  
      (SELECT MAX(Sodu)  
       FROM KHACH)
```

- Liệt kê hãng cung ứng gạo có DonGia <= DonGia trung bình :

```
SELECT Ten  
FROM CUNGUNG a  
WHERE (a.Hang = 'Gạo') AND DonGia <=
```

```
(SELECT AVG(b.DonGia)
FROM CUNGUNG b
WHERE b.Hang = 'Gạo')
```

❖ **Sử dụng các toán tử so sánh : <, <=, =, >, >=, <> với ANY / ALL**

- Liệt kê nhân viên từ bảng NHANVIEN có NgaySinh trùng với người khác

```
SELECT a.Manv, a.HoLot, a.Ten
FROM NHANVIEN a
WHERE a.NgaySinh = ANY
(SELECT b.NgaySinh
FROM NHANVIEN b
WHERE b.Manv <> a.Manv)
```

- Liệt kê nhân viên từ bảng NHANVIEN có NgaySinh không trùng với người khác

```
SELECT a.Manv, a.HoLot, a.Ten
FROM NHANVIEN a
WHERE a.NgaySinh <> ALL
(SELECT b.NgaySinh
FROM NHANVIEN b
WHERE b.Manv <> a.Manv)
```

❖ **Sử dụng toán tử IN (NOT IN) biểu diễn quan hệ**

- Liệt kê Ten, Hang và DonGia của các hãng cung ứng có ít nhất 1 mặt hàng trong HOPDONG :

```
SELECT DISTINCT Ten, Hang, DonGia
FROM CUNGUNG
WHERE Hang IN
(SELECT Hang
```

FROM HOPDONG)

- Liệt kê Ten, Hang và DonGia của các hãng cung ứng có ít nhất 1 mặt hàng mà khách hàng 'KS. Bạch Đằng' đặt hàng :

```
SELECT DISTINCT Ten, Hang, DonGia
FROM CUNGUNG
WHERE Hang IN
      (SELECT Hang
       FROM HOPDONG
       WHERE TenKh = 'KS. Bạch Đằng' )
```

- Liệt kê Ten, DiaChi các hãng cung ứng không bán Gạo :

```
SELECT DISTINCT Ten, DiaChi
FROM CUNGUNG
WHERE Ten NOT IN
      (SELECT Ten
       FROM CUNGUNG
       WHERE Hang = 'Gạo')
```

- Liệt kê hãng có ít nhất 1 mặt hàng giống Cty Lương thực :

```
SELECT DISTINCT a.Ten, a.DiaChi
FROM CUNGUNG a
WHERE a.Hang IN
      (SELECT b.Hang
       FROM CUNGUNG b
       WHERE b.Ten = 'Cty lương thực')
```

- Liệt kê hợp đồng có TenKh và Hang giống hợp đồng số 3 :

```
SELECT *
FROM HOPDONG a
WHERE (a.TenKh, a.Hang) IN
      (SELECT b.TenKh, b.Hang
```

```
FROM HOPDONG
WHERE b.SoHd = 3)
```

- Liệt kê Hãng không có mặt hàng trong HOPDONG :

```
SELECT DISTINCT Ten
FROM CUNGUNG
MINUS
SELECT DISTINCT Ten
FROM CUNGUNG
WHERE Hang IN
      (SELECT Hang
       FROM HOPDONG )
```

- Liệt kê Hãng có mặt hàng không có trong HOPDONG :

```
SELECT DISTINCT Ten
FROM CUNGUNG
WHERE Hang NOT IN
      (SELECT Hang
       FROM HOPDONG )
```

- Liệt kê Hãng, mà tất cả mặt hàng của nó đều có trong HOPDONG :

```
SELECT DISTINCT Ten
FROM CUNGUNG
GROUP BY Ten
      HAVING SET(Hang) IN
            (SELECT Hang
             FROM HOPDONG )
```

❖ **Sử dụng [NOT] EXISTS <bảng> biểu diễn sự tồn tại**

- Liệt kê Ten, DiaChi khách hàng đã đặt hàng :

```
SELECT Ten, DiaChi
FROM KHACH
```

```

WHERE EXISTS
    (SELECT Hang
     FROM HOPDONG
     WHERE TenKh = Ten)

```

- Liệt kê Ten, DiaChi khách hàng chưa có hợp đồng:

```

SELECT Ten, DiaChi
FROM KHACH
WHERE NOT EXISTS
    (SELECT Hang
     FROM HOPDONG
     WHERE TenKh = Ten)

```

6. Gom nhóm dữ liệu – mệnh đề Group by

Cú pháp:

```

SELECT [DISTINCT] <biểu thức 1> [AS <tên 1>] [...] | *
FROM <bảng 1> [<bí danh 1>] [...]
[WHERE <điều kiện nối > [AND / OR <điều kiện lọc>]]
GROUP BY <thuộc tính 1> [...]
        [HAVING <điều kiện nhóm>]

```

Công dụng:

Nhóm các bản ghi (thỏa mãn <điều kiện nối> hoặc <điều kiện lọc>, nếu có) theo danh sách thuộc tính sau GROUP BY , sau đó kết xuất kết quả tổng hợp (thỏa <điều kiện nhóm>).

Các thuộc tính sau GROUP BY phải có trong danh sách thuộc tính sau SELECT.

Ghi chú :

- Hàm tập hợp SET(...) (tập hợp các phần tử) chỉ sử dụng trong điều kiện nhóm.

- Các hàm tính toán có thể sử dụng sau SELECT hoặc trong điều kiện nhóm.

+ Ví dụ

- Liệt kê các hãng cùng tổng số mặt hàng mà nó cung ứng :

```
SELECT Ten, COUNT(Hang) AS SoHang
FROM CUNGUNG
GROUP BY Ten
```

- Liệt kê các mặt hàng cùng tổng SoLuong trong HOPDONG

```
SELECT Hang, SUM(SoLuong) AS TongCong
FROM HOPDONG
GROUP BY Hang
```

- Liệt kê các mặt hàng cùng giá bình quân trong CUNGUNG :

```
SELECT Hang, AVG(DonGia) AS DonGiaBQ
FROM CUNGUNG
GROUP BY Hang
```

- Liệt kê các mặt hàng cùng tổng số hãng cung ứng mặt hàng đó :

```
SELECT Hang, COUNT(Ten) AS SoCty
FROM CUNGUNG
GROUP BY Hang
```

- Liệt kê các mặt hàng có nhiều hãng cung ứng mặt hàng đó :

```
SELECT Hang, COUNT(Ten) AS SoCty
FROM CUNGUNG
GROUP BY Hang
```

```
HAVING COUNT(Ten) > 1
```

- Liệt kê các hãng cung ứng cả Đường và Sữa :

```
SELECT DISTINCT Ten
FROM CUNGUNG
GROUP BY Ten
```

HAVING SET(Hang) CONTAINS ('Đường', 'Sữa')

Bài tập và sản phẩm thực hành bài 3.1

I. Kiến thức:

Câu 1: SQL là gì? So sánh sự khác nhau giữa ngôn ngữ định nghĩa dữ liệu (DDL: Data Definition Language) và ngôn ngữ thao tác dữ liệu (DML: Data Manipulation Language)? Cho ví dụ?

Câu 2: Trình bày

- a) Các kiểu dữ liệu của SQL?
- b) Các toán tử logic của SQL?
- c) Hàm tập hợp của SQL?

Câu 3: Trình bày cú pháp và giải thích các thành phần cơ bản của lệnh SELECT?

II. Kỹ năng:

BÀI 1: Cho lược đồ cơ sở dữ liệu quản lý nhân viên của một công ty như sau:

Nhanvien(MANV,HOTEN, NU,NGAYSINH,LUONG,MAPB, MACV)

Mỗi nhân viên có một mã nhân viên (MANV) duy nhất, mỗi mã nhân viên xác định họ và tên nhân viên (HOTEN), giới tính (NU), lương (LUONG), mã phòng ban (MAPB), mã chức vụ (MACV).

Phongban(MAPB,TENPB,TRUSO,MANVPHUTRACH,KINHPHI,DOA
NHTHU)

Mỗi phòng ban có tên gọi phòng ban(TENPB), địa điểm đặt trụ sở (TRUSO), mã nhân viên phụ trách(MANVPHUTRACH), kinh phí hoạt động (KINHPHI), và doanh thu(DOANHTHU)

Chucvu(MACV,TENCV,LUONGTHAPNHAT,LUONGCAONHAT)

Mỗi chức vụ có tên gọi chức vụ (TENCV), mức lương tối thiểu(LUONGTHAPNHAT), mức lương tối đa (LUONGCAONHAT).

Hãy biểu diễn các câu hỏi sau bằng SQL

1.Lập danh sách gồm các thông tin về các phòng ban trong công ty như: mã số phòng ban, tên phòng ban, địa điểm trụ sở, mã số người phụ trách, kinh phí hoạt động, doanh thu.

2.Lập danh sách những nhân viên sinh nhật trong tháng 10

3.Lập danh sách gồm các thông tin mã số nhân viên, họ và tên và lương cả năm của các nhân viên (giả sử rằng lương cả năm =12*lương)

4.Lập những phòng ban có kinh phí hoạt động cao nhất.

5.Lập danh sách nhân viên của phòng ban có mã số phòng ban là 40.

6. Lập danh sách nhân viên của phòng có mã số phòng ban 10,30,50.

7.Lập danh sách các nhân viên có lương tháng từ 2.500.000 đến 4.000.000

8.Tìm những nhân viên có tuổi cao nhất thuộc phòng ban có MAPB là 10

9.Lập danh sách các nhân viên của phòng 10,30,50. kết quả in ra theo thứ tự tăng dần của mã phòng nếu trùng mã phòng thì sắp xếp giảm dần theo mức lương.

10.Lập danh sách các nhân viên phòng 10,30,50, chỉ in ra những người là lãnh đạo của mỗi phòng ban này.

11.lập danh sách gồm mã phòng mà người có mức lương cao nhất của phòng lớn hơn hoặc bằng 4.000.000

12.Lập mã phòng ban, tên phòng ban, họ và tên của lãnh đạo phòng tương ứng.

13.Lập danh sách những người làm việc cùng phòng với ông Nguyen Van Thanh

14.Lập biết mã số nhân viên, họ và tên, mức lương của người lãnh đạo ông Nguyen Van Thanh.

15.Lập danh sách nhân viên có mức lương lớn hơn hay bằng mức lương cao nhất của phòng ông Nguyen Van Thanh.

16.Cho biết mã số nhân viên, họ và tên , tổng số nhân viên, mức lương cao nhất, mức lương thấp nhất, mức lương trung bình của từng phòng ban.

17.Cho biết các nhân viên có mức lương cao nhất của các phòng ban.

Đánh giá kết quả học tập

TT	Tiêu chí đánh giá	Cách thức và phương pháp đánh giá	Điểm tối đa	Kết quả thực hiện của người học
I	Kiến thức			
1	Ngôn ngữ SQL	Vấn đáp, đối chiếu với nội dung bài học	3	
1.1	Định nghĩa SQL		1	
1.1	So sánh sự khác nhau ngôn ngữ DDL và DML		1	
1.2	Lấy ví dụ cụ thể		1	
2	Trình bày		4	
2.1	Các kiểu dữ liệu		1	

2.2	Các toán tử logic của SQL		1	
2.3	Hàm tập hợp của SQL	Vấn đáp, đối chiếu với nội dung bài học	0.5	
3	Trình bày câu lệnh Select	Vấn đáp, đối chiếu với nội dung bài học	3	
3.1	Cú pháp câu lệnh Select		1.5	
3.2	Giải thích các thành phần cơ bản của lệnh Select		1.5	
Cộng:			10 đ	
II	Kỹ năng			
1	Chuẩn bị máy tính, phần mềm đúng theo yêu cầu của bài thực hành	Kiểm tra công tác chuẩn bị, đối chiếu với kế hoạch đã lập	10	
2	Sử dụng thành thạo các dạng lệnh Select làm bài tập + Truy vấn đơn giản + Truy vấn có điều kiện + Truy vấn có sắp xếp + Truy vấn con lồng nhau + Truy vấn gom nhóm	Làm bài tập đối chiếu với nội dung bài học, thực hành phần mềm SQL Server	30	
Cộng:			40đ	
III	Thái độ			
1	Tác phong công nghiệp	Theo dõi việc thực hiện, đối chiếu với nội quy của trường.	4	
1.1	Đi học đầy đủ, đúng giờ		1,5	
1.2	Không vi phạm nội quy lớp học		1,5	
1.3	Tính cẩn thận, tỉ mỉ	Quan sát việc	1	

		thực hiện bài tập		
2	Đảm bảo thời gian thực hiện bài tập	Theo dõi thời gian thực hiện bài tập, đối chiếu với thời gian quy định.	2	
3	Đảm bảo an toàn lao động và vệ sinh công nghiệp	Theo dõi việc thực hiện, đối chiếu với quy định về an toàn và vệ sinh công nghiệp	4	
3.1	Tuân thủ quy định về an toàn		2	
3.3	Vệ sinh phòng thực hành đúng quy định		2	
Cộng:			10 đ	

KẾT QUẢ HỌC TẬP

Tiêu chí đánh giá	Kết quả thực hiện	Hệ số	Kết quả học tập
<i>Kiến thức</i>		0,3	
<i>Kỹ năng</i>		0,4	
<i>Thái độ</i>		0,3	
Cộng:			

Bài tập và sản phẩm thực hành bài 3.2

I. Kiến thức:

Câu 1: Trình bày cú pháp và công dụng câu lệnh truy vấn lồng nhau

Câu 2: Trình bày cú pháp và công dụng câu lệnh gom nhóm

Câu 3: Trình bày

a) Các toán tử tập hợp của SQL?

b) Các hàm tính toán của SQL?

II. Kỹ năng:

BÀI 1: Cho một lược đồ cơ sở dữ liệu của 1 thư viện như sau:

ThẻĐộcGiả(MãThẻDg, HọTênDg, Địachỉ)

Sách(Mãsách, Tênsách, NămXb, NhàXb, TênTácGiả)

Phiếumượn(MãPhiếu, NgàyMượn, MãSách, MãThẻDg, NgàyTrả, Trả)

Dùng SQL trả lời các câu hỏi sau đây:

1. Liệt kê các phiếu mượn trong ngày 20/11/2009
2. Liệt kê họ tên độc giả mượn sách trong tháng 3 năm 2009
3. Liệt kê các tên sách thuộc nhà xuất bản “Giáo Dục” hiện có trong thư viện
4. Liệt kê tên các độc giả hiện đang còn giữ sách có tên “Đồ Hoạ Máy Tính” chưa trả cho thư viện
5. Liệt kê tên và địa chỉ liên lạc của các độc giả hiện đang còn nợ sách của thư viện
6. Liệt kê tên các độc giả và tên sách mà các độc giả này đang giữ đã đến kỳ hạn trả cho thư viện. Biết rằng qui định là sau khi mượn 10 ngày thì phải trả sách lại cho thư viện
7. Thống kê xem mỗi thẻ độc giả đã mượn được bao nhiêu cuốn sách trong tháng 11/2009

8. Liệt kê các cuốn sách chưa được ai mượn lần nào
9. Liệt kê tất cả các cuốn sách được mượn nhiều nhất trong tháng 11/2009
10. Liệt kê các độc giả quen thuộc của thư viện. Biết rằng các độc giả quen thuộc là các độc giả đến mượn sách ít nhất là 3 lần

Đánh giá kết quả học tập

TT	Tiêu chí đánh giá	Cách thức và phương pháp đánh giá	Điểm tối đa	Kết quả thực hiện của người học
I	Kiến thức			
1	Trình bày cú pháp và công dụng câu lệnh truy vấn lồng nhau	Vấn đáp, đối chiếu với nội dung bài học	3	
2	Trình bày cú pháp và công dụng câu lệnh gom nhóm	Vấn đáp, đối chiếu với nội dung bài học	4	
3	Trình bày a) Các toán tử tập hợp của SQL? b) Các hàm tính toán	Vấn đáp, đối chiếu với nội dung bài học	3	
Cộng:			10 đ	
II	Kỹ năng			
1	Chuẩn bị máy tính, phần mềm đúng theo yêu cầu của bài thực hành	Kiểm tra công tác chuẩn bị, đối chiếu với kế	10	

		hoạch đã lập		
2	Sử dụng thành thạo các dạng lệnh Select làm bài tập + Truy vấn đơn giản + Truy vấn có điều kiện + Truy vấn có sắp xếp + Truy vấn con lồng nhau + Truy vấn gom nhóm	Làm bài tập đối chiếu với nội dung bài học, thực hành phần mềm SQL Server	30	
Cộng:			40 đ	
III	Thái độ			
1	Tác phong công nghiệp	Theo dõi việc thực hiện, đối	4	
1.1	Đi học đầy đủ, đúng giờ	chiếu với nội	1,5	
1.2	Không vi phạm nội quy lớp học	quy của trường.	1,5	
1.3	Tính cẩn thận, tỉ mỉ	Quan sát việc thực hiện bài tập	1	
2	Đảm bảo thời gian thực hiện bài tập	Theo dõi thời gian thực hiện bài tập, đối chiếu với thời gian quy định.	2	
3	Đảm bảo an toàn lao động và vệ sinh công nghiệp	Theo dõi việc thực hiện, đối	4	
3.1	Tuân thủ quy định về an toàn	chiếu với quy	2	
3.3	Vệ sinh phòng thực hành đúng	định về an toàn	2	

	quy định	và vệ sinh công nghiệp		
Cộng:			10 đ	

KẾT QUẢ HỌC TẬP

Tiêu chí đánh giá	Kết quả thực hiện	Hệ số	Kết quả học tập
<i>Kiến thức</i>		0,3	
<i>Kỹ năng</i>		0,4	
<i>Thái độ</i>		0,3	
Cộng:			

BÀI TẬP

R1=Nhân viên (MaNV, Ho_tên, Nsinh, nghề nghiệp, Địa chỉ, lương)

R2=Liên kết (MaNV, MaP)

R3=Phòng (Map, Tên_phong, tel)

1. Để hiện các thông tin về một nhân viên nào đó gồm (MaNV , Họ_tên, N_sinh, Chức_vụ, địa_chỉ, lương)

```
SELECT Distinc * FROM R1;
```

2. Đưa ra (họ_tên,Nsinh,chức_vụ,địa_chỉ,lương,tên_phòng) với điều kiện lương. 500.000 và địa_chỉ không ở Hà nội.

```
SELECT Ho_tên,Nsinh,chức_vụ,địa_chỉ,lương,tên_phòng
```

```
FROM Nhânviên R1,Liênkết R2,Phòng R3
```

```
WHERE (R1.lương. 500.000) and (not R1.địa_chỉ='Hà nội') and
```

```
(R1.MaNV=R2.MaNV) and (R2. MaP=R3.MaP);
```

- Trong lệnh trên ta đã dùng R1,R2,R3 làm bí danh cho Nhânviên, Liênkết, Phong

Các bí danh đó chỉ có tác dụng trong một câu lệnh

Các ví dụ sau này ta dùng R1,R2,R3 để thay cho các bảng trên cho gọn

Có 4 toán tử hay được dùng với các kiểu dữ liệu.Trong mệnh đề WHERE là:

In (not In)

Between..and..(not between..)

Like(not like)

Is null (not is Null).

+ Toán tử In (not In):Dùng để kiểm tra giá trị trong (không nằm trong) một danh sách được

chỉ ra.

3. Đưa ra những người có địa_chỉ ở Hà nội và Hà tây.

```
SELECT * FROM R1 WHERE địa_chỉ in ('Hà nội', 'Hà tây');
```

+Toán tử Between..and..(not ..) : kiểm tra giá trị nằm giữa (không nằm giữa) một phạm vi được chỉ ra.

4. Đưa ra những người có lương nằm trong không (500.000-:-1.000.000).

```
SELECT * FROM R1 WHERE lương between 500.000 and 1.000.000
```

BÀI TẬP TỰ GIẢI:

BÀI 1: Dựa vào lược đồ CSDL

Congtrinh(MACT, TENCT, DIADIEM, NGAYCAPGP,
NGAYKC, NGAYT)

Nhanvien(MANV, HOTEN, NGAYSINH, PHAI, DIACHI, MAPB)

Phongban(MAPB, TENPB)

Phancong(MACT, MANV, SLNGAYCONG)

Hãy thực hiện các câu hỏi sau bằng SQL

a. Danh sách những nhân viên có tham gia vào công trình có mã công trình(MACT) là X. Yêu cầu các thông tin: MANV, HOTEN, SLNGAYCONG, trong đó MANV được sắp tăng dần

b. Đếm số lượng ngày công của mỗi công trình. Yêu cầu các thông tin: MACT, TENCT, TONGNGAYCONG (TONGNGAYCONG là thuộc tính tự đặt)

c. Danh sách những nhân viên có sinh nhật trong tháng 08. yêu cầu các thông tin: MANV, TENNV, NGAYSINH, DIACHI, TENPB, sắp xếp quan hệ kết quả theo thứ tự tuổi giảm dần.

d. Đếm số lượng nhân viên của mỗi phòng ban. Yêu cầu các thông tin: MAPB, TENPB, SOLUONG. (SOLUONG là thuộc tính tự đặt.)

BÀI 2: Dựa vào lược đồ cơ sở dữ liệu

Giaovien(MAGV, HOTEN, MAKHOA)

Monhoc(MAMH, TENMH)

Phonghoc(PHONG,CHUCNANG)

Khoa(MAKHOA, TENKHOA)

Lop(MALOP, TENLOP, MAKHOA)

Lichday(MAGV, MAMH, PHONG, MALOP, NGAYDAY, TUTIET, ĐENTIET, BAIDAY, LYTHUYET, GHICHU)

Hãy thực hiện các câu hỏi sau bằng SQL

a.Xem lịch báo giảng tuần từ ngày 08/09/2003 đến ngày 14/09/2003 của giáo viên có MAGV (mã giáo viên) là TH3A040. Yêu cầu: MAGV,HOTEN, TENLOP,TENMH,PHONG, NGAYDAY, TUTIET, ĐENTIET, BAIDAY, GHICHU)

b.Xem lịch báo giảng ngày 08/09/2003 của các giáo viên có mã khoa là CNTT. Yêu cầu:MAGV,HOTEN,TENLOP,TENMH,PHONG, NGAYDAY, TUTIET, ĐENTIET,BAIDAY, GHICHU)

c.Cho biết số lượng giáo viên (SOLUONGGV) của mỗi khoa, kết quả cần sắp xếp tăng dần theo cột tên khoa. yêu cầu: TENKHOA ,SOLUONGGV (SOLUONGGV là thuộc tính tự đặt)

BÀI 3: Hàng năm, Trường X tổ chức kỳ thi giỏi nghề cho các học sinh- sinh viên của trường, mỗi thí sinh sẽ thi hai môn (chẳng hạn các thí sinh thi giỏi nghề công nghệ thông tin thi hai môn là Visual Basic và Cơ Sở Dữ Liệu).

Giả sử lược đồ cơ sở dữ liệu của bài toán quản lý các kỳ thi trên được cho như sau:

THISINH (MASV, HOTEN, NGAYSINH, MALOP)

LOP(MALOP,TENLOP,MAKHOA)

KHOA(MAKHOA,TENKHOA,ĐIENTHOAI)

MONTHI(MAMT,TENMONTHI)

KETQUA(MASV,MAMT,DIEMTHI)

(Phần giải thích các thuộc tính: HOTEN (họ tên thí sinh), NGAYSINH (ngày sinh), MALOP (mã lớp), MASV(mã sinh viên), TENLOP(tên lớp), MAKHOA(mã khoa), TENKHOA (tên khoa), DIENTHOAI(số điện thoại khoa), MAMT (mã môn thi), TENMONTHI (tên môn thi), DIEMTHI (điểm thi)).

Dựa vào lược đồ cơ sở dữ liệu trên, hãy thực hiện các yêu cầu sau bằng ngôn ngữ SQL:

a.Hãy cho biết số lượng thí sinh của mỗi khoa đăng ký thi giỏi nghề, cần sắp xếp kết quả theo chiều tăng dần của cột TENKHOA.

b.Lập danh sách những thí sinh đạt danh hiệu giỏi nghề (Thí sinh đạt danh hiệu giỏi nghề nếu thí sinh không có môn thi nào điểm dưới 8).

c.Lập danh sách những thí sinh nhỏ tuổi nhất có mã khoa là “CNTT” dự thi giỏi nghề.

BÀI 4:

Để quản lý kinh doanh dùng các bảng sau:

+ HH(hàng hoá): (MaHH CHAR(3), TenHH CHAR(20), Qcach CHAR(20), DVT CHAR(5), DGIA INT(10))

+ CH(cửa hàng): (MaCH CHAR(3), TenCH CHAR(20), DDiem CHAR(20), PTrach CHAR(4))

+ KH(khách hàng): (MaKH CHAR(4), TenKH CHAR(20), Loai CHAR(2), Diachi CHAR(20))

+ CT(chứng từ):(Sohieu CHAR(12), Ngay D, LoaiCT CHAR(1), MaKH CHAR(4), MaCH CHAR(3), MaHH CHAR(3), SoLuong INT(6)).

a) Xem trong bảng CT có những loại hàng hoá nào được xuất.

```
SELECT DISTINCT MaHH FROM CT
```

```
SELECT DISTINCT CT.MaHH, TenHH FROM CT, HH WHERE CT.MaHH =  
HH.MaHH
```

b) Đưa ra danh sách các nhân viên có lương ≥ 200000

```
SELECT * FROM NHANVIEN WHERE Luong  $\geq$  200000
```

c) Cho xem danh sách gồm 3 cột Mã đơn vị, họ tên, nhiệm vụ từ bảng nhân viên và được sắp xếp theo mã đơn vị, cùng đơn vị theo nhiệm vụ:

```
SELECT MaDV, Hoten, NHVu FROM NHANVIEN ORDER BY MaDV,  
NHVu
```


Chương 4. RÀNG BUỘC TOÀN VỆ

Mã chương MH16-04

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 vậy, luôn phải đảm bảo cơ sở dữ liệu thỏa ràng buộc toàn vẹn sau mỗi thao tác làm thay đổi tình trạng của cơ sở dữ liệu.

Mục tiêu:

- Trình bày được các thuật ngữ ràng buộc, ràng buộc toàn vẹn, hiểu được các khái niệm về ràng buộc toàn vẹn.

Nội dung:

1. Ràng buộc toàn vẹn

1.1. Khái niệm ràng buộc toàn vẹn

Trong mỗi CSDL luôn tồn tại nhiều mối liên hệ giữa các thuộc tính, giữa các bộ; sự liên hệ này có thể xảy ra trong cùng một quan hệ hoặc trong các quan hệ của một lược đồ CSDL. Các mối liên hệ này là những điều kiện bất biến mà tất cả các bộ của những quan hệ có liên quan trong CSDL đều phải thỏa mãn ở mọi thời điểm. Những điều kiện bất biến đó được gọi là ràng buộc toàn vẹn.

Trong thực tế ràng buộc toàn vẹn là các quy tắc quản lý được áp đặt trên các đối tượng của thế giới thực. Chẳng hạn mỗi sinh viên phải có một mã sinh viên duy nhất, hai thí sinh dự thi vào một trường phải có số báo danh khác nhau, một sinh viên dự thi một môn học không quá 3 lần,...

Ràng buộc toàn vẹn là một điều kiện bất biến không được vi phạm trong một CSDL.

1.2. Các yếu tố của ràng buộc toàn vẹn:

Mỗi ràng buộc toàn vẹn có bốn yếu tố: điều kiện, bối cảnh, bảng tâm ảnh hưởng và hành động phải cần thực hiện khi phát hiện có ràng buộc toàn vẹn bị vi phạm:

1.2.1. Điều kiện:

Điều kiện của ràng buộc toàn vẹn là sự mô tả, và biểu diễn hình thức nội dung của nó.

Điều kiện của một ràng buộc toàn vẹn R có thể được biểu diễn bằng ngôn ngữ tự nhiên, ngôn ngữ đại số quan hệ, ngôn ngữ mã giả, ngôn ngữ truy vấn SQL,... ngoài ra điều kiện của ràng buộc toàn vẹn cũng có thể được biểu diễn bằng phụ thuộc hàm.

Ví dụ sau đây chỉ ra một số ràng buộc toàn vẹn trên lược đồ CSDL quản lý sinh viên.

- Mỗi lớp học phải có một mã số duy nhất để phân biệt với các lớp học khác trong trường.
- Mỗi lớp học phải thuộc về một khoa của trường.
- Mỗi sinh viên có một mã số sinh viên duy nhất, không trùng với bất cứ sinh viên nào trong trường.
- Mỗi học viên phải đăng ký vào một lớp học trong trường. Mỗi học viên chỉ được thi tối đa 3 lần cho mỗi môn học.
- Tổng số học viên của một lớp phải lớn hơn hoặc bằng số lượng đếm được của một lớp tại một thời điểm nào đó.

1.2.2. Bối cảnh:

Bối cảnh của ràng buộc toàn vẹn là những quan hệ mà ràng buộc đó có hiệu lực hay nói một cách khác, đó là những quan hệ cần phải được kiểm tra khi tiến hành cập nhật dữ liệu. Bối cảnh của một ràng buộc toàn vẹn có thể là một hoặc nhiều quan hệ.

Chẳng hạn với ràng buộc toàn vẹn R trên thì bối cảnh của nó là quan hệ Sinhvien

1.2.3. Bảng tầm ảnh hưởng

Trong quá trình phân tích thiết kế một CSDL, người phân tích cần lập bảng tầm ảnh hưởng cho một ràng buộc toàn vẹn nhằm xác định thời điểm cần phải tiến hành kiểm tra khi tiến hành cập nhật dữ liệu.

Thời điểm cần phải kiểm tra ràng buộc toàn vẹn chính là thời điểm cập nhật dữ liệu.

Một bảng tầm ảnh hưởng của một ràng buộc toàn vẹn có dạng sau:

Tên RBTV	Thêm(T)	Sửa(S)	Xoá(X)
r_1	+		
r_2		-	
r_3			-(*)
r_n			

Bảng này chứa toàn các ký hiệu + , - hoặc -(*).

- Dấu + tại (dòng r_1 , cột Thêm) thì có nghĩa là khi thêm một bộ vào quan hệ r_1 thì RBTV bị vi phạm.

- Dấu - tại ô (dòng r_2 , cột Sửa) thì có nghĩa là khi sửa một bộ trên quan hệ r_2 thì RBTV không bị vi phạm.

Quy ước:

- Không được sửa thuộc tính khoá.

- Nếu không bị vi phạm do không được phép sửa đổi thì ký hiệu là -(*).

1.2.4. Hành động cần phải có khi phát hiện có RBTV bị vi phạm:

Khi một ràng buộc toàn vẹn bị vi phạm, cần có những hành động thích hợp. Thông thường có 2 giải pháp:

- Thứ nhất: Đưa ra thông báo và yêu cầu sửa chữa dữ liệu của các thuộc tính cho phù hợp với quy tắc đảm bảo tính nhất quán dữ liệu. Thông báo phải đầy đủ và phải thân thiện với người sử dụng. Giải pháp này là phù hợp cho việc xử lý thời gian thực.

- Thứ hai: Từ chối thao tác cập nhật. Giải pháp này là phù hợp đối với việc xử lý theo lô. Việc từ chối cũng phải được lưu lại bằng những thông báo đầy đủ, rõ ràng vì sao thao tác bị từ chối và cần phải sửa lại những dữ liệu nào?

Khóa nội, khoá ngoại, giá trị NOT NULL là những ràng buộc toàn vẹn miền giá trị của các thuộc tính. Những ràng buộc toàn vẹn này là những ràng buộc toàn vẹn đơn giản trong CSDL.

Các hệ quản trị cơ sở dữ liệu thường có các cơ chế tự động kiểm tra các ràng buộc toàn vẹn về miền giá trị của khóa nội, khoá ngoại, giá trị NOT NULL.

Việc kiểm tra ràng buộc toàn vẹn có thể tiến hành vào những thời điểm sau đây:

- Thứ nhất: Kiểm tra ngay sau khi thực hiện một thao tác cập nhật CSDL. Thao tác cập nhật chỉ được xem là hợp lệ nếu như nó không vi phạm bất cứ một ràng buộc toàn vẹn nào, nghĩa là nó không làm mất tính toàn vẹn của CSDL. Nếu vi phạm ràng buộc toàn vẹn, thao tác cập nhật bị coi là không hợp lệ và sẽ bị hệ thống huỷ bỏ (hoặc có một xử lý thích hợp nào đó).

- Thứ hai: Kiểm tra định kỳ hay đột xuất, nghĩa là việc kiểm tra ràng buộc toàn vẹn được tiến hành độc lập với thao tác cập nhật dữ liệu. Đối với những trường hợp vi phạm ràng buộc toàn vẹn, hệ thống có những xử lý ngầm định hoặc yêu cầu người sử dụng xử lý những sai sót một cách tường minh.

2. Phân loại các ràng buộc toàn vẹn:

Xét ví dụ sau đây: Cho một CSDL C dùng để quản lý việc đặt hàng và giao hàng của một công ty. Lược đồ CSDL C gồm các lược đồ quan hệ như sau:

Q1: Khách (MAKH, TENKH, DIACHIKH, DIENTHOAI)

Tên từ:

Mỗi khách hàng có một mã khách hàng (MAKH) duy nhất, mỗi MAKH xác định tên khách hàng (TENKH), địa chỉ (DIACHIKH), số điện thoại (DIENTHOAI).

Q2: Hang(MAHANG,TENHANG,QUYCACH, DVTINH)

Tên từ:

Mỗi mặt hàng có một mã hàng (MAHANG) duy nhất, mỗi MAHANG xác định tên hàng (TENHANG), quy cách hàng (QUYCACH), đơn vị tính (DVTINH).

Q3: Dathang(SODH,MAHANG, SLDAT, NGAYDH, MAKH)

Tên từ:

Mỗi mã số đặt hàng (SODH) xác định một ngày đặt hàng (NGAYDH) và mã khách hàng tương ứng (MAKH). Biết mã số đặt hàng và mã mặt hàng thì biết được số lượng đặt hàng(SLDAT). Mỗi khách hàng trong một ngày có thể có nhiều lần đặt hàng

Q4: Hoadon(SOHD, NGAYLAP, SODH, TRIGIAHD, NGAYXUAT)

Tên từ:

Mỗi hoá đơn tổng hợp có một mã số duy nhất là SOHD, mỗi hoá đơn bán hàng có thể gồm nhiều mặt hàng. Mỗi hoá đơn xác định ngày lập hoá đơn (NGAYLAP), ứng với số đặt hàng nào (SODH). Giả sử rằng hoá đơn bán hàng theo yêu cầu của chỉ một đơn đặt hàng có mã số là SODH và ngược lại, mỗi đơn đặt hàng chỉ được giải quyết chỉ trong một hoá đơn. Do điều kiện khách quan có thể công ty không giao đầy đủ các mặt hàng cũng như số lượng từng mặt hàng như yêu cầu trong đơn đặt hàng nhưng không bao giờ giao vượt ngoài yêu cầu. Mỗi hóa đơn xác định một trị giá của nhưng các mặt hàng trong hoá đơn (TRIGIAHD) và một ngày xuất kho giao hàng cho khách (NGAYXUAT)

Q5: Chitiethd (SOHD, MAHANG, GIABAN, SLBAN)

Tên từ:

Mỗi SOHD, MAHANG xác định giá bán (GIABAN) và số lượng bán (SLBAN) của một mặt hàng trong một hoá đơn.

Q6: Phiethu(SOPT, NGAYTHU, MAKH, SOTIEN)

Tên từ:

Mỗi phiếu thu có một số phiếu thu (SOPT) duy nhất, mỗi SOPT xác định một ngày thu (NGAYTHU) của một khách hàng có mã khách hàng là MAKH và số tiền thu là SOTIEN. Mỗi khách hàng trong một ngày có thể có nhiều số phiếu thu.

2.1. Ràng buộc toàn vẹn có bối cảnh là một quan hệ

2.1.1. Ràng buộc toàn vẹn liên bộ:

+ Ràng buộc toàn vẹn về khoá chính:

Đây là một trường hợp đặc biệt của Ràng Buộc toàn Vẹn liên bộ, RBTV này rất phổ biến và thường được các hệ quản trị CSDL tự động kiểm tra.

Ví dụ:

Với r là một quan hệ trên lược đồ quan hệ Khách ta có RBTV sau:

$$R_1: \quad \forall t_1, t_2 \in r$$

$$t_1. \text{MAKH} \neq t_2. \text{MAKH}$$

Cuối \forall

R_1	Thêm	Sửa	Xoá
Khach	+	-	-

+ Ràng buộc toàn vẹn về tính duy nhất

Ví dụ: mỗi phòng ban phải có một tên gọi duy nhất

+ Ngoài ra nhiều khi ta còn gặp những RBTV khác chẳng hạn như các RBTV trong quan hệ sau đây.

Ví dụ: KETQUA(MASV,MAMH,LANTHI,DIEM)

Mỗi sinh viên chỉ được đăng thi mỗi môn tối đa là 3 lần.

2.1.2. Ràng buộc toàn vẹn về miền giá trị:

Ràng buộc toàn vẹn có liên quan đến miền giá trị của các thuộc tính trong một quan hệ. Ràng buộc này thường gặp. Thông thường các hệ quản trị CSDL đã tự động kiểm tra (một số) ràng buộc loại này.

Ví dụ: Với r là một quan hệ của Hoadon ta có ràng buộc toàn vẹn sau

$$R_3: \forall t \in r$$

$$t.TRIGIAHD > 0$$

Cuối \forall

R ₃	Thêm	Sửa	Xoá
Hoadon	+	+	-

2.1.3. Ràng buộc toàn vẹn liên thuộc tính:

Ràng buộc toàn vẹn liên thuộc tính (một quan hệ) là mối liên hệ giữa các thuộc tính trong một lược đồ quan hệ.

Ví dụ: Với r là một quan hệ của Hoadon ta có ràng buộc toàn vẹn sau:

$$R_4: \forall t \in r$$

$$t.NGAYLAP \leq t.NGAYXUAT$$

Cuối \forall

R ₄	Thêm	Sửa	Xoá
Hoadon	+	+	-

2.2. Ràng

buộc toàn vẹn có

bối cảnh là nhiều quan hệ:

2.2.1. Ràng buộc toàn vẹn về khoá ngoại:

Ràng buộc toàn vẹn về khoá ngoại còn được gọi là ràng buộc toàn vẹn phụ thuộc tồn tại. Cũng giống như ràng buộc toàn vẹn về khoá nội, loại ràng buộc toàn vẹn này rất phổ biến trong các CSDL.

Ví dụ:

$$R_2. \text{ dathang[MAKH]} \subseteq \text{khach[MAKH]}$$

R ₂	Thêm	Sửa	Xoá
dathang	+	+	-
Khach	-	-	+

2.2.2. Ràng buộc toàn vẹn liên thuộc tính liên quan hệ:

Ràng buộc loại này là mối liên hệ giữa các thuộc tính trong nhiều lược đồ quan hệ.

Ví dụ: Với r,s lần lượt là quan hệ của Dathang và Hoadon. Ta có RBTV R_5 như sau:

$$R_5: \forall t_1 \in r, t_2 \in s$$

$$\text{Nếu } t_1.SODH=t_2.SODH \text{ thì}$$

$$t_1.NGAYDH \leq t_2.NGAYXUAT$$

Cuối \forall

R_5	Thêm	Sửa	Xoá
Dathang	+	-	-
Hoandon	+	+	-

2.2.3. Ràng buộc toàn vẹn liên bộ liên quan hệ:

Ràng buộc loại này là mối liên hệ giữa các bộ trong một lược đồ cơ sở dữ liệu. Chẳng hạn như tổng số tiền phải trả trong mỗi hoá đơn (chitiethd) phải bằng TRỊ GIÁ HOÁ ĐƠN của hoá đơn đó trong quan hệ Hoadon. Hoặc số lượng học viên trong một lớp phải bằng SOHOCVIEN của lớp đó.

Ngoài ra còn có một số loại RBTV khác như :RBTV về thuộc tính tổng hợp, RBTV do tồn tại chu trình ,RBTV về giá trị thuộc tính theo thời gian.

Bài tập và sản phẩm thực hành bài 4.1

I. Kiến thức:

Câu 1: Ràng buộc toàn vẹn là gì? Các yếu tố của ràng buộc toàn vẹn?

Câu 2: Phân loại và cho ví dụ minh họa các ràng buộc toàn vẹn

II. Kỹ năng:

Bài 1: . Việc tổ chức kỳ thi tốt nghiệp của một khoa như sau:

Mỗi thí sinh có một Mã số sinh viên duy nhất (MASV), mỗi MASV xác định được các thông tin: họ và tên (HOTEN), ngày sinh (NGAYSINH), nơi sinh, phái, dân tộc.

Mỗi lớp có một mã lớp (MALOP) duy nhất, mỗi mã lớp xác định các thông tin: tên lớp (TENLOP), mỗi lớp chỉ thuộc sự quản lý của một khoa nào đó. Mỗi khoa có một mã khoa duy nhất (MAKHOA), mỗi mã khoa xác định tên khoa (TENKHOA).

Mỗi thí sinh đều phải dự thi tốt nghiệp ba môn. Mỗi môn thi có một mã môn thi (MAMT) duy nhất, mỗi mã môn thi xác định các thông tin: tên môn thi (TENMT), thời gian làm bài – được tính bằng phút (PHUT), ngày thi (NGAYTHI), buổi thi (BUOITHI), môn thi này là môn lý thuyết hay thực hành (LYTHUYET). Chú ý rằng, nếu một môn học được cho thi ở nhiều hệ thì được đặt MAMT khác nhau (chẳng hạn cả trung cấp và cao đẳng ngành công nghệ thông tin đều thi môn Cơ Sở Dữ Liệu), để diễn tả điều này, mỗi mã môn học cần phải được ghi chú (GHICHU) để cho biết môn thi đó dành cho khối nào trung cấp, hay cao đẳng). Mỗi thí sinh ứng với một môn thi có một điểm thi (DIEMTHI) duy nhất, điểm thi được chấm theo thang điểm 10 và có lấy điểm lẻ đến 0.5. Một thí sinh được coi là đậu tốt nghiệp nếu điểm thi của tất cả các môn của thí sinh đó đều lớn hơn hoặc bằng 5.

Trong một phòng thi có thể có thí sinh của nhiều lớp. Trong một kỳ thi, mỗi thí sinh có thể thi tại những phòng thi (PHONGTHI) khác nhau, chẳng hạn một thí sinh thi tốt nghiệp ba môn là Cơ sở dữ liệu, Lập trình C và Visual

Basic thì môn Cơ Sở Dữ Liệu và Lập Trình C thi tại phòng A3.4, còn môn thực hành Visual Basic thi tại phòng máy H6.1

Qua phân tích sơ bộ trên, ta có thể lập một lược đồ cơ sở dữ liệu như sau:

THISINH (MASV, HOTEN, NGAYSINH, MALOP)

LOP (MALOP, TENLOP)

MONTHI (MAMT, TENMT, LYTHUYET, PHUT, NGAYTHI, BUOITHI, GHICHU)

KETQUA (MASV, MAMT, DIEMTHI)

a. Tìm khoá cho mỗi lược đồ quan hệ trên.

b. Hãy phát biểu các ràng buộc toàn có trong cơ sở dữ liệu trên.

Đánh giá kết quả học tập

TT	Tiêu chí đánh giá	Cách thức và phương pháp đánh giá	Điểm tối đa	Kết quả thực hiện của người học
I	Kiến thức			
1	Ràng buộc toàn vẹn	Vấn đáp, đối chiếu với nội dung bài học	4	
1.1	Khái niệm RBTV		2	
1.2	Các yếu tố của RBTV		2	
2	Các loại RBTV	Vấn đáp, đối chiếu với nội dung bài học	6	
2.1	Phân loại: liệt kê, trình bày đặc điểm của từng loại		3	

2.2	Ví dụ Minh họa cho từng loại RBTV		3	
Cộng:			10 đ	
II	Kỹ năng			
1	+ Phân tích được yêu cầu bài toán. + Xác định được các khóa chính cho các quan hệ đã chỉ ra trong lược đồ. + Phát biểu được các RBTV có trong cơ sở dữ liệu đã chỉ ra.	Làm bài tập đối chiếu với nội dung bài học.	10	
Cộng:			10đ	
III	Thái độ			
1	Tác phong công nghiệp	Theo dõi việc	4	
1.1	Đi học đầy đủ, đúng giờ	thực hiện, đối	1,5	
1.2	Không vi phạm nội quy lớp học	chiếu với nội quy của trường.	1,5	
1.3	Tính cẩn thận, tỉ mỉ	Quan sát việc thực hiện bài tập	1	
2	Đảm bảo thời gian thực hiện bài tập	Theo dõi thời gian thực hiện bài tập, đối chiếu với thời gian quy định.	2	
3	Đảm bảo an toàn lao động và	Theo dõi việc	4	

	vệ sinh công nghiệp	thực hiện, đối		
3.1	Tuân thủ quy định về an toàn	chiếu với quy	2	
3.3	Vệ sinh phòng học đúng quy định	định về an toàn và vệ sinh công nghiệp	2	
Cộng:			10 đ	

KẾT QUẢ HỌC TẬP

Tiêu chí đánh giá	Kết quả thực hiện	Hệ số	Kết quả học tập
<i>Kiến thức</i>		0,3	
<i>Kỹ năng</i>		0,4	
<i>Thái độ</i>		0,3	
Cộng:			

BÀI TẬP TỰ GIẢI:

Bài 2: Quản lý đăng ký chuyên đề

Phòng giáo vụ tại một trường đại học muốn tin học hóa việc quản lý học các chuyên đề của sinh viên. Sau đây là kết quả của việc phân tích thiết kế ứng dụng trên.

Mỗi sinh viên có một mã số duy nhất, một họ tên, thuộc một phái, có một ngày sinh, một địa chỉ và học một ngành duy nhất.

Mỗi ngành có một mã ngành duy nhất, có một tên ngành duy nhất. Ngoài ra cũng cần lưu lại một con số cho biết số chuyên đề mà một sinh viên theo học một ngành cụ thể phải học, và cũng cần lưu lại tổng số sinh viên đã từng theo học ngành này.

Sinh viên phải học các chuyên đề khác nhau. Mỗi chuyên đề có một mã duy nhất và có một tên duy nhất. Cần lưu lại tên về số sinh viên tối đa có thể chấp nhận được mỗi khi có một lớp mở cho chuyên đề cụ thể.

Mỗi chuyên đề có thể được học bởi sinh viên thuộc nhiều ngành và sinh viên thuộc mỗi ngành phải học nhiều chuyên đề. Mỗi ngành học tối đa là 8 chuyên đề.

Vào mỗi học kỳ của mỗi năm học, ta cần lưu lại các chuyên đề nào được mở ra cho học kỳ của năm đó để sinh viên có thể đăng ký. Sinh viên chỉ được đăng ký những chuyên đề có mở.

Khi sinh viên đăng ký học, lưu lại việc đăng ký học một chuyên đề của một sinh viên vào một năm của một học kỳ nào đó.

Một sinh viên chỉ được đăng ký vào các chuyên đề thuộc ngành học của sinh viên đó mà thôi. Mỗi năm có 2 học kỳ. Sinh viên chỉ được đăng ký tối đa là 3 chuyên đề trong một học kỳ mà thôi.

1. Hãy thiết kế mô hình ER cho ứng dụng trên.
2. Chuyển mô hình ER sang mô hình quan hệ. Xác định khóa chính, khóa ngoại và liệt kê có phân loại tất cả ràng buộc toàn vẹn nhận diện được.

Chương 5. LÝ THUYẾT THIẾT KẾ CƠ SỞ DỮ LIỆU

Mã chương MH16-05

Giới thiệu:

Trong chương này trình bày những khái niệm cơ bản nhất về mô hình dữ liệu quan hệ của E.F.Codd. Những khái niệm cơ bản này gồm các khái niệm về quan hệ, phụ thuộc hàm, hệ tiên đề Armstrong, bao đóng, khoá, các dạng chuẩn của quan hệ,...

Những khái niệm này đóng vai trò rất quan trọng trong mô hình dữ liệu quan hệ. Chúng được dùng nhiều trong việc thiết kế các hệ quản trị cơ sở dữ liệu (CSDL) hiện nay.

Mục tiêu:

- Mô tả được khái niệm cơ sở của lý thuyết cơ sở dữ liệu như khoá, phụ thuộc hàm, bao đóng, các dạng chuẩn,..
- Trình bày và thiết kế được dữ liệu ở mức tốt nhất (có thể ứng dụng được) bằng các phép tách, giải thuật chuẩn hóa lược đồ.

Nội dung:

1. Các vấn đề gặp phải khi tổ chức dữ liệu:

Khi thiết kế, tổ chức cơ sở dữ liệu quan hệ ta thường đứng trước vấn đề lựa chọn các lược đồ quan hệ: lược đồ nào tốt hơn? Tại sao? Mục này sẽ nghiên cứu một số tiêu chuẩn đánh giá lược đồ quan hệ và các thuật toán giúp chúng ta xây dựng được lược đồ cơ sở dữ liệu quan hệ có cấu trúc tốt.

Có thể nói tổng quát, một lược đồ quan hệ có cấu trúc tốt là lược đồ không chứa sự dư thừa dữ liệu và các dị thường dữ liệu.

- Dư thừa dữ liệu là sự trùng lặp thông tin trong cơ sở dữ liệu.
- Dị thường dữ liệu là các sự cố xảy ra khi cập nhật dữ liệu (lập, dị thường chèn bộ, dị thường xóa bộ, dị thường sửa bộ) làm cho dữ liệu không tương thích, bất định hoặc mất mát.

+ Dị thường do dữ liệu lặp: một số thông tin có thể bị lặp lại một cách vô ích.

+ Dị thường chèn bộ: không thể chèn bộ mới vào quan hệ, nếu không có đầy đủ dữ liệu.

+ Dị thường xóa bộ: ngược lại với dị thường chèn bộ, việc xóa bộ có thể dẫn đến mất thông tin.

+ Dị thường sửa bộ: việc sửa đổi dữ liệu dư thừa có thể dẫn đến sự không tương thích dữ liệu.

Cơ sở lý thuyết của việc thiết kế lược đồ cơ sở dữ liệu quan hệ tốt là khái niệm phụ thuộc dữ liệu. Phụ thuộc dữ liệu biểu diễn các quan hệ nhân quả giữa các thuộc tính trong quan hệ. Cũng dựa trên khái niệm phụ thuộc dữ liệu người ta định nghĩa các dạng chuẩn của lược đồ quan hệ. Còn quá trình biến đổi lược đồ thành lược đồ tương đương thỏa mãn dạng chuẩn gọi là quá trình chuẩn hóa lược đồ quan hệ.

2. Phụ thuộc hàm

2.1. Định nghĩa phụ thuộc hàm

Cho lược đồ quan hệ $R=(A_1, A_2, \dots, A_n)$ và X, Y là các tập con của $R^+ = \{A_1, A_2, \dots, A_n\}$. Ta nói rằng X xác định hàm Y hay Y phụ thuộc hàm X , ký hiệu $X \twoheadrightarrow Y$, nếu mọi quan hệ bất kỳ r của lược đồ R thỏa mãn:

$$u, v \in r : u(X) = v(X) \implies u(Y) = v(Y)$$

Phụ thuộc hàm $X \twoheadrightarrow Y$ gọi là phụ thuộc hàm *tầm thường* nếu $Y \subseteq X$ (hiển nhiên là nếu $Y \subseteq X$ thì theo định nghĩa ta có $X \twoheadrightarrow Y$).

Phụ thuộc hàm $X \twoheadrightarrow Y$ gọi là phụ thuộc hàm *nguyên tố* nếu không có tập con thực sự $Z \subseteq X$ thỏa $Z \twoheadrightarrow Y$.

Tập thuộc tính $K \subseteq R$ gọi là *khóa* nếu nó xác định hàm tất cả các thuộc tính và $K \subseteq R$ là phụ thuộc hàm nguyên tố.

2.2. Cách xác định phụ thuộc hàm cho lược đồ quan hệ

Cách duy nhất để xác định đúng các phụ thuộc thích hợp cho một lược đồ quan hệ là xem xét nội dung tân từ của lược đồ quan hệ đó.

Ví dụ một số phụ thuộc hàm ứng với từng lược đồ quan hệ được xác định như sau:

MASV \rightarrow HOTENSV, NGAYSINH, MALOP, GIOITINH

MALOP \rightarrow TENLOP, MAKHOA

2.3. Một số tính chất của phụ thuộc hàm – hệ luật dẫn Armstrong

Để có thể xác định được các phụ thuộc hàm khác từ tập phụ thuộc hàm đã có, ta sử dụng các quy tắc suy diễn đơn giản để kiểm tra xem một phụ thuộc hàm có được suy diễn logic từ F hay không.

Một trong các quy tắc suy diễn đó gọi là hệ tiên đề Armstrong(1974), gồm các luật sau:

1. Luật phản xạ (reflexivity) $X \rightarrow Y \Rightarrow X \rightarrow Y$

2. Luật tăng trưởng (augmentation) $X \rightarrow Y \Rightarrow XZ \rightarrow YZ$

3. Luật bắc cầu (transitivity) $X \rightarrow Y, Y \rightarrow Z \Rightarrow X \rightarrow Z$

Các quy tắc suy rộng:

4. Luật hợp (the union rule) Cho $X \rightarrow Y, X \rightarrow Z \Rightarrow X \rightarrow YZ$

5. Luật bắc cầu giả (the pseudotransitivity rule)

Cho $X \rightarrow Y, WY \rightarrow Z \Rightarrow XW \rightarrow Z$

6. Luật phân rã (the decomposition rule)

Cho $X \rightarrow Y, Z \rightarrow Y \Rightarrow X \rightarrow Z$

Với $X, Y, Z, W \in R^+$

Ví dụ:

Cho lược đồ R(ABC) và $F = \{AB \rightarrow C, C \rightarrow A\}$. Dùng các quy tắc Armstrong ta chứng minh rằng $(B,C) \rightarrow (A,B,C)$.

Thật vậy, ta có

	C	A	(theo giả thiết)
	BC	AB	(theo luật tăng trưởng)
	C	C	(theo luật phản xạ)
=>	BC	ABC (đccm)	(theo luật hợp)

3. Bao đóng của tập phụ thuộc hàm và bao đóng của tập thuộc tính

3.1. Bao đóng của tập phụ thuộc hàm F

Bao đóng của tập phụ thuộc hàm F , ký hiệu là F^+ , là tập hợp tất cả các phụ thuộc hàm suy diễn logic từ F :

$$F^+ = \{X \rightarrow Y \mid F \models X \rightarrow Y\}$$

Hay nói cách khác: Bao đóng (closure) của tập phụ thuộc hàm F (ký hiệu là F^+) là tập hợp tất cả các phụ thuộc hàm có thể suy ra từ F dựa vào các tiên đề Armstrong. Rõ ràng $F \subseteq F^+$

Ví dụ: Cho $R=(A,B,C)$ và $F = \{A \rightarrow B, B \rightarrow C\}$. Khi đó bao đóng F^+ gồm các phụ thuộc hàm $X \rightarrow Y$ thoả

(i) X chứa A , Y bất kỳ:

A,B,C \rightarrow A,B,C;	A,B,C \rightarrow A,B;	A,B,C \rightarrow A,C;	A,B,C \rightarrow B,C;
A,B,C \rightarrow A;	A,B,C \rightarrow B;	A,B,C \rightarrow B;	A,B,C \rightarrow C;
A,B \rightarrow A,B,C;	A,B \rightarrow A,B;	A,B \rightarrow A,C;	A,B \rightarrow B,C;
A,B \rightarrow A;	A,B \rightarrow B;	A,B \rightarrow B;	A,B \rightarrow C;
A,C \rightarrow A,B,C;	A,C \rightarrow A,B;	A,C \rightarrow A,C;	A,C \rightarrow B,C;
A,C \rightarrow A;	A,C \rightarrow B;	A,C \rightarrow B;	A,C \rightarrow C;
A \rightarrow A,B,C;	A \rightarrow A,B;	A \rightarrow A,C;	A \rightarrow B,C;
A \rightarrow A;	A \rightarrow B;	A \rightarrow B;	A \rightarrow C;

(ii) X chứa B nhưng không chứa A , Y không chứa A :

BC \rightarrow BC;	BC \rightarrow B;	BC \rightarrow C
----------------------	---------------------	--------------------

B BC; B B; B C

(iii) C C

Về mặt lý thuyết ta hoàn toàn có thể xây dựng thủ tục tính bao đóng F^+ của tập phụ thuộc hàm F , nhưng trên thực tế bài toán xác định F^+ là không khả thi vì với số thuộc tính và phụ thuộc hàm lớn sẽ dẫn đến bùng nổ tổ hợp.

Thay vào đó chúng ta sẽ xem xét một bài toán khác: "Kiểm tra xem một phụ thuộc hàm có thuộc bao đóng F^+ hay không?". Bài toán này gọi là *bài toán thành viên*. Bài toán thành viên thiết thực hơn bài toán tính bao đóng vì trong thực tế rất hiếm khi phải tìm tất cả các phụ thuộc hàm suy diễn lô-gic từ F .

Bài toán thành viên liên quan mật thiết với khái niệm bao đóng của tập thuộc tính.

3.2. Bao đóng của tập thuộc tính X

Bao đóng của tập thuộc tính X R (đối với tập phụ thuộc hàm F), ký hiệu là X_F^+ (X^+), là tập hợp tất cả các thuộc tính phụ thuộc hàm vào X :

$$X^+ = \{A \mid X \vdash A \mid F^+\}$$

Từ định nghĩa dễ dàng suy ra: $X \subseteq X^+$ và $X \cup Y \subseteq X^+ \cup Y^+$. Nghĩa là X^+ là tập thuộc tính lớn nhất phụ thuộc hàm vào X .

Ví dụ: Cho $R(ABC)$ và $F = \{A \rightarrow B, B \rightarrow C\}$. Khi đó ta dễ dàng thấy bao đóng của thuộc tính B là $B^+ = \{B, C\}$

vì $B \rightarrow \{B, C\}$ và B không xác định A .

3.3. Bài toán thành viên

Qua phần trên ta nhận thấy X^+ được định nghĩa thông qua F^+ . Vấn đề nảy sinh khi nghiên cứu lý thuyết CSDL là: Cho trước tập các phụ thuộc hàm F và một phụ thuộc hàm f , bài toán kiểm tra có hay không $f \in F^+$ gọi là *bài toán thành viên*.

Để giải quyết bài toán bài toán thành viên thật sự không đơn giản;

vì mặc dù F là rất nhỏ nhưng F^+ thì có thể rất lớn. Tuy nhiên ta có thể giải bằng cách tính X^+ và so sánh X^+ với tập Y . Dựa vào tính chất $X \rightarrow Y \iff X^+ \subseteq Y$, ta có ngay câu trả lời $X \rightarrow Y$ hay không? Như vậy thay vì giải bài toán thành viên ta đưa về giải bài toán tìm bao đóng của tập thuộc tính.

3.4. Thuật toán tìm bao đóng của một tập thuộc tính

Thuật toán tìm bao đóng với độ phức tạp $O(N^2)$, với N là số lượng thuộc tính của lược đồ quan hệ Q .

Dữ Liệu Vào	Q, F, X	Q^+
Dữ Liệu Ra	X^+	

```

Bước 1:   Đặt  $X^+ = X$ 
Bước 2:   Temp =  $X^+$ 
           $\forall f \ U \rightarrow V \in F$ 
            if ( $U \subseteq X^+$ )
               $X^+ = X^+ \cup V.$ 
               $F = F - f;$ 
Bước 3:   if ( $X^+ = \text{Temp}$ )
            "X+ chính là kết quả cần tìm"
            Dừng thuật toán
          else
            trở lại Bước 2:
  
```

Ví dụ:

Cho lược đồ quan hệ $Q(ABCDEFGH)$ và tập phụ thuộc hàm

$F = \{B \rightarrow A, DA \rightarrow CE, D \rightarrow H, GH \rightarrow C, AC \rightarrow D\}.$

Tìm bao đóng của các tập $X = \{AC\}$ dựa trên F .

Giải:

- $X^+ = AC$

- Đặt $Temp = X^+$

+ Xét $AC \rightarrow D$, có $AC \quad X^+: X^+ = X^+ \quad D = ACD$.

Loại $AC \rightarrow D$ khỏi F. Lặp bước 2

+ Xét $DA \rightarrow CE$, có $DA \quad X^+: X^+ = X^+ \quad CE = ACDE$.

Loại $DA \rightarrow CE$ khỏi F. Lặp bước 2

+ Xét $D \rightarrow H$, có $D \quad X^+: X^+ = X^+ \quad H = ACDEH$.

Loại $D \rightarrow H$ khỏi F Lặp bước 2

Vì các phụ thuộc hàm $U \rightarrow V$ còn lại không thỏa điều kiện $U \quad X^+$ nên $X^+ = Temp$. Thuật toán dừng.

Vậy $X^+ = \{ACDEH\}$

4. Khóa của lược đồ quan hệ - một số thuật toán tìm khóa

4.1. Định nghĩa khóa của quan hệ

Cho quan hệ $R(A_1, A_2, \dots, A_n)$ được xác định bởi tập thuộc tính R^+ và tập phụ thuộc hàm F định nghĩa trên R, cho $K \subseteq R^+$.

K là một khoá của R nếu thỏa đồng thời cả hai điều kiện sau:

1. $K \subseteq R^+ \quad F^+ \quad (hay \quad K_F^+ = R^+)$

(K chỉ thỏa điều kiện 1 thì được gọi là siêu khoá)

2. Không tồn tại $K' \subseteq K$ sao cho $K'^+ = R^+$

Tập $S \subseteq \{A_1, \dots, A_n\}$ là *siêu khoá* của R nếu S chứa khoá. Một lược đồ quan hệ có thể có nhiều siêu khoá, nhiều khoá.

4.2. Thuật toán tìm một khóa của một lược đồ quan hệ

$K = Q^+$;

While A \subseteq K do

if $(K - A)^+ = Q^+$ then $K = K - A$

K còn lại chính là một khoá cần tìm.

Nếu muốn tìm các khoá khác (nếu có) của lược đồ quan hệ, ta có thể thay đổi thứ tự loại bỏ các phần tử của K.

Ví dụ:

Cho lược đồ quan hệ R(ABC) và tập phụ thuộc hàm

$$F = \{ A \rightarrow B; A \rightarrow C; B \rightarrow A \}$$

Hãy tìm một khóa của R.

Giải:

$$K = \{ A, B, C \}$$

Loại thuộc tính A, do $(K - A)^+ = R^+$ nên $K = \{ B, C \}$

thuộc tính B không loại được do $(K - B)^+ \neq R^+$ nên $K = \{ B, C \}$

Loại thuộc tính C, do $(K - C)^+ = R^+$ nên $K = \{ B \}$. Vậy một khóa của R là B.

4.3. Thuật toán tìm tất cả các khóa của một lược đồ quan hệ

Một số khái niệm hỗ trợ cho thuật toán tìm tất cả các khóa sau đây:

- Tập nguồn (TN): chứa tất cả thuộc tính chỉ xuất hiện ở vế trái mà không xuất hiện ở vế phải của tập phụ thuộc hàm và tập các thuộc tính không tham gia vào tập phụ thuộc hàm F.

- Tập đích (TD): chứa tất cả các thuộc tính chỉ xuất hiện ở vế phải mà không xuất hiện ở vế trái của tập phụ thuộc hàm.

- Tập trung gian (TG): chứa tất cả các thuộc tính tham gia vào cả 2 vế của tập phụ thuộc hàm.

Dữ liệu vào: Lược đồ quan hệ R và tập phụ thuộc hàm F.

Dữ liệu ra: Tất cả các khóa K của quan hệ.

Thuật toán:

Bước 0: Tìm tập thuộc tính nguồn (TN), tập thuộc tính trung gian

(TG).

Tìm tất cả các tập con của tập trung gian gọi là X_i (bằng phương pháp duyệt nhị phân)

if TG = then

$K = TN$; kết thúc.

Ngược lại

Qua bước 1

Bước 1 Tìm tất cả các tập con của TG: X_i

$S =$

$X_i \quad TG$

if $(TN \quad X_i)^+ = R^+$ then

$S = S \cup \{TN \quad X_i\}$

{S là tập các siêu khoá cần tìm}

Bước 2: Tính $TN \quad X_i$

Bước 3: Tính $(TN \quad X_i)^+$

Bước 4: Nếu $X_i^+ = R^+$ thì X_i là siêu khoá

Nếu một tập con $TN \quad X_i$ có bao đóng đúng bằng R^+ thì $TN \quad X_i$ là một siêu khoá của R.

Giả sử sau bước này có m siêu khoá: $S = \{S_1, S_2, \dots, S_m\}$

Bước 5 : Xây dựng tập chứa tất cả các khoá của R từ tập S

Xét mọi S_i, S_j con của S ($i \neq j$), nếu $S_i \supset S_j$ thì ta loại S_j ($i, j = 1..m$), kết quả còn lại chính là tập tất cả các khoá cần tìm.

Ví dụ: Cho lược đồ quan hệ R(ABC) và tập phụ thuộc hàm

$F = \{A \rightarrow B; A \rightarrow C; B \rightarrow A\}$

Hãy tìm tất cả các khóa của R.

Giải: Áp dụng thuật tìm tất cả các khóa đã cho ở trên ta có:

$TN = \{ \}; TG = \{A, B\}$

Gọi X_i là tập con của tập trung gian. Ta lập bảng như sau:

X_i	TN X_i	$(TN X_i)^+$	Siêu khóa	Khóa
			-	-
A	A	ABC	A	A
B	B	ABC	B	B
AB	AB	ABC	AB	-

Vậy lược đồ quan hệ R có hai khóa $K_1 = \{A\}$, $K_2 = \{B\}$

5. Phủ tối thiểu

5.1. Tập phụ thuộc hàm tương đương

Cho F và G là hai tập phụ thuộc hàm, ta nói F và G tương đương (hay F phủ G hoặc G phủ F) và ký hiệu là $F^+ = G^+$ nếu và chỉ nếu mỗi phụ thuộc hàm thuộc F đều thuộc G^+ và mỗi phụ thuộc hàm thuộc G đều thuộc F^+ .

Ta nói F phủ G nếu $G^+ \subseteq F^+$

Chẳng hạn cho lược đồ quan hệ Q(ABCDEFGH), thì hai tập phụ thuộc hàm F và G (xác định trên Q) là tương đương.

$F = \{B \rightarrow A; DA \rightarrow CE; D \rightarrow H; GH \rightarrow C; AC \rightarrow D; DG \rightarrow C\}$

$G = \{B \rightarrow A; DA \rightarrow CE; D \rightarrow H; GH \rightarrow C; AC \rightarrow D; BC \rightarrow AC; BC \rightarrow D; DA \rightarrow AH; AC \rightarrow DEH\}$

(Việc kiểm tra các phụ thuộc hàm trong G có được suy diễn từ F và ngược lại xem như bài tập dành cho bạn đọc).

5.2. Phủ tối thiểu

F_t được gọi là tập phụ thuộc hàm tối thiểu (hay phủ tối thiểu) nếu F thỏa đồng thời ba điều kiện sau:

1. F là tập phụ thuộc hàm có vế trái không dư thừa.
2. F là tập phụ thuộc hàm có vế phải một thuộc tính.
3. F là tập phụ thuộc hàm không dư thừa.

5.2.1. Phụ thuộc hàm có vế trái dư thừa:

F là tập phụ thuộc hàm trên lược đồ quan hệ Q , Z là tập thuộc tính, $Z \rightarrow Y \in F$. Nói rằng phụ thuộc hàm $Z \rightarrow Y$ có vế trái dư thừa (phụ thuộc không đầy đủ) nếu có một $A \in Z$ sao cho:

$$F \equiv F - \{Z \rightarrow Y\} \cup \{(Z-A) \rightarrow Y\}$$

Ngược lại $Z \rightarrow Y$ là phụ thuộc hàm có vế trái không dư thừa hay Y phụ thuộc hàm đầy đủ vào Z (phụ thuộc hàm đầy đủ).

Ta nói F là tập phụ thuộc hàm có vế trái không dư thừa nếu F không chứa phụ thuộc hàm có vế trái dư thừa.

Thuật toán loại khỏi F các phụ thuộc hàm có vế trái dư thừa:

Bước 1: - Xét lần lượt các phụ thuộc hàm $X \rightarrow Y$ của F .

Bước 2: - Với mọi tập con thực sự $X' \neq \emptyset$ của X .

- Nếu $X' \rightarrow Y \in F^+$ thì thay $X \rightarrow Y$ trong F bằng $X' \rightarrow Y$.

- Lặp lại bước 2.

5.2.2. Tập phụ thuộc hàm có vế phải một thuộc tính:

Mỗi tập phụ thuộc hàm F đều tương đương với tập phụ thuộc hàm G mà vế phải của các phụ thuộc hàm trong G chỉ gồm một thuộc tính.

G được gọi là tập phụ thuộc hàm có vế phải một thuộc tính.

Ví dụ:

$F = \{A \rightarrow BC, B \rightarrow C, AB \rightarrow D\}$ ta suy ra

$$F \equiv \{A \rightarrow B, A \rightarrow C, B \rightarrow C, AB \rightarrow D\} = G$$

5.2.3. Tập phụ thuộc hàm không dư thừa:

Nói rằng F là tập phụ thuộc hàm không dư thừa nếu không tồn tại $F' \subset F$ sao cho $F' \equiv F$. Ngược lại F là tập phụ thuộc hàm dư thừa.

Thuật toán loại khỏi F các phụ thuộc hàm dư thừa:

Bước 1: - Lần lượt xét các phụ thuộc hàm $X \rightarrow Y$ của F

Bước 2: - Nếu $X \rightarrow Y$ là thành viên của $F - \{X \rightarrow Y\}$ thì loại $X \rightarrow Y$ khỏi F .

Bước 3: - lặp lại bước 2 cho các phụ thuộc hàm tiếp theo của F.

5.3. Thuật toán tìm phủ tối thiểu

Từ điều kiện xác định phủ tối thiểu, ta có thuật toán tìm phủ tối thiểu như sau:

Thuật toán:

Bước 1: - Loại khỏi F các phụ thuộc hàm có vế trái dư thừa.

Bước 2: - Tách các phụ thuộc hàm có vế phải trên một thuộc tính thành các phụ thuộc hàm có vế phải một thuộc tính.

Bước 3: - Loại khỏi F các phụ thuộc hàm dư thừa.

Chú ý: Theo thuật toán trên, có thể tìm được nhiều hơn một phủ tối thiểu F_{tt} để $F \equiv F_{tt}$ và nếu thứ tự loại các phụ thuộc hàm khác nhau sẽ thu được các phủ tối thiểu khác nhau.

Ví dụ: cho $R(\text{MSCD}, \text{MSSV}, \text{CD}, \text{HG})$ và tập phụ thuộc hàm F:

$F = \{ \text{MSCD} \rightarrow \text{CD}; \text{CD} \rightarrow \text{MSCD}; \text{CD}, \text{MSSV} \rightarrow \text{HG}; \text{MSCD}, \text{HG} \rightarrow \text{MSSV}; \text{CD}, \text{HG} \rightarrow \text{MSSV}; \text{MSCD}, \text{MSSV} \rightarrow \text{HG} \}$

Hãy tìm một F_{tt} của F?

Kết quả ta có được một phủ tối thiểu sau:

$F_{tt} = \{ \text{MSCD} \rightarrow \text{CD}; \text{CD} \rightarrow \text{MSCD}; \text{CD}, \text{HG} \rightarrow \text{MSSV}; \text{MSCD}, \text{MSSV} \rightarrow \text{HG} \}$

6. Dạng chuẩn của lược đồ quan hệ

6.1. Một số khái niệm liên quan đến các dạng chuẩn

Thuộc tính khóa/thuộc tính không khóa: A là thuộc tính khóa nếu A có tham gia vào bất kỳ một khóa nào của quan hệ. Ngược lại A gọi là thuộc tính không khóa.

Thuộc tính phụ thuộc đầy đủ/ Phụ thuộc hàm đầy đủ: A là một thuộc tính phụ thuộc đầy đủ vào tập thuộc tính X nếu $X \rightarrow A$ là một phụ thuộc hàm đầy đủ (tức là không tồn tại $X' \subset X$ sao cho $X' \rightarrow A$ F^+)

Chú ý rằng một phụ thuộc hàm mà vế trái chỉ có một thuộc tính là phụ thuộc hàm đầy đủ.

6.2. *Dạng chuẩn 1 (First Normal Form)*

Định nghĩa: Lược đồ quan hệ R đạt dạng chuẩn 1 (1NF) nếu và chỉ nếu toàn bộ các thuộc tính của mọi bộ trên R đều mang giá trị đơn.

Ví dụ:

Xét quan hệ KETQUA sau:

MASV	HOVATEN	KHOA	TENMONHOC	DIEMTHI
01234	Nguyễn Văn An	CNTT	Cơ sở dữ liệu	6
			Toán rời rạc	8
			Lập trình web	7
02345	Lê Văn Thịnh	CNTT	Cơ sở dữ liệu	7

Quan hệ này không đạt chuẩn 1NF vì các thuộc tính TENMONHOC, DIEMTHI của bộ thứ nhất không mang giá trị đơn. Ta có thể đưa quan hệ trên về quan hệ KETQUA1 đạt chuẩn 1 như sau:

MASV	HOVATEN	KHOA	TENMONHOC	DIEMTHI
01234	Nguyễn Văn An	CNTT	Cơ sở dữ liệu	6
01234	Nguyễn Văn An	CNTT	Toán rời rạc	8
01234	Nguyễn Văn An	CNTT	Lập trình web	7
02345	Lê Văn Thịnh	CNTT	Cơ sở dữ liệu	7

Chú ý rằng khi xét các dạng chuẩn, nếu không xét gì thêm thì mặc định quan hệ đang xét ít nhất đạt dạng chuẩn 1.

6.3. *Dạng chuẩn 2 (Second Normal Form)*

Định nghĩa: Một lược đồ quan hệ R ở dạng chuẩn 2 (2NF) nếu R đạt dạng chuẩn 1 và mọi thuộc tính không khóa của R đều phụ thuộc đầy đủ vào khóa.

Hệ quả:

1. Nếu R đạt dạng chuẩn 1 và tập thuộc tính không khóa của R bằng rỗng thì R đạt chuẩn 2.

2. Nếu tất cả các khóa quan hệ chỉ gồm một thuộc tính thì quan hệ đó ít nhất đạt chuẩn 2.

Thuật toán kiểm tra dạng chuẩn 2:

Vào: lược đồ quan hệ R, tập phụ thuộc hàm F

Ra: Khẳng định R đạt hoặc không đạt chuẩn 2.

Bước 1: Tìm tất cả các khóa của R.

Bước 2: Với mỗi khóa K, tìm bao đóng của tất cả tập con thực sự của K.

Bước 3: Nếu có bao đóng S⁺ chứa thuộc tính không khóa thì R không đạt chuẩn 2. Ngược lại thì đạt chuẩn 2.

Ví dụ:

Cho lược đồ quan hệ R(ABCD) và tập phụ thuộc hàm F={AB→C; B→D; BC→A}. Hỏi R có đạt chuẩn 2 hay không?

Giải:

- Tìm tất cả các khóa của R:

$$TN = \{B\}, TG = \{AC\}$$

X _i	TN X _i	(TN X _i) ⁺	Siêu khóa	Khóa
	B	BD	-	-
A	BA	BACD	BA	BA
C	BC	BCAD	BC	BC
AC	BAC	BACD	BAC	-

Tất cả các khóa của R là K₁ = {BA}, K₂ = {BC}. Gọi Z là tập thuộc tính khóa, X là tập thuộc tính không khóa, ta có:

$$Z = K_1 \cup K_2 = \{BAC\}$$

$$X = R^+ \setminus Z = \{ABCD\} \setminus \{BAC\} = \{D\}$$

Ta thấy B⊂K₁, B→D, mà D là thuộc tính không khóa. Vì thuộc tính không khóa D không phụ thuộc đầy đủ vào khóa nên R không đạt chuẩn 2.

6.4. Dạng chuẩn 3 (Third Normal Form)

Định nghĩa: Một lược đồ quan hệ R đạt chuẩn 3 (3NF) nếu mọi phụ thuộc hàm $X \rightarrow A \in F^+$ với $A \notin X$ đều có.

- Hoặc X là siêu khóa.
- Hoặc A là thuộc tính khóa

Hệ quả:

1. Nếu R đạt chuẩn 3 thì R đạt chuẩn 2.
2. Nếu R không có thuộc tính không khóa thì R đạt chuẩn 3.

Định lý:

R là một lược đồ quan hệ.

F là tập các phụ thuộc hàm có vẻ phải một thuộc tính.

R đạt chuẩn 3 nếu và chỉ nếu mọi phụ thuộc hàm $X \rightarrow A \in F^+$ với $A \notin X$ đều có.

- Hoặc X là siêu khóa.
- Hoặc A là thuộc tính khóa

(Việc chứng minh định lý xem như là một bài tập nâng cao)

Thuật toán kiểm tra dạng chuẩn 3:

Vào: lược đồ quan hệ R, tập phụ thuộc hàm F.

Ra: Khẳng định R đạt hoặc không đạt chuẩn 3.

Bước 1: Tìm tất cả các khóa của R.

Bước 2: Từ F tạo tập phụ thuộc hàm tương đương F_{tt} có vẻ phải một thuộc tính.

Bước 3: Nếu mọi phụ thuộc hàm $X \rightarrow A \in F_{tt}$ với $A \notin X$ đều có X là siêu khóa hoặc A là thuộc tính khóa thì R đạt chuẩn 3. Ngược lại R không đạt chuẩn 3.

Ví dụ: Cho lược đồ quan hệ R(ABCD), $F = \{AB \rightarrow C; D \rightarrow B; C \rightarrow ABD\}$. Hỏi R có đạt chuẩn 3 hay không?

Giải:

- Tìm tất cả các khóa của R:

$$TN = \{\emptyset\} \quad TG = \{ABCD\}$$

X_i	TN X_i	$(TN \ X_i)^+$	Siêu khóa	Khóa
			-	-
A	A	A	-	-
B	B	B	-	-
C	C	CABD	C	C
D	D	DB	-	-
AB	AB	ABCD	AB	AB
AC	AC	ACBD	AC	-
AD	AD	ADBC	AD	AD
BC	BC	BCAD	BC	-
BD	BD	BD	-	-
CD	CD	CDAB	CD	-
ABC	ABC	ABCD	ABC	-
ABD	ABD	ABDC	ABD	-
ACD	ACD	ACDB	ACD	-
BCD	BCD	BCDA	BCD	-
ABCD	ABCD	ABCD	ABCD	-

Tất cả các khóa của R là $K_1 = \{C\}$, $K_2 = \{AB\}$, $K_3 = \{AD\}$. Gọi Z là tập thuộc tính khóa, X là tập thuộc tính không khóa, ta có:

$$Z = K_1 \cup K_2 \cup K_3 = \{CABD\}$$

$$X = R^+ \setminus Z = \{ABCD\} \setminus \{CABD\} = \{\emptyset\}$$

Vì tập thuộc tính không khóa $X = \{\emptyset\}$ nên R đạt chuẩn 3 (theo hệ quả 2).

6.5. Dạng chuẩn BC (Boyce Codd Normal Form)

Định nghĩa: Một lược đồ quan hệ R đạt dạng chuẩn BC nếu mọi phụ thuộc hàm $X \rightarrow A \in F^+$ với $A \notin X$ đều có X là siêu khóa.

Hệ quả:

1. Nếu R đạt chuẩn BC thì R đạt chuẩn 3 (hiển nhiên do định nghĩa).
2. Mỗi lược đồ có hai thuộc tính đều đạt chuẩn BC (xét phụ thuộc hàm có thể có của R).

Định lý:

R là lược đồ quan hệ.

F là tập phụ thuộc hàm có vẻ phải một thuộc tính.

R đạt chuẩn BC nếu và chỉ nếu mọi phụ thuộc hàm $X \rightarrow A \in F^+$ với $A \notin X$ đều có X là siêu khóa.

(Việc chứng minh định lý xem như là một bài tập nâng cao).

Thuật toán kiểm tra dạng chuẩn BC:

Vào: lược đồ quan hệ R, tập phụ thuộc hàm F.

Ra: Khẳng định R đạt hoặc không đạt chuẩn BC.

Bước 1: Tìm tất cả các khóa của R.

Bước 2: Từ F tạo tập phụ thuộc hàm tương đương F_{tt} có vẻ phải một thuộc tính.

Bước 3: Nếu mọi phụ thuộc hàm $X \rightarrow A \in F_{tt}$ với $A \notin X$ đều có X là siêu khóa thì R đạt chuẩn BC. Ngược lại R không đạt chuẩn BC.

Ví dụ:

Cho lược đồ quan hệ R(ABCDEI), $F = \{ACD \rightarrow EBI; CE \rightarrow AD\}$. Hỏi R có đạt chuẩn BC hay không?

Giải:

- Tìm tất cả các khóa của R:

TN={C} TG={ADE}

X_i	TN X_i	$(TN \ X_i)^+$	Siêu khóa	Khóa
	C	C	-	-
A	CA	CA	-	-
D	CD	CD	-	-
E	CE	CEADBI	CE	CE
AD	CAD	CADEBI	CAD	CAD
AE	CAE	CAEDBI	CAE	-
DE	CDE	CDEABI	CDE	-
ADE	CADE	CADEBI	CADE	-

Tất cả các khóa của R là $K_1 = \{CE\}$, $K_2 = \{CAD\}$. Gọi Z là tập thuộc tính khóa, X là tập thuộc tính không khóa, ta có:

$$Z = K_1 \quad K_2 = \{CEAD\}$$

$$X = R^+ \setminus Z = \{ ABCDEI \} \setminus \{ CEAD \} = \{ BI \}$$

- Tìm F_{tt} có vẻ phải một thuộc tính

$$F_{tt} = \{ ACD \rightarrow E; ACD \rightarrow B; ACD \rightarrow I; CE \rightarrow A; CE \rightarrow D \}$$

Ta nhận thấy mọi phụ thuộc hàm trong F_{tt} đều có vế trái là một siêu khóa nên R đạt chuẩn BC.

Thuật toán kiểm tra dạng chuẩn của một lược đồ quan hệ.

Vào : lược đồ quan hệ R, tập phụ thuộc hàm F.

Ra : khẳng định R đạt chuẩn gì?

Bước 1: Tìm tất cả các khóa của R.

Bước 2: Kiểm tra chuẩn BC, nếu đúng R đạt chuẩn BC, kết thúc thuật toán. Ngược lại qua bước 3.

Bước 3: Kiểm tra chuẩn 3, nếu đúng R đạt chuẩn 3, kết thúc thuật toán, ngược lại qua bước 4.

Bước 4: Kiểm tra dạng chuẩn 2, nếu đúng R đạt chuẩn 2, kết thúc thuật toán, ngược lại R đạt chuẩn 1.

Định nghĩa: Dạng chuẩn của một lược đồ cơ sở dữ liệu là dạng chuẩn thấp nhất trong các dạng chuẩn của các lược đồ quan hệ con.

Bài tập và sản phẩm thực hành bài 5.1

I. Kiến thức:

Câu 1: Phụ thuộc hàm: Định nghĩa? Cách xác định phụ thuộc hàm? Nêu tính chất của phụ thuộc hàm?

Câu 2: Trình bày thuật toán tìm bao đóng của tập thuộc tính?

Câu 3: Định nghĩa khóa của lược đồ quan hệ? Trình bày giải thuật tìm tất cả các khóa của lược đồ quan hệ?

Câu 4: Nêu định nghĩa các dạng chuẩn của lược đồ quan hệ? Trình bày các giải thuật kiểm tra dạng chuẩn ?

II. Kỹ năng:

Bài 1: Cho lược đồ quan hệ $R(B,C,D,E,F,G,I,J)$ và tập các phụ thuộc hàm

$$P = \{ I \rightarrow B; DE \rightarrow GJ; D \rightarrow C; CF \rightarrow J; B \rightarrow I; C \rightarrow G; F \rightarrow J \}$$

- Tìm tất cả các khoá của lược đồ quan hệ R.
- Tìm một phủ tối thiểu của tập phụ thuộc hàm P trên.

Bài 2: Cho lược đồ quan hệ $S(A,B,C,D,E,F,G)$ và tập các phụ thuộc hàm

$$P = \{ B \rightarrow FG; F \rightarrow CE; G \rightarrow BD; B \rightarrow A \}$$

Chứng tỏ rằng phụ thuộc hàm $AB \rightarrow D$ được suy diễn từ P nhờ hệ luật dẫn Armstrong? (Nêu rõ là áp dụng luật gì)

Câu 3: Cho lược đồ quan hệ $Q(A,B,C,D,E)$ và tập các phụ thuộc hàm

$$F = \{ A \rightarrow BC; C \rightarrow DE \}$$

- Lược đồ quan hệ Q đạt dạng chuẩn nào?
- Nếu chưa đạt dạng chuẩn 3 (3NF) hãy phân rã Q thành lược đồ quan hệ đạt tối thiểu dạng chuẩn 3 bảo toàn thông tin.

Đánh giá kết quả học tập

TT	Tiêu chí đánh giá	Cách thức và phương pháp	Điểm tối đa	Kết quả thực
----	-------------------	--------------------------	-------------	--------------

		đánh giá		hiện của người học
I	Kiến thức			
1	Phụ thuộc hàm	Vấn đáp, đối chiếu với nội dung bài học	2	
1.1	Định nghĩa			
1.2	Cách xác định phụ thuộc hàm			
1.3	Tính chất phụ thuộc hàm			
2	Thuật toán tìm bao đóng của tập thuộc tính	Vấn đáp, đối chiếu với nội dung bài học	1	
3	Khóa của lược đồ	Vấn đáp, đối chiếu với nội dung bài học	3	
3.1	Định nghĩa về khóa			
3.2	Giải thuật tìm tất cả các khóa			
4	Các dạng chuẩn	Vấn đáp, đối chiếu với nội dung bài học	4	
4.1	Định nghĩa các dạng chuẩn			
4.2	Giải thuật kiểm tra các dạng chuẩn			
Cộng:			10 đ	
II	Kỹ năng			
1	+ Phân tích được yêu cầu bài toán. + Xác định được các khóa	Làm bài tập đối chiếu với nội dung bài học.	10	

	<p>chính cho các quan hệ đã chỉ ra trong lược đồ.</p> <p>+ Xác định được phủ tối thiểu.</p> <p>+ Xác định được dạng chuẩn của lược đồ quan hệ.</p> <p>+ Chuẩn hóa bằng cách phân rã (nâng cao)</p>			
Cộng:			10đ	
III	Thái độ			
1	Tác phong công nghiệp	Theo dõi việc	4	
1.1	Đi học đầy đủ, đúng giờ	thực hiện, đối	1,5	
1.2	Không vi phạm nội quy lớp học	chiếu với nội quy của trường.	1,5	
1.3	Tính cẩn thận, tỉ mỉ	Quan sát việc thực hiện bài tập	1	
2	Đảm bảo thời gian thực hiện bài tập	Theo dõi thời gian thực hiện bài tập, đối chiếu với thời gian quy định.	2	
3	Đảm bảo an toàn lao động và vệ sinh công nghiệp	Theo dõi việc thực hiện, đối	4	
3.1	Tuân thủ quy định về an toàn	chiếu với quy	2	
3.3	Vệ sinh phòng học đúng quy định	định về an toàn và vệ sinh công	2	

		nghiệp		
Cộng:			10 đ	

KẾT QUẢ HỌC TẬP

Tiêu chí đánh giá	Kết quả thực hiện	Hệ số	Kết quả học tập
<i>Kiến thức</i>		0,3	
<i>Kỹ năng</i>		0,4	
<i>Thái độ</i>		0,3	
Cộng:			

BÀI TẬP TỰ GIẢI:

5.1. a) Cho lược đồ quan hệ $Q(ABCD)$, r là một quan hệ trên Q .

r			
A	B	C	D
a1	b1	c1	d1
a1	b2	c1	d1
a1	b3	c2	d1
a2	b2	c2	d2

phụ thuộc hàm nào sau đây không thỏa r

a) $D \rightarrow A$; b) $AC \rightarrow D$; c) $CD \rightarrow A$; d) $D \rightarrow B$;

b. Cho lược đồ quan hệ $Q(ABCD)$, r là quan trên Q được cho như sau:

A	B	C	D
a1	b1	c1	d2
a3	b1	c2	d1
a1	b1	c2	d2

Những phụ thuộc hàm nào sau đây thỏa r ?

$AB \rightarrow D$; $C \rightarrow B$; $B \rightarrow C$; $BC \rightarrow A$; $BD \rightarrow A$.

c. Cho lược đồ quan hệ $Q(ABCD)$, r là quan hệ được cho như sau:

A	B	C	D
x	u	x	y
y	x	z	x
z	y	y	y
y	z	w	z

Những phụ thuộc hàm nào sau đây không thỏa r ?

$A \rightarrow B$; $A \rightarrow C$; $B \rightarrow A$; $C \rightarrow D$; $D \rightarrow C$; $D \rightarrow A$

5.2. a. Cho lược đồ quan hệ $Q(ABCD)$ và tập phụ thuộc hàm $F = \{A \rightarrow B; BC \rightarrow D\}$.

Những phụ thuộc hàm nào sau đây thuộc F^+ ?

$C \rightarrow D$; $A \rightarrow D$; $AD \rightarrow C$; $AC \rightarrow D$; $BC \rightarrow A$; $B \rightarrow CD$.

b. Cho lược đồ quan hệ $Q(ABCDEFGH)$ và tập phụ thuộc hàm

$F = \{AB \rightarrow C; B \rightarrow D; CD \rightarrow E; CE \rightarrow GH; G \rightarrow A\}$

Những phụ thuộc hàm nào sau đây không thuộc vào F^+ ?

$AB \rightarrow E$; $AB \rightarrow GH$; $CGH \rightarrow E$; $CB \rightarrow E$; $GB \rightarrow E$.

c) Cho lược đồ quan hệ Q, F như sau: với $Q(ABCD)$ $F = \{A \rightarrow B; A \rightarrow C\}$.

Trong các phụ thuộc hàm sau, những phụ thuộc hàm được suy ra từ F ?

$A \rightarrow D$; $C \rightarrow D$; $AB \rightarrow B$; $BC \rightarrow A$; $A \rightarrow BC$

5.3. Cho lược đồ quan hệ $Q(ABCD)$ và tập phụ thuộc hàm

$F = \{A \rightarrow D; D \rightarrow A; AB \rightarrow C\}$

a. Tính AC^+

b. Chứng minh $BD \rightarrow C$

5.4. a) $Q(ABCDEG)$ Cho $F = \{AB \rightarrow C; C \rightarrow A; BC \rightarrow D; ACD \rightarrow B; D \rightarrow EG; BE \rightarrow C; CG \rightarrow BD; CE \rightarrow AG\}$

$X = [BD], X^+ = ?$

$Y = [CG], Y^+ = ?$

b. Cho lược đồ quan hệ Q và tập phụ thuộc hàm $F = \{AB \rightarrow E; AG \rightarrow I; BE \rightarrow I; E \rightarrow G; GI \rightarrow H\}$. Chứng minh rằng $AB \rightarrow GH$.

c. Tương tự cho tập phụ thuộc hàm

$F = \{AB \rightarrow C; B \rightarrow D; CD \rightarrow E; CE \rightarrow GH; G \rightarrow A\}$

Chứng minh rằng $AB \rightarrow E; AB \rightarrow G$

d. $Q(ABCDEFGH)$

$F = \{B \rightarrow A; DA \rightarrow CE; D \rightarrow H; GH \rightarrow C; AC \rightarrow D\}$ Hãy tìm một khoá của Q ?

5.5. Hãy tìm tất cả các khoá cho lược đồ quan hệ sau:
 $Q(\text{BROKER}, \text{OFFICE}, \text{STOCK}, \text{QUANTITY}, \text{INVESTOR}, \text{DIVIDENT})$

$F = \{\text{STOCK} \rightarrow \text{DIVIDENT}, \text{INVESTOR} \rightarrow \text{BROKER}, \text{INVESTOR}, \text{STOCK} \rightarrow \text{QUANTITY}, \text{BROKER} \rightarrow \text{OFFICE}\}$

5.6. Cho $Q(A, B, C, D)$, $F = \{AB \rightarrow C; D \rightarrow B; C \rightarrow ABD\}$

Hãy tìm tất cả các khoá của Q

5.7. Cho lược đồ quan hệ $Q(\text{MSCD}, \text{MSSV}, \text{CD}, \text{HG})$ và tập phụ thuộc F như sau:
 $F = \{\text{MSCD} \rightarrow \text{CD}; \text{CD} \rightarrow \text{MSCD}; \text{CD}, \text{MSSV} \rightarrow \text{HG}; \text{MSCD}, \text{HG} \rightarrow \text{MSSV}; \text{CD}, \text{HG} \rightarrow \text{MSSV}; \text{MSCD}, \text{MSSV} \rightarrow \text{HG}\}$

Hãy tìm phủ tối thiểu của F .

5.8. Xác định phủ tối thiểu của tập phụ thuộc hàm sau: $Q(ABCDEG)$

$F = \{AB \rightarrow C; C \rightarrow A; BC \rightarrow D; ACD \rightarrow B; D \rightarrow EG; BE \rightarrow C; CG \rightarrow BD; CE \rightarrow AG\}$

5.9. Các nhận xét sau đúng (Đ) hay sai (S) ? (kẻ bảng sau và ghi Đ hoặc S cho mỗi câu trên)

a	b	c	d	e	f	g	H

- a. Cho Q và $F = \{AB \rightarrow C; A \rightarrow B\}$ thì Q đạt dạng chuẩn 1.
- b. Một lược đồ quan hệ Q luôn tìm được ít nhất một khoá.
- c. Nếu $XY \rightarrow Z$ thì $X \rightarrow Z$ và $Y \rightarrow Z$.
- d. Các thuộc tính không tham gia vào vế phải của bất kỳ phụ thuộc hàm nào thì phải là thuộc tính tham gia vào khoá.
- e. Nếu $X \rightarrow Y$ và $YZ \rightarrow W$ thì $XZ \rightarrow W$
- f. Nếu Q đạt dạng chuẩn một và khoá của Q chỉ có một thuộc tính thì Q đạt dạng chuẩn ba.
- g. Một tập phụ thuộc hàm F có thể có nhiều tập phủ tối thiểu.
- h. Nếu $X \rightarrow Y$ và $U \rightarrow V$ thì $XU \rightarrow YV$.

5.10 a. Cho $Q(ABCD)$ và $F = \{AB \rightarrow C; D \rightarrow B; C \rightarrow ABD\}$.

Hãy kiểm tra xem $AB \rightarrow D$ có thuộc F^+ hay không ?

Hãy tìm tất cả các khoá của lược đồ quan hệ Q . Xác định dạng chuẩn của Q .

b. Cho $Q(A,B,C,D)$ và $F = \{C \rightarrow A; A \rightarrow C; AD \rightarrow B; BC \rightarrow D; AB \rightarrow D; CD \rightarrow B\}$
 Hãy tìm phủ tối thiểu của F .

5.11. Cho biết dạng chuẩn của các lược đồ quan hệ sau:

a. $Q(ABCDEG); F = [A \rightarrow BC, C \rightarrow DE, E \rightarrow G]$

b. $Q(ABCDEGH); F = [C \rightarrow AB, D \rightarrow E, B \rightarrow G]$

c. $Q(ABCDEGH); F = [A \rightarrow BC, D \rightarrow E, H \rightarrow G]$

d. $Q(ABCDEG); F = [AB \rightarrow C; C \rightarrow B; ABD \rightarrow E; G \rightarrow A]$

e.Q(ABCDEFGHI);F=[AC→B; BI→ACD; ABC→D; H→I; ACE→BCG,
CG→AE]

TÀI LIỆU THAM KHẢO

- [1]. Lê Tiến Vương, Nhập môn cơ sở dữ liệu quan hệ, Nhà xuất bản Giáo dục, 2000.
- [2]. Vũ Đức Thi, Cơ sở dữ liệu kiến thức và thực hành, Nhà xuất bản thống kê 1997.
- [3]. Nguyễn An Tế, Giáo trình nhập môn cơ sở dữ liệu, ĐHKHTN-ĐHQGTPHCM 1996.
- [4]. Đỗ Trung Tuấn, Cơ sở dữ liệu, Nhà xuất bản giáo dục, 1998.

PHỤ LỤC