

BỘ LAO ĐỘNG THƯƠNG BINH VÀ XÃ HỘI
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT VĨNH LONG



GIÁO TRÌNH
CƠ SỞ DỮ LIỆU
(DATABASE SYSTEMS)

BIÊN SOẠN
NGUYỄN NGỌC NGA

Vĩnh Long - 9/2013

Chương 1: TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU

Trong chương này trình bày những khái niệm cơ bản về các hệ cơ sở dữ liệu do E.F Codd đề xuất. Những khái niệm này bao gồm mục tiêu của một hệ cơ sở dữ liệu. Sự cần thiết phải tổ chức dữ liệu dưới dạng cơ sở dữ liệu. Tính độc lập của dữ liệu thể hiện mô hình hình kiến trúc 3 mức. Vì vậy có thể nói cơ sở dữ liệu phản ánh tính trung thực, khách quan của thế giới dữ liệu không dư thừa và cũng không thiếu thông tin. Nội dung của chương bao gồm các phần:

- Tìm hiểu về cơ sở dữ liệu là gì?
- Sự cần thiết của hệ cơ sở dữ liệu.
- Các mô hình cơ sở dữ liệu.
- Hệ quản trị & người quản trị cơ sở dữ liệu.
- Tổ chức lưu trữ dữ liệu.
- Các ràng buộc toàn vẹn trên cơ sở dữ liệu quan hệ.

1.1. Mở đầu

1.1.1. Định nghĩa cơ sở dữ liệu

Cơ sở dữ liệu (DataBase – DB) là một tập hợp chứa các thông tin có cấu trúc để mô tả các đối tượng trong thế giới thực có liên quan với nhau.

Ví dụ: Cơ sở dữ liệu quản lý điểm (QLDIEM) của sinh viên sẽ lưu trữ các thông tin của ngành học, sinh viên, môn học, kết quả ...

1.1.2. Định nghĩa môn cơ sở dữ liệu

Môn cơ sở dữ liệu (CSDL) là lĩnh vực của tin học chuyên nghiên cứu các cơ chế, nguyên lý, phương pháp tổ chức các nhóm dữ liệu trên các bộ nhớ ngoài nhằm phục vụ cho việc quản lý dữ liệu được tốt hơn.

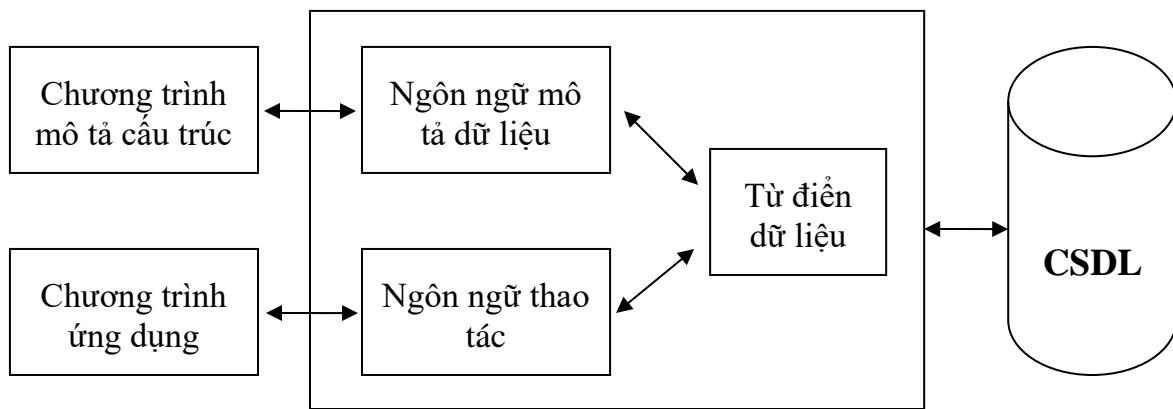
1.1.3. Hệ quản trị cơ sở dữ liệu (DataBase Management System - DBMS)

Hệ quản trị cơ sở dữ liệu (HQTCSDL) là hệ thống các chương trình để người dùng thực hiện khai thác cơ sở dữ liệu theo ba chức năng: Mô tả, thao tác, và truy xuất dữ liệu.

Hệ quản trị cơ sở dữ liệu gồm các thành phần:

- Ngôn ngữ giao tiếp giữa người dùng (User) và CSDL.
- Từ điển dữ liệu (Data Dictionary).
- Biện pháp bảo mật.
- Cơ chế giải quyết tranh chấp dữ liệu.
- Cơ chế sao lưu (Backup) và phục hồi (Restore) dữ liệu.
- Cung cấp một giao diện (Interface) tốt.
- Bảo đảm tính độc lập giữa dữ liệu và chương trình.

Một số HQTCSDL hiện nay: Visual Foxpro, Microsoft Access, MySQL, SQL Server, Oracle,...



Hình 1.1 Sơ đồ tổng quát của HQTCSDL

1.1.4. Ưu điểm của cơ sở dữ liệu

- Giảm sự trùng lắp thông tin xuống mức thấp nhất, 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.

1.1.5. Các mức biểu diễn một CSDL

Để hệ thống có thể dùng được thì phải tìm dữ liệu một cách hiệu quả. Điều này dẫn đến việc cần phải thiết kế những cấu trúc dữ liệu phức tạp đối với việc biểu diễn dữ liệu trong cơ sở dữ liệu. Vì nhiều người sử dụng các hệ thống cơ sở dữ liệu không phải là người hiểu biết nhiều về máy tính nên những người phát triển hệ thống phải che dấu đi sự phức tạp khỏi người sử dụng thông qua một số mức trừu tượng, nhằm làm đơn giản hóa sự tương tác của người sử dụng đối với hệ thống.

Có 3 mức biểu diễn dữ liệu như sau:

a. Mức trong (mức vật lý - Physical level)

- Nói đến cách thức dữ liệu được lưu trữ thực sự trên bộ nhớ? Trả lời câu hỏi dữ liệu gì và lưu trữ chúng như thế nào? Cần các chỉ mục gì?
- Những người thao tác tại mức này là người quản trị CSDL và những người dùng chuyên môn.

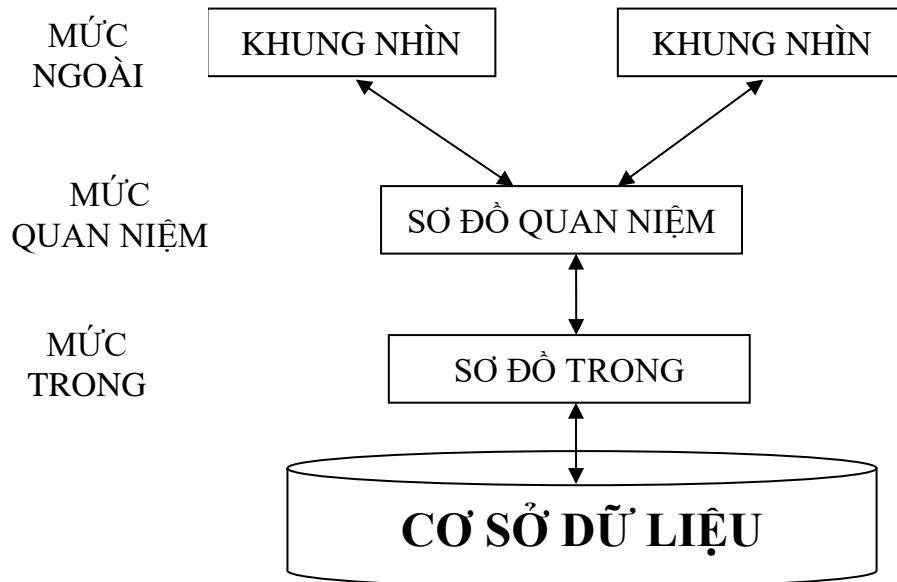
b. Mức quan niệm hay mức logic (conception level, logical level)

Trả lời câu hỏi cần phải lưu trữ bao nhiêu loại dữ liệu? Đó là những dữ liệu gì? Mối quan hệ giữa chúng như thế nào?

c. Mức ngoài hay mức khung nhìn (view level)

- Là mức của người sử dụng và các chương trình ứng dụng.
- Làm việc tại mức này là: Các nhà chuyên môn, kỹ sư tin học, người sử dụng không chuyên.
- Mỗi người sử dụng hay mỗi chương trình ứng dụng có thể được nhìn CSDL theo một góc độ khác nhau. Tuỳ thuộc vào sự phân quyền của người quản lý đối với người sử dụng mà có thể thấy tất cả hoặc một phần của CSDL.

- Người sử dụng hay chương trình ứng dụng có thể không được biết về cấu trúc tổ chức lưu trữ thông tin trong CSDL. Họ chỉ có thể làm việc trên một phần CSDL theo cách nhìn do người quản trị hay chương trình ứng dụng quy định.



Hình 1.2 Sơ đồ các mức trong CSDL

1.1.6. Tính độc lập của dữ liệu

Tính độc lập dữ liệu là tính bất biến của các chương trình ứng dụng đối với các thay đổi trong cấu trúc lưu trữ và chiến lược truy xuất dữ liệu. Hay nói cách khác, độc lập dữ liệu là việc tách các mô tả dữ liệu ra khỏi các chương trình ứng dụng sử dụng dữ liệu đó.

Sự phụ thuộc dữ liệu của các ứng dụng hiện nay: một hệ thống ứng dụng được gọi là phụ thuộc dữ liệu khi không thể thay đổi được cấu trúc lưu trữ hoặc chiến lược truy xuất mà không ảnh hưởng đến chương trình ứng dụng.

Yêu cầu của các hệ ứng dụng: các hệ ứng dụng khác nhau cần có những cái nhìn khác nhau đối với những dữ liệu như nhau. Người quản trị cơ sở dữ liệu phải có khả năng thay đổi cấu trúc lưu trữ hoặc chiến lược truy xuất để đáp ứng các yêu cầu thay đổi mà không cần sửa đổi các chương trình ứng dụng hiện có.

Phân loại tính độc lập dữ liệu:

Độc lập dữ liệu mức vật lý: các thay đổi về cách tổ chức vật lý của CSDL có thể ảnh hưởng đến tính hiệu quả của các chương trình ứng dụng, nhưng chúng không yêu cầu viết lại các chương trình.

Độc lập dữ liệu mức logic: khi sử dụng CSDL, có thể thay đổi sơ đồ quan niệm, như thêm thông tin vào các kiểu khác nhau của thực thể, hay xóa bỏ thông tin về các thực thể. Nhiều thay đổi của sơ đồ quan niệm không làm ảnh hưởng đến các sơ đồ đang có, hoặc có thể thực hiện bằng các định nghĩa lại cách ánh xạ từ sơ đồ con đến sơ đồ quan niệm. Do đó, không cần phải sửa đổi lại các chương trình ứng dụng.

1.2. Các mô hình của cơ sở dữ liệu

Mô hình cơ sở dữ liệu là sự trừu tượng hóa môi trường thế giới thực. Mỗi loại mô hình cơ sở dữ liệu đặc trưng cho một cách tiếp cận dữ liệu của những nhà phân tích thiết kế CSDL.

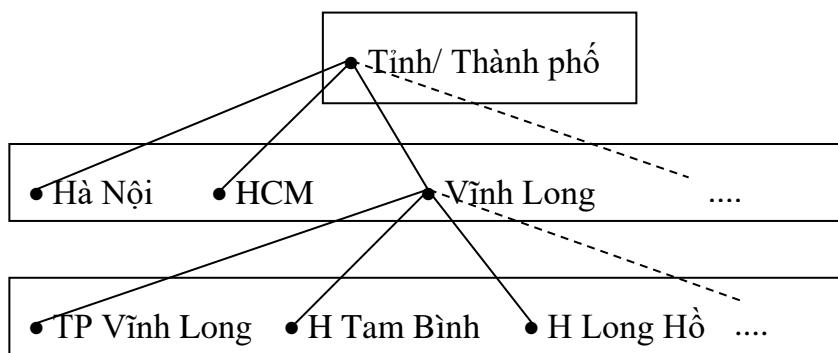
Mỗi loại mô hình cơ sở dữ liệu đều có những ưu điểm và những mặt hạn chế riêng của nó, nhưng vẫn có những mô hình dữ liệu nổi trội và được nhiều người quan tâm nghiên cứu. Sau đây, chúng ta sẽ tìm hiểu sơ lược về các mô hình dữ liệu.

1.2.1. Mô hình phân cấp (Hierarchical Model):

Mô hình phân cấp ra đời khoảng những năm 60 của thế kỷ XX, được biểu diễn dưới dạng cấu trúc cây, trong đó mỗi nút của cây biểu diễn một thực thể, giữa nút con và nút cha được liên hệ với nhau theo một mối liên hệ xác định.

Ví dụ: Mô hình dữ liệu phân cấp thể hiện mối quan hệ giữa các thực thể:

Tỉnh/ Thành phố – Quận/ Huyện

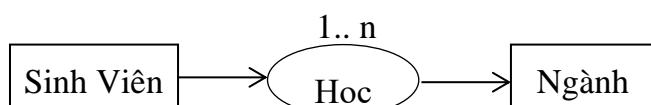


....

1.2.2. Mô hình mạng (Network Model)

Mô hình nở ra đời từ những năm 60 của thế kỷ XX và được định nghĩa lại những năm 70, được biểu diễn dưới dạng đồ thị có hướng, mỗi nút là một thực thể, mỗi nút có thể kết hợp với các nút khác để biểu diễn mối liên hệ giữa chúng.

Ví dụ: Mô hình mạng thể hiện mối liên hệ giữa Sinh Viên và Ngành



Ý nghĩa: một sinh viên theo học 1 ngành, một ngành có nhiều sinh viên.

1.2.3. Mô hình quan hệ (Relational Model)

Cơ sở dữ liệu được xây dựng theo lý thuyết do E.F.Codd giới thiệu năm 1970. Thuật ngữ ‘quan hệ’ là bảng dữ liệu hai chiều được Codd gọi là bảng quan hệ. Mô hình quan hệ khác hẳn với các mô hình trước nó và từ năm 1980 đã trở thành mô hình được dùng rộng rãi để phát triển hệ quản trị CSDL.

Mô hình quan hệ là mô hình được sử dụng xuyên suốt trong giáo trình này vì thế những đặc điểm liên quan đến mô hình này sẽ được trình bày khá chi tiết.

a. Bảng (Table)

Là hình ảnh trực quan của một quan hệ, bao gồm các cột (mô tả thông tin) và các dòng (lưu trữ thông tin).

- Thuộc tính: Là thông tin mô tả một đặc điểm nào đó của quan hệ.

Giáo trình Cơ sở dữ liệu

ví dụ quan hệ SINHVIEN có các thuộc tính: Mã số sinh viên, họ tên, giới tính, ngày sinh, ...

- Bộ: Các dòng dữ liệu cụ thể của một quan hệ.

Ví dụ quan hệ SINHVIEN có 2 bộ như sau:

1500115 Nguyễn Văn Đạt Nam 10/12/1997

1500280 Lê Văn Nữ 12/10/1998

Ví dụ quan hệ SINHVIEN

SINHVIEN	MSSV	HoTen	GioiTinh	NgaySinh
	1500115	Nguyễn Văn Đạt	Nam	10/12/1997
	1500280	Lê Văn	Nữ	12/10/1998

Có sơ đồ quan hệ: SINHVIEN=(MSSV, HoTen, GioiTinh, NgaySinh)

hay

SINHVIEN(MSSV, HoTen, GioiTinh, NgaySinh)

Trong đó MSSV: mã số sinh viên

HoTen: họ và tên sinh viên

GioiTinh: giới tính

NgaySinh: ngày sinh

Lưu ý: Trong một quan hệ, không được chứa hai bộ giống nhau.

b. Giá trị rỗng (NULL)

Giá trị rỗng là giá trị đặc biệt, dùng để gán cho các thuộc tính khi chưa dùng đến, hay khi người ta chưa biết giá trị cần gán.

NULL là giá trị đặc biệt, với ý nghĩa ‘không giống bất kỳ giá trị nào’, thậm chí không giống với các giá trị NULL khác.

Ví dụ: hai sinh viên có mã ngành (MaNganh) để trống, thì hai giá trị rỗng đó cũng không giống nhau.

Sinh Viên	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	
	SV002	Đương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1993	

Mã ngành (MaNganh) bị bỏ trống, hai vùng bị bỏ trống là khác nhau.

c. Khóa của quan hệ

- Khóa (khóa chính – Primary Key)

Giáo trình Cơ sở dữ liệu

Khóa của quan hệ là một tập nhỏ nhất các thuộc tính dùng để phân biệt các bộ với nhau. Khóa chính không được bỏ trống và các bộ không được có khóa chính trùng nhau.

Ví dụ 1: Quan hệ SINHVIEN (MaSV, HTSV, NS, MaNganh) có dữ liệu như sau:

SINHVIEN	MaSV	HTSV	NS	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV010	Lê Lan	12/07/1995	CNOT
	SV020	Nguyễn Tài	10/07/1991	CNTT

Trong quan hệ SINHVIEN thì MaSV là khóa, MaSV sẽ dùng để phân biệt các sinh viên với nhau. Khi đó MaSV không được bỏ trống và không cho phép hai MaSV trùng nhau.

Ví dụ 2: quan hệ sinh viên (SINHVIEN) sau là vi phạm ràng buộc toàn vẹn về khóa chính.

SINHVIEN	MaSV	HTSV	NS	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT
		Phạm Thị Ái	02/10/1993	CNTT
	SV001	Lê Lan	12/07/1995	CNOT
	SV020	Nguyễn Tài	10/07/1991	CNTT

- Bị trùng ở giá trị ‘SV01’ ở mẫu tin thứ nhất và mẫu tin thứ tư.

- Bị trống MaSV ở mẫu tin thứ 3.

Trong một quan hệ có thể có nhiều khóa, ta gọi là khóa ứng viên. Nhưng khi thiết kế bảng dữ liệu, ta chỉ chọn một khóa ứng viên làm khóa chính. Các tập thuộc tính trong khóa sẽ đặt ở đầu và được gạch dưới.

- Siêu khóa (Supper Key)

Là một tập thuộc tính có chứa khóa chính. Như vậy, khóa chính là một trường hợp của siêu khóa.

Ví dụ: Cho quan hệ SINHVIEN (MaSV, HTSV, NS, MaNganh)

Sinh Viên	MaSV	HTSV	NS	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT

	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV010	Lê Lan	12/07/1995	CNOT
	SV020	Nguyễn Tài	10/07/1991	CNTT

Với quan hệ SINHVIEN thì MaSV là khóa chính và có nhiều siêu khóa, ví dụ như (MaSV, HTSV); (MaSV, NS); (MaSV, MaNganh);...

- Khóa ngoại (Foreign Key - FK)

Khóa ngoại của quan hệ R là một tập thuộc tính FK sao cho FK đóng vai trò là khóa chính của một quan hệ S khác. Lúc này quan hệ S gọi là quan hệ cha, R gọi là quan hệ con.

Khóa ngoại thường được đặt cuối quan hệ, có thêm ký hiệu # ở phía trước (có tài liệu qui ước khóa ngoại bằng cách gạch dưới các thuộc tính khóa ngoại bởi nét khuất - - -).

Ví dụ: Cho hai quan hệ NGANH (MaNganh, TenNganh) và SINHVIEN (MaSV, HTSV, NgaySinh, #MaNganh)

NGANH	MaNganh	TenNganh
	CNTT	Công Nghệ Thông Tin
	CKCT	Cơ Khí Chế Tạo
	CNOT	Công Nghệ Ô Tô

SINHVIEN	MaSV	HTSV	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT

Ta dễ dàng nhận ra mã ngành (MaNganh) của SINHVIEN là khóa ngoại vì MaNganh là khóa của quan hệ Nganh. Ta có thể gọi cách khác quan hệ NGANH là cha của SINHVIEN. Giữa MaNganh của SINHVIEN và MaNganh của NGANH có ràng buộc toàn vẹn về khóa ngoại.

Các thuộc tính của khóa ngoại của bảng con phải cùng cấu trúc với các thuộc tính khóa chính bên bảng cha.

c. Ràng buộc toàn vẹn trên cơ sở dữ liệu quan hệ

Ràng buộc toàn vẹn là các qui tắc được thiết kế để kiểm tra từng bộ dữ liệu có hợp lệ hay không trước khi thực hiện thao tác (thêm, xóa, cập nhật) trên cơ sở dữ liệu.

- Ràng buộc về khóa chính (Primary key)

Các tập thuộc tính trong khóa chính phải là nhỏ nhất, duy nhất, không được rỗng và phải xác định các thuộc tính còn lại trên mỗi bộ.

Giáo trình Cơ sở dữ liệu

- Ràng buộc về khóa ngoại (Foreign key)

Tập thuộc tính P (có các bộ là X) là khóa chính của quan hệ S. Nhưng P là khóa ngoại của quan hệ R (có các bộ là Y). Khi đó đó giữa X và Y có ràng buộc trên P như sau:

- Thực hiện xóa, cập nhật trên P của X thì ảnh hưởng đến P của Y.
- Thêm P của Y thì phụ thuộc vào P của X.

Có các loại ràng buộc toàn vẹn trên khóa ngoại khi ta tiến hành cập nhật (Update) và xóa (Delete):

No action: Không cho phép thay đổi (Cập nhật, xóa) dữ liệu bảng cha khi bảng con còn sử dụng dữ liệu mà bảng cha đang thao tác.

Cascade: cho phép thay đổi dữ liệu bảng cha và thay đổi cả dữ liệu của bảng con có liên quan đến dữ liệu mà bảng cha đang thao tác.

Set default: cho phép thay đổi dữ liệu bảng cha và thay đổi cả dữ liệu của bảng con liên quan đến dữ liệu mà bảng cha đang thao tác thành giá trị mặc định.

Set null: cho phép thay đổi dữ liệu bảng cha và thay đổi cả dữ liệu của bảng con có liên quan đến dữ liệu mà bảng cha đang thao tác thành rỗng (null).

Ví dụ: Cho hai quan hệ Ngành (NGANH) và Sinh Viên (SINHVIEN)

NGANH (MaNganh, TenNganh)

NGANH	MaNganh	TenNganh
	CNTT	Công Nghệ Thông Tin
	CKCT	Cơ Khí Chế Tạo
	CNOT	Công Nghệ Ô Tô

SINHVIEN (MaSV, HoTen, NgaySinh, #MaNganh)

SINHVIEN	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Đương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT

Quan hệ SINHVIEN có khóa ngoại là MaNganh.

Yêu cầu 1: xóa ngành ‘CKCT’ từ quan hệ NGANH với cài đặt:

No action: Không thực hiện xóa được vì trong sinh viên còn có mã ngành là ‘CKCT’.

Cascade: xóa ngành ‘CKCT’ và tất cả sinh viên có mã ngành là ‘CKCT’.

Kết quả:

NGANH	MaNganh	TenNganh
	CNTT	Công Nghệ Thông Tin
	CNOT	Công Nghệ Ô Tô

SINHVIEN	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV003	Phạm Thị Ái	02/10/1993	CNTT

Set default: Xóa ngành ‘CKCT’ trong quan hệ Ngành và gán lại mã ngành cho các sinh viên có mã ngành ‘CKCT’ thành giá trị mặc định. (giả sử giá trị mặc định là ‘CNTT’).

Kết quả:

NGANH	MaNganh	TenNganh
	CNTT	Công Nghệ Thông Tin
	CNOT	Công Nghệ Ô Tô

SINHVIEN	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CNTT
	SV003	Phạm Thị Ái	02/10/1993	CNTT

Set null: Xóa ngành ‘CKCT’ trong quan hệ Ngành và gán lại mã ngành cho các sinh viên có mã ngành ‘CKCT’ thành giá trị null.

Kết quả:

NGANH	MaNganh	TenNganh
	CNTT	Công Nghệ Thông Tin
	CNOT	Công Nghệ Ô Tô

SINHVIEN	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT

SV002	Dương Tùng Lê	12/4/1994	
SV003	Phạm Thị Ái	02/10/1993	CNTT

Yêu cầu 2: thêm sinh viên vào có mã ngành (MaNganh) ‘CNTD’.

Không thêm được vì trong quan hệ NGANH không có mã ngành ‘CNTD’.

Yêu cầu 3: thêm một sinh viên ngành ‘CNTT’?

- Ràng buộc rỗng (null | not null)

Cho phép một thuộc tính được phép chứa giá trị rỗng (null) hay không được phép chứa giá trị rỗng (not null).

Ràng buộc rỗng đặt sau mỗi mô tả cột.

Ví dụ: Cài đặt cho họ tên (HoTen) không được để trống.

HoTen not null

- Ràng buộc về tính duy nhất (Unique)

Ràng buộc này không cho phép thuộc tính chứa giá trị trùng nhau. Có thể áp dụng ràng buộc này trên một thuộc tính hay tập thuộc tính.

Ví dụ: giả sử quan hệ NGANH có dữ liệu đã được cài đặt ràng buộc duy nhất trên thuộc tính (TenNganh):

NGANH(MaNganh, TenNganh)

NGANH	MaNganh	TenNganh
	CNTT	Công nghệ thông tin
	CKCT	Cơ khí chế tạo
	CNOT	Công nghệ ô tô

Yêu cầu: có thể thêm một bộ để quan hệ NGANH có được dữ liệu như sau hay không? Tại sao?

NGANH	MaNganh	TenNganh
	CNTT	Công nghệ thông tin
	CKCT	Cơ khí chế tạo
	CNOT	Công nghệ ô tô
	HTTT	Công nghệ thông tin

Trả lời: không thực hiện được vì giá trị định thêm vào có TenNganh là ‘Công nghệ thông tin’ vi phạm ràng buộc duy nhất.

- Ràng buộc về luận lý (Check)

Ràng buộc này qui định giá trị nhận vào phải thỏa một điều kiện luận lý. Điều kiện luận lý là biểu thức có thể sử dụng một thuộc tính hoặc nhiều thuộc tính. Ràng buộc này có thể áp dụng để kiểm tra sự thay đổi của một thuộc tính hoặc nhiều thuộc tính trên một bộ.

Ví dụ: cho quan hệ KetQua (MaSV, MaMon, Diem, LanThi).

Yêu cầu: thiết kế ràng buộc toàn vẹn để giá trị của thuộc tính Diem chỉ nhận giá trị từ 0 đến 10.

Thiết kế: Check (Diem>=0 and Diem<=10)

1.2.4. Mô hình thực thể quan hệ (Entity Relationship Model):

Là mô hình cho phép mô tả các thực thể và mối kết hợp (mối liên hệ) giữa các thực thể. Mỗi thực thể sẽ được miêu tả chi tiết thông qua các thuộc tính. Mỗi kết hợp giữa các thực thể chỉ rõ số lượng tối đa và tối thiểu của các thực thể. Một trong những cách biểu diễn mô hình thực thể kết hợp là dùng đồ thị và sơ đồ khối.

a. Thực thể (Entity)

Thực thể là một đối tượng của thế giới thực và được phân biệt với các thực thể khác.

Ví dụ: Con người là một thực thể có các thông tin như: họ tên, giới tính (phái), ngày sinh, địa chỉ, số chứng minh nhân dân.

b. Tập thực thể (Entity set)

Tập thực thể là một nhóm các thực thể có các thuộc tính ‘tương tự’ nhau.

Ví dụ: Tập hợp các con người làm chung một cơ quan sẽ tạo thành một tập thực thể NHANVIEN (nhân viên).

c. Thuộc tính và khóa

Thuộc tính (Attribute): là đặc tính (thông tin) của tập thực thể, mỗi thuộc tính có miền giá trị, hoặc các ràng buộc mô tả.

Ví dụ: tập các thực thể làm chung một cơ quan ta gọi chung là nhân viên. Mỗi nhân viên có các thông tin (thuộc tính) cơ bản: họ tên, ngày sinh, phái, số chứng minh nhân dân, địa chỉ, nơi sinh,...

Khóa (Key): là tập hợp các thuộc tính nhỏ nhất để phân biệt sự khác nhau giữa hai thực thể trong một tập thực thể.

Ví dụ: trong tập thực thể nhân viên. Mỗi nhân viên có các thông tin: họ tên, ngày sinh, giới tính, số chứng minh nhân dân,... Thuộc tính phân biệt giữa các nhân viên là số chứng minh nhân dân (vì không có hai nhân viên có cùng số chứng minh nhân dân) hoặc để thuận tiện cho việc quản lý trong nội bộ thì mỗi thực thể nhân viên sẽ được nhận một mã số riêng đóng vai trò là khóa chính.

d. Mối kết hợp (Relationship)

Mối kết hợp là mô tả mối liên hệ giữa hai tập thực thể.

Ví dụ: Tập thực thể sinh viên (SINHVIEN) và tập thực thể ngành (NGANH) có mối kết hợp là học (Hoc). Ta thường gọi là SINHVIEN Hoc NGANH. Mỗi thực thể sinh viên có các thuộc tính: mã số (MaSV) là khóa, họ và tên (HoTen), ngày sinh (NgaySinh). Mỗi thực thể ngành có các thuộc tính: mã ngành (MaNganh) là khóa, tên ngành (TenNganh).

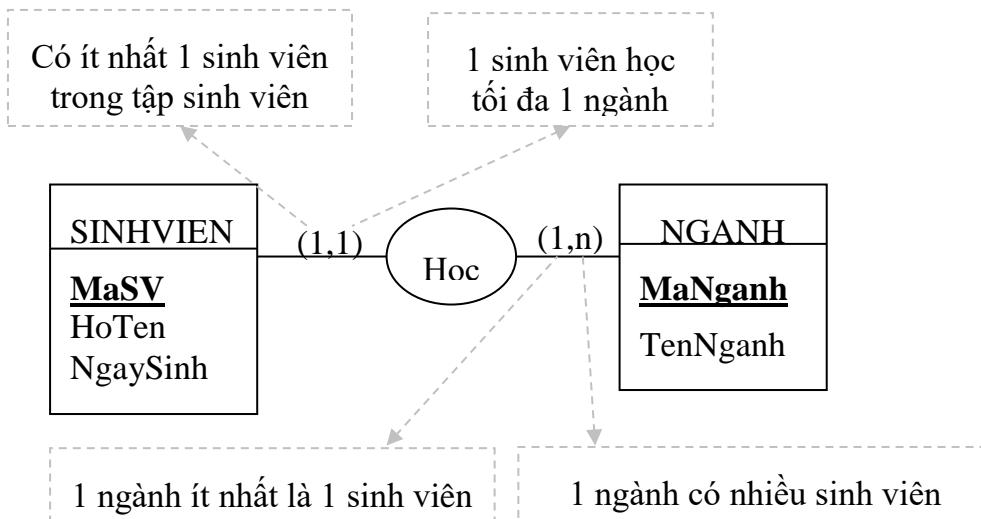
e. Bản số

Bản số là một cặp số ($<\text{min}>$, $<\text{max}>$) trong đó:

$<\text{min}>$: qui định số lượng ít nhất của các thực thể trong tập thực thể.

$<\text{max}>$: qui định số lượng lớn nhất các thực thể trong tập thực thể kia có quan hệ với một thực thể trong tập thực thể này.

Ví dụ: sơ đồ thực thể quan hệ của Sinh viên và Ngành

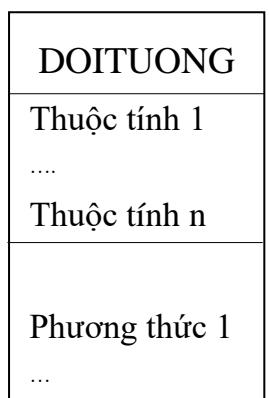


1.2.5. Mô hình dữ liệu hướng đối tượng (Object Oriented Data Model)

Mô hình cung cấp các đặc tính nhận dạng đối tượng. Trong đó mỗi đối tượng đặc trưng bởi hai yếu tố:

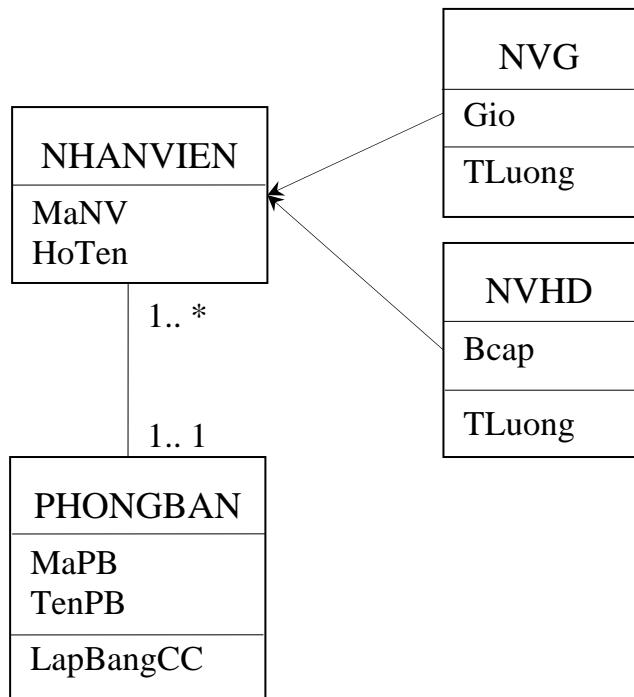
Tập các thuộc tính: dùng để mô tả các chi tiết về đối tượng.

Tập các phương thức: dùng để thao tác (hành động) trên đối tượng.



Mô hình tổng quát của đối tượng

Ví dụ: cho hai đối tượng nhân viên (NHANVIEN) và phòng ban (PHONGBAN). Nhân viên có hai loại: nhân viên làm theo giờ (NVG) và nhân viên hợp đồng (NVHD). Mỗi nhân viên sẽ được phân công làm việc tại một phòng ban hay một đơn vị (PHONGBAN) nào đó.



Ý nghĩa của mô hình:

NHANVIEN: có các thuộc tính (MaNV, HoTen).

NVG và NVHD thừa kế các thuộc tính từ NHANVIEN. Mỗi loại nhân viên đều có phương thức TLuong. Nhưng cách thực hiện ở mỗi loại nhân viên có thể khác nhau.

PHONGBAN: có các thuộc tính (MaPB, TenPB) và phương thức (LapBangCC)

Đây là loại mô hình tiên tiến, đang được nghiên cứu và sử dụng rộng rãi, dựa trên cách tiếp cận hướng đối tượng. Mô hình này sử dụng các khái niệm như lớp (class), sự kế thừa (inheritance), đa kế thừa (multi - inheritance). Đặc trưng cơ bản của cách tiếp cận này là tính đóng gói (encapsulation), tính đa hình (polymorphism) và tính tái sử dụng (reusability).

1.3. Ngôn ngữ trên dữ liệu

1.3.1. Ngôn ngữ tự nhiên và ngôn ngữ hình thức

Ngôn ngữ tự nhiên: là ngôn ngữ giao tiếp giữa người với người, được phát triển và hoàn thiện kể từ khi con người xuất hiện, ngôn ngữ tự nhiên được thể hiện qua tiếng nói, chữ viết, động tác. Tất cả đều phải tuân thủ theo nguyên tắc nhất định gọi là ngữ pháp.

Ngôn ngữ hình thức: là tập hợp các ký hiệu và quy định do con người đặt ra. Trong tin học ngôn ngữ hình thức là công cụ để giao tiếp giữa người với máy hoặc máy với máy.

Các hệ quản trị cơ sở dữ liệu bao giờ cũng cung cấp một ngôn ngữ hình thức làm công cụ để giao tiếp giữa người và hệ thống. Đó là ngôn ngữ dữ liệu.

Trong các ngôn ngữ lập trình bình thường, các mệnh đề khai báo và khai thác đều là các thành phần của một ngôn ngữ. Nhưng về mặt cơ sở dữ liệu, thường tách ra hai chức năng: khai báo và tính toán là hai ngôn ngữ phân biệt.

1.3.2. Ngôn ngữ mô tả dữ liệu (Data Definition Language - DDL)

Định nghĩa: là ngôn ngữ dùng để đặc tả các lược đồ cơ sở dữ liệu gồm các thao tác:

- Khai báo tên, cấu trúc và các quan hệ giữa các đối tượng.
- Đổi tên thành phần, đổi cấu trúc của các đối tượng.
- Thêm, xóa, sửa các thành phần trong cấu trúc.
- Thêm, xoá các quan hệ giữa các bảng.

Đặc điểm:

- Mỗi DDL thuộc HQTCSDL riêng, và ‘nói’ theo cú pháp của của mô hình dữ liệu tương ứng.
- DDL được dùng khi thiết kế hoặc thay đổi cấu trúc CSDL.

Trong đa phần các hệ quản trị cơ sở dữ liệu, DDL thường có bộ phận đặc biệt gọi là ngôn ngữ lưu trữ và định nghĩa dữ liệu: là một tập hợp các định nghĩa chính xác về cách thức dữ liệu lưu trữ trên bộ nhớ ngoài và cách truy xuất chúng. Thiết kế chi tiết của CSDL vật lý (thường được che đi đối với mức người dùng) được thực hiện bởi các chương trình con của HQTCSDL, sau khi các khối này biên dịch các mệnh đề trong DDL.

Ngôn ngữ mô tả dữ liệu có thể sử dụng trực tiếp trong cửa sổ lệnh, hoặc được thực thi trong một chương trình, hoặc được thực hiện bằng thao tác thông qua giao diện đồ họa.

1.3.3. Ngôn ngữ thao tác dữ liệu (Data Manipulation Language - DML)

Là ngôn ngữ cho phép thực hiện các thao tác làm thay đổi dữ liệu, có các chức năng như:

- Thêm, xóa, cập nhật dữ liệu vào cơ sở dữ liệu.
- Thực hiện các phép tính số học và logic trên các thuộc tính.
- Kết nối các chuỗi thao tác.
- Gán, nhận kết quả để in hoặc để xây dựng các đối tượng khác.

Ở mức vật lý, chúng ta cần định ra các giải thuật để truy xuất hiệu quả dữ liệu. Ở mức trừu tượng cao hơn, ta cần nhấn mạnh đến tính dễ sử dụng. Mục đích là cung cấp sự tương tác giữa người và máy.

Cách dùng ngôn ngữ thao tác dữ liệu: các lệnh DML có thể sử dụng theo nhiều cách:

- Tương tác trực tiếp đến HQTCSDL từ hệ thống menu và giao diện đồ họa.
- Sử dụng các lệnh cơ bản của HQTCSDL.
- Thông qua trình ứng dụng.

1.3.4. Các dạng thể hiện của ngôn ngữ

a. Kiểu đại số

Sử dụng rộng rãi các biểu thức đại số và logic. Khi đó đòi hỏi phải biểu diễn các phát biểu (yêu cầu) thành các biểu thức toán học.

Ví dụ: chọn những học sinh có xếp loại là ‘Khá’

Biểu diễn: SINHVIEN(XepLoai = ‘Khá’)[MSSV, HoTen]

Ý nghĩa của cách biểu diễn trên, chọn ra những sinh viên có xếp loại là khá, nhưng chỉ lấy các thông tin về mã số sinh viên và họ tên.

b. Kiểu logic

Sử dụng các vị từ.

Ví dụ: có hay không sinh viên ‘Nguyễn Hoài Ân’

?t.SinhVien : t.HoTen = ‘Nguyễn Hoài Ân’

c. Kiểu thủ tục

Ngôn ngữ kiểu thủ tục yêu cầu người dùng chỉ rõ dữ liệu cần và cách tìm ra dữ liệu mong muốn.

Ví dụ: với ví dụ trong kiểu logic thì ở kiểu thủ tục phải:

Chỉ rõ dữ liệu: chọn ra những bộ từ quan hệ sinh viên thỏa điều kiện họ và tên là ‘Nguyễn Hoài Ân’

Cách tìm: chiết các thuộc tính HoTen từ quan hệ sinh viên và chọn trên hàng để lấy thông tin của sinh viên ‘Nguyễn Hoài Ân’.

d. Kiểu phi thủ tục

Thường được gọi là kiểu mô tả, ngôn ngữ phi thủ tục thường dễ đọc và dễ sử dụng tuy nhiên so về tính hiệu quả thì không bằng ngôn ngữ kiểu thủ tục.

e. Ngôn ngữ tự nhiên

Là dạng ngôn ngữ đang được nghiên cứu, gắn liền với lĩnh vực nhận dạng và phân tích tiếng nói.

Một vài ngôn ngữ tự nhiên đã được cài đặt thông qua sự hỗ trợ của các thiết bị ngoại vi (chuột, bàn phím) để tiến hành giao tiếp giữa người dùng và hệ thống.

1.3.6. Quản lý giao tác

a. Định nghĩa

Một giao tác là tập hợp các thao tác thực hiện một chức năng logic trong một trình ứng dụng.

b. Ví dụ

Gọi T1 là giao tác chuyển 100\$ từ tài khoản A sang tài khoản B. Giao tác T1 có tập hợp các thao tác như sau:

Kiểm tra số tiền trong tài khoản A không nhỏ hơn 100\$ thì

- Giảm 100\$ trong tài khoản A
- Tăng thêm 100\$ vào tài khoản B

c. Các tính chất

- **Tính nguyên tố:** giao tác được gọi là thành công khi và chỉ khi thực hiện thành công mọi thao tác bên trong ngược lại là giao tác thất bại. (hoặc là tất cả hoặc là không có gì)
- **Tính chất nhất quán:** giao tác phải đảm bảo tính nhất quán trong cơ sở dữ liệu. Nếu cơ sở dữ liệu nhất quán trước khi bắt đầu giao tác thì sau khi kết thúc giao tác cơ sở dữ liệu vẫn đảm bảo tính nhất quán.

- **Tính chất riêng biệt:** mặc dù nhiều giao tác có thể được khai thác cạnh tranh nhau, nhưng phải đảm bảo đối với mỗi cặp thao tác T_i và T_j bất kỳ thì hoặc là T_i hoàn thành giao tác trước khi T_j bắt đầu, hoặc T_j phải hoàn thành giao tác trước khi T_i bắt đầu thực hiện giao tác.
- **Tính chất bền vững:** kết quả hoàn tất của một giao tác phải luôn được bảo lưu, cho dù hệ thống có bị trục trặc nào đó.

d. **Trạng thái giao tác:**

Để đảm bảo 4 tính chất của giao tác thì giao tác sẽ có 3 trạng thái:

- **Hoàn tất (commit):** sau khi thực hiện giao tác xong, HQTCSDL sẽ kiểm tra, nếu thấy thỏa các ràng buộc toàn vẹn và các ràng buộc khác thì sẽ xác nhận là trạng thái hoàn tất là đúng.
- **Thoái bộ (rollback):** các ràng buộc chưa thỏa, hoặc do người dùng không muốn thực hiện tiếp thao tác mặc dù giao tác đã thực hiện xong nhưng cơ sở dữ liệu phải trả về trạng thái ban đầu (trước khi thực hiện giao tác).
- **Thất bại (fail):** khi giao tác không được thực hiện trọn vẹn. Do lỗi hệ thống hoặc người dùng muốn ngừng giao tác.

e. Các bộ phận phụ trách giao tác

Bộ phận quản lý giao tác: tùy theo kiểu hư hỏng của hệ thống, một giao tác có thể thực hiện không thành công. Muốn đảm bảo tính nguyên tố của giao tác, thì một giao tác thất bại không ảnh hưởng đến tình trạng của CSDL khi đó CSDL phải được trả về trạng thái trước khi thực hiện giao tác.

Bộ phận điều khiển cạnh tranh: khi nhiều giao tác cập nhật cơ sở dữ liệu có sự cạnh tranh, tính nhất quán của cơ sở dữ liệu có thể không được bảo toàn, mặc dù mỗi giao tác đều thực hiện đúng. Bộ phận điều khiển cạnh tranh sẽ kiểm soát và điều khiển sự tương tác giữa các thao tác để đảm bảo tính nhất quán của CSDL.

BÀI TẬP CHƯƠNG 1

Chọn câu trả lời đúng nhất

1. Cơ sở dữ liệu là:

- Một bộ sưu tập các loại dữ liệu mô tả các đối tượng của thế giới thực.
- Một tập các đối tượng của thế giới thực.
- Một tập các chương trình ứng dụng và dữ liệu.
- Hệ quản trị cơ sở dữ liệu.

2. Cơ sở dữ liệu là tài nguyên thông tin.....

- Truy xuất theo chế độ trực tuyến.
- Chia sẻ chung cho nhiều người sử dụng.
- Cài đặt trên hệ quản trị cơ sở dữ liệu.
- Trên các thiết bị đầu cuối.

3. Đối tượng nghiên cứu của CSDL là.....

- A. Các thực thể.
- B. Các mối quan hệ.
- C. Các thực thể và mối quan hệ giữa các thực thể.
- D. Các mối liên kết giữa các thực thể.

4. Cơ sở dữ liệu có những ưu điểm:

- A. Giảm bớt dư thừa dữ liệu trong lưu trữ.
- B. Tránh không nhất quán trong lưu trữ dữ liệu.
- C. Có khả năng lưu trữ dữ liệu lớn.
- D. Cả ba ý trên đều đúng.

5. Mô hình kiến trúc 3 mức của hệ CSDL gồm:

- A. Mức trong, mức vật lý, mức dữ liệu.
- B. Mức dữ liệu, mức trong, mức ngoài
- C. mức ngoài, mức quan niệm, mức trong.
- D. Mức trong, mức ngoài, mức khung nhìn.

6. Mô hình dữ liệu là cách nhìn....

- A. Toàn bộ nội dung thông tin của CSDL.
- B. Khung nhìn của người sử dụng.
- C. Là cách tổ chức lưu trữ.
- D. Là cách biểu diễn cơ sở dữ liệu trừu tượng ở mức thấp nhất.

7. Mục tiêu của các hệ quản trị cơ sở dữ liệu là:

- A. Cấu trúc lưu trữ dữ liệu và chiến lược truy nhập tới các hệ cơ sở dữ liệu.
- B. Dữ liệu chỉ được biểu diễn, mô tả một cách duy nhất.
- C. Cấu trúc dữ liệu và các chương trình ứng dụng không phụ thuộc lẫn nhau.
- D. Cấu trúc lưu trữ dữ liệu và các chương trình ứng dụng phụ thuộc lẫn nhau.

8. Có rất nhiều cách nhìn dữ liệu ở..... người sử dụng khác nhau có cách nhìn dữ liệu khác nhau và các hệ ứng dụng khác nhau có những cách nhìn dữ liệu cũng khác nhau.

- A. Mô hình ngoài.
- B. Mô hình quan niệm.
- C. Mô hình trong.
- D. Mô hình dữ liệu

Chương 2: ĐẠI SỐ QUAN HỆ

Trong chương này, ta sẽ đề cập đến cách tiếp cận các phương pháp thiết kế ngôn ngữ hỏi trên các quan hệ bằng các phép toán đại số. Vì trong thực tế, các hệ quản trị cơ sở dữ liệu có những cú pháp riêng, bắt buộc người dùng phải tuân thủ khi thiết kế ngôn ngữ hỏi vì thế để thuận tiện cho việc xây dựng ngôn ngữ hỏi mà không phụ thuộc vào hệ quản trị cơ sở dữ liệu thì việc sử dụng đại số quan hệ được xem là phù hợp nhất.

Đại số quan hệ là thực hiện câu hỏi trên cơ sở dữ liệu như:

- Chiếu lấy các cột (thuộc tính) cần thiết từ các quan hệ.
- Chọn các bộ thỏa điều kiện từ các quan hệ.
- Thực hiện tích Decartes giữa các quan hệ.
- Xem các quan hệ như các tập hợp, thực hiện các phép toán trên tập hợp: hợp, giao, trừ, chia, ...
- Thực hiện tính toán, thống kê trên nhóm.

2.1. Giới thiệu

Đại số quan hệ là tập hợp các phép toán thao tác trên các thuộc tính và các bộ của các quan hệ trong một cơ sở dữ liệu cho trước, để trả lời các câu hỏi liên quan các thông tin về dữ liệu. Các phép toán bao gồm: chiếu, chọn, hợp, hiệu, giao và một số phép tính khác để trả lời các câu truy vấn.

Gọi R là một quan hệ trên tập thuộc tính (A_1, A_2, \dots, A_n).

Ký hiệu: $R(A_1, A_2, \dots, A_n)$

Lưu ý:

Toán hạng của đại số quan hệ cũng là một quan hệ.

Kết quả của phép toán đại số quan hệ cũng là một quan hệ.

2.2. Các phép toán cơ bản

2.2.1. Phép chọn (Selection)

a. Mục đích

Dùng để lấy các bộ con của quan hệ đã cho (R), thỏa điều kiện F cho trước.

Ký hiệu: $\sigma_F(R)$

Trong đó F là:

- Biểu thức so sánh sử dụng các toán tử: $>$, \leq , $=$, $>=$, \neq
- hoặc
- Biểu thức logic sử dụng các toán tử: \neg (not), \wedge (and), \vee (or)

Vậy F là biểu thức tổ hợp của biểu thức so sánh và biểu thức logic. Kết quả của F là hoặc ‘đúng’ hoặc ‘sai’.

Ví dụ: cho quan hệ NHANVIEN(MaNV, HoTen, Phai, Tuoi). Chọn các nhân viên có tuổi từ 30.

Điều kiện F là: ($Tuoi \geq 30$)

Quan hệ R là: NHANVIEN

$\rightarrow \sigma_{Tuoi \geq 30}(NHANVIEN)$

b. Định nghĩa

Cho R(A_1, A_2, \dots, A_n) và F.

Thực hiện phép chọn trên R thỏa điều kiện F được ký hiệu như sau: $\sigma_F(R)$.

$\sigma_F(R) = \{t \mid (t \in R) \wedge F(t)\}$ t: các bộ có trong R; F(t): các bộ t thỏa điều kiện F.

c. Ví dụ

Ví dụ 1: cho quan hệ SINHVIEN(MaSV, HoTen, NgaySinh, MaNganh) có các bộ được mô tả như sau:

SINHVIEN	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1991	CNTT
	SV002	Đương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1995	CNTT
	SV004	Nguyễn Dương Thái	10/7/1990	CKCT
	SV005	Lê Thị Lan	07/12/1987	CNOT

Yêu cầu: chọn các sinh viên (SINHVIEN) có ngày sinh (NgaySinh) sau ngày ‘01/10/1992’

Thực hiện:

- Xác định điều kiện F: NgaySinh > ‘01/10/1992’
- Thực hiện chọn: $\sigma_F(R) = \sigma_{NgaySinh > '01/10/1992'}(SINHVIEN)$

Kết quả của phép chọn trên là quan hệ như sau:

	MaSV	HoTen	NgaySinh	MaNganh
	SV002	Đương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1995	CNTT

Lưu ý: khi thực hiện so sánh trên kiểu dữ liệu ngày tháng năm thì năm được ưu tiên so sánh trước rồi mới đến tháng và sau cùng là ngày.

Hằng ngày tháng năm thường được đặt trong cặp dấu {} hay cặp ‘’.

Ví dụ 2: sử dụng lại quan hệ SINHVIEN. Chọn ra những sinh viên theo học ngành có mã ngành là ‘CNOT’ và có tuổi trên 25.

Thực hiện:

- Xác định điều kiện F:

$(\text{year}(\text{GetDate}()) - \text{year}(\text{NgaySinh}) > 25) \wedge (\text{MaNganh} = 'CNOT')$.

- Thực hiện chọn:

$\sigma_F(R) = \sigma_{((\text{year}(\text{GetDate}()) - \text{year}(\text{NgaySinh}) > 25) \wedge (\text{MaNganh} = 'CNOT'))}(\text{SINHVIEN})$

Kết quả của phép chọn trên là quan hệ như sau: (năm hiện tại là 2014)

	MaSV	HoTen	NgaySinh	MaNganh
	SV005	Lê Thị Lan	07/12/1987	CNOT

2.2.2. Phép chiếu (Projection)

a. Mục đích

Dùng để lấy tất cả các bộ trên quan hệ R nhưng chỉ lấy các thuộc tính X (loại bỏ một vài thuộc tính không phải X trên quan hệ R).

Ký hiệu: $\Pi_X(R)$

b. Định nghĩa

Cho $R(A_1, A_2, \dots, A_n)$ có tập thuộc tính là A. Thực hiện phép chiếu lấy các thuộc tính $X \in A$ trên quan hệ R có: ký hiệu là $\Pi_X(R)$; Kết quả là một quan hệ có tập thuộc tính X.

$\Pi_X(R) = \{X(t) \mid t \in R\}$ X(t) các bộ t chỉ có thuộc tính X.

c. Ví dụ

Ví dụ 1: cho quan hệ SINHVIEN(MaSV, HoTen, NgaySinh, MaNganh) có các bộ được mô tả như sau:

SINHVIEN	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1991	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1995	CNTT
	SV004	Nguyễn Dương Thái	10/7/1990	CKCT
	SV005	Lê Thị Lan	07/12/1987	CNOT
	SV006	Dương Trọng Đạt	7/12/1999	CNOT

Yêu cầu: chiếu các thông tin Mã số (MaSV), họ tên (HoTen), mã ngành (MaNganh) của sinh viên.

Thực hiện:

- Xác định các thuộc tính cần chiếu X: MaSV, HoTen, MaNganh
- Thực hiện chiếu: $\Pi_X(R) = \Pi_{\text{MaSV}, \text{HoTen}, \text{MaNganh}}(\text{SINHVIEN})$

Kết quả của phép chọn trên là quan hệ như sau:

	MaSV	HoTen	MaNganh
	SV001	Nguyễn Văn An	CNTT
	SV002	Dương Tùng Lê	CKCT
	SV003	Phạm Thị Ái	CNTT
	SV004	Nguyễn Dương Thái	CKCT
	SV005	Lê Thị Lan	CNOT
	SV006	Dương Trọng Đạt	CNOT

2.2.3. Quan hệ tương thích

Hai quan hệ R và S được gọi là tương thích nếu có cùng sơ đồ quan hệ (có cùng các thuộc tính) hoặc các thuộc tính của từng quan hệ là như nhau, không quan tâm đến dữ liệu của các bộ.

R(A₁, ..., A_n) tương thích S(A₁, ..., A_n) hoặc

R(A₁, ..., A_n) tương thích S(B₁, ..., B_n) sao cho miền giá trị A_i tương đương miền giá trị của B_i (i= 1.. n).

Ví dụ 1: cho quan hệ MATHANG và HANGHOA

MATHANG (MaMH, TenMH)

HANGHOA (MaHH, TenHH)

Ví dụ 2: quan hệ SINHVIEN và SV quan hệ được gọi là tương thích.

SINHVIEN (MSSV, HT, NS, MaNganh)

SINHVIEN	MSSV	HT	MaNganh
	SV001	Nguyễn Văn An	CNTT
	SV002	Dương Tùng Lê	CKCT
	SV003	Phạm Thị Ái	CNTT

và SV (MaSV, HoTen, NgaySinh, MaNganh)

SV	MaSV	HoTen	MaNganh
	SV004	Nguyễn Dương Thái	CKCT
	SV005	Lê Thị Lan	CNOT

2.2.4. Phép trừ (Minus)

a. Mục đích

Lấy các bộ thuộc quan hệ R mà không thuộc quan hệ S (với R và S là hai quan hệ tương thích cho trước).

Ký hiệu: $R - S$

b. Định nghĩa

Cho hai quan hệ tương thích $R(A_1, A_2, \dots, A_n)$ và $S(A_1, A_2, \dots, A_n)$. Hiệu của quan hệ R với quan hệ S là một quan hệ có tập thuộc tính (A_1, A_2, \dots, A_n) , có các bộ t thuộc các bộ của R và không thuộc các bộ của S.

Ký hiệu của phép trừ giữa hai quan hệ R và S: $R - S$

$$R - S = \{t \mid (t \in R) \wedge (t \notin S)\}$$

c. Ví dụ

Ví dụ: cho quan hệ SINHVIEN₁ và SINHVIEN₂ như sau:

SINHVIEN₁ (MaSV, HoTen, NgaySinh, MaNganh)

SINHVIEN ₁	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV004	Trần Thái Văn	02/10/1994	CNOT
	SV005	Lê Dương	12/10/1993	CNTT

SINHVIEN₂ (MaSV, HoTen, NgaySinh, MaNganh)

SINHVIEN ₂	MaSV	HoTen	NgaySinh	MaNganh
	SV010	Lý Thanh Hải	10/8/1993	CNTT
	SV004	Trần Thái Văn	02/10/1994	CNOT

Phép trừ của SINHVIEN₁ với SINHVIEN₂: SINHVIEN₁ – SINHVIEN₂

Kết quả:

	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT

SV005	Lê Dương	12/10/1993	CNTT
-------	----------	------------	------

Nếu quan hệ R không có bộ nào hoặc các bộ của quan hệ S trùng với các bộ trong quan hệ S thì kết quả R-S sẽ như thế nào?

2.2.5. Phép hợp (Union)

a. Mục đích

Lấy các bộ thuộc quan hệ R và các bộ thuộc quan hệ S với R và S là hai quan hệ tương thích cho trước.

Ký hiệu: $R \cup S$

b. Định nghĩa

Là một quan hệ $T(A_1, A_2, \dots, A_n)$ khi thực hiện lấy những bộ của quan hệ $R(A_1, A_2, \dots, A_n)$ vào các bộ của quan hệ $S(A_1, A_2, \dots, A_n)$

Ký hiệu phép hợp giữa R và S: $R \cup S$.

$$R \cup S = \{ t \mid (t \in R) \vee (t \in S) \}$$

c. Ví dụ

Ví dụ: thực hiện phép giao cho hai quan hệ SINHVIEN₁ và SINHVIEN₂:

SINHVIEN₁ (MaSV, HoTen, NgaySinh, MaNganh)

SINHVIEN ₁	MaSV	HoTen	NgaySinh	MaNganh
SINHVIEN ₁	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV004	Trần Thái Văn	02/10/1994	CNOT
	SV005	Lê Dương	12/10/1993	CNTT
	SV009	Hồ Thành Toại	10/12/1989	CNOT
	SV010	Lý Thanh Hải	10/8/1993	CNTT

SINHVIEN₂ (MaSV, HoTen, NgaySinh, MaNganh)

SINHVIEN ₂	MaSV	HoTen	NgaySinh	MaNganh
SINHVIEN ₂	SV010	Lý Thanh Hải	10/8/1993	CNTT
	SV012	Phan Hoài Anh	22/4/1993	CKCT
	SV015	Phan Thành Tài	10/12/1989	CNOT

Thực hiện kết nối tự nhiên: SINHVIEN₁ \cup SINHVIEN₂

Kết quả phép kết nối tự nhiên giữa SINHVIEN₁ và SINHVIEN₂:

	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV004	Trần Thái Văn	02/10/1994	CNOT
	SV005	Lê Dương	12/10/1993	CNTT
	SV009	Hồ Thành Toại	10/12/1989	CNOT
	SV010	Lý Thanh Hải	10/8/1993	CNTT
	SV012	Phan Hoài Anh	22/4/1993	CKCT
	SV015	Phan Thành Tài	10/12/1989	CNOT

2.2.6. Phép tích Decartes

Cho hai quan hệ R(A₁, A₂, ..., A_n) có I bộ và S(B₁, B₂, ..., B_m) có J bộ. Tích Decartes của hai quan hệ R và S là quan hệ T(A₁, A₂, ..., A_n, B₁, B₂, ..., B_m) có I.J bộ do từng bộ trong R kết hợp với J bộ có trong S.

Ký hiệu phép tích Decartes giữa R và S: R x S

$$R \times S = \{ tq \mid (t \in R) \wedge (q \in S) \}$$

Ví dụ: Cho quan hệ SINHVIEN và quan hệ NGANH như sau:

SINHVIEN2 (MaSV, HoTen, NgaySinh)

SINHVIEN	MaSV	HoTen	NgaySinh
	SV010	Lý Thanh Hải	10/8/1993
	SV012	Phan Hoài Anh	22/4/1993

NGANH (MaN, TenN)

NGANH	MaN	TenN
	HTTT	Hệ thống thông tin
	CNPM	Công nghệ phần mềm

Tích Decartes giữa SINHVIEN và NGANH: SINHVIEN x NGANH

Kết quả:

	MaSV	HoTen	NgaySinh	MaN	TenN
	SV010	Lý Thanh Hải	10/8/1993	HTTT	Hệ thống thông tin
	SV010	Lý Thanh Hải	10/8/1993	CNPM	Công nghệ phần mềm
	SV012	Phan Hoài Anh	22/4/1993	HTTT	Hệ thống thông tin
	SV012	Phan Hoài Anh	22/4/1993	CNPM	Công nghệ phần mềm

2.3. Các phép toán khác

2.3.1. Phép giao (Intersection)

Ký hiệu: \cap

Là một quan hệ $T(A_1, A_2, \dots, A_n)$ khi thực hiện phép giao là chọn ra những bộ có trong quan hệ $R(A_1, A_2, \dots, A_n)$ và cũng có trong các bộ của quan hệ $S(A_1, A_2, \dots, A_n)$.

Ký hiệu phép giao giữa R và S: $R \cap S$

$$R \cap S = \{ t \mid (t \in R) \wedge (t \in S) \}$$

Ví dụ: Thực hiện phép giao giữa hai quan hệ sau:

SINHVIEN₁ (MaSV, HoTen, NgaySinh, MaNganh)

SINHVIEN₁	MaSV	HoTen	NgaySinh	MaNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV004	Trần Thái Văn	02/10/1994	CNOT
	SV005	Lê Dương	12/10/1993	CNTT
	SV006	Mai Thành Vân	22/10/1995	CKCT

SINHVIEN₂ (MaSV, HoTen, NgaySinh, MaNganh)

SINHVIEN₂	MaSV	HoTen	NgaySinh	MaNganh
	SV010	Lý Thanh Hải	10/8/1993	CNTT
	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV012	Phan Hoài Anh	22/4/1993	CKCT
	SV005	Lê Dương	12/10/1993	CNTT

SV006	Mai Thanh Vân	22/10/1995	CKCT
-------	---------------	------------	------

Thực hiện phép giao giữa SINHVIEN₁ với SINHVIEN₂:

SINHVIEN₁ ∩ SINHVIEN₂

Kết quả phép giao SINHVIEN₁ và SINHVIEN₂:

	MaSV	HoTen	NgaySinh	MaNganh
	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV005	Lê Dương	12/10/1993	CNTT
	SV006	Mai Thanh Vân	22/10/1995	CKCT

2.3.2. Phép chia (Division)

Ký hiệu: \div

Cho quan hệ R(A₁, A₂, ..., A_n) có tập Thuộc tính là A và các bộ là r, quan hệ S(B₁, B₂, ..., B_m) có tập thuộc tính là B và các bộ là s. Khi đó phép chia của quan hệ R cho quan hệ S là một quan hệ có tập thuộc tính là A\B (các thuộc tính chỉ có trong A và không có trong B), có các bộ là các bộ thuộc r khi và chỉ khi bộ s ∈ r.

Ký hiệu: quan hệ R chia quan hệ S: R \div S

$$R \div S = \{r \mid r \in R, r \text{ chứa thuộc tính của } R \text{ nhưng không thuộc } S\}$$

Ví dụ: cho lược đồ của hai quan hệ sau

SV(MaSV, HoTen, NgaySinh, MaNganh, TenNganh)

SV	MaSV	HoTen	NgaySinh	MaNganh	TenNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT	Công Nghệ Thông Tin
	SV002	Dương Tùng Lê	12/4/1994	CKTT	Cơ Khí Thực Tập
	SV001	Nguyễn Văn An	10/6/1993	CNCT	Cơ Khí Chế Tạo
	SV004	Trần Thái Văn	02/10/1994	CNOT	Công Nghệ Ô Tô
	SV005	Lê Dương	12/10/1993	CNTT	Công Nghệ Thông Tin

NGANH (MaNganh, TenNganh)

NGANH	MaNganh	TenNganh
	CNTT	Công Nghệ Thông Tin
	CKCT	Cơ Khí Chế Tạo

Ký hiệu phép chia của SV cho NGANH: SV \div NGANH

Kết quả

	MaSV	HoTen	NgaySinh
	SV001	Nguyễn Văn An	10/6/1993

2.3.3. Phép kết nối (Join)

Ký hiệu: \bowtie

Cho quan hệ R(A_1, A_2, \dots, A_n), quan hệ S(B_1, B_2, \dots, B_m). Gọi θ là một trong các phép toán: $>$, $<=$, $=$, $>=$, \neq . Phép kết nối giữa quan hệ R trên thuộc tính A với quan hệ S trên thuộc tính B là những bộ t của tích Decartes sao cho $t[A_i] \theta t[B_j]$. ($i=1-n ; j=1-m$)

Ký hiệu phép kết nối giữa quan hệ R và quan hệ S: $R \bowtie_{A_i \theta B_j} S$

$$R \bowtie_{A_i \theta B_j} S = \{v = (t, u) | (t \in R) \wedge (u \in S) \wedge (t.A_i \theta u.B_j)\}$$

Ví dụ: Cho quan hệ SINHVIEN và NHANVIEN như sau:

SINHVIEN (MaSV, HoTenSV, NgaySinh)

SINHVIEN	MaSV	HoTenSV	NgaySinh
	SV001	Nguyễn Văn An	10/6/1986
	SV002	Dương Tùng Lê	12/4/1994
	SV003	Phạm Thị Ái	02/10/1988
	SV004	Trần Thái Văn	02/10/1994
	SV005	Lê Dương	12/10/1987

NHANVIEN (HoTenNV, NS)

NHANVIEN	HoTenNV	NS
	Nguyễn Hoài Bảo	12/6/1983
	Phan Thanh Tài	22/4/1988
	Dương Trọng Bình	02/11/1989

Yêu cầu: Thực hiện kết nối sao cho NgaySinh của sinh viên (SINHVIEN) lớn hơn NS của nhân viên (NV).

Thực hiện phép kết nối:

Ta có:

R: SINHVIEN

S: NHANVIEN

A: NgaySinh

B: NS

$\theta: <$

Phép kết nối: SINHVIEN \bowtie NHANVIEN
NgaySinh > NS

Kết quả :

	MaSV	HoTenSV	NgaySinh	HoTenNV	NS
	SV001	Nguyễn Văn An	10/6/1986	Phan Thanh Tài	22/4/1988
	SV001	Nguyễn Văn An	10/6/1986	Dương Trọng Bình	02/11/1989
	SV003	Phạm Thị Ái	02/10/1988	Dương Trọng Bình	02/11/1989
	SV005	Lê Dương	12/10/1987	Dương Trọng Bình	02/11/1989

2.3.4. Phép kết nối tự nhiên (Natural Join)

Ký hiệu: *

Là một quan hệ $T(A_1, A_2, \dots, J, A_n, X_1, X_2, \dots, X_m)$ khi thực hiện kết nối quan hệ $R(A_1, A_2, \dots, J, A_n)$ với quan hệ $S(X_1, X_2, \dots, J, X_m)$ thông qua cột có cùng tên là J .

Phép kết nối tự nhiên còn gọi là kết nối nội (Inner Join)

Ký hiệu phép kết nối tự nhiên giữa R và S: $R * S$

$$R * S = \{v = (t, u) | (t \in R) \wedge (u \in S) \wedge (t.J = u.J)\}$$

Ví dụ: Cho quan hệ sinh viên (SINHVIEN) và quan hệ ngành (NGANH):

SINHVIEN (MaSV, HoTen, NgaySinh, MaN)

SINHVIEN	MaSV	HoTen	NgaySinh	MaN
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Dương Tùng Lê	12/4/1994	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV004	Trần Thái Văn	02/10/1994	CNOT
	SV005	Lê Dương	12/10/1993	CNTT
	SV006	Mai Thanh Vân	22/10/1995	CKCT
	SV010	Lý Thanh Hải	10/8/1993	CNTT

NGANH (MaN, TenNganh)

NGANH	MaN	TenNganh
	CNTT	Công Nghệ Thông Tin
	CKCT	Cơ Khí Chế Tạo
	CNOT	Công Nghệ Ô Tô

Kết quả kết nối tự nhiên giữa quan hệ SINHVIEN và quan hệ NGANH:

Thực hiện:

SINHVIEN * NGANH

Kết quả:

	MaSV	HoTen	NgaySinh	MaN	TenNganh
	SV001	Nguyễn Văn An	10/6/1993	CNTT	Công Nghệ Thông Tin
	SV002	Dương Tùng Lê	12/4/1994	CKCT	Cơ Khí Chế Tạo
	SV003	Phạm Thị Ái	02/10/1993	CNTT	Công Nghệ Thông Tin
	SV004	Trần Thái Văn	02/10/1994	CNOT	Công Nghệ Ô Tô
	SV005	Lê Dương	12/10/1993	CNTT	Công Nghệ Thông Tin
	SV006	Mai Thanh Vân	22/10/1995	CKCT	Cơ Khí Chế Tạo
	SV010	Lý Thanh Hải	10/8/1993	CNTT	Công Nghệ Thông Tin

2.3.5. Phép kết nối mở rộng (Outer Join)

Cho hai quan hệ R (A_1, A_2, \dots, A_n) và quan hệ S(X_1, X_2, \dots, J, X_m).

a. Định nghĩa:

Là phép toán cho phép làm việc với thông tin bị thiếu, tức là vẫn thực hiện phép kết nối tự nhiên trên các trị trống của thuộc tính dùng để kết nối.

Có ba loại kết nối mở rộng: Mở rộng trái, mở rộng phải và mở rộng toàn bộ

Ví dụ: Cho hai quan hệ NHANVIEN và NVLDAO

NHANVIEN (MaNV, HoTenNV, NgaySinh, MaDV)

NHANVIEN	MaNV	HoTenNV	NgaySinh	MaDV
	001	Nguyễn Văn An	10/6/1993	TCHC
	002	Dương Tùng Lê	12/4/1994	TCHC
	003	Phạm Thị Ái	02/10/1993	QLNS

	004	Trần Thái Văn	02/10/1994	QLNS
	010	Lý Thanh Hải	10/8/1993	KTTV

NVLDao (MaNV, HoTenLD)

NVLDao	MaNV	HoTenLD
	001	Nguyễn Văn An
	002	Dương Tùng Lê
	003	Phạm Thị Ái
	015	Lê Thanh Hoài
	020	Dương Hoàng

b. Phép kết nối mở rộng trái (Left Join)

Ký hiệu \bowtie

Là một quan hệ sau khi thực hiện kết nối tự nhiên giữa R và S (R^*S) hợp với các bộ (r,s) mà r không thuộc S khi đó s gán thành giá trị rỗng.

Ký hiệu quan hệ R kết nối mở rộng trái với S: $R \bowtie S$.

$$R \bowtie S = \{ (R^*S) \cup (r,s) | (r \in R) \wedge (s \notin S) ; s \text{ gán thành null} \}$$

Ví dụ: Thực hiện kết nối mở rộng trái cho quan hệ NHANVIEN và NVLDao (NHANVIEN \bowtie NVLDao) có kết quả như sau:

	MaNV	HoTenNV	NgaySinh	MaDV	HoTenLD
	001	Nguyễn Văn An	10/6/1993	TCHC	Nguyễn Văn An
	002	Dương Tùng Lê	12/4/1994	TCHC	Dương Tùng Lê
	003	Phạm Thị Ái	02/10/1993	QLNS	Phạm Thị Ái
	004	Trần Thái Văn	02/10/1994	QLNS	
	010	Lý Thanh Hải	10/8/1993	KTTV	

c. Phép kết nối mở rộng phải (Right Join)

Ký hiệu: $\bowtie\sqsubset$

Là một quan hệ sau khi thực hiện kết nối tự nhiên R và S (R^*S) hợp với các bộ (r,s) mà s không thuộc R khi đó r gán thành giá trị rỗng.

Ký hiệu kết nối mở rộng phải giữa quan hệ R và S: $R \bowtie\sqsubset S$

$$R \bowtie\sqsubset S = \{ (R^*S) \cup (r,s) | (s \in S) \wedge (r \notin R) ; r \text{ gán thành null} \}$$

Giáo trình Cơ sở dữ liệu

Ví dụ: Thực hiện kết nối mở rộng phải của quan hệ NHANVIEN và quan hệ NVLDao có kết quả như sau:

Thực hiện:

NHANVIEN $\bowtie\bowtie$ NVLDao

Kết quả:

	MaNV	HoTenNV	NgaySinh	MaDV	HoTenLD
	001	Nguyễn Văn An	10/6/1993	TCHC	Nguyễn Văn An
	002	Đương Tùng Lê	12/4/1994	TCHC	Đương Tùng Lê
	003	Phạm Thị Ái	02/10/1993	QLNS	Phạm Thị Ái
	015				Lê Thanh Hoài
	020				Đương Hoàng

d. Phép kết nối mở rộng toàn phần (Full Join)

Ký hiệu: $\bowtie\bowtie$

Là phép toán thực hiện phép toán hợp giữa kết quả của phép kết nối tự nhiên với các bộ (r,s) mà r không thuộc S khi đó s gán thành giá trị rỗng và hợp với các bộ (r,s) mà s không thuộc R khi đó u gán thành giá trị rỗng..

Ký hiệu phép kết nối mở rộng hai vé của quan hệ R và S: R $\bowtie\bowtie$ S

R $\bowtie\bowtie$ S = (R * S) \cup { (r,s) | (r \in R) \wedge (r \notin S) ; u gán thành null } \cup
{ (r,s) | (s \in S) \wedge (s \notin R) ; s gán thành null } }

Ký hiệu phép kết nối mở rộng hai vé giữa quan hệ NHANVIEN và NVLDao có kết quả như sau:

Thực hiện:

NHANVIEN $\bowtie\bowtie$ NVLDao

Kết quả:

	MaNV	HoTenNV	NgaySinh	MaDV	HoTenLD
	001	Nguyễn Văn An	10/6/1993	TCHC	Nguyễn Văn An
	002	Đương Tùng Lê	12/4/1994	TCHC	Đương Tùng Lê
	003	Phạm Thị Ái	02/10/1993	QLNS	Phạm Thị Ái
	004	Trần Thái Văn	02/10/1994	QLNS	
	010	Lý Thanh Hải	10/8/1993	KTTV	

015				Lê Thanh Hoài
020				Dương Hoàng

2.3.6. Phép đặt tên lại cho biểu thức và thuộc tính

Nhu cầu đặt tên lại

Các kết quả của biểu thức đại số quan hệ không có tên để chúng ta gọi vì thế đặt tên là điều cần thiết.

Trong quá trình thực hiện biểu thức đại số quan hệ, có những biểu thức được sử dụng được nhiều lần, thay vì định nghĩa lại ta nên đặt cho các biểu thức.

Gán (hoặc thay đổi) các biểu thức đại số ban đầu.

Định nghĩa

Phép đặt tên là phép toán cho phép đổi tên thuộc tính hoặc tên quan hệ.

Ký hiệu: ρ

Ý nghĩa: Đổi tên kết quả biểu thức đại số E thành x: $\rho_x(E)$

Ví dụ: Cho sơ đồ quan hệ SINHVIEN

SINHVIEN (MaSV, HoTen, NgaySinh, MaN, TenNganh)

SINHVIEN	MaSV	HoTen	NgaySinh	MaN	TenNganh
SV001	Nguyễn Văn An	10/6/1993	CNOT	Công Nghệ Ô Tô	
SV002	Dương Tùng Lê	12/4/1994	CKCT	Cơ Khí Chế Tạo	
SV005	Lê Dương	12/10/1992	CNTT	Công Nghệ Thông Tin	
SV006	Mai Thanh Vân	22/10/1995	CKCT	Cơ Khí Chế Tạo	
SV010	Lý Thanh Hải	10/8/1992	CNOT	Công Nghệ Ô Tô	

Yêu cầu: Chọn ra những sinh viên ngành ‘CNOT’ có tuổi trên 20. Giả sử năm hiện tại 2014 và đặt tên cho kết quả là SV20

Thực hiện:

$\rho_{SV20(\sigma((year(date()) - year(NgaySinh)) > 20) \wedge (MaNhanh = 'CNOT'))}(SINHVIEN)$

Kết quả:

SV20	MaSV	HoTen	NgaySinh	MaN	TenNganh
SV001	Nguyễn Văn An	10/6/1993	CNOT	Công Nghệ Ô Tô	
SV010	Lý Thanh Hải	10/8/1992	CNOT	Công Nghệ Ô Tô	

2.3.7. Các hàm kết tập

Hàm kết tập là các hàm thực hiện việc tính toán trên một tập giá trị để trả về một kết quả duy nhất.

a. Hàm SUM

Ý nghĩa: Tính tổng một tập giá trị số trên các bộ thỏa điều kiện cho trước. Đối số của hàm SUM có thể là một thuộc tính số hoặc biểu thức số có chứa các thuộc tính.

Cú pháp: $\text{Sum}_{<\text{Biểu thức thuộc tính số}>}(<\text{Quan hệ}>)$

Ví dụ: Cho quan hệ nhân viên NHANVIEN (MSNV, HT, HSL).

NHANVIEN	MSNV	HT	HSL
	NV01	Nguyễn Văn An	3.3
	NV02	Dương Tùng Lê	3.0
	NV03	Lê Dương	3.6
	NV04	Mai Thanh Vân	2.43
	NV05	Lý Thanh Hải	3.9

Yêu cầu: Tính tổng lương của các nhân viên trong cơ quan (giả sử lương cơ bản 1.500.000 đồng). Biết rằng lương của từng nhân viên được tính theo công thức sau: Luong = HSL * lương cơ bản.

Thực hiện:

$\rho_{\text{TongLuong}}(\text{Sum}_{HSL*1500000}(\text{NHANVIEN}))$

Kết quả:

TongLuong	
	20325000

b. Hàm AVG

Ý nghĩa: Tính giá trị trung bình của một tập giá trị số trên các bộ thỏa điều kiện cho trước. Đối số của hàm AVG có thể là một thuộc tính số hoặc biểu thức số có chứa các thuộc tính.

Cú pháp: $\text{Avg}_{<\text{Biểu thức thuộc tính số}>}(<\text{Quan hệ}>)$

Ví dụ: Cho bộ KETQUA(MSSV, MSM, Diem).

KETQUA	MSSV	MSM	Diem
	SV05	M01	7.5
	SV10	M01	9.0

	SV10	M02	8.5
	SV09	M01	5.0
	SV10	M01	7.5

Yêu cầu: Tính điểm trung bình của sinh viên có mã số ‘SV10’.

Thực hiện:

$\rho_{DTB} (\text{Avg Diem} (\sigma_{MSSV='SV10'}(\text{KETQUA})))$

Kết quả:

DTB	
	8.3

c. Hàm COUNT

Ý nghĩa: Đếm số bộ có trong một quan hệ thỏa điều kiện cho trước. Đôi số của hàm COUNT là một thuộc tính hoặc danh sách các thuộc tính. Chúng ta có thể dùng ký hiệu * để đại diện cho tất cả các thuộc tính. Biết rằng hàm COUNT chỉ đếm số mẫu tin nên khi ta chọn đôi số là một thuộc tính hoặc tất cả các thuộc tính, thì kết quả đếm cũng như nhau.

Cú pháp: Count * (<Quan hệ>)

Hoặc Count<thuộc tính> | <danh sách các thuộc tính> (<Quan hệ>)

Ví dụ: Cho bộ NHANVIEN(MSNV, HT, HSL) như sau.

NHANVIEN	MSNV	HT	HSL
	NV01	Nguyễn Văn An	3.3
	NV02	Dương Tùng Lê	3.0
	NV03	Lê Dương	3.6
	NV04	Mai Thanh Vân	2.43
	NV05	Lý Thanh Hải	3.9

Yêu cầu: Cho biết số lượng nhân viên có hệ số lương trên 3.0

Thực hiện:

Count *($\sigma_{HSL>3.0}(\text{NHANVIEN})$)

Kết quả:

	3

d. Hàm MAX

Ý nghĩa: Trả về phần tử có giá trị lớn nhất trong một tập giá trị của các bộ thỏa điều kiện cho trước. Đôi số của hàm MAX là một thuộc tính hoặc biểu thức thuộc tính.

Cú pháp: Max <Biểu thức thuộc tính> (<Quan hệ>)

Ví dụ: KETQUA (MSSV, MSM, Diem).

KETQUA	MSSV	MSM	Diem
KETQUA	SV05	M01	7.5
	SV10	M01	9.0
	SV10	M02	8.5
	SV09	M01	5.0
	SV10	M01	7.5
	SV10	M04	9.0

Yêu cầu: Cho biết sinh viên có mã số ‘SV10’ có điểm cao nhất bao nhiêu.

Thực hiện:

MAX Diem ($\sigma_{\text{MSSV}=\text{'SV40010'}}(\text{KETQUA})$)

Kết quả:

	9.0

e. Hàm MIN

Ý nghĩa: Trả về phần tử có giá trị nhỏ nhất trong một tập giá trị trên các bộ thỏa điều kiện cho trước. Đôi số của hàm MIN là một thuộc tính hoặc biểu thức thuộc tính.

Cú pháp: Min <Biểu thức thuộc tính> (<Quan hệ>)

Ví dụ: Cho quan hệ SINHVIEN(MSSV, HT, NgaySinh, MSL).

	MSSV	HoTen	NgaySinh	MaN
SINHVIEN	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Đương Tùng Lê	12/4/1996	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV004	Trần Thái Văn	02/10/1995	CNOT
	SV005	Lê Dương	12/10/1993	CNTT

Giáo trình Cơ sở dữ liệu

SV006	Mai Thanh Vân	22/10/1995	CKCT
SV010	Lý Thanh Hải	10/8/1993	CNTT

Yêu cầu: Cho biết tuổi nhỏ nhất của sinh viên là bao nhiêu. Giả sử năm hiện tại 2014.

Thực hiện:

Min $\text{year}(\text{date}()) - \text{year}(\text{NgaySinh})$ (SINHVIEN)

Kết quả:

18	

Ví dụ: Cho quan hệ kết quả (KQ) như sau:

KQ	MTS	Nganh	TongDiem	KetQua
TS001	CNTT		15.0	TT
TS002	CKCT		17.5	TT
TS003	CNTT		29.0	TT
TS004	CNOT		7.0	KT
TS005	CNTT		12.0	KT
TS006	CNTT		15.0	TT
TS007	CNOT		7.0	KT
TS008	CKCT		17.5	TT

Yêu cầu: Thực hiện quá trình thống kê sau:

Tổng số thí sinh

Số thí sinh trúng tuyển

Số thí sinh không trúng tuyển

Điểm thi cao nhất

Điểm thi thấp nhất

Điểm thi trung bình

Thực hiện

Tổng số thí sinh

Count*(KQ)

Số thí sinh trúng tuyển

$\text{Count}^*(\sigma_{\text{KetQua}=\text{'TT'}}(\text{KQ}))$

Số thí sinh không trúng tuyển

$\text{Count}^*(\sigma_{\text{KetQua}=\text{'KTT'}}(\text{KQ}))$

Điểm thi cao nhất

$\text{Max}_{\text{TongDiem}}(\text{KQ})$

Điểm thi thấp nhất

$\text{Min}_{\text{TongDiem}}(\text{KQ})$

Điểm thi trung bình

$\text{Avg}_{\text{TongDiem}}(\text{KQ})$

f. Phép gom nhóm

Là phép toán gom các giá trị giống nhau trong cùng một thuộc tính thành một nhóm.

Ký hiệu: G

Ví dụ: Cho quan hệ SINHVIEN(MaSV, HoTen, MaNganh) như sau:

SINHVIEN	MaSV	HoTen	Phai	MaNganh
	SV001	Nguyễn Văn An	1	CNTT
	SV002	Dương Tùng Lê	1	CKCT
	SV003	Phạm Thị Ái	0	CNTT
	SV004	Trần Thái Văn	1	CNOT
	SV005	Lê Dương	1	CKCT
	SV006	Nguyễn Thành An	1	CNOT
	SV007	Dương Hoàng Minh	1	CKCT
	SV008	Trần Hoài Bảo	1	CNTT
	SV009	Nguyễn Minh Nam	1	CNOT
	SV010	Dương Thị Mai	0	CNTT

Với quan hệ trên nếu gom nhóm theo thuộc tính MaNganh có thể gom thành ba nhóm như sau:

Nhóm CKCT

SV002	Dương Tùng Lê
SV005	Lê Dương

Giáo trình Cơ sở dữ liệu

SV007	Dương Hoàng Minh
-------	------------------

Nhóm CNTT

SV001	Nguyễn Văn An
SV003	Phạm Thị Ái
SV008	Trần Hoài Bảo
SV010	Dương Thị Mai

Nhóm CNOT

SV004	Trần Thái Văn
SV006	Nguyễn Thành An
SV009	Nguyễn Minh Nam

Việc gom nhóm là công việc thường được thực hiện cùng với các hàm kết tập (Sum, Max, Min, Avg,...).

Cú pháp: $<\text{nhóm}>G (<\text{Hàm kết tập}>)$

Ví dụ : Sử dụng quan hệ SINHVIEN(MaSV, HoTen, MaNganh).

Yêu cầu 1: Đếm số sinh viên theo từng ngành (MaNganh)

Thực hiện:

MaNganh G (Count * (SINHVIEN))

Kết quả:

MaNganh	SoLuong
CNTT	4
CKCT	3
CNOT	3

Yêu cầu 2: Đếm số lượng sinh viên nam và sinh viên nữ.

Thực hiện: MaNganh G (Count * (SINHVIEN))

Kết quả:

Phai	SoLuong
1	8
0	2

Giáo trình Cơ sở dữ liệu

Ví dụ: Cho quan hệ HOCVIEN như sau:

HOCVIEN	MaHV	DiemLT	DiemTH	Lop	KetQua
	HV001	9	8.5	CCA	Giỏi
	HV002	8	7	CCA	Khá
	HV003	9	9	CCA	Xuất sắc
	HV004	10	10	CCA	Xuất sắc
	HV005	9	7	CCA	Giỏi
	HV006	8	6	CCB	Khá
	HV007	5	5	CCB	TB
	HV008	8	9	CCB	Giỏi
	HV009	8	8	CAD	Giỏi
	HV010	7	6	CAD	Khá

Thực hiện các yêu cầu sau:

Đếm số lượng học viên

Đếm số lượng học viên của từng lớp (Lop)

Đếm số lượng học viên của từng loại (kết quả)

Tìm điểm cao nhất của từng lớp.

Gợi ý:

- Nếu gom nhóm theo lớp: 3 nhóm (CCA, CCB, CAD)
- Nếu gom nhóm theo kết quả: 4 nhóm (Xuất sắc, giỏi, khá, TB)

Thực hiện:

Đếm số lượng học viên

Thực hiện: $\rho_{SL}(\text{Count(HOCVIEN)})$

Kết quả:

SL	
	10

Đếm số lượng học viên của từng lớp

Thực hiện: $\rho_{SL}(\text{LopG}(\text{Count} * (\text{HOCVIEN})))$

Kết quả:

SL	Lop	
	CCA	5
	CCB	3
	CAD	2

Đếm số lượng học viên của từng loại

Thực hiện: $\rho_{SL} (\text{KetQuaG}(\text{Count} * (\text{HOCVIEN})))$

Kết quả:

SL	KetQua	
	Giỏi	4
	Khá	3
	Xuất sắc	2
	TB	1

Tìm điểm cao nhất của từng lớp

Thực hiện: $\rho_{SL} (\text{LopG} (\text{Max DiemLT + DiemTH} (\text{HOCVIEN})))$

Kết quả:

SL	Lop	
	CCA	20
	CCB	17
	CAD	16

Trong quá trình tính toán trên nhóm: Chỉ được chiếu các thuộc tính đã được nhóm, khóa của các quan hệ và các hàm kết tập.

2.4. Một số khái niệm bổ sung

2.4.1. Đại số hệ

Đại số hệ là một cặp $\langle M, F \rangle$ trong đó tập M được gọi là tập nền, tập F là tập phép toán trên M.

Mỗi phép toán $f \in F$ là một ánh xạ bộ phận:

$F: M^n \rightarrow M$

Và được gọi là phép toán n ngôi.

Ví dụ:

Với $F = \{ +, -, *, / \}$ và tập nền là tập số thực R

Ta sẽ có $\langle R, F \rangle$ là một đại số hệ.

2.4.2. Đại số quan hệ

Đại số quan hệ là tập hợp các phép toán đã xét cùng với cơ sở dữ liệu cho trước tạo thành một đại số hệ, gọi là đại số quan hệ.

2.4.3. Biểu thức quan hệ

Biểu thức quan hệ là biểu thức được xây dựng từ các quan hệ trong cơ sở dữ liệu cho trước và các phép toán quan hệ.

Thứ tự ưu tiên trong biểu thức quan hệ:

Thứ tự ưu tiên cao nhất thuộc phép toán một ngôi, sau đó đến phép toán hai ngôi. Thứ tự ưu tiên của phép toán là không tuyệt đối vì tùy thuộc vào những hệ quản trị khác nhau thì có những qui định cụ thể riêng. Ví dụ phép toán giao có thể được ưu tiên cao hơn phép toán hôi

Ví dụ: Tìm mã mặt hàng và tên mặt hàng được nhập và xuất trong cùng ngày ‘4/10/2000’.

Cho các quan hệ:

HANGHOA (MaH, TenHang)

NHAPHANG(MaH, NgayNhap, SL)

XUATHANG (MaH, NgayXuat, SL)

$$\begin{aligned} & \pi_{MaH, TenH}(HANGHOA) * (\pi_{MaH}(\sigma_{NgayNhap='4/10/2013'}(NHAPHANG)) \\ & \quad \cup \pi_{MaH}(\sigma_{NgayXuat='4/10/2013'}(XUATHANG))) \end{aligned}$$

BÀI TẬP CHƯƠNG 2

Bài 1:

Cho hai quan hệ R và S có các thuộc tính :A, B.

R		S	
A	B	A	B
a1	b1	a1	b2
a1	b2	a1	b1
a2	b1	a2	b1
a2	b2		

- Thực hiện phép giao R và S. sau đó cho biết kết quả thực hiện.
- Thực hiện phép hội R và S. sau đó cho biết kết quả thực hiện.
- Thực hiện phép trừ R và S, sau đó cho biết kết quả thực hiện.

Bài 2:

Cho quan hệ R(A, B) và S(C, D)

R

A	B
a1	b1
a1	b2
a2	b1
a2	b2

S

C	D
c1	d2
c1	d1
c2	d1

Thực hiện tích Decartes giữa R và S, hiển thị kết quả.

Bài 3:

Cho hai quan hệ R(A, B, C, D) và S(C, D)

R

A	B	C	D
a1	b1	c2	d1
a1	b1	c1	d1
a2	b1	c2	d1
a2	b2	c2	d1

S

C	D
c1	d1
c2	d1

Thực hiện phép chia R cho S.

Bài 4 : Cho hai quan hệ R(A, B, C) và S (C, D, E)

R

A	B	C
1	b1	c1
3	b2	c1
2	b1	c2
2	b2	c4

S

C	D	E
c1	d1	3
c2	d1	1
c3	d2	2

Cho biết kết quả thực hiện của các yêu cầu sau:

- Thực hiện kết nối với điều kiện Theta θ (θ : join) $R.A \geq S.E$
- Kết nối tự nhiên (natural join).
- Phép kết nối mở rộng trái.

- Phép kết nối mở rộng phải.
- Phép kết nối mở rộng đầy đủ.

Bài 5

Cho lược đồ cơ sở dữ liệu dùng để quản lý hồ sơ sinh viên bao gồm các quan hệ SV(sinh viên), LOP (Lớp), KHOA(khoa), MONHOC (môn học), KQ (kết quả) được mô tả bởi các lược đồ quan hệ như sau:

KHOA (MaKhoa, TenKhoa, SoNV)

Ý nghĩa: Mỗi khoa có mỗi MaKhoa duy nhất. Mỗi MaKhoa xác định tất cả các thuộc tính còn lại của khoa đó.

LOP (MaLop, TenLop, SiSo, MaKhoa)

Ý nghĩa: Mỗi lớp có một mã lớp duy nhất, mỗi lớp chỉ thuộc về một khoa nào đó.

SV (MASV, HoTen, Phai, NgaySinh,Tinh, HocBong, MaLop)

Ý nghĩa: Mỗi sinh viên có mỗi MASV duy nhất. Mỗi MASV xác định tất cả các thuộc tính còn lại của sinh viên đó.

MONHOC(MaMH, TenMH, SoTiet)

Ý nghĩa: Mỗi Môn học có một MAMH duy nhất. Mỗi MAMH xác định tất cả các thuộc tính còn lại của môn học đó.

KQ (MaSV, MAMH, DiemThi)

Ý nghĩa: Mỗi sinh viên cùng với một môn học xác định duy nhất một điểm thi.

YÊU CẦU:

1. Cho biết họ tên, tên lớp, tên khoa của các sinh viên.
2. Lập danh sách sinh viên gồm MASV, HoTen, TenKhoa, HocBong.
3. Lập danh sách sinh viên nữ khoa ‘CNTT’, danh sách cần MASV, HoTen, HocBong.
4. Lập bảng điểm cho tất cả sinh viên khoa ‘CNTT’, bảng điểm gồm các cột MASV, HoTen, TenMH, DienThi.
5. Lập phiếu điểm cho sinh viên có MASV=’1500123’.
6. Lập danh sách sinh viên gồm MASV, HoTen, TenLop, TenKhoa.
7. Lập bảng điểm môn học có mã môn học là CSDL cho tất cả sinh viên có mã lớp là ‘1CTT16A’.
8. Lập danh sách sinh viên của lớp có mã lớp là ‘2CTT15A’ và có điểm thi môn học từ 8 trở lên.

Bài 6:

Cho lược đồ cơ sở dữ liệu

Khoa(MaKhoa, TenKhoa)

Lop(MaLop, TenLop, MaKhoa)

Sinhvien(MaSV, HoTenSV, Phai, NgaySinh, NoiSinh, Tinh, MaLop)

Giáo trình Cơ sở dữ liệu

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:

1. 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.
2. Lập danh sách các sinh viên của lớp có MaLop là ‘CDTH2A’, danh sách cần các thông tin: MaSV, HoTenSV, NgaySinh, Tinh.
3. 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.
4. Lập bảng điểm thi lần 1 môn học ‘TH1208’ cho tất cả sinh viên thuộc hai lớp có MaLop là ‘1CTT14A’ và ‘2CTT14A’, danh sách cần: MaSV, HoTenSV, DiemThi.
5. Lập danh sách các giảng viên đã dạy lớp ‘2CTT5A’, danh sách cần các thông tin: MaGV, HoTenGV, TenKhoa, HocVi, TenMH.
6. Lập danh sách các môn mà lớp CDTH2A đã học, danh sách cần các thông tin: MAMH, TENMH, DONVIHT, HOTENGV.
7. Lập danh sách những giảng viên đã dạy sinh viên có MaSV là ‘17001256’, danh sách cần MaGV, HoTenGV, HocVi, ChuyenNganh, TenKhoa, TenMH.
8. Lập danh sách các sinh viên có mã khoa ‘CNTT’ có điểm thi lần 1 môn học ‘TH1213’ lớn hơn hoặc bằng 8, danh sách cần MaSV, HoTenSV, DiemThi, TenLop.

Chương 3: NGÔN NGỮ SQL

Trong chương trước, chúng ta đã tìm hiểu về đại số quan hệ, trong đại số quan hệ chúng ta chỉ thực hiện các thao tác tìm dữ liệu còn việc mô tả và cập nhật dữ liệu chưa được nhắc đến, bên cạnh đó đại số quan hệ vẫn còn mang tính hình thức vì thế trong chương này, chúng ta sẽ đề cập đến các cách tiếp cận với ngôn ngữ truy vấn của SQL Server để thực hiện các thao tác trên cơ sở dữ liệu, phân quyền sử dụng cơ sở dữ liệu. Trong chương này, chúng ta sẽ tìm hiểu ba thao tác đặc trưng trên cơ sở dữ liệu:

- Mô tả cấu trúc.
- Thao tác trên dữ liệu.
- Truy vấn dữ liệu.
- Phân quyền sử dụng cơ sở dữ liệu.

3.1. Giới thiệu ngôn ngữ SQL

3.1.1. Sự xuất hiện của SQL

Mỗi hệ quản trị CSDL đều phải có ngôn ngữ giao tiếp giữa người sử dụng với cơ sở dữ liệu. Ngôn ngữ giao tiếp CSDL gồm các loại sau:

Ngôn ngữ mô tả dữ liệu (Data Definition Language –DDL): Cho phép khai báo cấu trúc các bảng của CSDL, khai báo các mối liên hệ của dữ liệu (relationship) và các quy tắc áp đặt lên các dữ liệu đó.

Ngôn ngữ thao tác dữ liệu (Data Manipulation Language- DML) cho phép người sử dụng có thể thêm (insert), xoá (delete), sửa (update) dữ liệu trong CSDL. Ngôn ngữ truy vấn dữ liệu (hay ngôn ngữ hỏi đáp có cấu trúc (Structured Query Language-SQL)): Cho phép người sử dụng khai thác CSDL để truy vấn các thông tin cần thiết trong CSDL.

Ngôn ngữ quản lý dữ liệu (Data Control Language- DCL): 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 khai thác CSDL cho người sử dụng.

Những năm 1975-1976, IBM lần đầu tiên đưa ra hệ quản trị CSDL quan hệ mang tên SYSTEM-R với ngôn ngữ giao tiếp CSDL là SEQUEL (Structured English Query Language). Năm 1976 ngôn ngữ SEQUEL được cải tiến thành SEQUEL-2, khoảng năm 1978-1979 SEQUEL-2 được cải tiến và đổi tên thành ngôn ngữ truy vấn có cấu trúc (Structured Query Language - SQL). Cuối năm 1979 hệ quản trị CSDL được cải tiến thành SYSTEM-R*. Năm 1986 viện tiêu chuẩn quốc gia Mỹ (American National Standards Institute –ANSI) đã công nhận và chuẩn hóa ngôn ngữ SQL và sau đó tổ chức tiêu chuẩn thế giới (International Standards Organization -ISO) cũng đã công nhận ngôn ngữ này. Đó là chuẩn SQL-86. Tới nay SQL đã qua 3 lần chuẩn hóa (1989, 1992, 1996) để mở rộng các phép toán và tăng cường khả năng bảo mật và tính toàn vẹn dữ liệu.

Trong chương này chúng ta chỉ nghiên cứu về ngôn ngữ SQL. Ngôn ngữ truy vấn SQL có tập lệnh khá phong phú 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 chỉ trình bày các câu lệnh quan trọng nhất của SQL đó là câu lệnh mô tả cơ sở dữ liệu, thao tác trên cơ sở dữ liệu, thực hiện sắp xếp và tìm kiếm. Để đơn giản trong cách trình bày, ta xem các cách gọi sau là tương đương.

SQL sử dụng các thuật ngữ:

Bảng ~ quan hệ ~ table

Cột ~ thuộc tính ~ trường ~ field

Dòng ~ bộ ~ mẫu tin ~ bản ghi.

3.1.2. Các tiện ích SQL

SQL DDL: cung cấp các lệnh để thao tác trên cấu trúc quan hệ.

SQL DML: là phần quan trọng của SQL, dùng để thực hiện các thao tác trên dữ liệu theo hai cơ chế dùng lệnh hoặc có giao diện đồ họa hỗ trợ.

DML: hỗ trợ lập trình nhúng từ các ngôn ngữ lập trình cấp cao.

View: tạo khung nhìn tới cơ sở dữ liệu nhằm hạn chế các thao tác cập nhật dữ liệu của những người dùng không có quyền quản trị.

Quản lý tác quyền: các lệnh đặc tả các quyền hạn của người dùng trên cơ sở dữ liệu.

Ràng buộc toàn vẹn: cung cấp các lệnh thiết lập ràng buộc toàn vẹn.

Điều khiển giao tác: các lệnh đặc tả thời gian bắt đầu thực hiện giao tác (công việc) và thời gian kết thúc công việc.

3.2 Ngôn ngữ mô tả dữ liệu

Là ngôn ngữ mô tả:

- Sơ đồ cho mỗi quan hệ.
- Miền giá trị cho từng thuộc tính.
- Ràng buộc toàn vẹn.
- Chỉ mục trên các quan hệ.
- Liên kết giữa các quan hệ.

Các ràng buộc toàn vẹn:

- NOT NULL: không cho phép cột chứa giá trị rỗng.
- NULL: cho phép cột chứa giá trị rỗng.
- UNIQUE: giá trị trong cột phải duy nhất.
- DEFAULT: nhận giá trị <Giá trị> khi không đưa dữ liệu.
- PRIMARY KEY: khóa chính.
- FOREIGN KEY: khóa ngoại.
- CHECK: kiểm tra dữ liệu trước khi đưa vào (luận lý).

Các kiểu dữ liệu thường dùng trong SQL

Tên kiểu	Số byte	Diễn giải
Bit	1	Chỉ nhận một trong hai giá trị hoặc 0 hoặc 1
TinyInt	1	Nhận giá trị từ 0..255
SmallInt	2	-32168..32767
Int	4	-32768..32767
BigInt	8	-2 ⁶³ .. 2 ⁶³ -1
Numeric	Var	-10 ³⁸ .. 10 ³⁸ -1
Float	Var	-1.79+E308.. 1.79+E308
SmallDateTime	4	01/01/1900.. 31/12/2079
DateTime	8	01/01/1900.. 31/12/9999
Char	Var	Chiều dài cố định, tối đa 8000 ký tự
Varchar	Var	Chiều dài biến đổi tùy theo chiều dài sử dụng thực, tối đa 8000 ký tự
NChar	Var	Chiều dài cố định, tối đa 4000 ký tự thuộc bảng mã Unicode (cho phép ký tự tiếng Việt có dấu)
NVarChar	Var	Chiều dài biến đổi tùy theo chiều dài sử dụng thực, tối đa 4000 ký tự thuộc bảng mã Unicode
Text	Var	Cho phép chứa tối đa 2147483647 ký tự
NText	Var	Cho phép chứa tối đa 2147483647 ký tự thuộc bảng mã Unicode
Money	8	922337203685477.5808.. 22,337,203,685,477.5807
SmallMoney	4	- 214,748.3648.. 214,748.3647
....		

Cho CSDL gồm các bảng sau:

KHOA(MSK, TenKhoa)

Ý nghĩa

MSK: mã số khoa là khóa chính, chứa 4 ký tự.

TenKhoa: tên khoa chứa tối đa 50 ký tự.

MONHOC (MSMH, TenMH, SoTC)

Ý nghĩa

MSMH: mã số môn học là khóa chính, chứa 8 ký tự.

TenMH: tên môn học, chứa tối đa 50 ký tự

Giáo trình Cơ sở dữ liệu

SoTC: số tín chỉ, có miền giá trị từ 1.. 20

LOP(MSL, TenLop, SiSo, #MSK)

Ý nghĩa

MSL: mã số lớp là khóa chính, chứa 8 ký tự.

TenLop: tên lớp, chứa tối đa 50 ký tự.

SiSo: số, có miền giá trị từ 15 đến 500.

MSK: mã số khoa, khóa phụ (khóa ngoại) chứa 4 ký tự, tham chiếu tới MSK của bảng KHOA. (KHOA.MSK)

SINHVIEN(MSSV, Ho, Ten, Phai, NgaySinh, QueQuan, DiaChi, NgayNH,

NgayDK, NgayToiDa, #MSL)

Ý nghĩa

MSSV: mã số sinh viên là khóa chính, chứa 10 ký tự

Ho: họ và chữ lót của sinh viên, chứa 50 tối đa ký tự

Ten: tên của sinh viên, chứa tối đa 10 ký tự

Phai: giới tính của sinh viên, nhận giá trị 0 hoặc 1.

NgaySinh: ngày sinh, dữ liệu kiểu ngày tháng.

QueQuan: quê quán của sinh viên.

DiaChi: địa chỉ của sinh viên.

NgayNH: ngày nhập học của sinh viên, có dữ liệu kiểu ngày tháng

NgayDK: ngày ra trường theo dự kiến.

NgayToiDa: ngày ra trường tối đa.

MSL: mã số lớp, khóa ngoại, chứa 8 ký tự, tham chiếu tới LOP.MSL

KETQUA(MSSV, MSMH, Diem)

Ý nghĩa

MSSV: vừa là thành phần khóa chính vừa là khóa ngoại, chứa 10 ký tự, tham chiếu tới SINHVIEN.MSSV.

MSMH: vừa thành phần khóa chính vừa là khóa ngoại, chứa 8 ký tự, tham chiếu tới MONHOC.MSMH.

Diem: điểm của sinh viên đạt được cho từng môn học, là số thực lớn hơn hoặc bằng 0 và nhỏ hơn hoặc bằng 10, nhận tối đa 2 số lẻ.

3.2.1 Tạo cơ sở dữ liệu

Cú pháp

CREATE DATABASE <Tên _CSDL>

Ý nghĩa: Tạo thêm một cơ sở dữ liệu.

Lưu ý:

- <Tên_CSDL> phải là một tên mới chưa có trong hệ quản trị cơ sở dữ liệu.
- <Tên_CSDL> phải được đặt ngắn gọn, gợi nhớ và tuân thủ theo qui tắc đặt tên của hệ quản trị (Không trùng từ khóa, ký tự bắt đầu của tên phải là chữ cái, trong tên chỉ chứa ký tự chữ cái hoặc chữ số,...).
- Trong hệ quản trị có nhiều cơ sở dữ liệu vì thế trước khi thao tác trên cơ sở dữ liệu nào phải đảm bảo rằng cơ sở dữ liệu cần sử dụng là cơ sở dữ liệu hiện hành.

Dùng lệnh: USE <Tên_CSDL> để chọn cơ sở dữ liệu cần thao tác.

3.2.2 Tạo cấu trúc cho bảng

Cú pháp:

CREATE TABLE <tên_bảng>

(

```
<tên_cột_1> <Mô tả cột 1> [Ràng buộc cho cột 1]
[ [, <tên_cột_2> < Mô tả cột 2> [Ràng buộc cho cột 2] ]
...
, <tên_cột_n> <Mô tả cột n> [Ràng buộc toàn vẹn cho cột n]
]
[Ràng buộc toàn vẹn khóa chính]
[Ràng buộc toàn vẹn khóa ngoại]
[Ràng buộc toàn vẹn luận lý]
[Ràng buộc toàn vẹn duy nhất]
```

)

Ý nghĩa: Tạo một bảng (table) vào trong cơ sở dữ liệu đang sử dụng. Bảng được tạo không được trùng tên với các bảng đã tồn tại trong cơ sở dữ liệu.

Trong đó:

- <Tên_cột_1>: dùng để đặt tên cho cột 1. Đặt tên không nên gõ dấu tiếng việt, đặt tên ngắn gọn và gợi nhớ, không sử dụng các ký tự đặc biệt.
- <Mô tả cột 1>: qui định kiểu dữ liệu và miền giá trị hoặc độ rộng cho cột 1.
- [Ràng buộc cho cột 1]: thiết lập các ràng buộc toàn vẹn cho cột 1. Các ràng buộc trên cột thường được sử dụng là: Default, Null hoặc Not Null, Unique, Primary key, Check).
 - + *Default (<giá trị>)*: cột sẽ nhận giá trị mặc định là <giá trị> nếu người dùng không nhập giá trị.
 - + *Not Null / Null*: qui định cột không cho phép bỏ trống (bắt buộc đưa giá trị vào cho cột) hoặc được phép bỏ trống. Nhưng không được cài đặt null cho cột đóng vai trò là khóa chính.
 - + *Unique*: qui định các giá trị trong cột phải duy nhất.

+ *Primary key*: xác định cột là khóa chính. Chỉ áp dụng được cho trường hợp khóa chính chỉ có 1 thuộc tính. Trong trường hợp khóa chính có nhiều cột thì phải tạo bằng ràng buộc toàn vẹn khóa chính.

+ *Check(<biểu thức ràng buộc>)*: qui định dữ liệu muốn đưa vào trong cột phải thỏa điều kiện của <biểu thức ràng buộc>.

Trong bảng có nhiều cột, thì các định nghĩa cột cách nhau bởi dấu phẩy (,). Trên một cột có thể thực hiện nhiều ràng buộc. Các ràng buộc trên cột đặt sau phần mô tả cột.

Ví dụ: qui định cột ngày sinh (NgaySinh) để tuổi không nhỏ hơn 18 và không quá 35.

NgaySinh Date Check ((Year(GetDate())-Year(NgaySinh)>=18) and (Year(GetDate())-Year(NgaySinh)<=35))

- Ràng buộc khóa chính: qui định một cột hoặc một danh sách cột đóng vai trò là khóa chính.

Constraint <Tên ràng buộc> primary key (<danh sách cột làm khóa>)

Ràng buộc khóa chính: Các thành phần của khóa chính không được bỏ trống và tập giá trị của khóa chính không được trùng lặp.

- Ràng buộc khóa ngoại: thiết lập mối quan hệ giữa bảng con và bảng cha và qui định sự ảnh hưởng của bảng con lên bảng cha khi bảng cha có sự thay đổi dữ liệu.

Constraint <Tên ràng buộc> foreign key (<các thuộc tính khóa ngoại>)

references <tên bảng cha> (<khóa chính của bảng cha>)

[on delete <cascade | no action | set default | null >]

[on update < cascade | no action | set default | null >]

Ý nghĩa:

* **on delete**: thực hiện xóa dữ liệu trên bảng cha thì ảnh hưởng đến các bảng con như sau: nếu chọn:

- *cascade*: xóa dữ liệu bảng cha và tiến hành xóa luôn dữ liệu trên bảng con có khóa ngoại giống với dữ liệu khóa chính đang tiến hành xóa.

- *no action*: nếu dữ liệu khóa ngoại của bảng con còn trùng với khóa chính của bảng cha thì không thực hiện được việc xóa dữ liệu ở cha, ngược lại khóa ngoại của bảng con không trùng với khóa chính của bảng đang xóa thì tiến hành xóa dữ liệu bảng cha (dữ liệu bảng con không bị thay đổi).

- *set default*: xóa dữ liệu bảng cha, đồng thời cập nhật khóa ngoại của bảng con trùng với khóa chính bị xóa thành giá trị mặc định được cài đặt trong quá trình tạo bảng con (phải phù hợp với ràng buộc toàn vẹn về khóa ngoại)

- *set null*: xóa dữ liệu bảng cha, cập nhật khóa ngoại của bảng con trùng với khóa chính bị xóa thành null.

* **on update**: thực hiện cập nhật bảng cha thì ảnh hưởng đến các bảng con như các tùy chọn *on delete*

Lưu ý:

Giáo trình Cơ sở dữ liệu

Giữa khóa ngoại của bảng con và khóa chính của bảng cha không cần phải trùng tên nhưng phải cùng số lượng cột và mỗi cột tương ứng cũng phải cùng mô tả.

Tất cả các cột làm khóa chính của bảng cha sẽ được khóa ngoại của bảng con tham chiếu.

Bảng cha (bảng chứa khóa chính) phải được tạo trước bảng con.

Khi nhiều bảng có liên kết vòng tròn thì cần thiết kế ràng buộc khóa ngoại phù hợp để không bị xung đột khi tiến hành xóa hoặc cập nhật dữ liệu. Lúc đó có thể thay thế các ràng buộc khóa ngoại bằng các TRIGGER.

- Ràng buộc luận lý: thực hiện việc kiểm tra tính hợp lệ của dữ liệu trên nhiều cột.

Constraint <Tên ràng buộc> Check (<Biểu thức ràng buộc>)

- Ràng buộc duy nhất: qui định trên cùng một cột không cho phép chứa giá trị trùng nhau.

Constraint <Tên ràng buộc> Unique <các cột cần ràng buộc duy nhất>

Ví dụ: Tạo bảng KHOA, LOP, SINHVIEN

Bảng KHOA(MSK, TenKhoa)

Create table KHOA

```
(  
    MSK Char(4),  
    TenKhoa NvarChar (50) not null  
    Constraint PK_KHOA Primary key(MSK),  
    Constraint U_TenK Unique TenKhoa  
)
```

Bảng LOP(MSL, TenLop, SiSo, #MSK)

Create table LOP

```
(  
    MSL Char(8) primary key,  
    TenLop NvarChar(30) not null,  
    SiSo int,  
    MSK Char(4)  
    Constraint FK_KHOA_LOP
```

Foreign key (MSK) references KHOA(MSK)

On delete cascade

On update cascade,

Constraint C_SS check (SiSo >=15 and SiSo <=150)

)

Bảng SINHVIEN (MSSV, Ho, Ten, Phai, NgaySinh, QueQuan, DiaChi,

Create table SINHVIEN

```
(  
    MSSV Char(10),  
    Ho NvarChar(50) not null,  
    Ten NvarChar(10) not null,  
    Phai bit check (Phai IN (0,1)),  
    NgaySinh Date Check (NgaySinh < GetDate()),  
    QueQuan NvarChar(50) not null,  
    DiaChi NVarChar(50),  
    NgayNH Date,  
    NgayDK Date,  
    NgayToiDa Date,  
    Constraint PK_SV Primary key (MSSV),  
    Constraint FK_SV_L foreign key (MSL ) references LOP(MSL)  
        on delete cascade on update cascade,  
    Constraint C_NgayNH1 Check (NgayNH<NgayDK),  
    Constraint C_NgayDK Check (NgayDK<NgayToiDa)  
)
```

Minh họa các ràng buộc trên khóa ngoại:

Ví dụ: Cho hai bảng KHOA và NGANH

Bảng KHOA

MaK	TenK
CKCT	Cơ khí chế tạo
CNOT	Công nghệ ô tô
CNTT	Công nghệ thông tin
DTCN	Điện tử điện công nghiệp
CNH	Công nghệ hàn

Bảng NGANH

MaNganh	TenNganh	MaK
HTTT	Hệ thống thông tin	CNTT

QTMT	Mạng máy tính	CNTT
CNPM	Công nghệ phần mềm	CNTT
KHMT	Khoa học máy tính	CNTT
CNH1	Công nghệ hàn 1	CNH
CNH2	Công nghệ hàn 2	CNH

Xóa khoa ‘Công nghệ hàn’, nếu cài đặt tùy chọn:

- Cascade, thì xóa các ngành thuộc khoa ‘Công nghệ hàn’.

Kết quả của việc xóa

Bảng KHOA

MaK	TenK
CKCT	Cơ khí chế tạo
CNOT	Công nghệ ô tô
CNTT	Công nghệ thông tin
DTCN	Điện tử điện công nghiệp

Bảng NGANH

MaNganh	TenNganh	MaK
HTTT	Hệ thống thông tin	CNTT
QTMT	Mạng máy tính	CNTT
CNPM	Công nghệ phần mềm	CNTT
KHMT	Khoa học máy tính	CNTT

- No action, thì không xóa được khoa ‘Công nghệ hàn’ vì trong table NGANH đang sử dụng dữ liệu liên quan đến khoa ‘Công nghệ hàn’. Vì thế table NGANH cũng không bị thay đổi.

- Set default, thì xóa ngành ‘Công nghệ hàn’, cài đặt ngành thuộc khoa ‘Công nghệ hàn’ thành ‘CNCT’. (giả sử giá trị mặc định được cài đặt cho MaK là ‘CNCT’).

Kết quả của việc xóa

Bảng KHOA

MaK	TenK

Giáo trình Cơ sở dữ liệu

CKCT	Cơ khí chế tạo
CNOT	Công nghệ ô tô
CNTT	Công nghệ thông tin
DTCN	Điện tử điện công nghiệp

Bảng NGANH

MaNganh	TenNganh	MaK
HTTT	Hệ thống thông tin	CNTT
QTMT	Mạng máy tính	CNTT
CNPM	Công nghệ phần mềm	CNTT
KHMT	Khoa học máy tính	CNTT
CNH1	Công nghệ hàn 1	CNCT
CNH2	Công nghệ hàn 2	CNCT

- Set null → xóa khoa ‘Công nghệ hàn’, cài đặt khoa của những ngành thuộc ‘Công nghệ hàn’ thành rỗng.

Kết quả của việc xóa

Bảng KHOA

MaK	TenK
CKCT	Cơ khí chế tạo
CNOT	Công nghệ ô tô
CNTT	Công nghệ thông tin
DTCN	Điện tử điện công nghiệp

Bảng NGANH

MaNganh	TenNganh	MaK
HTTT	Hệ thống thông tin	CNTT
QTMT	Mạng máy tính	CNTT
CNPM	Công nghệ phần mềm	CNTT

KHMT	Khoa học máy tính	CNTT
CNH1	Công nghệ hàn 1	
CNH2	Công nghệ hàn 2	

3.2.3. Thay đổi cấu trúc bảng

a. Thêm cột

Cú pháp

ALTER TABLE <Tên_bảng>

ADD <Tên_cột> <Mô tả cột> [Ràng buộc toàn vẹn]

Ý nghĩa: Thêm một hoặc nhiều cột vào bảng (Bảng phải có trong cơ sở dữ liệu và cột được thêm phải chưa có trong bảng)

Ví dụ: thêm địa chỉ Email chứa khoảng 50 ký tự bảng sinh viên (SINHVIEN).

Alter table SINHVIEN Add Email VarChar(50)

b. Xóa cột

Cú pháp

ALTER TABLE <Tên_bảng> DROP COLUMN <Tên_cột>

Ý nghĩa: Xóa cột có trong bảng. Trước khi thực hiện xóa cột phải xóa các tham chiếu đến nó.

Ví dụ: xóa cột SiSo trong bảng LOP(MSL, TenLop, SiSo, MSK)

Alter table LOP Drop column SiSo

c. Thay đổi mô tả cột

Cú pháp

ALTER TABLE <Tên_bảng>

ALTER COLUMN <Tên_cột> <Mô tả mới cho cột>

Ý nghĩa: Thay đổi mô tả cho <Tên_cột>, trong trường hợp cột có tham gia vào các tham chiếu thì việc thay đổi phải phù hợp với các ràng buộc toàn vẹn đã được cài đặt trước đó.

Ví dụ: thay đổi độ rộng họ (Ho) trong bảng SINHVIEN còn 40 ký tự

Alter table SINHVIEN alter column Ho NvarChar(40)

d. Thêm ràng buộc toàn vẹn (RBTW)

ALTER TABLE <Tên_bảng> ADD

CONSTRAINT <Ten_RBTW1> <RBTW>

[, ...]

CONSTRAINT<Ten_RBTWn> <RBTW>

]

Ý nghĩa: Thêm các ràng buộc toàn vẹn cho bảng. Việc thêm các ràng buộc toàn vẹn phải phù hợp với các mô tả trước đó. Chẳng hạn, trước đó cột đã được phép chứa giá trị rỗng thì không thể là khóa chính.

Giả sử bảng KETQUA được tạo như sau

Create table KETQUA (MSSV Char(10), MSMH Char(10),

Diem Numeric (4,2))

Với lệnh trên thì bảng KETQUA chưa có khóa chính, khóa ngoại và chưa thực hiện ràng buộc luận lý trên điểm. Vậy ta sẽ thêm ràng buộc cho khóa chính, khóa ngoại và ràng buộc luận lý trên điểm như sau:

Alter table KETQUA Add

Constraint PK_ketqua Primary key(MSSV, MSMH),

Constraint FK_sv_kq Foreign key(MSSV)

 references SINHVIEN (MSSV) on delete cascade on update cascade,

Constraint FK_monhoc_kq Foreign key(MSMH)

 references MONHOC (MSMH) on delete cascade on update cascade,

Constraint C_Diem check(Diem>=0 and Diem<=10)

e. Xóa ràng buộc toàn vẹn (BTW)

Cú pháp:

ALTER TABLE <Tên_bảng> DROP CONSTRAINT <Ten_RBTV>

Thực hiện xóa ràng buộc toàn vẹn <Ten_RBTV> trong bảng.

Ví dụ: xóa ràng buộc khóa chính trên bảng KETQUA

Alter table KETQUA Drop Constraint PK_KETQUA

Lưu ý:

- Nếu bổ sung thêm một cột vào bảng mà trong bảng đã có ít nhất một bản ghi thì cột mới cần bổ sung phải cho phép chấp nhận giá trị NULL hoặc phải thiết lập giá trị mặc định (Default).

- Không thể xóa một cột khi đang được tham chiếu bởi các bảng con, muốn xóa cột đó thì phải xóa ràng buộc khóa ngoại từ các bảng con sao cho không còn bất kỳ khóa ngoại nào tham chiếu tới nó.

Cú pháp tổng quát thay đổi cấu trúc bảng

```
ALTER TABLE <Tên_bảng>
    ALTER COLUMN <Tên_cột> <Mô tả cột>
    ADD <Tên_cột> <Mô tả cột> [Ràng buộc toàn vẹn]
    DROP COLUMN <Tên_cột>
    ADD CONSTRAINT <Ten_RBTV> <RBTV>
    DROP CONSTRAINT <Tên_RBTV>
```

3.2.4 . Xóa các thành phần trong cơ sở dữ liệu

a. Xóa cơ sở dữ liệu

Cú pháp:

DROP DATABASE <Tên CSDL>

Ý nghĩa: xóa một cơ sở dữ liệu đã tồn tại trong hệ quản trị CSDL.

Ví dụ: Xóa cơ sở dữ liệu QLSV

DROP DATABASE QLSV

Lưu ý: Chỉ xóa được cơ sở dữ liệu khi không thành phần nào trong cơ sở dữ liệu đang được sử dụng.

b. Xóa bảng

Cú pháp:

DROP TABLE <tên_bảng>

Ý nghĩa: Xóa bảng <tên_bảng> vì thế xóa cấu trúc và toàn bộ dữ liệu trong bảng.

Ví dụ 1: xóa bảng KETQUA

Drop table KETQUA

Lưu ý: Không thể xóa bảng nếu như có một bảng nào đó đang tham chiếu tới. Vì thế, nếu thực sự muốn xóa bảng đang được tham chiếu thì phải xóa các tham chiếu trước.

Ví dụ 2: với ràng buộc giữa KHOA và LOP, giữa LOP và SINHVIEN. Khi tiến hành xóa KHOA hoặc LOP cần chú ý tới các ràng buộc toàn vẹn.

- Xóa bảng SINHVIEN thì không ảnh hưởng đến bảng LOP và KHOA. Nên có thể dùng lệnh Drop để xóa.

DROP table SINHVIEN

Xóa bảng KHOA, không được vì bảng LOP đang tham chiếu. Vì thế phải xóa ràng buộc giữa LOP với KHOA.

ALTER table LOP drop Constraint FK_KHOA_LOP

DROP table KHOA

3.3. Ngôn ngữ thao tác

Là ngôn ngữ có chức năng tác động đến dữ liệu trong cơ sở dữ liệu như: thêm mẫu tin, xóa mẫu tin, thay đổi giá trị của mẫu tin của các bảng.

Cho các bảng có các dữ liệu như sau:

KHOA	MSK	TenKhoa
	CNTT	Công nghệ thông tin
	CKCT	Cơ khí chế tạo
	CNOT	Công nghệ ô tô
	KTTC	Kế toán – Tài chính

SINHVIEN	MSSV	Ho	Ten	Phai	NgaySinh	MSL
	CNTT037001	Lê Hoài	Anh	1	10/12/1993	CNTT2011
	CNTT037005	Văn Kim	Tài	1	10/02/1993	CNTT2011
	CNTT038004	Nguyễn Dương	Thanh	1	12/12/1994	CNTT2012
	CNTT039005	Lê Thị	Lan	0	15/11/1995	CNTT2013
	CNTT039006	Trần Thị	Liên	0	10/05/1995	CNTT2013
	CKCT038007	Mai Quốc	Thanh	1	17/12/1993	CKCT2012
	CKCT038047	Lê Bảo	Toàn	1	09/12/1994	CKCT2012
	CKCT040030	Nguyễn Chí	Lâm	1	10/01/1994	CKCT2014

LOP	MSL	TenLop	MSK
	CNTT2011	Công nghệ thông tin khóa 37	CNTT
	CNTT2012	Công nghệ thông tin khóa 38	CNTT
	CNTT2013	Công nghệ thông tin khóa 39	CNTT
	CKCT2012	Cơ khí chế tạo khóa 38	CKCT
	CKCT2014	Cơ khí chế tạo khóa 40	CKCT

3.3.1. Thêm dữ liệu vào bảng

Thêm dữ liệu vào bảng có nhiều cách. Thêm trực tiếp một mẫu tin hoặc tổng hợp dữ liệu từ các bảng có sẵn.

a. Thêm một dòng vào bảng

Cú pháp

INSERT INTO <tên bảng> [(<danh sách cột>)]

VALUES (<danh sách giá trị>)

Thêm một mẫu tin giá trị <danh sách giá trị> vào bảng <tên bảng>.

Trong đó:

- <danh sách các cột>: Danh sách các cột sẽ được chèn thêm dữ liệu.

Nếu không có <danh sách cột> thì dữ liệu được thêm vào tất cả các cột.

Ngược lại dữ liệu chỉ được thêm vào các cột nằm trong <danh sách cột>, các cột không có trong <danh sách cột> nhận giá trị rỗng (nếu không được ràng buộc not null) hoặc giá trị mặc định của ràng buộc DEFAULT. Việc thêm dữ liệu sẽ thất bại nếu tồn tại cột nào đó không có trong <danh sách cột> không có ràng buộc DEFAULT mà có ràng buộc NOT NULL.

- <danh sách giá trị>: Là các giá trị sẽ được thêm vào tương ứng với <danh sách cột>. Số lượng giá trị phải bằng với số lượng cột được chỉ định và mỗi giá trị phải có kiểu tương ứng với kiểu của cột.

Ví dụ 1: thêm một dòng dữ liệu vào bảng KHOA có giá trị sau:

MSK là ‘DCTH’; TenK ‘Đại cương tổng hợp’

Thực hiện yêu cầu:

Insert into KHOA values(‘DCTH’, N’Đại cương tổng hợp’)

Kết quả:

KHOA	MSK	TenKhoa
	CNTT	Công nghệ thông tin
	CKCT	Cơ khí chế tạo
	CNOT	Công nghệ ô tô
	DCTH	Đại cương tổng hợp
	KTTC	Kế toán – Tài chính

Ví dụ 2: thêm một sinh viên có dữ liệu như sau vào bảng SINHVIEN

MSSV: ‘CKCT038050’;

họ và tên: ‘Phạm Thành Công’;

MSL: ‘CKCT2012’

Thực hiện yêu cầu:

Insert into table SINHVIEN (MSSV, Ho, Ten, MSL)

Values ('CKCT038050', N'Phạm Thành', N'Công', 'CKCT2012')

Kết quả:

SINHVIEN	MSSV	Ho	Ten		MSL
	CNTT037001	Lê Hoài	Anh	...	CNTT2011
	CNTT037005	Văn Kim	Tài		CNTT2011
	CNTT038004	Nguyễn Dương	Thanh		CNTT2012
	CNTT039005	Lê Thị	Lan		CNTT2013
	CNTT039006	Trần Thị	Liên		CNTT2013
	CKCT038007	Mai Quốc	Thanh		CKCT2012
	CKCT038047	Lê Bảo	Toàn		CKCT2012
	CKCT038050	Phạm Thành	Công		CKCT2012
	CKCT040030	Nguyễn Chí	Lâm		CKCT2014

Lưu ý: dữ liệu thêm vào phải thỏa các điều kiện của ràng buộc toàn vẹn.

Có thể thêm một sinh viên có dữ liệu như sau vào bảng SINHVIEN hay không? Tại sao?.

MSSV: 'CKCT038050'; họ và tên: 'Phạm Thành Công';

MSL: 'CKCT2013'

b. Thêm dữ liệu từ kết quả truy vấn

Cú pháp:

INSERT INTO <tên bảng> [<danh sách cột>] <Đâu lệnh truy vấn dữ liệu Select>

Ý nghĩa: Thêm vào trong bảng <tên bảng> các dòng dữ liệu từ kết quả của <Đâu lệnh truy vấn dữ liệu Select>. Vì vậy câu truy vấn phải có cấu trúc như một bảng và có các cột tương thích với <danh sách cột>. Câu lệnh truy vấn sẽ được tìm hiểu kỹ trong phần truy vấn dữ liệu.

Lưu ý: Khi thêm dữ liệu vào bảng, các dữ liệu của từng cột phải thỏa điều kiện của ràng buộc toàn vẹn.

3.3.2. Xóa dữ liệu ra khỏi bảng

Cú pháp:

DELETE <Tên Bảng> [**FROM** <danh sách bảng>]

[**WHERE** <điều kiện>]

Giáo trình Cơ sở dữ liệu

Ý nghĩa: Xóa các dòng dữ liệu trong bảng <Tên Bảng> thỏa điều kiện <điều kiện> từ các bảng có trong <danh sách bảng>.

Trong đó:

- <Tên bảng>: Bảng cần xóa các mẫu tin.
- FROM <danh sách bảng>: Nếu việc xóa dữ liệu cần tổng hợp từ các bảng trong <danh sách bảng>, ngược lại không cần [FROM <danh sách bảng>] khi việc xóa chỉ cần dữ liệu từ duy nhất bảng <Tên bảng>.
- WHERE <điều kiện>: điều kiện để xóa các mẫu tin. Nếu có [FROM <danh sách bảng>] thì phải thực hiện kết nối giữa các bảng.

Ví dụ: Xóa các sinh viên nam sinh năm 1993 trong bảng SINHVIEN có dữ liệu như sau:

SINHVIEN	MSSV	Ho	Ten	Phai	NgaySinh	MSL
	CNTT037001	Lê Hoài	Anh	1	10/12/1993	CNTT2011
	CNTT037005	Văn Kim	Tài	1	10/02/1993	CNTT2011
	CNTT038004	Nguyễn Dương	Thanh	1	12/12/1994	CNTT2012
	CNTT039005	Lê Thị	Lan	0	15/11/1995	CNTT2013
	CNTT039006	Trần Thị	Liên	0	10/05/1995	CNTT2013
	CKCT038007	Mai Quốc	Thanh	1	17/12/1993	CKCT2012
	CKCT038047	Lê Bảo	Toàn	1	09/12/1994	CKCT2012
	CKCT038050	Phạm Thành	Công	1	10/03/1988	CKCT2012
	CKCT040030	Nguyễn Chí	Lâm	1	10/01/1994	CKCT2014

Do cột phái (Phai) được qui định kiểu dữ liệu là bit nên người ta thường qui ước: phái là nam tương đương (Phai=1) ngược lại là phái nữ (Phai =0).

Ngày sinh (NgaySinh) là kiểu ngày tháng năm nên việc tính tuổi phải sử dụng hàm YEAR để tính năm sinh. Từ đó tính tuổi của sinh viên.

Thực hiện yêu cầu:

Delete SINHVIEN where phai=1 and year(NgaySinh)=1993

Kết quả

SINHVIEN	MSSV	Ho	Ten	Phai	NgaySinh	MSL
	CNTT038004	Nguyễn Dương	Thanh	1	12/12/1994	CNTT2012
	CNTT039005	Lê Thị	Lan	0	15/11/1995	CNTT2013
	CNTT039006	Trần Thị	Liên	0	10/05/1995	CNTT2013

	CKCT038047	Lê Bảo	Toàn	1	09/12/1994	CKCT2012
	CKCT038050	Phạm Thành	Công	1	10/03/1988	CKCT2012
	CKCT040030	Nguyễn Chí	Lâm	1	10/01/1994	CKCT2014

Lưu ý:

- Nếu câu lệnh DELETE mà không có WHERE thì sẽ xóa toàn bộ dữ liệu trong bảng <tên bảng>.
- Khi xóa dữ liệu cần chắc chắn biết rõ các ràng buộc toàn vẹn về khóa ngoại từ các bảng con đang tham chiếu đến bảng <tên bảng> để tránh xóa các dữ liệu không mong muốn hoặc thực hiện việc xóa không thành công.

Giả sử cần xóa dữ liệu trong bảng KHOA, cho biết với các loại ràng buộc khóa ngoại nào trên bảng LOP thì việc xóa dữ liệu của bảng KHOA không thể thực hiện được hoặc không ảnh hưởng đến LOP?

3.3.3. Cập nhật dữ liệu trong bảng

UPDATE <tên bảng>

SET <tên cột 1> = <bíểu thức 1>

[, ...

[, <tên cột n> = <bíểu thức n>]

]

[**FROM** <danh sách bảng>]

[**WHERE** <điều kiện>]

Ý nghĩa: Cập nhật lại các giá trị cho các cột <tên cột 1>, ..., <tên cột n> thành các giá trị tương ứng <bíểu thức 1>, ..., <bíểu thức n> của những mẩu tin thỏa điều kiện <điều kiện>. Trong đó <tên cột i> và <bíểu thức i> phải cùng mô tả. ($1 \leq i \leq n$)

Ví dụ 1: Cập nhật thông tin của KHOA để mã khoa (MSK) là ‘KTTC’ thành ‘TCKT’ và tên khoa (TenKhoa) thành ‘Tài chính – Kế toán’.

KHOA	MSK	TenKhoa
	CNTT	Công nghệ thông tin
	CKCT	Cơ khí chế tạo
	CNOT	Công nghệ ô tô
	DCTH	Đại cương tổng hợp
	KTTC	Kế toán – Tài chính

- Thực hiện yêu cầu:

Update KHOA

Set MSK ='TCKT',

TenKhoa=N'Tài chính – Kế toán'

Where MSK ='KTTC'

- Kết quả:

KHOA	MSK	TenKhoa
	CNTT	Công nghệ thông tin
	CKCT	Cơ khí chế tạo
	CNOT	Công nghệ ô tô
	DCTH	Đại cương tổng hợp
	KTTC	'Tài chính – Kế toán'

Lưu ý:

- Trước khi cập nhật dữ liệu trong bảng cần xác định ảnh hưởng của việc cập nhật đến các bảng tham chiếu.
- Các cột được cập nhật cần thỏa các điều kiện ràng buộc trước đó.

Nếu cập nhật khóa chính của một bảng thì có ảnh hưởng đến các bảng khác hay không?
Giải thích?

Việc cập nhật một dòng có tương đương với thao tác xóa dòng đó rồi thêm mới một có giá trị giống với yêu cầu cập nhật)? Giải thích?

3.4. Truy vấn dữ liệu

3.4.1. Tìm kiếm đơn giản

Cú pháp:

SELECT <danh sách cột>

FROM <danh sách bảng>

WHERE <điều kiện>

Ý nghĩa:

Trả về cấu trúc dạng bảng (quan hệ) có các cột là <danh sách cột> được tổng hợp từ các table trong <danh sách bảng>, các dòng là các mẫu tin thỏa <điều kiện>

Trong đó:

- <danh sách cột>: Là các cột sẽ được hiển thị trong kết quả; được chiết từ các cột thuộc <danh sách bảng>.
- <danh sách bảng>: Là các bảng cung cấp dữ liệu cho câu truy vấn.
- <điều kiện>: Là biểu thức luận lý để chọn lại những dòng thỏa điều kiện.

Giáo trình Cơ sở dữ liệu

- Có thể đổi tên cho biểu thức cột hoặc cột thông qua từ khóa AS <Tên mới>
- Trong From nếu có nhiều bảng thì các bảng cần phải thực hiện phép kết nối để lấy dữ liệu theo đúng yêu cầu, thường thì sử dụng phép kết nối ‘=’.
- Trước khi thực hiện lệnh truy vấn bao giờ cũng đòi hỏi trả lời ba câu hỏi sau:
 - i: Các cột nào cần hiển thị trong kết quả trả về?
 - ii: Điều kiện lọc dữ liệu là gì? Có cần kết nối hay không?
 - iii: Các cột trong kết quả trả về và các cột có tham gia trong điều kiện được lấy từ những bảng nào?

Trả lời câu hỏi

i là xác định được <danh sách cột>.
ii là xác định được <điều kiện>.
iii: là xác định được <danh sách bảng>.

Ví dụ 1: Tìm mã số sinh viên, họ và tên của những sinh viên nữ.

Thực hiện:

Trả lời câu hỏi:

- i:Các cột cần hiển thị? →MSSV, Ho, Ten.
- ii: Điều kiện lấy dữ liệu? sinh viên nữ → Phai =0.
- iii: MSSV, Ho, Ten, Phai: Đều thuộc bảng SINHVIEN.

Thực hiện yêu cầu:

Select MSSV, Ho, Ten

From SINHVIEN

Where phai =0

Cho bảng sinh viên có dữ liệu như sau:

SINHVIEN	MSSV	Ho	Ten	Phai	NgaySinh	MSL
CNTT038004	Nguyễn Dương	Thanh	1	12/12/1994	CNTT2012	
CNTT039005	Lê Thị	Lan	0	15/11/1995	CNTT2013	
CNTT039006	Trần Thị	Liên	0	10/05/1995	CNTT2013	
CKCT038047	Lê Bảo	Toàn	1	09/12/1994	CKCT2012	
CKCT038050	Phạm Thành	Công	1	10/03/1988	CKCT2012	
CKCT040030	Nguyễn Chí	Lâm	1	10/01/1994	CKCT2014	

Khi thực hiện câu truy vấn trên sẽ có kết quả như sau:

	MSSV	Ho	Ten

CNTT039005	Lê Thị	Lan
CNTT039006	Trần Thị	Liên

Ví dụ 2: Tìm sinh viên của lớp ‘Công nghệ thông tin 39’.

Thực hiện:

- Trả lời câu hỏi:

i: Các cột cần hiển thị? Tìm sinh viên thì ta sẽ hiển thị các thông tin cơ bản của sinh viên như: mã số sinh viên, họ tên, phái, ngày sinh → MSSV, Ho, Ten, Phai, NgaySinh.

ii: Điều kiện lấy dữ liệu? Lớp ‘Công nghệ thông tin K39’ → xác định ‘Công nghệ thông tin K39’ là tên lớp hay mã số lớp → tên lớp → TenLop = ‘Công nghệ thông tin K39’

iii: MSSV, Ho, Ten, Phai, NgaySinh: các cột của bảng SINHVIEN; TenLop: là cột thuộc bảng LOP → SINHVIEN, LOP.

Với yêu cầu này ta phải lấy dữ liệu từ hai bảng SINHVIEN, LOP nên cần phải thực hiện phép kết nối sao cho MSL trong bảng SINHVIEN = MSL trong bảng LOP.

- Thực hiện yêu cầu:

Select MSSV, Ho + ‘ ‘ + Ten as HoTen, Phai, NgaySinh

From SINHVIEN, LOP

Where TenLop = N’Công nghệ thông tin K39’ and
SINHVIEN.MSL = LOP.MSL

Mô phỏng cách thực hiện câu lệnh SQL trên:

Cho dữ liệu của bảng SINHVIEN, LOP như sau

LOP	MSL	TenLop
CNTT2013		Công nghệ thông tin K39
CKCT2012		Cơ khí chế tạo K38
CKCT2014		Cơ khí chế tạo K40

SINHVIEN	MSSV	Ho	Ten	Phai	NgaySinh	MSL
	CNTT039004	Nguyễn Dương	Thanh	1	12/12/1994	CNTT2013
	CNTT039005	Lê Thị	Lan	0	15/11/1995	CNTT2013
	CKCT038047	Lê Bảo	Toàn	1	09/12/1994	CKCT2012
	CKCT040030	Nguyễn Chí	Lâm	1	10/01/1994	CKCT2014

Mô phỏng từng bước thực hiện câu lệnh truy vấn trên:

Giáo trình Cơ sở dữ liệu

- Tích Decartes hai bảng ta được

MSSV	Ho	Ten	Phai	NgaySinh	MSL	TenLop
CNTT038004	Nguyễn Dương	Thanh	1	12/12/1994	CNTT2012	Công nghệ thông tin K39
CNTT039005	Lê Thị	Lan	0	15/11/1995	CNTT2012	Công nghệ thông tin K39

- Chiếu lấy các cột MSSV, Ho, Ten, Phai, NgaySinh

MSSV	HoTen	Phai	NgaySinh
CNTT038004	Nguyễn Dương Thanh	1	12/12/1994
CNTT039005	Lê Thị Lan	0	15/11/1995

Lưu ý:

- Nếu truy vấn được thực hiện trên nhiều bảng và trong các bảng có các trường trùng tên thì khi sử dụng các trường đó phải xác định tên bảng chứa trường thông qua cú pháp <tên bảng>.<tên trường>.

Ví dụ: cho biết mã lớp và tên lớp của sinh viên ‘Lê Thị Lan’.

Select LOP.MSL, TenLop From LOP, SINHVIEN

Where LOP.MSL = SINHVIEN.MSL and
HoTen=N’Lê Thị Lan’

- Trong khi thực hiện truy vấn nếu thấy tên các bảng quá dài hoặc không phù hợp với cách sử dụng của người lập trình thì có thể đặt thêm bí danh cho bảng theo cú pháp sau <Tên bảng> <Bi danh>.

Ví dụ: Cho biết mã ngành, tên ngành của sinh viên ‘Nguyễn Thành Tài’

Select SV.MaNganh, N.TenNganh

From SINHVIEN SV, NGANH N

Where SV.MaNganh and N.MaNganh and

Ho + ‘ ‘ + Ten =N’Nguyễn Thành Tài’

Trong ví dụ trên thì bảng SINHVIEN được đặt với bí danh SV, và bảng NGANH được đặt bí danh nói N.

- Tuy nhiên, đừng làm dụng việc đặt bí danh mà làm sai đi ý nghĩa của bảng ban đầu.

Select SV.MaNganh, SV.TenNganh

From SINHVIEN N, NGANH SV

Where SV.MaNganh and N.MaNganh and

Ho + ‘ ‘ + Ten =N’Nguyễn Thành Tài’

3.4.2. Tìm kiếm phần tử duy nhất (Loại bỏ các phần tử trùng nhau)

Cú pháp

SELECT DISTINCT <danh sách cột cần tổng hợp>

FROM <danh sách bảng>

WHERE <điều kiện>

Ý nghĩa:

Trả về một cấu trúc dạng bảng có các cột là <danh sách cột> được tổng hợp từ các table trong <danh sách bảng>, các dòng là các mẫu tin thỏa <điều kiện> nhưng sẽ loại bỏ những mẫu tin bị trùng.

Ví dụ: Cho bảng KETQUA như sau:

KETQUA	MSSV	MSMH	Diem
	CNTT039004	CNTT0034	7.0
	CNTT039004	CNTT0050	6.7
	CNTT039004	CNTT0100	7.9
	CKCT038047	CKCT0067	5.9
	CKCT038047	CKCT0070	6.9
	CKCT040030	CKCT0072	7.0
	CKCT040030	CKCT0055	8.0
	CKCT040030	CNTT0050	8.5

Yêu cầu: Tìm mã số sinh viên của các sinh viên có tham gia học.

Thực hiện: Sinh viên có tham gia học thì MSSV sẽ có xuất hiện trong bảng KETQUA, sinh viên có thể học nhiều môn vì thế một MSSV có thể xuất hiện nhiều lần. Nhưng ta chỉ lấy một MSSV là biết sinh viên có tham gia học.

Select distinct MSSV From KetQua.

Kết quả trả về

MSSV
CNTT039004
CKCT038047
CKCT040030

3.4.3. Tìm kiếm tất cả các phần tử

Cú pháp

SELECT *

FROM <danh sách bảng>

WHERE <điều kiện>

Ý nghĩa:

Giáo trình Cơ sở dữ liệu

Trả về một cấu trúc dạng bảng (quan hệ) có các cột là tất cả các cột có trong các bảng tham gia truy vấn, các dòng là các mẫu tin thỏa <điều kiện>.

Ví dụ 1: cho bảng sinh viên và lớp như sau, liệt kê tất cả thông tin của sinh viên (SINHVIEN) và lớp (LOP) của sinh viên.

LOP	MSL	TenLop
CNTT2013	Công nghệ thông tin K39	
CKCT2012	Cơ khí chế tạo K38	
CKCT2014	Cơ khí chế tạo K40	

SINHVIEN	MSSV	Ho	Ten	Phai	MSL
	CNTT039004	Nguyễn	Thanh	1	CNTT2013
	CNTT039005	Lê Thị	Lan	0	CNTT2013
	CKCT038047	Lê Bảo	Toàn	1	CKCT2012
	CKCT040030	Nguyễn Chí	Lâm	1	CKCT2014

Thực hiện:

Select * From SINHVIEN SV, LOP L

Where SV.MSL = L.MSL

Kết quả trả về

MSSV	Ho	Ten	Phai	SV.MSL	L.MSL	TenLop
CNTT039004	Nguyễn	Thanh	1	CNTT2013	CNTT2013	Công nghệ thông tin K39
CNTT039005	Lê Thị	Lan	0	CNTT2013	CNTT2013	Công nghệ thông tin K39
CKCT038047	Lê Bảo	Toàn	1	CKCT2012	CKCT2012	Cơ khí chế tạo K38
CKCT040030	Nguyễn Chí	Lâm	1	CKCT2014	CKCT2014	Cơ khí chế tạo K40

Ví dụ 2: sử dụng lại bảng SINHVIEN, tìm họ tên, phái của sinh viên của lớp ‘Công nghệ thông tin K39’.

Thực hiện:

Select Ho + ‘ ‘ + Ten as HoTen, Phai

From SINHVIEN SV, LOP L
 Where SV.MSL = L.MSL and
 TenLop =N ‘Công nghệ thông tin K39’

Kết quả:

HoTen	Phai
Nguyễn Thanh	1
Lê Thị Lan	0

3.4.4. Tìm kiếm có sắp xếp

Cú pháp:

SELECT <danh sách cột cần tổng hợp>

FROM <danh sách bảng>

WHERE <điều kiện>

ORDER BY <Cột 1 cần sắp xếp> [**Asc | Desc**]

[[, ...], <cột n cần sắp xếp> [**Asc | Desc**]]

Ý nghĩa: Sau khi thực hiện câu lệnh SQL thì phát biểu Order By dùng để sắp xếp kết quả theo thứ tự ưu tiên các cột từ trái qua phải. Asc: tăng dần; Desc: giảm dần. Nếu không chỉ định Asc hay Desc thì mặc nhiên hệ thống chọn Asc.

Ví dụ: Cho bảng SINHVIEN như sau, In danh sách sinh viên sắp xếp tăng dần theo tên

SINHVIEN	MSSV	Ho	Ten	Phai	MSL
	CNTT039004	Nguyễn	Thanh	1	CNTT2013
	CNTT039005	Lê Thị	Lan	0	CNTT2013
	CKCT038047	Lê Bảo	Toàn	1	CKCT2012
	CKCT040030	Nguyễn Chí	Lâm	1	CKCT2014

Thực hiện:

Select * From SINHVIEN Order By Ten

Kết quả:

SINHVIEN	MSSV	Ho	Ten	Phai	MSL
	CNTT039005	Lê Thị	Lan	0	CNTT2013
	CKCT040030	Nguyễn Chí	Lâm	1	CKCT2014
	CNTT039004	Nguyễn	Thanh	1	CNTT2013

	CKCT038047	Lê Bảo	Toàn	1	CKCT2012
--	------------	--------	------	---	----------

3.4.5. Hạn chế số lượng mẫu tin

Cú pháp

SELECT TOP <Số mẫu tin> <danh sách cột cần tổng hợp>

FROM <danh sách bảng>

WHERE <điều kiện>

ORDER BY <Cột 1 cần sắp xếp> [Ascending |Descending]

Ý nghĩa: trả về <Số mẫu tin> đầu tiên của phát biểu Select

Ví dụ 1: Cho bảng KETQUA có cấu trúc và dữ liệu như sau:

KETQUA	MSSV	MSMH	Diem
	CNTT039004	CNTT0034	7.0
	CNTT039004	CNTT0050	6.7
	CNTT039004	CNTT0100	7.9
	CKCT038047	CKCT0067	5.9
	CKCT038047	CKCT0070	6.9
	CKCT040030	CKCT0072	7.0
	CKCT040030	CKCT0055	8.0
	CKCT040030	CNTT0050	8.5

Yêu cầu: Hiển thị 3 mẫu tin đầu tiên của bảng KETQUA

Thực hiện:

Select top 3 * From KETQUA

Kết quả:

MSSV	MSMH	Diem
CNTT039004	CNTT0034	7.0
CNTT039004	CNTT0050	6.7
CNTT039004	CNTT0100	7.9

Ví dụ 2: sử dụng lại bảng KETQUA, cho biết MSSV, MSMH, Diem của sinh viên có Diem lớn nhất

Thực hiện:

Select Top 1 MSSV

From KETQUA

Order By Diem desc

Kết quả:

MSSV	MSMH	Diem
CKCT040030	CNTT0050	8.5

Lưu ý: Chúng ta có thể sử dụng kết hợp phát biểu TOP với phát biểu ORDER BY để lấy các giá trị cao nhất hoặc thấp nhất.

Ví dụ: cho bảng SINHVIEN như sau:

SINHVIEN	MSSV	Ho	Ten	Phai	NgaySinh
	CNTT039005	Lê Thị	Lan	0	10/10/1899
	CKCT040030	Nguyễn Chí	Lâm	1	12/5/1992
	CNTT039004	Nguyễn	Thanh	1	17/10/1990
	CKCT038047	Lê Bảo	Toàn	1	2/3/1899

Yêu cầu: cho biết tất cả thông tin của sinh viên có tuổi nhỏ nhất.

Thực hiện: để lấy tất cả thông tin thì ta dùng * thay cho danh sách các cột, chọn ra sinh viên nhỏ tuổi nhất thì ta sẽ sắp xếp NgaySinh giảm dần để sinh viên nhỏ tuổi nhất sẽ là mẫu tin đầu tiên.

Select top 1 * From SINHVIEN Order By NgaySinh Desc

Kết quả:

MSSV	Ho	Ten	Phai	NgaySinh
CKCT040030	Nguyễn Chí	Lâm	1	12/5/1992

3.5. Các hàm tổng hợp

3.5.1. Hàm SUM

Tính tổng một tập giá trị số của các mẫu tin thỏa điều kiện cho trước, đối số của hàm SUM có thể là một cột số hoặc biểu thức có chứa cột số.

Cú pháp: SUM(<biểu thức cột số>) [AS <Tên cột chứa giá trị tính tổng>]

Trong đó AS dùng để đặt tên lại cho <biểu thức cột số>. Việc đặt tên lại là không bắt buộc. Nhưng, ta nên đặt tên lại cho biểu thức để dễ nhận biết và thuận tiện cho việc sử dụng khi cần thiết.

Ví dụ: cho bảng NHANVIEN(MSNV, HT, HSL).

NHANVIEN	MSNV	HT	HSL
	NV01	Nguyễn Văn An	3.3
	NV02	Dương Tùng Lê	3.0
	NV03	Lê Dương	3.6
	NV04	Mai Thanh Vân	2.43
	NV05	Lý Thanh Hải	3.9

Yêu cầu: Tính tổng lương của các nhân viên trong cơ quan (giả sử lương cơ bản 1.500.000)

Thực hiện:

Select SUM (HSL * 1500000) as TongLuong From NHANVIEN

Kết quả:

TongLuong
20325000

3.5.2. Hàm AVG

Tính trung bình của một tập giá trị số của các mẫu tin thỏa điều kiện cho trước, đối số của hàm AVG có thể là một thuộc tính số hoặc biểu thức số có chứa các thuộc tính số.

Cú pháp: AVG (<biểu thức số>) [AS <tên cột chứa giá trị tính của Avg>]

Ví dụ: cho bảng KETQUA(MSSV, MSM, Diem).

KETQUA	MSSV	MSM	Diem
	SV05	M01	7.5
	SV10	M01	9.0
	SV10	M02	8.5
	SV09	M01	5.0
	SV10	M01	7.5

Yêu cầu: Tính điểm trung bình của sinh viên có mã số ‘SV10’

Thực hiện:

Select AVG(Diem) as DTB From KETQUA where MSSV = ‘SV10’

Kết quả:

DTB

8.3

3.5.3. Hàm COUNT

Đếm số mẫu tin có trong phát biểu SELECT.

Cú pháp: COUNT(*) AS <tên cột chứa giá trị đếm>

Ví dụ: cho bảng NHANVIEN(MSNV, HT, HSL)

NHANVIEN	MSNV	HT	HSL
	NV01	Nguyễn Văn An	3.30
	NV02	Dương Tùng Lê	3.00
	NV03	Lê Dương	3.60
	NV04	Mai Thanh Vân	2.43
	NV05	Lý Thanh Hải	3.90

Yêu cầu: cho biết số lượng nhân viên có hệ số lương từ 3.0 trở lên.

Thực hiện:

Select Count(*) as SL From NHANVIEN where HSL>=3.0

Kết quả là :

SL
4

3.5.4. Hàm MAX

Trả về phần tử có giá trị lớn nhất trong một tập giá trị.

Cú pháp: MAX (<biểu thức cần tìm giá trị lớn nhất>) AS <tên cột max>

Ví dụ: Cho bảng KETQUA(MSSV, MSM, Diem) có dữ liệu như sau.

KETQUA	MSSV	MSM	Diem
	SV05	M01	7.5
	SV10	M01	9.0
	SV10	M02	8.5
	SV09	M01	5.0
	SV10	M01	7.5

Yêu cầu: cho biết sinh viên có mã số ‘SV10’ có điểm cao nhất là bao nhiêu.

Thực hiện:

Select MAX (Diem) as MaxDiem From KETQUA

Where MSSV ='SV10'

Kết quả:

MaxDiem
9.0

3.5.5. Hàm MIN

Trả về phần tử có giá trị nhỏ nhất trong một tập giá trị.

Cú pháp: MIN (<biểu thức cần tìm giá trị nhỏ nhất> AS <tên cột min>)

Ví dụ 1: cho bảng SINHVIEN(MSSV, HoTen, NgaySinh, MaN) có dữ liệu như sau:

SINHVIEN	MSSV	HoTen	NgaySinh	MaN
	SV001	Nguyễn Văn An	10/6/1993	CNTT
	SV002	Đương Tùng Lê	12/4/1996	CKCT
	SV003	Phạm Thị Ái	02/10/1993	CNTT
	SV004	Trần Thái Văn	02/10/1995	CNOT
	SV005	Lê Dương	12/10/1993	CNTT
	SV006	Mai Thanh Vân	22/10/1995	CKCT
	SV010	Lý Thanh Hải	10/8/1993	CNTT

Yêu cầu: cho biết tuổi nhỏ nhất của sinh viên là bao nhiêu. Giả sử năm hiện tại 2014

Thực hiện:

Select Min (year(date()) – year(NgaySinh)) as TMin From SINHVIEN

Kết quả

TMin
18

Ví dụ 2: Cho bảng kết quả KQ (MSTS, Nganh, TongDiem, KetQua):

Yêu cầu: thực hiện việc tính toán và thống kê sau:

- Cho biết tổng số thí sinh.
- Cho biết số lượng thí sinh trúng tuyển.
- Cho biết số lượng thí sinh không trúng tuyển.
- Cho biết điểm thi (Diem) cao nhất là bao nhiêu?

Giáo trình Cơ sở dữ liệu

- Cho biết điểm thi thấp nhất là bao nhiêu?
- Tính điểm thi trung bình của các thí sinh.

KQ	MTS	Nganh	TongDiem	KetQua
TS001	CNTT		15.0	TT
TS002	CKCT		17.5	TT
TS003	CNTT		29.0	TT
TS004	CNOT		7.0	KTT
TS005	CNTT		12.0	KTT
TS006	CNTT		15.0	TT
TS007	CNOT		7.0	KTT
TS008	CKCT		17.5	TT

Thực hiện

- Tổng số thí sinh

Select Count (*) as SLTS From KQ

- Số thí sinh trúng tuyển

Select Count(*) as SLTT From KQ From KetQua='TT'

- Số thí sinh không trúng tuyển

Select Count(*) as SLKTT From KQ From KetQua='KTT'

- Điểm thi cao nhất

Select Max(TongDiem) as DiemMax From KQ

- Điểm thi thấp nhất

Select Min(TongDiem) as DiemMin From KQ

- Điểm thi trung bình

Select Avg (TongDiem) DTB From KQ

3.5.6. Tính toán theo nhóm

Thực hiện tính toán: Tổng, trung bình, điểm, giá trị lớn nhất, giá trị nhỏ nhất theo từng nhóm.

Cú pháp:

GROUP BY <danh sách cột cần nhóm>

HAVING <điều kiện>

Having <điều kiện>: thiết lập điều kiện lọc trên nhóm.

Ví dụ: cho bảng KQ

KQ	MTS	Nganh	TongDiem	KetQua
TS001	CNTT	15.0	TT	
TS002	CKCT	17.5	TT	
TS003	CNTT	29.0	TT	
TS004	CNOT	7.0	KT	
TS005	CNTT	12.0	KT	
TS006	CKCT	15.0	TT	
TS007	CNOT	17.0	TT	
TS008	CKCT	17.5	TT	
TS009	CKCT	15.5	TT	
TS010	CNTT	20.0	TT	

Yêu cầu 1: đếm số thí sinh theo từng ngành (Nganh)

Thực hiện:

Select Nganh, Count(*) as SL From KQ Group By Nganh

Kết quả:

Nganh	SL
CNTT	4
CKCT	4
CNOT	2

Yêu cầu 2: tính điểm trung bình từng ngành của các thí sinh trúng tuyển.

Thực hiện:

Cách 1:

Select Avg(Tongdiem) as DTB From KQ

where KetQua = 'TT' Group By Nganh

Với cách thực hiện trên thì việc lọc điều kiện trên mẫu tin chứ không phải lọc trên nhóm vì thế tốc độ thực hiện sẽ chậm hơn khi thực hiện lọc trên nhóm.

Giáo trình Cơ sở dữ liệu

Cách 2: thực hiện lọc điều kiện trên nhóm.

Select Avg(Tongdiem) as DTB From KQ

Group By Nganh Having KetQua = 'TT'

Kết quả:

Nganh	DTB
CNTT	21.3
CKCT	16.0
CNOT	17.0

Ví dụ 2: cho bảng HOCVIEN như sau:

HOCVIEN	MaHV	DiemLT	DiemTH	Lop	KetQua
	HV001	9	8.5	CCA	Giỏi
	HV002	8	7	CCA	Khá
	HV003	9	9	CCA	Xuất sắc
	HV004	10	10	CCA	Xuất sắc
	HV005	9	7	CCA	Giỏi
	HV006	8	6	CCB	Khá
	HV007	5	5	CCB	TB
	HV008	8	9	CCB	Giỏi
	HV009	8	8	CAD	Giỏi
	HV010	7	6	CAD	Khá

Thực hiện các yêu cầu sau:

- Đếm số lượng học viên
- Đếm số lượng học viên của từng lớp (Lop)
- Tìm điểm cao nhất của từng lớp.

Thực hiện:

- Đếm số lượng học viên

Select Count(*) as SL From HOCVIEN

Kết quả:

SL
10

b. Đếm số lượng học viên của từng lớp

Select Lop, Count(*) as LopSL

From HOCVIEN Group By Lop

Kết quả:

Lop	LopSL
CCA	5
CCB	3
CAD	2

c. Tìm điểm cao nhất của từng lớp

Select Lop, Max(DiemLT + DiemTH) as DMax

From HOCVIEN Group By Lop

Hoặc

Select Top 1 Lop, (DiemLT + DiemTH) as DMax

From HOCVIEN Group By Lop

Kết quả:

Lop	DMax
CCA	20
CCB	17
CAD	16

Lưu ý:

- Khi thực hiện việc tính toán trên nhóm thông qua phát biểu Group By thì danh sách các cột chỉ được phép chứa các khóa chính của các bảng, các hàm tính toán trên nhóm, danh sách các cột được nhóm.
- Cú pháp tổng quát của phát biểu SELECT

```
SELECT [TOP <Số mẫu tin> ] [Distinct]
      <danh sách biểu thức cột cần tổng hợp> | *
  FROM <danh sách bảng>
 WHERE <điều kiện>
  [
    GROUP BY <danh sách cột cần nhóm>
    HAVING <điều kiện trên nhóm>
  ]
  [ ORDER BY <Cột 1 cần sắp xếp> [Asc|Desc]
  [ [, ...], <cột n cần sắp xếp> [Asc |Desc] ]
```

3.5.7. Phát biểu SELECT lồng nhau

Trong nhiều trường hợp, tổng hợp dữ liệu một phát biểu SELECT không thể thực hiện được hiện được yêu cầu. Khi đó đòi hỏi việc tổng hợp dữ liệu từ những phát biểu SELECT khác. Trường hợp này ta gọi là SELECT lồng nhau.

Cú pháp:

```
SELECT [DISTINCT] [TOP <số mẫu tin>] <danh sách biểu thức cột cần tổng hợp>*
  FROM <danh sách bảng 1>
 WHERE [<điều kiện 1> AND | OR ] <biểu thức cột 1> <ĐK so sánh>
  (
    SELECT <biểu thức cột 1>
      FROM <danh sách bảng 2>
      WHERE <điều kiện 2>
      [GROUP BY <danh sách cột cần nhóm 2>
      [HAVING <điều kiện trên nhóm 2>] ]
      [ORDER BY <Cột 1 cần sắp xếp 2> [Asc |Desc]
      [ [, ...], <cột n cần sắp xếp 2> [Asc |Desc] ]
    )
    [ GROUP BY <danh sách cột cần nhóm 1>
    [HAVING <điều kiện trên nhóm 1>] ]
    [ORDER BY <Cột 1 cần sắp xếp 1> [Asc |Desc]
    [ [, ...], <cột n cần sắp xếp 1> [Asc |Desc] ]
```

Ý nghĩa:

- Tổng hợp dữ liệu từ <danh sách bảng 1> thỏa <điều kiện 1>.

Giáo trình Cơ sở dữ liệu

- <biểu thức cột 1> từ <danh sách bảng 1> thỏa <điều kiện chọn> với <biểu thức cột 2> được tổng hợp từ <danh sách bảng 2>. Vì thế <biểu thức cột 1> phải cùng kiểu với <biểu thức cột 2>.
- <ĐK so sánh> có thể chấp nhận các phép toán: <, <=, >, >=, =, in, not in

Ví dụ: cho các bảng như sau

Bảng MONHOC

MONHOC	MSMH	TenMon	SoTC
	CNTT0034	Cơ sở dữ liệu	3
	CNTT0050	Lập trình căn bản	4
	CNTT0100	Cấu trúc dữ liệu	3
	CNTT0025	Toán rời rạc	2
	CNTT0015	Đồ họa máy tính	2
	CNTT0020	Cấu trúc máy tính	3
	CNTT0024	Điện toán căn bản	3

Bảng SINHVIEN

SINHVIEN	MSSV	Ho	Ten	Phai	NgaySinh
	SV038004	Nguyễn Dương	Thanh	1	12/12/1994
	SV039005	Lê Thị	Lan	0	15/11/1995
	SV038047	Lê Bảo	Toàn	1	09/12/1994
	SV038050	Phạm Thành	Công	1	10/03/1988

Bảng KETQUA

KETQUA	MSSV	MSMH	Diem
	SV038004	CNTT0034	7.0
	SV038004	CNTT0050	6.7
	SV038004	CNTT0100	7.9
	SV039005	CNTT0015	5.9
	SV039005	CNTT0020	6.9

	SV039005	CNTT0024	7.0
	SV038050	CNTT0024	8.0

Yêu cầu 1: cho biết sinh viên có họ và tên là ‘Nguyễn Dương Thanh’ có điểm cao nhất là bao nhiêu.

Thực hiện:

Select Diem From KETQUA, SINHVIEN

Where KETQUA.MSSV = SINHVIEN.MSSV and

Ho + ‘ ‘ + Ten = N’Nguyễn Dương Thanh’ and

Diem = (Select Max(Diem) From KETQUA, SINHVIEN

Where KETQUA.MSSV = SINHVIEN.MSSV and

Ho + ‘ ‘ + Ten = N’Nguyễn Dương Thanh’

)

Kết quả: 7.9

Yêu cầu 2: cho biết môn học nào chưa có sinh viên đăng ký học.

Thực hiện:

Select * From MONHOC

where MSMH Not In (Select MSMH From KETQUA)

Kết quả:

MSMH	TenMon	SoTC
CNTT0025	Toán rời rạc	2

3.5.8. Truy vấn trên tập hợp

Mỗi phát biểu SELECT sẽ trả về cấu trúc dạng bảng. Ta xem, kết quả của mỗi truy vấn là một tập hợp. Vì thế có thể thực hiện các phép toán trên tập hợp như: hợp (Union), giao (Interselect), trừ (Minus).

Cú pháp:

(Phát biểu SELECT thứ 1)

<Phép toán tập hợp UNION | INTERSELECT | MINUS ... >

(Phát biểu SELECT thứ 2)

[....

<Phép toán tập hợp UNION | INTERSELECT | MINUS ... >

(Phát biểu SELECT thứ n)

]

Lưu ý:

Giáo trình Cơ sở dữ liệu

- Kết quả của các phát biểu SELECT phải cùng cấu trúc (tương thích).
- Chỉ có phát biểu SELECT sau cùng mới được sử dụng mệnh đề Order by.
- Hạn chế sử dụng phát biểu SELECT lồng.

Ví dụ: Cho hai bảng sau:

SINHVIEN	MSSV	Ho	Ten	Phai	NgaySinh
	SV038004	Nguyễn Dương	Thanh	1	12/12/1994
	SV039005	Lê Thị	Lan	0	15/11/1995
	SV038047	Lê Bảo	Toàn	1	09/12/1994
	SV038050	Phạm Thành	Công	1	10/03/1988
	SV040015	Nguyễn Thị	Ái	0	12/03/1999

KETQUA	MSSV	MSMH	Diem
	SV038004	CNTT0034	7.0
	SV038004	CNTT0050	6.7
	SV038004	CNTT0100	7.9
	SV039005	CNTT0015	5.9
	SV039005	CNTT0020	6.9
	SV039005	CNTT0024	7.0
	SV038050	CNTT0024	8.0

- Tìm những SINHVIEN chưa học môn nào

(Select MSSV From SINHVIEN)

Minus

(Select Distinct MSSV From KETQUA)

Kết quả:

MSSV
SV038047
SV040015

- Tìm những SINHVIEN có tham gia học

(Select MSSV From SINHVIEN)

Interselect

(Select Distinct MSSV From KETQUA)

Kết quả:

MSSV
SV038004
SV039005
SV038050
SV040015

3.6. Khung nhìn (View)

View có cấu trúc giống với bảng nhưng không chứa dữ liệu mà sẽ tổng hợp dữ liệu từ nơi khác. Dữ liệu mà ta nhìn thấy trong View được tổng hợp từ các bảng hoặc từ các view khác vì thế chức năng chính của View là tổng hợp dữ liệu, không cho phép thay đổi dữ liệu. Cú pháp của lệnh như sau:

3.6.1. Tạo khung nhìn

Cú pháp

CREATE VIEW <TenView> (<Danh sách cột>)

AS <Phát biểu SELECT>

Ý nghĩa: Tạo ra một view để lấy <Danh sách cột> có trong các cột của <Phát biểu SELECT>, tất cả các dòng thuộc kết quả của <Phát biểu SELECT>

Ví dụ 1: tạo khung nhìn gồm các mặt hàng và giá của nhà cung cấp ‘Thanh Hoàng’.

Cho các bảng sau

MATHANG (MaMH, TenMH)

NHASX (NSX, TenNSX)

KHACHHANG(MaKH, TenKH)

CUNGCAP(MaMH, NSX, Gia)

DATHANG(MaKH, MaMH, Ngay, SL)

Thực hiện:

CREATE VIEW VIEW_TH(TenMH, GIA) AS

SELECT TenMH, GIA

FROM MATHANG, NHASX, CUNGCAP

WHERE MATHANG.MaMH = CUNGCAP.MaMH and

NHASX.NSX=CUNGCAP.NSX and

TenNSX='Thanh Hoàng'

Ví dụ 2: Tạo khung nhìn khách hàng đặt các đơn hàng.

Thực hiện:

CREATE VIEW DH(TenKH, Ngay, TenMH, SL)

AS

```
SELECT TenKH, Ngay, TenMH, SL  
      FROM MATHANG, KHACHHANG, DATHANG and  
      WHERE MATHANG.MaMH = DATHANG.Mamh and  
            KHACHHANG.MaKH =DATHANG.MaKH
```

Lý do mà chúng ta cần phải tạo khung nhìn vì:

- An toàn dữ liệu.
- Tổng hợp dữ liệu từ các bảng (hoặc các view) khác.
- Che dấu tính phức tạp của tổ chức dữ liệu.
- Đơn giản hóa các câu truy vấn.

3.6.2. Xóa khung nhìn

Cú pháp

DROP VIEW <Tên View>

Ý nghĩa: Thực hiện việc xóa view <Tên view> khỏi cơ sở dữ liệu

Ví dụ: Xóa view DH

DROP View DH

3.7. Các lệnh về an toàn dữ liệu

3.7.1. Tạo quyền truy xuất

Cú pháp

GRANT <Thao tác truy xuất 1> (<danh sách cột>)

[<Thao tác truy xuất 2> (<danh sách cột>)]

...

ON <bảng> TO (<danh sách các giá trị khóa>)

WITH GRANT OPTION

Ý nghĩa: Cho phép thực hiện các thao tác truy xuất: READ, INSERT, DELETE, UPDATE lên các cột <danh sách cột> của <bảng> tại các dòng có khóa chính thuộc <danh sách các giá trị khóa>

Ví dụ: Cho phép đọc, thêm, sửa lên cột Ho, Ten, Phai, NgaySinh của sinh viên có mã số là 'SV038040'.

GRANT READ, INSERT, UPDATE (Ho, Ten, Phai, NgaySinh)

ON SINHVIEN TO ('SV038040')
WITH GRANT OPTION

3.7.2. Hủy bỏ quyền truy xuất

Cú pháp

REVOKE <Thao tác truy xuất 1> (<danh sách cột>)
[<Thao tác truy xuất 2> (<danh sách cột>)]

...

ON <bảng> **FROM** (<danh sách các giá trị khóa>)

Ý nghĩa: Hủy bỏ các quyền truy xuất (INSERT, UPDATE, DELETE, READ) trên cơ sở dữ liệu.

Ví dụ: hủy quyền thêm, cập nhật lên cột Ho, Ten, Phai, NgaySinh của sinh viên có mã số là 'SV038040'.

REVOKE INSERT, UPDATE (Ho, Ten, Phai, NgaySinh)

ON SINHVIEN FROM ('SV038040')

3.7.3. Các hàm thông dụng

- **ASCII(<ký tự>)** → Trả về vị trí của ký tự trong bảng mã ASCII I

Ví dụ: ASCI I('A') → 65

- **Char (<số>)** → Trả về ký tự ứng với số trong bảng mã ASCII

Ví dụ: Char(66) → 'B'

- **Upper (<chuỗi>)** → Trả về chuỗi chữ hoa

Ví dụ: Upper('lAn') → 'LAN'

- **Lower(<chuỗi>)** → Trả về chuỗi chữ thường

Ví dụ: Lower('LaN') → 'lan'

- **Len (<chuỗi>)** → Trả về số ký tự trong chuỗi

Ví dụ: Len('CNTT') → 4

- **Ltrim(<chuỗi>)** → Trả về chuỗi bị cắt tất cả khoảng trắng bên trái

Ví dụ: Ltrim(' Công Nghê ') → 'Công Nghê '

- **Rtrim(<chuỗi>)** → Trả về chuỗi bị cắt tất cả khoảng trắng bên phải

Ví dụ: Rtrim(' Công Nghê ') → ' Công Nghê '

- **Left(<chuỗi>, n)** → Trả về n ký tự bên trái

Ví dụ: Left('Phan Thanh', 4) → 'Phan'

- **Right(<chuỗi>, n)** → Trả về n ký tự bên phải

Ví dụ: Right('Phan Thanh', 5) → 'Thanh'

- **CharIndex(<ch1>, <ch2>)** → Trả về vị trí xuất hiện đầu tiên của chuỗi <ch1> trong <ch2>

Ví dụ: CharIndex('Lan', 'Ngoc Lan') → 6

- Replace (<ch>, <ch1>, <ch2>) → Trả về chuỗi sau khi thay thế <ch2> vào vị trí của <ch1> trong chuỗi <ch>

Ví dụ: Replace('Nguyen Van Thanh', 'Van', 'Thi') → 'Nguyen Thi Thanh'

- GetDate() → Trả về ngày tháng năm hiện tại của hệ thống

Ví dụ: GetDate() → 2014-10-17 14:25:36.234

- Datepart([d|m|yy],<ngày>) → Trả về một thành phần trong biểu thức ngày: ngày (d) hoặc tháng (m) hoặc năm (yy).

Ví dụ: DatePart(m,'20/12/2014') → 12

- Cast(<bíểu thức> as <kiểu>) Dùng để chuyển đổi kiểu giá trị của <bíểu thức>

Ví dụ: Cast(2+4 as Char) → '6'

Lưu ý:

- Với các hằng chuỗi thuộc mã Unicode thì phải thêm N phía trước.

- Khi thực hiện trên các lệnh trên hằng ngày tháng cần biết dạng ngày tháng của hệ thống, hoặc cài đặt định dạng ngày cho hệ thống.

BÀI TẬP CHƯƠNG 3

PHẦN 1: NGÔN NGỮ MÔ TẢ

DATA DEFINITION LANGUAGE - DDL

I. NỘI DUNG THỰC HIỆN

1. Sơ lược hệ quản trị Cơ sở dữ liệu SQL Server 2008 và các thao tác cơ bản trên CSDL.

2. Các kiểu dữ liệu (data type) thông dụng.
3. Thao tác trên bảng và ràng buộc toàn vẹn.
4. Thay đổi cấu trúc bảng.

II. LÝ THUYẾT

1. Giới thiệu sơ lược hệ quản trị Cơ sở dữ liệu SQL Server 2008

1.1. Kết nối với Server

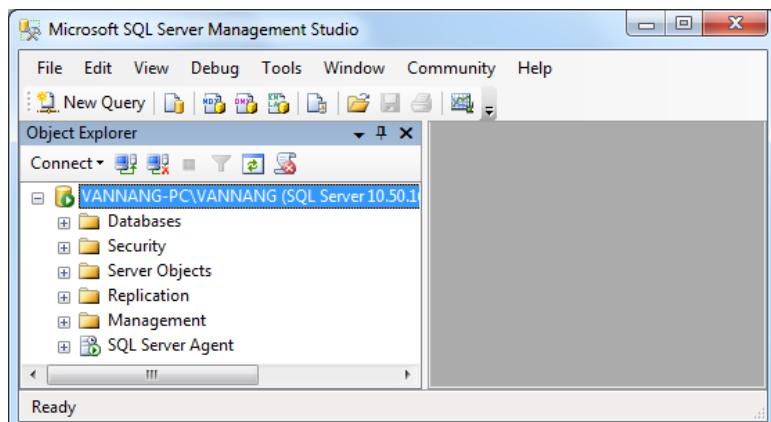
Chọn **Start/All Programs/Microsoft SQL Server 2008/ SQL Server Management Studio**. Cửa sổ đăng nhập sẽ xuất hiện như sau:



- + **Windows Authentication:** Tài khoản mặc định của SQL Server.
- + **SQL Server Authentication:** Tài khoản người dùng tự tạo khi cài đặt. Người sử dụng cần nhập User name và Password khi đăng nhập.

Người dùng chọn **Windows Authentication**, **SQL Server Authentication** để đăng nhập. Sau đó chọn **Connect** để kết nối Server.

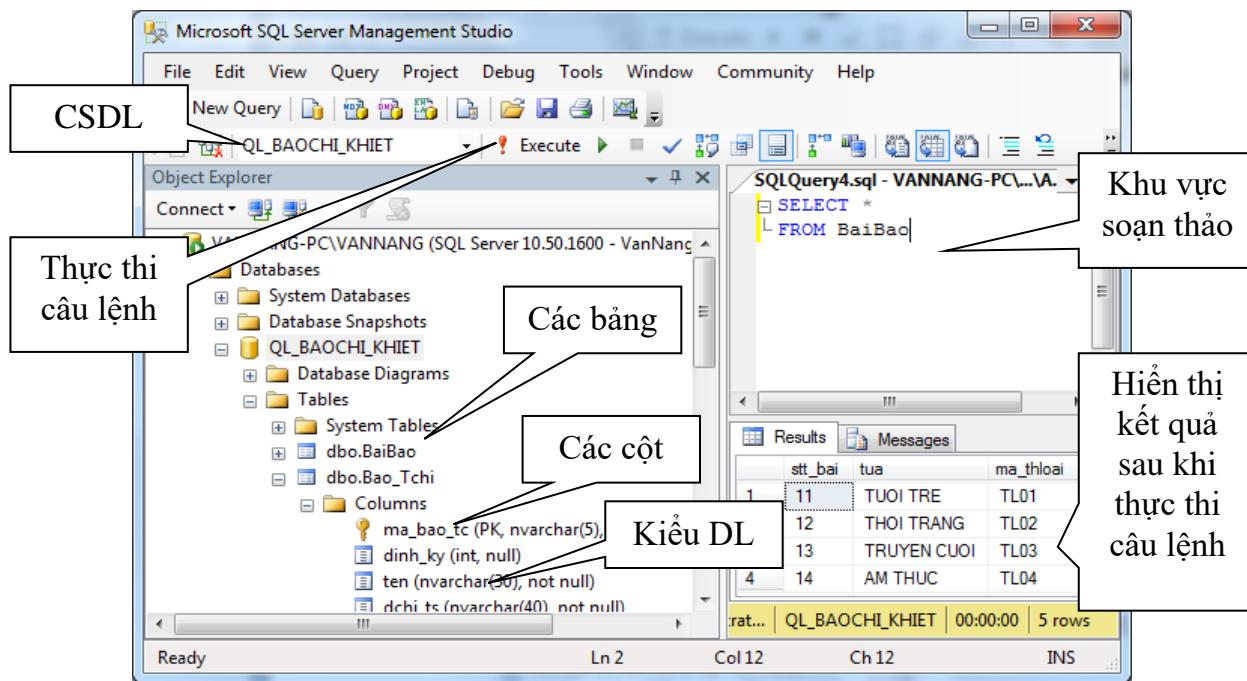
Sau khi chọn **Connect**, cửa sổ **SQL Server Management Studio** sẽ xuất hiện như sau:



1.2. Quan sát CSDL bằng Object Browser

- Ta có thể quan sát lược đồ của các CSDL mà SQL Server đang quản lý bằng cách nhấp vào **View → Object Explorer**. Cửa sổ dọc **Object Explorer** xuất hiện như sau:

Giáo trình Cơ sở dữ liệu



- Thành phần quan trọng trong Object Explore: Databases

Chứa CSDL của hệ thống và CSDL của người dùng. Mỗi cơ sở dữ liệu chứa các thành phần:

- a. Database Diagrams: Mô hình quan hệ giữa các table
- b. Tables: Danh sách các table trong CSDL.
- c. View: Danh sách các khung nhìn

...

- Các bước thực hiện lệnh trên SQL Server

1. Mở cửa sổ truy vấn mới: Click vào chức năng New Query trên thanh công cụ chuẩn (Standard)
2. Soạn thảo lệnh
3. Chọn lệnh muốn thực hiện: Quét khói các lệnh
4. Thực hiện lệnh: F5 hoặc click vào chức năng Execute
5. Quan sát thông báo và kết quả thực hiện ở Messages và Results

2. Các kiểu dữ liệu (data type) thông dụng

Tên kiểu	Giải thích	Biểu diễn hằng
Bit	Số nguyên 0 hoặc 1	0, 1
Int	Số nguyên từ -2^{31} (-2.147.483.648) đến $2^{31} - 1$ (2.147.483.647)	-1000, 23455

Tên kiểu	Giải thích	Biểu diễn hằng
SmallInt	Số nguyên từ -2^{15} (- 32.768) đến $2^{15} - 1$ (32.767)	32124, -764
TinyInt	Số nguyên từ 0 đến 255	31, 45
Decimal/ Numeric	Các số thập phân từ $-10^{38} - 1$ đến $10^{38} - 1$	1894.1204
Money	Tiền tệ từ -2^{63} đến $2^{63} - 1$	234, -87.65 \$12, \$542023.14
SmallMoney	Tiền tệ từ -214.748,3648 đến 214.748,3647	-\$45.56 -34.54
Float	Số thực từ $-1,79 \times 10^{-308}$ đến $1,79 \times 10^{308}$	101.5E5 0.5E-2
Real	Số thực từ -3.40×10^{-38} đến 3.40×10^{38}	245.21E-10
DateTime	Ngày giờ từ 1/1/1753 đến 31/10/9999	'April 15, 1998' , '15 April, 1998' , '980415' , '04/15/98'
SmallDateTime	Ngày giờ từ 1/1/1900 đến 06/6/2079	như DateTime
Char	Ký tự có độ dài ô nhớ cố định và tối đa là 8000 kí tự (không hỗ trợ Unicode)	'50% complete.' 'O'Brien' 'O'Brien'
Nchar	Ký tự có độ dài ô nhớ cố định và tối đa là 4000 kí tự (hỗ trợ Unicode)	'Nguyễn Văn Minh' 'Lương Tâm'
Varchar	Ký tự có độ dài ô nhớ không cố định và tối đa là 8000 kí tự (không hỗ trợ Unicode)	'50% complete.' 'O'Brien' 'O'Brien'
Nvarchar	Ký tự có độ dài ô nhớ không cố định và tối đa là 4000 kí tự (hỗ trợ Unicode)	'Nguyễn Văn Minh' 'Lương Tâm'

3. Tạo và sử dụng CSDL

3.1. Tạo CSDL

Create Database <Tên_CSDL>

Ví dụ: tạo CSDL Quản lý sinh viên thì soạn thảo lệnh

Create Database QL_SINHVIEN

Lưu ý: <Tên_CSDL> phải là một tên chưa có trong Databases. Được viết ngắn gọn gọi nhó, chỉ nên sử dụng tiếng Việt không dấu, không dùng ký tự đặc biệt.

3.2. Sử dụng CSDL

CSDL mặc định thường là master (CSDL của hệ thống), để chuyển sang sử dụng CSDL tự tạo, ta dùng lệnh:

Use <Tên_CSDL>

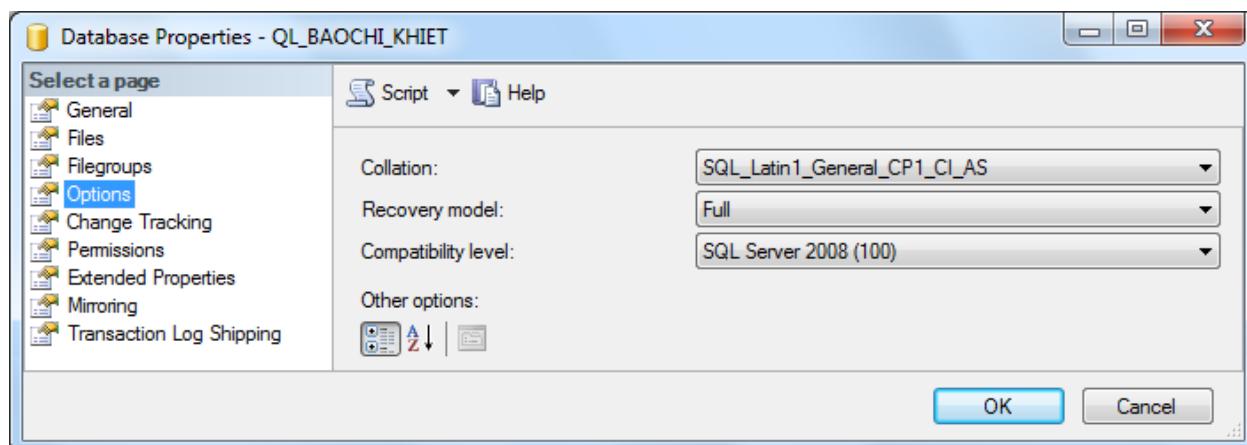
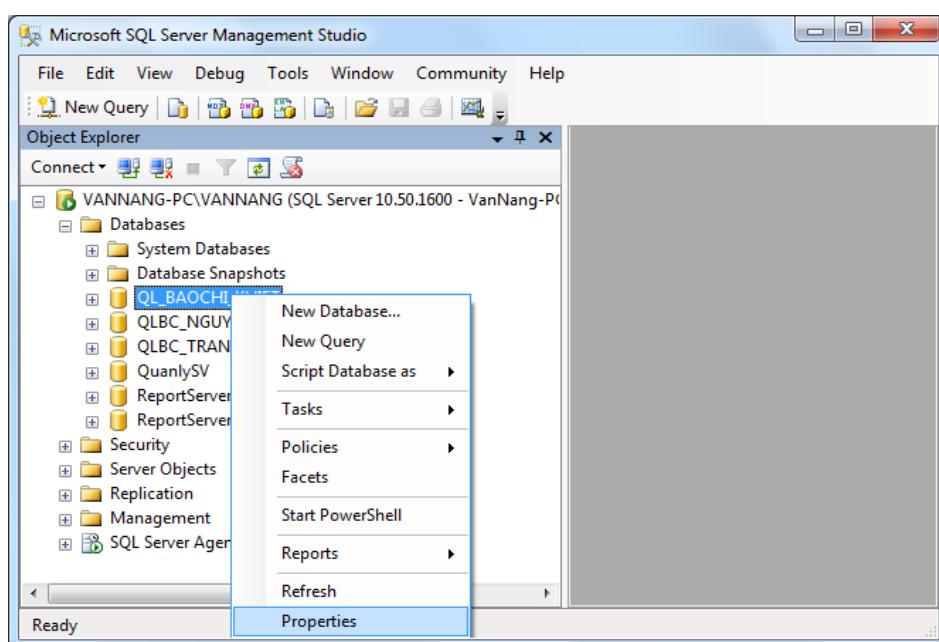
Ví dụ: CSDL mặc định thường là master (CSDL của hệ thống), để chuyển sang sử dụng CSDL tự tạo, ta dùng lệnh:

Use QL_SINHVIEN

Lưu ý: <Tên_CSDL> là cơ sở dữ liệu đã có trong Databases.

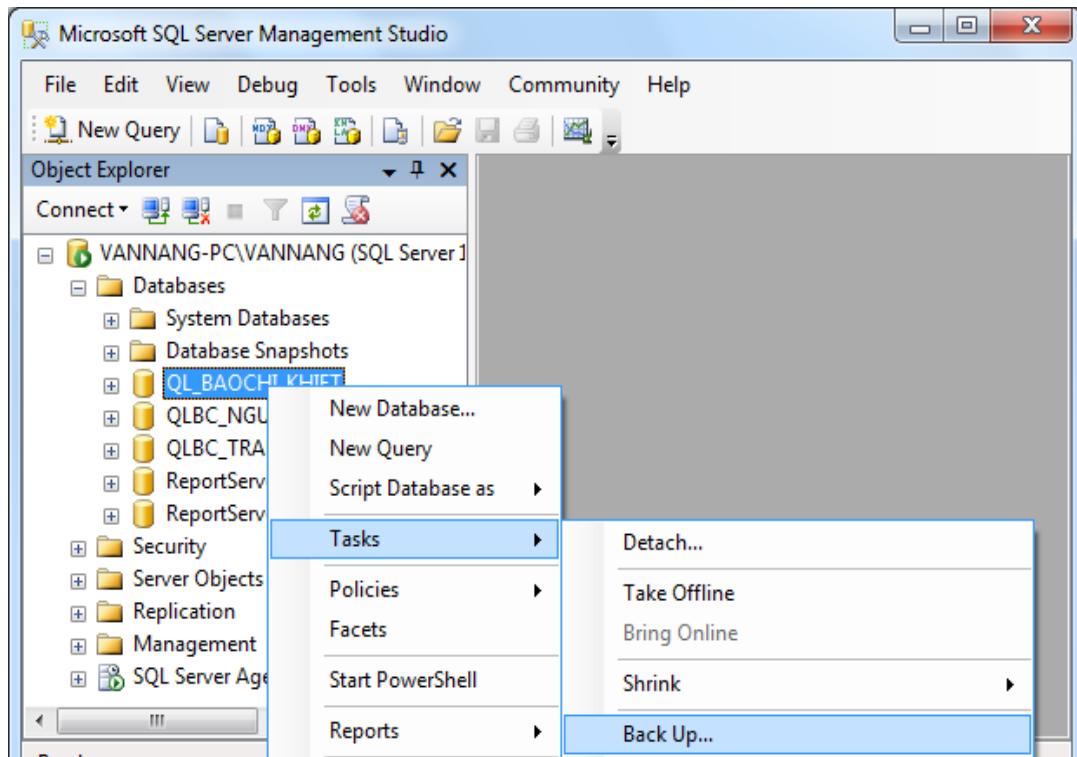
3.3. Sao lưu CSDL

Bước 1: Từ nút ‘Databases’, click phải lên CSDL muốn sao lưu chọn ‘Properties’

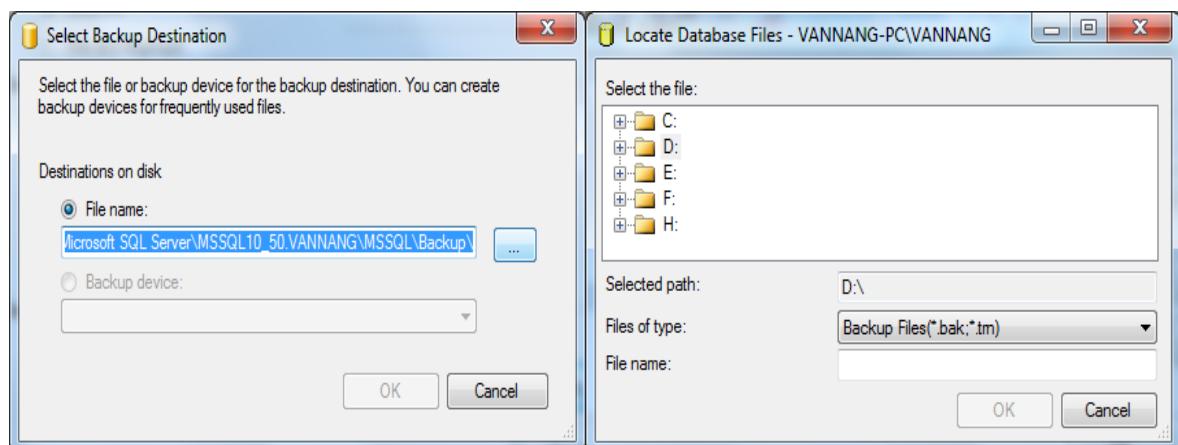


Giáo trình Cơ sở dữ liệu

Bước 2: Từ nút ‘Databases’, click phải lên CSDL muôn sao lưu sau đó chọn ‘Tasks’ và chọn ‘Backup’.



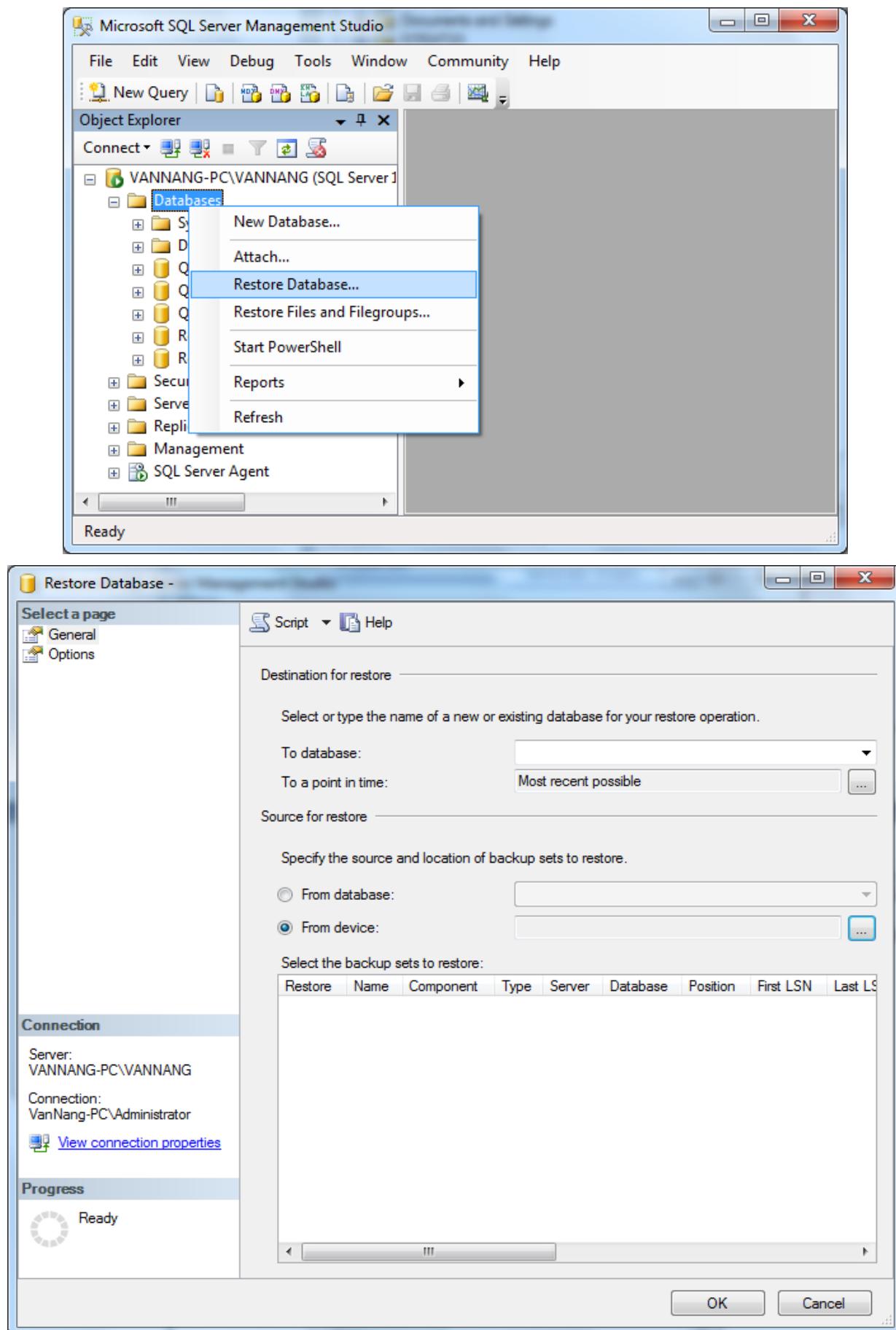
Bước 3: Nhấn chọn ‘Add...’ và nhập vào đường dẫn và tên file sao lưu, nhấn OK để quá trình sao lưu hoàn tất.

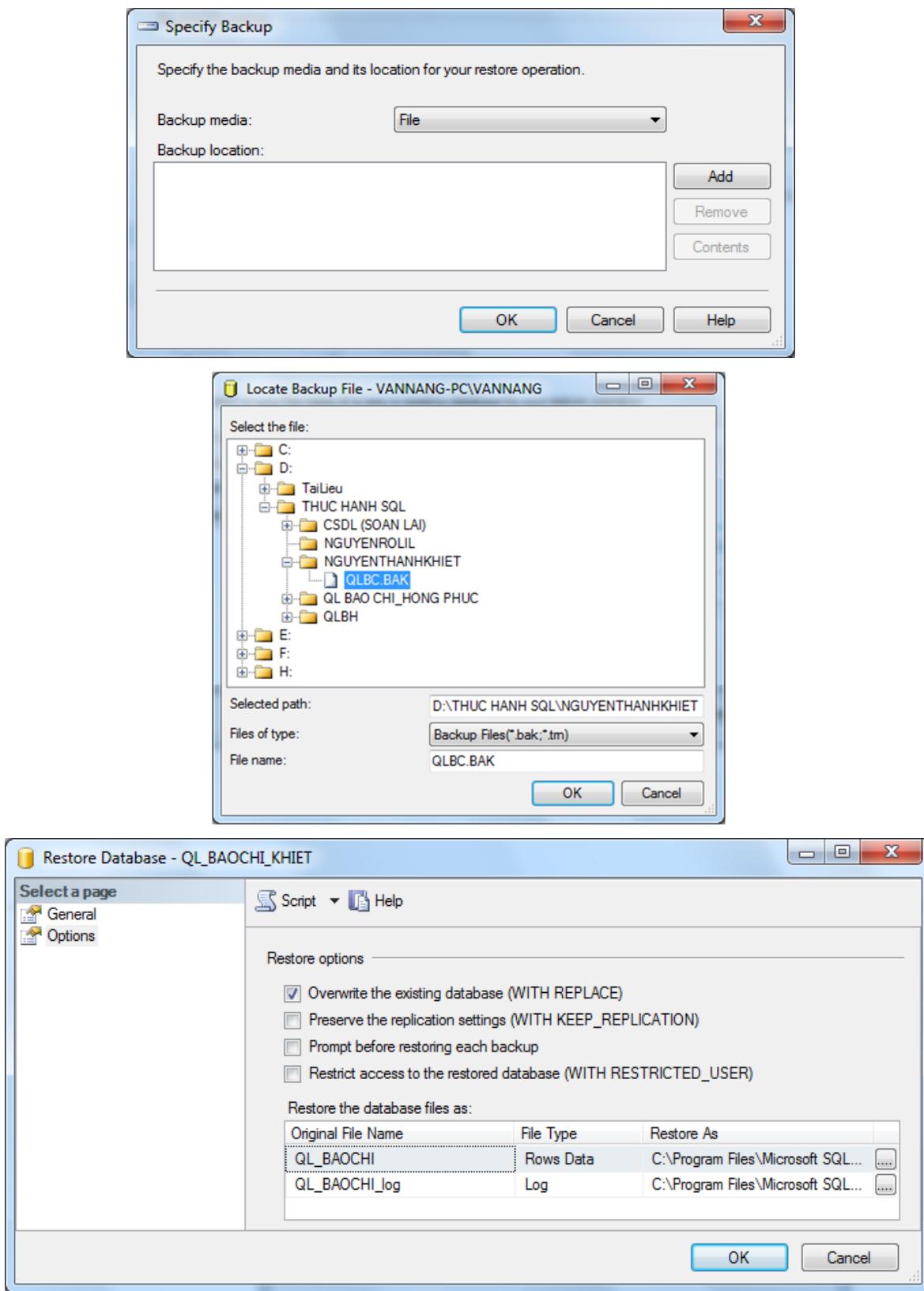


3.4. Phục hồi CSDL

Bước 1: Click chuột phải lên ‘Databases’ sau đó ‘Restore Database’

Giáo trình Cơ sở dữ liệu





4. Thao tác trên bảng

4.1 Tạo bảng và các ràng buộc toàn vẹn

```
CREATE TABLE <tên_bảng>
(
    <tên_cột_1> <Mô tả cột 1> [Ràng buộc cho cột 1]
    [, <tên_cột_2> <Mô tả cột 2> [Ràng buộc cho cột 2] ]
    ...
    [Các ràng buộc toàn vẹn trên bảng]
)
```

4.2 Thay đổi cấu trúc bảng

a. Thêm cột

```
ALTER TABLE <Tên_bảng>
    ADD <Tên_cột> <Mô tả cột> [Ràng buộc toàn vẹn]
```

b. Xóa cột

```
ALTER TABLE <Tên_bảng>
    DROP COLUMN <Tên_cột>
```

c. Thay đổi mô tả cột

```
ALTER TABLE <Tên_bảng>
    ALTER COLUMN <Tên_cột> <Mô tả mới cho cột>
```

c. Thêm ràng buộc toàn vẹn (RBTV)

```
ALTER TABLE <Tên_bảng> ADD
    CONSTRAINT <Ten_RBTV1> <RBTV>
    [, ...]
    CONSTRAINT <Ten_RBTVn> <RBTV>
]
```

d. Xóa ràng buộc toàn vẹn (BTV)

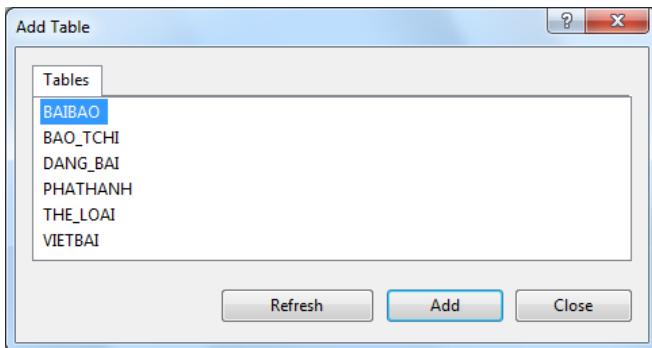
```
ALTER TABLE <Tên_bảng>
    DROP CONSTRAINT <Ten_RBTV>
```

4.3 Xóa bảng

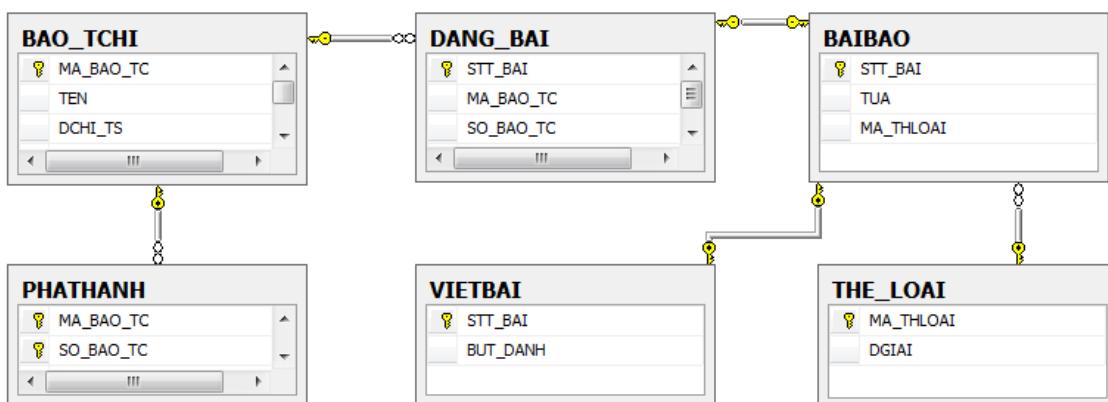
```
DROP TABLE <Tên_table>
```

4.3. Xem mô hình ràng buộc khóa ngoại

- Chọn CSDL → Database Diagram → New Database Diagram (Right Click)

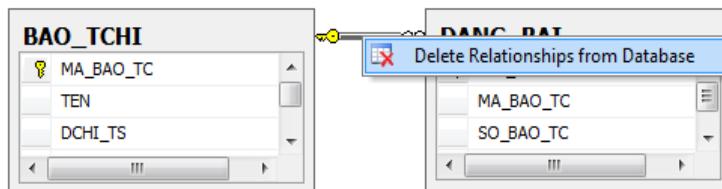


- Chọn các bảng trong CSDL để thêm vào Diagram
- Nếu giữa các bảng trong CSDL đi có các ràng buộc khóa ngoại, các ràng buộc này sẽ tự động được hiển thị trong Diagram.



Ta có thể xóa ràng buộc khóa ngoại trên Diagram như sau:

- Chọn ràng buộc cần xóa → Chọn Delete Relationship from Database (Right Click)



5. Hướng dẫn thực hành

5.1. Mở Query Analyzer đăng nhập vào SQL Server trên máy chủ

5.2. Quan sát CSDL

Sử dụng Object Browser để xem cấu trúc của các bảng cũng như dữ liệu của từng trường và mở từng bảng dữ liệu để quan sát.

Khi quan sát CSDL cần lưu ý:

- Sự liên thông giữa các bảng dữ liệu. Vẽ sơ đồ liên thông để có thể tham khảo trong quá trình truy vấn thông tin từ nhiều bảng.
- Kiểu dữ liệu của từng trường, để đặt điều kiện trên các cột cho đúng kiểu hoặc áp dụng các hàm xử lý dữ liệu phù hợp.

5.3. Mở CSDL muốn truy vấn

5.4. Soạn thảo và thực hiện câu truy vấn

5.5. Quan sát kết quả và kiểm chứng

Cần phải kiểm tra lại kết quả truy vấn xem có đúng hay không bằng cách mở lại các bảng để kiểm tra xem kết quả đạt được có đúng như mong muốn.

5.6. Lưu lại các câu truy vấn vào tập tin *.sql

Lưu ý: Các bước tạo một CSDL

- + Bước 1: Tạo bảng dữ liệu bằng lệnh CREATE TABLE.
- + Bước 2: Tạo liên kết giữa các bảng dữ liệu thông qua các khóa ngoại và khóa chính.
- + Bước 3: Nhập thông tin cho các bảng dữ liệu. Thông tin của các khóa chính được nhập trước, sau đó nhập thông tin cho các khóa ngoại.

III. THỰC HÀNH

Bài 1. Cơ sở dữ liệu về Quản lý giải bóng đá (QLBongDa)

1. MÔ TẢ CƠ SỞ DỮ LIỆU

• CAUTHU(MACT, HOTEN, VITRI, NGAYSINH, DIACHI, MACLB, MAQG, SO)

Mỗi cầu thủ có một mã số duy nhất (MACT), một họ tên (HOTEN), một ngày sinh (NGAYSINH), địa chỉ (DIACHI), thuộc một câu lạc bộ (MACLB), thuộc một quốc gia (MAQG), giữ một vị trí trên sân (VITRI), vị trí trên sân của cầu thủ chỉ thuộc một trong các vị trí sau: Thủ môn, Tiền đạo, Tiền vệ, Trung vệ, Hậu vệ; số áo (SO), cầu thủ thuộc cùng một câu lạc bộ phải có số áo khác nhau.

• QUOCGIA(MAQG, TENQG)

Mỗi quốc gia có một mã số duy nhất (MAQG) để phân biệt với các quốc gia khác, một tên duy nhất (TENQG).

• CAULACBO(MACLB, TENCLB, MASAN, MATINH)

Mỗi câu lạc bộ có một mã duy nhất (MACLB), một tên (TENCLB), quản lý một sân bóng (MASAN), thuộc một tỉnh thành (MATINH). Mỗi câu lạc bộ chỉ được phép đăng ký tối đa 8 cầu thủ nước ngoài.

• TINH(MATINH, TENTINH)

Mỗi tỉnh/thành có một duy nhất (MATINH), tên tỉnh/thành (TENTINH) duy nhất.

• SANVD(MASAN, TENSAN, DIACHI)

Mỗi sân bóng có một mã số duy nhất (MASAN), một tên sân duy nhất (TENSAN), một địa chỉ (DIACHI).

• HUANLV(MAHLV, TENHLV, NGAYSINH, DIACHI, DIENTHOAI, MAQG)

Mỗi huấn luyện viên có một mã số duy nhất (MAHLV), một họ tên (TENHLV), một ngày sinh (NGAYSINH), một địa chỉ (DIACHI), một điện thoại, thuộc một quốc gia (MAQG).

• HLV_CLB(MAHLV, MACLB, VAITRO)

Một câu lạc bộ có thể có nhiều huấn luyện viên (MAHLV), một huấn luyện viên có thể tham gia huấn luyện cho nhiều câu lạc bộ (MACLB) với một vai trò (VAITRO) cụ thể, vai trò của huấn luyện viên trong một câu lạc bộ thuộc một trong các vai trò sau: HLV chính, HLV phụ, HLV thể lực, HLV thủ môn. Mỗi câu lạc bộ chỉ có tối đa 2 HLV chính.

- TRANDAU(**MATRAN**, NAM, VONG, NGAYTD, MACLB1, MACLB2, MASAN, KETQUA)

Mỗi trận đấu của một vòng đấu (VONG) trong mùa giải (NAM) có một mã số duy nhất (MATRAN), được tổ chức thi đấu trên sân (MASAN) vào một ngày (NGAYTD) với sự tham gia của 2 câu lạc bộ có mã MACLB1, MACLB2, có kết quả (KETQUA).

- BANGXH(**MACLB**, NAM, VONG, SOTRAN, THANG, HOA, THUA, HIEUSO, DIEM, HANG)

Bảng xếp hạng cho biết kết quả xếp hạng vòng đấu (VONG) trong mùa giải (NAM) của đội bóng (MACLB) với số trận đã đấu (SOTRAN), trong đó số trận thắng (THANG), số trận hòa (HOA), số trận thua (THUA), có hiệu số bàn thắng - thua (HIEUSO), tổng điểm (DIEM), xếp hạng (HANG).

- CAUTHU(**MACT**, HOTEN, VITRI, NGAYSINH, DIACHI, MACLB, MAQG, SO)

STT	Thuộc tính	Kiểu dữ liệu	Ghi chú
1	MACT	NUMERIC	KHÓA CHÍNH, TĂNG TỰ ĐỘNG ¹
2	HOTEN	NVARCHAR(100)	BẮT BUỘC
3	VITRI	NVARCHAR(20)	BẮT BUỘC
4	NGAYSINH	DATETIME	
5	DIACHI	NVARCHAR(200)	
6	MACLB	VARCHAR (5)	BẮT BUỘC
7	MAQG	VARCHAR (5)	BẮT BUỘC
8	SO	INT	BẮT BUỘC

Tăng tự động: Ở phần Column Properties, ở thuộc tính Identity Specification:

- (Is Identity): Yes
- Identity Increment: (mức tăng, mặc định là 1)
- Identity Seed: (giá trị đầu, mặc định là 1)

- QUOCGIA(**MAQG**, TENQG)

STT	Thuộc tính	Kiểu dữ liệu	Ghi chú
1	MAQG	VARCHAR (5)	KHÓA CHÍNH
2	TENQG	NVARCHAR(60)	BẮT BUỘC

- CAULACBO(**MACLB**, TENCLB, MASAN, MATINH)

STT	Thuộc tính	Kiểu dữ liệu	Ghi chú
1	MACLB	VARCHAR (5)	KHÓA CHÍNH
2	TENCLB	NVARCHAR(100)	BẮT BUỘC
3	MASAN	VARCHAR (5)	BẮT BUỘC
4	MATINH	VARCHAR (5)	BẮT BUỘC

- TINH(**MATINH**, TENTINH)

Giáo trình Cơ sở dữ liệu

STT	Thuộc tính	Kiểu dữ liệu	Ghi chú
1	MATINH	VARCHAR (5)	KHÓA CHÍNH
2	TENTINH	NVARCHAR(100)	BẮT BUỘC

- SANVD(MASAN, TENSAN, DIACHI)

STT	Thuộc tính	Kiểu dữ liệu	Ghi chú
1	MASAN	VARCHAR (5)	KHÓA CHÍNH
2	TENSAN	NVARCHAR(100)	BẮT BUỘC
3	DIACHI	NVARCHAR(200)	

- HUANLV(MAHLV, TENHLV, NGAYSINH, DIACHI, DIENTHOAI, MAQG)

STT	Thuộc tính	Kiểu dữ liệu	Diễn giải
1	MAHLV	VARCHAR (5)	KHÓA CHÍNH
2	TENHLV	NVARCHAR(100)	BẮT BUỘC
3	NGAYSINH	DATETIME	
4	DIACHI	NVARCHAR(200)	
5	DIENTHOAI	NVARCHAR(20)	
6	MAQG	VARCHAR (5)	BẮT BUỘC

- HLV_CLB(MAHLV, MACLB, VAITRO)

STT	Thuộc tính	Kiểu dữ liệu	Diễn giải
1	MAHLV	VARCHAR (5)	KHÓA CHÍNH
2	MACLB	VARCHAR (5)	KHÓA CHÍNH
3	VAITRO	NVARCHAR(100)	BẮT BUỘC

- TRANDAU(MATRAN, NAM, VONG, NGAYTD, MACLB1, MACLB2, MASAN, KETQUA)

STT	Thuộc tính	Kiểu dữ liệu	Diễn giải
1	MATRAN	NUMERIC	KHÓA CHÍNH, TĂNG TỰ ĐỘNG
2	NAM	INT	BẮT BUỘC
3	VONG	INT	BẮT BUỘC
4	NGAYTD	DATETIME	BẮT BUỘC
5	MACLB1	VARCHAR (5)	BẮT BUỘC
6	MACLB2	VARCHAR (5)	BẮT BUỘC
7	MASAN	VARCHAR (5)	BẮT BUỘC
8	KETQUA	VARCHAR (5)	BẮT BUỘC

- BANGXH(MACLB, NAM, VONG, SOTRAN, THANG, HOA, THUA, HIEUSO, DIEM, HANG)

STT	Thuộc tính	Kiểu dữ liệu	Diễn giải
1	MACLB	VARCHAR (5)	KHÓA CHÍNH
2	NAM	INT	KHÓA CHÍNH
3	VONG	INT	KHÓA CHÍNH
4	SOTRAN	INT	BẮT BUỘC
5	THANG	INT	BẮT BUỘC
6	HOA	INT	BẮT BUỘC
7	THUA	INT	BẮT BUỘC
8	HIEUSO	VARCHAR (5)	BẮT BUỘC
9	DIEM	INT	BẮT BUỘC
10	HANG	INT	BẮT BUỘC

2. YÊU CẦU

1. Tạo CSDL có tên QLBongDa.
2. Tạo các Table CAUTHU, QUOCGIA, CAULACBO, TINH, HUANLV, SANVD, HLV_CLB, TRANDAU, BANGXH.
3. Tạo liên kết cho các Table trong CSDL trên.

Bài 2. Cơ sở dữ liệu QLCongTrinh (Quản lý công trình)

1. MÔ TẢ CƠ SỞ DỮ LIỆU

KTRUCSU(**HOTEN_KTS**, NAMS_KTS, PHAI, NOI_TN, DCHI_LL_KTS)

Mỗi kiến trúc sư có một họ tên, một năm sinh, phái, nơi tốt nghiệp, địa chỉ liên lạc. Giả sử không có hai kiến trúc sư trùng tên.

CHUTHAU(**TEN_THAU**, TEL, DCHI_THAU)

Mỗi chủ thầu xây dựng công trình đều có một tên thầu, một số điện thoại và một địa chỉ. Giả sử không có hai chủ thầu trùng tên.

CHUNHAN(**TEN_CHU**, DCHI_CHU)

Mỗi chủ nhân của một công trình đều có một tên chủ nhân và một địa chỉ. Giả sử không có hai chủ nhân trùng tên.

CONGNHAN(**HOTEN_CN**, NAMS_CN, NAM_VAO_N, CH_MON)

Mỗi công nhân tham gia xây dựng các công trình đều có một họ tên, một năm sinh, một năm vào nghề, và một chuyên môn nào đó. Giả sử không có hai công nhân trùng họ tên.

CONGTRINH(**STT_CTR**, **TEN_CTR**, **DCHI_CTR**, **TINH_THANH**, **KINH_PHI**, **TEN_CHU**, **TEN_THAU**, **NGAY_BD**)

Mỗi công trình có một số thứ tự, một tên công trình, một địa chỉ, thuộc về một tỉnh thành nào đó, được xây dựng với một kinh phí (đơn vị tính là triệu đồng), thuộc về sở hữu của một chủ nhân, do một chủ thầu xây dựng, và được khởi công xây từ một ngày nào đó.

THAMGIA(**HOTEN_CN**, **STT_CTR**, **NGAY_TGIA**, **SO_NGAY**)

Một công nhân có thể tham gia xây dựng nhiều công trình, và một công trình cũng có nhiều công nhân tham gia. Khi một công nhân tham gia vào một công trình nào đó sẽ được ghi nhận lại ngày bắt đầu tham gia và số ngày tham gia.

THIETKE(**HOTEN_KTS**, **STT_CTR**, **THU_LAO**)

Một kiến trúc sư có thể thiết kế nhiều công trình, mỗi công trình cũng có thể do nhiều kiến trúc sư cùng thiết kế. Khi một kiến trúc sư thiết kế một công trình sẽ có một thù lao tương ứng.

2. YÊU CẦU

1. Tạo CSDL có tên là QL_CONGTRINH.
2. Tạo các Table trong cơ sở dữ liệu QL_CONGTRINH.
3. Tạo liên kết giữa các Table trong CSDL trên.

4. Xem mô hình liên kết giữa các table

Bài 3. Cơ sở dữ liệu QLGiangDay như sau:

GIAOVIEN(MaGV, HoTen, #MaKhoa)

MONHOC (MaMH, TenMH)

PHONGHOC(Phong, ChucNang)

KHOA (MaKhoa, TenKhoa)

LOP(MaLop, TenLop, #MaKhoa)

LICHDAY (#MaGV, #MaMH, #Phong, #MaLop, NgayDay, TuTiet, DenTiet, BaiDay, LyThuyet, GhiChu)

Hãy thực hiện các yêu cầu sau:

1. Tạo CSDL QLGiangDay.
2. Tạo các table nhu trên vào CSDL QLGV có thiết lập ràng buộc khoá chính, khoá ngoại.
3. Xem mô hình liên kết giữa các table.
4. Thêm cột ngày sinh, giới tính, tuổi vào GIAOVIEN.
5. Thêm cột SucChua (sức chứa) vào PHONGHOC.
6. Thêm SoTC, số TCLT, số TCTH vào MONHOC.
7. Thêm SiSo vào LOP.
8. Thêm ràng buộc toàn vẹn cho cột SiSo, SoTC, TCLT, TCTH, SucChua, TuoiGV.

BÀI 4 : CSDL quản lý hàng hóa như sau :

1. MÔ TẢ CƠ SỞ DỮ LIỆU

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

Điễn giải: Quan hệ khách hàng sẽ lưu trữ thông tin của khách hàng thành viên gồm có các thuộc tính: mã khách hàng, họ tên, địa chỉ, số điện thoại, ngày sinh, ngày đăng ký và doanh số (tổng trị giá các hóa đơn của khách hàng thành viên này).

NHANVIEN (MANV,HOTEN, NGVL, SODT)

Điễn giải: Mỗi nhân viên bán hàng cần ghi nhận họ tên, ngày vào làm, điện thoại liên lạc, mỗi nhân viên phân biệt với nhau bằng mã nhân viên.

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

Điễn giải: Mỗi sản phẩm có một mã số, một tên gọi, đơn vị tính, nước sản xuất và một giá bán.

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

Điễn giải: Khi mua hàng, mỗi khách hàng sẽ nhận một hóa đơn tính tiền, trong đó sẽ có số hóa đơn, ngày mua, nhân viên nào bán hàng, trị giá của hóa đơn là bao nhiêu và mã số của khách hàng nếu là khách hàng thành viên.

CTHD (SOHD,MASP,SL)

Điễn giải: Diễn giải chi tiết trong mỗi hóa đơn gồm có những sản phẩm gì với số lượng là bao nhiêu.

BẢNG THUỘC TÍNH

Quan hệ	Thuộc tính	Điễn giải	Kiểu dữ liệu
KHACHHANG	MAKH	Mã khách hàng	Varchar(4)
	HOTEN	Họ tên	nvarchar(40)
	DCHI	Địa chỉ	nvarchar(50)
	SODT	Số điện thoại	varchar(20)
	NGSINH	Ngày sinh	smalldatetime
	NGDK	Ngày đăng ký thành viên	smalldatetime
	DOANHSO	Tổng trị giá các hóa đơn khách hàng đã mua	money
NHANVIEN	MANV	Mã nhân viên	Varchar(4)
	HOTEN	Họ tên	nvarchar(40)
	SODT	Số điện thoại	varchar(20)
	NGVL	Ngày vào làm	smalldatetime
SANPHAM	MASP	Mã sản phẩm	Varchar(4)
	TENSP	Tên sản phẩm	nvarchar(40)
	DVT	Đơn vị tính	nvarchar(20)
	NUOCSX	Nước sản xuất	nvarchar(40)
	GIA	Giá bán	money
HOADON	SOHD	Số hóa đơn	int
	NGHD	Ngày mua hàng	smalldatetime
	MAKH	Mã khách hàng nào mua	Varchar(4)
	MANV	Nhân viên bán hàng	Varchar(4)
	TRIGIA	Trị giá hóa đơn	money
CTHD	SOHD	Số hóa đơn	int
	MASP	Mã sản phẩm	Varchar(4)
	SL	Số lượng	int

2. YÊU CẦU

- Tạo các quan hệ và khai báo các khóa chính, khóa ngoại của quan hệ.
- Thêm vào thuộc tính GHICHU có kiểu dữ liệu *varchar(20)* cho quan hệ SANPHAM.
- Thêm vào thuộc tính LOAIKH có kiểu dữ liệu là *tinyint* cho quan hệ KHACHHANG.
- Sửa kiểu dữ liệu của thuộc tính GHICHU trong quan hệ SANPHAM thành *varchar(100)*.
- Xóa thuộc tính GHICHU trong quan hệ SANPHAM.
- Cài đặt thuộc tính LOAIKH trong quan hệ KHACHHANG có thể lưu các giá trị là: ‘Vang lai’, ‘Thuong xuyen’, ‘Vip’, ...
- Đơn vị tính của sản phẩm chỉ có thể là (‘Cây’, ‘hop’, ‘cai’, ‘quyen’, ‘chuc’)
- Giá bán của sản phẩm từ 500 đồng trở lên.
- Mỗi lần mua hàng, khách hàng phải mua ít nhất 1 sản phẩm.
- Ngày khách hàng đăng ký là khách hàng thành viên phải lớn hơn ngày sinh của người đó.

11. Ngày mua hàng (NGHD) của một khách hàng thành viên sẽ lớn hơn hoặc bằng ngày khách hàng đó đăng ký thành viên (NGDK).
12. Ngày bán hàng (NGHD) của một nhân viên phải lớn hơn hoặc bằng ngày nhân viên đó vào làm.
13. Mỗi một hóa đơn phải có ít nhất một chi tiết hóa đơn.
14. Trị giá của một hóa đơn là tổng thành tiền (số lượng*đơn giá) của các chi tiết thuộc hóa đơn đó.
15. Doanh số của một khách hàng là tổng trị giá các hóa đơn mà khách hàng thành viên đó đã mua.

Bài 5 : Mô tả CSDL quản lý học viên như sau

1. MÔ TẢ CƠ SỞ DỮ LIỆU

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

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

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

Điễn giải: mỗi lớp gồm có mã lớp, tên lớp, học viên làm lớp trưởng của lớp, số số lớp và giáo viên chủ nhiệm.

KHOA (MAKHOA, TENKHOA, NGTLAP, TRGKHOA)

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

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

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

DIEUKIEN (MAMH, MAMH_TRUOC)

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

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

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

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

Điễn giải: mỗi học kỳ của năm học sẽ phân công giảng dạy lớp nào học môn gì, giáo viên nào phụ trách.

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

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

Quan hệ	Thuộc tính	Diễn giải	Kiểu dữ liệu
KHOA	MAKHOA	Mã khoa (khóa chính)	varchar(4)
	TENKHOA	Tên khoa	nvarchar(40)
	NGTLAP	Ngày thành lập khoa	smalldatetime
	TRGKHOA	Trưởng khoa (mã giáo viên)	Varchar(4)
MONHOC	MAMH	Mã môn học (khóa chính)	varchar(10)
	TENMH	Tên môn học	nvarchar(40)
	TCLT	Số tín chỉ lý thuyết	tinyint
	TCTH	Số tín chỉ thực hành	tinyint
	MAKHOA	Môn học thuộc khoa nào	varchar(4)
DIEUKIEN	MAMH	Mã môn học (thuộc tính khóa)	varchar(10)
	MAMH_TRUOC	Mã môn học phải học trước (thuộc tính khóa)	varchar(10)
GIAOVIEN	MAGV	Mã giáo viên (khóa chính)	Varchar(4)
	HOTEN	Họ tên	nvarchar(40)
	HOCVI, HOCHAM	Học vị, học hàm	nvarchar(10)
	GIOITINH	Giới tính	nvarchar(3)
	NGSINH, NGVL	Ngày sinh, ngày vào làm việc	smalldatetime
	HESO	Hệ số lương	numeric(4,2)
	MUCLUONG	Mức lương	money
	MAKHOA	Thuộc khoa nào	varchar(4)
LOP	MALOP	Mã lớp (khóa chính)	Varchar(3)
	TENLOP	Tên lớp	nvarchar(40)
	TRGLOP	Lớp trưởng (mã học viên)	Varchar(5)
	SISO	Sĩ số lớp	tinyint
	MAGVCN	Mã giáo viên chủ nhiệm	Varchar(4)
HOCVIEN	MAHV	Mã học viên (khóa chính)	Varchar(5)
	HO	Họ và tên lót	nvarchar(40)
	TEN	Tên	nvarchar(10)
	NGSINH	Ngày sinh	smalldatetime
	GIOITINH	Giới tính	nvarchar(3)
	NOISINH	Nơi sinh	nvarchar(40)
	MALOP	Mã lớp	Varchar(3)
GIANGDAY	MALOP	Mã lớp (thuộc tính khóa)	Varchar(3)
	MAMH	Mã môn học (thuộc tính khóa)	varchar(10)
	MAGV	Mã giáo viên	Varchar(4)
	HOCKY	Học kỳ	tinyint
	NAM	Năm học	smallint
	TUNGAY, DENNGAY	Ngày bắt đầu và ngày kết thúc môn học	smalldatetime
KETQUATHI	MAHV	Mã học viên (thuộc tính khóa)	Varchar(5)
	MAMH	Mã môn học (thuộc tính khóa)	varchar(10)
	LANTHI	Lần thi (thuộc tính khóa)	tinyint
	NGTHI	Ngày thi	smalldatetime
	DIEM	Điểm thi	numeric(4,2)
	KQUA	Kết quả thi	varchar(10)

2. YÊU CẦU

1. Tạo bảng và khai báo tất cả các ràng buộc khóa chính, khóa ngoại. Thêm vào 3 thuộc tính GHICHU, DIEMTB, XEPOAI cho quan hệ HOCVIEN.
2. Mã học viên là một chuỗi 5 ký tự, 3 ký tự đầu là mã lớp, 2 ký tự cuối cùng là số thứ tự học viên trong lớp. VD: ‘K1101’
3. Thuộc tính GIOITINH chỉ có giá trị là ‘Nam’ hoặc ‘Nữ’.
4. Điểm số của một lần thi có giá trị từ 0 đến 10 và cần lưu đến 2 số lẻ (VD: 6.22).
5. Kết quả thi là ‘Đạt’ nếu điểm từ 5 đến 10 và ‘K Đạt’ nếu điểm nhỏ hơn 5.
6. Học viên thi một môn tối đa 3 lần.
7. Học kỳ chỉ có giá trị từ 1 đến 3.
8. Học vị của giáo viên chỉ có thể là ‘CN’, ‘KS’, ‘Ths’, ‘TS’.
9. Lớp trưởng của một lớp phải là học viên của lớp đó.
10. Trưởng khoa phải là giáo viên thuộc khoa và có học vị ‘TS’ hoặc ‘PTS’.
11. Học viên ít nhất là 18 tuổi.
12. Giảng dạy một môn học ngày bắt đầu (TUNGAY) phải nhỏ hơn ngày kết thúc (DENNGAY).
13. Giáo viên khi vào làm ít nhất là 22 tuổi.
14. Tất cả các môn học đều có số tín chỉ lý thuyết và tín chỉ thực hành chênh lệch nhau không quá 3.
15. Học viên chỉ được thi một môn học nào đó khi lớp của học viên đã học xong môn học này.
16. Mỗi học kỳ của một năm học, một lớp chỉ được học tối đa 3 môn.
17. Sỉ số của một lớp bằng với số lượng học viên thuộc lớp đó.
18. Trong quan hệ DIEUKIEN giá trị của thuộc tính MAMH và MAMH_TRUOC trong cùng một bộ không được giống nhau ('A', 'A') và cũng không tồn tại hai bộ ('A', 'B') và ('B', 'A').
19. Các giáo viên có cùng học vị, học hàm, hệ số lương thì mức lương bằng nhau.
20. Học viên chỉ được thi lại (lần thi >1) khi điểm của lần thi trước đó dưới 5.
21. Ngày thi của lần thi sau phải lớn hơn ngày thi của lần thi trước (cùng học viên, cùng môn học).
22. Học viên chỉ được thi những môn mà lớp của học viên đó đã học xong.
23. Khi phân công giảng dạy một môn học, phải xét đến thứ tự trước sau giữa các môn học (sau khi học xong những môn học phải học trước mới được học những môn liền sau).
24. Giáo viên chỉ được phân công dạy những môn thuộc khoa giáo viên đó phụ trách.

PHẦN II NGÔN NGỮ THAO TÁC DATA MANIPULATION LANGUAGE - DML

I. NỘI DUNG THỰC HIỆN

- Tìm hiểu cách thêm dữ liệu và các lưu ý khi thêm dữ liệu
- Thực hiện xóa dữ liệu và các lưu ý khi xóa dữ liệu
- Thực hiện cập nhật dữ liệu và các lưu ý khi cập nhật dữ liệu

II LÝ THUYẾT

1. Thêm dữ liệu

```
INSERT INTO <tên bảng> [<danh sách cột> ]  
    VALUES (<danh sách giá trị>)
```

2. Xóa dữ liệu

```
DELETE <Tên Bảng>  
    [FROM <danh sách bảng>]  
    [WHERE <điều kiện>]
```

3 Cập nhật dữ liệu trong bảng

```
UPDATE <tên bảng>  
    SET <tên cột 1> = <bíểu thức 1>  
        [, ...  
        , <tên cột n> = <bíểu thức n>]  
    ]  
    [FROM <danh sách bảng>]  
    [WHERE <điều kiện>]
```

III THỰC HÀNH

Bài 1: Sử dụng CSDL QLBongDa (ở phần I) thực hiện các yêu cầu sau:

1. Nhập dữ liệu vào các table như sau:

CAUTHU(MACT, HOTEN, VITRI, NGAYSINH, DIACHI, MACLB, MAQG, SO)

Giáo trình Cơ sở dữ liệu

MACT	HOTEN	VITRI	NGAYSINH	DIACHI	MACLB	MAQG	SO
1	Nguyễn Vũ Phong	Tiền vệ	20/02/1990	NULL	BBD	VN	17
2	Nguyễn Công Vinh	Tiền đạo	10/03/1992	NULL	HAGL	VN	9
4	Trần Tân Tài	Tiền vệ	12/11/1989	NULL	BBD	VN	8
5	Phan Hồng Sơn	Thủ môn	10/06/1991	NULL	HAGL	VN	1
6	Ronaldo	Tiền vệ	12/12/1989	NULL	SDN	BRA	7
7	Robinho	Tiền vệ	12/10/1989	NULL	SDN	BRA	8
8	Vidic	Hậu vệ	15/10/1987	NULL	HAGL	ANH	3
9	Trần Văn Santos	Thủ môn	21/10/1990	NULL	BBD	BRA	1
10	Nguyễn Trường Sơn	Hậu vệ	26/8/1993	NULL	BBD	VN	4

- QUOCGIA(MAQG, TENQG)

MAQG	TENQG
VN	Việt Nam
ANH	Anh Quốc
TBN	Tây Ban Nha
BDN	Bồ Đào Nha
BRA	Bra-xin
ITA	Ý
THA	Thái Lan

- CAULACBO(MACLB, TENCLB, MASAN, MATINH)

MACLB	TENCLB	MASAN	MATINH
BBD	BECAMEX BÌNH DƯƠNG	GD	BD
HAGL	HOÀNG ANH GIA LAI	PL	GL
SDN	SHB ĐÀ NẴNG	CL	DN
KKH	KHATOCO KHÁNH HÒA	NT	KH
TPY	THÉP PHÚ YÊN	TH	PY
GDT	GẠCH ĐÔNG TÂM LONG AN	LA	LA

- TINH(MATINH, TENTINH)

MATINH	TENTINH
BD	Bình Dương
GL	Gia Lai
DN	Đà Nẵng
KH	Khánh Hòa
PY	Phú Yên
LA	Long An

- SANVD(MASAN, TENSAN, DIACHI)

MASAN	TENSAN	DIACHI
GD	Gò Đậu	123 QL1, TX Thủ Dầu Một, Bình Dương
PL	Pleiku	22 Hồ Tùng Mậu, Thống Nhất, Thị xã Pleiku, Gia Lai
CL	Chi Lăng	127 Võ Văn Tần, Đà Nẵng
NT	Nha Trang	128 Phan Chu Trinh, Nha Trang, Khánh Hòa
TH	Tuy Hòa	57 Trường Chinh, Tuy Hòa, Phú Yên
LA	Long An	102 Hùng Vương, Tp Tân An, Long An

- HUANLV(MAHLV, TENHLV, NGAYSINH, DIACHI, DIENTHOAI, MAQG)

Giáo trình Cơ sở dữ liệu

MAHLV	TENHLV	NGAYSINH	DIACHI	DIENTHOAI	MAQG
HLV01	Vital	15/10/1955	NULL	0918011075	BDN
HLV02	Lê Huỳnh Đức	20/05/1972	NULL	01223456789	VN
HLV03	Kiatisuk	11/12/1970	NULL	01990123456	THA
HLV04	Hoàng Anh Tuấn	10/06/1970	NULL	0989112233	VN
HLV05	Trần Công Minh	07/07/1973	NULL	0909099990	VN
HLV06	Trần Văn Phúc	02/03/1965	NULL	01650101234	VN

• HLV_CLB(MAHLV, MACLB, VAITRO)

MAHLV	MACLB	VAITRO
HLV01	BBD	HLV Chính
HLV02	SDN	HLV Chính
HLV03	HAGL	HLV Chính
HLV04	KKH	HLV Chính
HLV05	GDT	HLV Chính
HLV06	BBD	HLV thủ môn

• TRANDAU(MATRAN, NAM, VONG, NGAYTD, MACLB1, MACLB2, MASAN, KETQUA)

MATRAN	NAM	VONG	NGAYTD	MACLB1	MACLB2	MASAN	KETQUA
1	2009	1	7/2/2009	BBD	SDN	GD	3-0
2	2009	1	7/2/2009	KKH	GDT	NT	1-1
3	2009	2	16/2/2009	SDN	KKH	CL	2-2
4	2009	2	16/2/2009	TPY	BBD	TH	5-0
5	2009	3	1/3/2009	TPY	GDT	TH	0-2
6	2009	3	1/3/2009	KKH	BBD	NT	0-1
7	2009	4	7/3/2009	KKH	TPY	NT	1-0
8	2009	4	7/3/2009	BBD	GDT	GD	2-2

• BANGXH(MACLB, NAM, VONG, SOTRAN, THANG, HOA, THUA, HIEUSO, DIEM, HANG)

MACLB	NAM	VONG	SOTRAN	THANG	HOA	THUA	HIEUSO	DIEM	HANG
BBD	2009	1	1	1	0	0	3-0	3	1
KKH	2009	1	1	0	1	0	1-1	1	2
GDT	2009	1	1	0	1	0	1-1	1	3
TPY	2009	1	0	0	0	0	0-0	0	4
SDN	2009	1	1	0	0	1	0-3	0	5
TPY	2009	2	1	1	0	0	5-0	3	1
BBD	2009	2	2	1	0	1	3-5	3	2
KKH	2009	2	2	0	2	0	3-3	2	3
GDT	2009	2	1	0	1	0	1-1	1	4
SDN	2009	2	2	1	1	0	2-5	1	5
BBD	2009	3	3	2	0	1	4-5	6	1
GDT	2009	3	2	1	1	0	3-1	4	2
TPY	2009	3	2	1	0	1	5-2	3	3
KKH	2009	3	3	0	2	1	3-4	2	4
SDN	2009	3	2	1	1	0	2-5	1	5
BBD	2009	4	4	2	1	1	6-7	7	1
GDT	2009	4	3	1	2	0	5-1	5	2
KKH	2009	4	4	1	2	1	4-4	5	3
TPY	2009	4	3	1	0	2	5-3	3	4
SDN	2009	4	2	1	1	0	2-5	1	5

Giáo trình Cơ sở dữ liệu

2. Xóa cầu thủ ‘Nguyễn Vũ Phong’.
3. Xóa các cầu thủ có tuổi trên 40.
4. Xóa các cầu thủ có tuổi trên 35 ở câu lạc bộ ‘BBD’.
5. Xóa các cầu thủ có tuổi trên 35 ở câu lạc bộ ‘BECAMEX BÌNH DƯƠNG’.
6. Đổi vị trí đấu của cầu thủ ‘Nguyễn Trường Sơn’ thành ‘Tiền đạo’.
7. Ghi nhận địa chỉ của huấn luyện viên ‘Lê Huỳnh Đức’ là ‘TP.HCM’
8. Đổi vị trí của cầu thủ ‘Nguyễn Công Vinh’ với cầu thủ ‘Trần Tấn Tài’.
9. Thay huấn luyện viên mới của ‘GẠCH ĐỒNG TÂM LONG AN’ là ‘Nguyễn Đức Thắng’.
10. Đổi tên câu lạc bộ ‘BECAMEX BÌNH DƯƠNG’ thành ‘BÌNH DƯƠNG’.

BÀI 2 : Sử dụng lại CSDL quản lý hàng hóa để thực hiện :

Cho dữ liệu của các bảng như sau

NHANVIEN

MANV	HOTEN	DTHOAI	NGVL
NV01	Nguyen Nhu Nhut	0927345678	13/4/2006
NV02	Le Thi Phi Yen	0987567390	21/4/2006
NV03	Nguyen Van B	0997047382	27/4/2006
NV04	Ngo Thanh Tuan	0913758498	24/6/2006
NV05	Nguyen Thi Truc Thanh	0918590387	20/7/2006

KHACHHANG

MAKH	HOTEN	DCHI	SODT	NGSINH	DOANHSO	NGDK
KH01	Nguyen Van A	731 Tran Hung Dao, Q5, TpHCM	08823451	22/1/1960	1,060,000	2/07/2006
KH02	Tran Ngoc Han	23/5 Nguyen Trai, Q5, TpHCM	0908256478	3/4/1974	280,000	3/07/2006
KH03	Tran Ngoc Linh	45 Nguyen Canh Chan, Q1, TpHCM	0938776266	12/6/1980	3,860,000	5/08/2006
KH04	Tran Minh Long	50/34 Le Dai Hanh, Q10, TpHCM	0917325476	9/3/1965	250,000	2/10/2006
KH05	Le Nhat Minh	34 Truong Dinh, Q3, TpHCM	08246108	10/3/1950	21,000	8/10/2006
KH06	Le Hoai Thuong	227 Nguyen Van Cu, Q5, TpHCM	08631738	31/1/1981	915,000	4/11/2006
KH07	Nguyen Van Tam	32/3 Tran Binh Trong, Q5, TpHCM	0916783565	6/4/1971	12,500	1/12/2006
KH08	Phan Thi Thanh	45/2 An Duong Vuong, Q5, TpHCM	0938435756	10/1/1971	365,000	3/12/2006
KH09	Le Ha Vinh	873 Le Hong Phong, Q5, TpHCM	08654763	3/9/1979	70,000	4/01/2007
KH10	Ha Duy Lap	34/34B Nguyen Trai, Q1, TpHCM	08768904	2/5/1983	67,500	6/01/2007

SANPHAM

MASP	TENSP	DVT	NUOCSX	GIA
BC01	But chi	Cây	Singapore	3,000
BC02	But chi	Cây	Singapore	5,000
BC03	But chi	Cây	Viet Nam	3,500
BC04	But chi	hop	Viet Nam	30,000
BB01	But bi	Cây	Viet Nam	5,000
BB02	But bi	Cây	Trung Quốc	7,000

Giáo trình Cơ sở dữ liệu

BB03	But bi	hop	Thai Lan	100,000
TV01	Tap 100 giay mong	quyen	Trung Quốc	2,500
TV02	Tap 200 giay mong	quyen	Trung Quốc	4,500
TV03	Tap 100 giay tot	quyen	Viet Nam	3,000
TV04	Tap 200 giay tot	quyen	Viet Nam	5,500
TV05	Tap 100 trang	chuc	Viet Nam	23,000
TV06	Tap 200 trang	chuc	Viet Nam	53,000
TV07	Tap 100 trang	chuc	Trung Quốc	34,000
ST01	So tay 500 trang	quyen	Trung Quốc	40,000
ST02	So tay loai 1	quyen	Viet Nam	55,000
ST03	So tay loai 2	quyen	Viet Nam	51,000
ST04	So tay	quyen	Thai Lan	55,000
ST05	So tay mong	quyen	Thai Lan	20,000
ST06	Phan viet bang	hop	Viet Nam	5,000
ST07	Phan khong bui	hop	Viet Nam	7,000
ST08	Bong bang	cai	Viet Nam	1,000
ST09	But long	Cây	Viet Nam	5,000
ST10	But long	Cây	Trung Quốc	7,000

HOADON

SOHD	NGHD	MAKH	MANV	TRIGIA
1001	23/07/2006	KH01	NV01	320,000
1002	12/08/2006	KH01	NV02	840,000
1003	23/08/2006	KH02	NV01	100,000
1004	01/09/2006	KH02	NV01	180,000
1005	20/10/2006	KH01	NV02	3,800,000
1006	16/10/2006	KH01	NV03	2,430,000
1007	28/10/2006	KH03	NV03	510,000
1008	28/10/2006	KH01	NV03	440,000
1009	28/10/2006	KH03	NV04	200,000
1010	01/11/2006	KH01	NV01	5,200,000
1011	04/11/2006	KH04	NV03	250,000
1012	30/11/2006	KH05	NV03	21,000
1013	12/12/2006	KH06	NV01	5,000
1014	31/12/2006	KH03	NV02	3,150,000
1015	01/01/2007	KH06	NV01	910,000
1016	01/01/2007	KH07	NV02	12,500
1017	02/01/2007	KH08	NV03	35,000
1018	13/01/2007	KH08	NV03	330,000
1019	13/01/2007	KH01	NV03	30,000

Giáo trình Cơ sở dữ liệu

1020	14/01/2007	KH09	NV04	70,000
1021	16/01/2007	KH10	NV03	67,500
1022	16/01/2007	Null	NV03	7,000
1023	17/01/2007	Null	NV01	330,000

CTHD

SOHD	MASP	SL
1001	TV02	10
1001	ST01	5
1001	BC01	5
1001	BC02	10
1001	ST08	10
1002	BC04	20
1002	BB01	20
1002	BB02	20
1003	BB03	10
1004	TV01	20
1004	TV02	10
1004	TV03	10
1004	TV04	10
1005	TV05	50
1005	TV06	50
1006	TV07	20

CTHD

SOHD	MASP	SL
1006	ST01	30
1006	ST02	10
1007	ST03	10
1008	ST04	8
1009	ST05	10
1010	TV07	50
1010	ST07	50
1010	ST08	100
1010	ST04	50
1010	TV03	100
1011	ST06	50
1012	ST07	3
1013	ST08	5
1014	BC02	80
1014	BB02	100
1014	BC04	60

CTHD

SOHD	MASP	SL
1014	BB01	50
1015	BB02	30
1015	BB03	7
1016	TV01	5
1017	TV02	1
1017	TV03	1
1017	TV04	5
1018	ST04	6
1019	ST05	1
1019	ST06	2
1020	ST07	10
1021	ST08	5
1021	TV01	7
1021	TV02	10
1022	ST07	1
1023	ST04	6

- Nhập dữ liệu cho các quan hệ trên.
- Tạo quan hệ SANPHAM1 chứa toàn bộ dữ liệu của quan hệ SANPHAM. Tạo quan hệ KHACHHANG1 chứa toàn bộ dữ liệu của quan hệ KHACHHANG.
- Cập nhật giá tăng 5% đối với những sản phẩm do ‘Thai Lan’ sản xuất (cho quan hệ SANPHAM1)
- Cập nhật giá giảm 5% đối với những sản phẩm do ‘Trung Quốc’ sản xuất có giá từ 10.000 trở xuống (cho quan hệ SANPHAM1).
- Cập nhật giá trị LOAIKH là ‘Vip’ đối với những khách hàng đăng ký thành viên trước ngày 1/1/2007 có doanh số từ 10.000.000 trở lên hoặc khách hàng đăng ký thành viên từ 1/1/2007 trở về sau có doanh số từ 2.000.000 trở lên (cho quan hệ KHACHHANG1).

BÀI 3 : Sử dụng CSDL quản lý học viên (ở phần I)

Cho dữ liệu ở các table như sau :

KHOA

MAKHOA	TENKHOA	NGTLAP	TRGKHOA
KHMT	Khoa hoc may tinh	7/6/2005	GV01

Giáo trình Cơ sở dữ liệu

HTTT	He thong thong tin	7/6/2005	GV02
CNPM	Cong nghe phan mem	7/6/2005	GV04
MTT	Mang va truyen thong	20/10/2005	GV03
KTMT	Ky thuat may tinh	20/12/2005	Null

DIEUKIEN

MAMH	MAMH_TRUOC
CSDL	CTR
CSDL	CTDLGT
CTDLGT	THDC
PTTKTT	THDC
PTTKTT	CTDLGT
DHMT	THDC
LTHDT	THDC
PTTKHTTT	CSDL

LOP

MALOP	TENLOP	TRGLOP	SISO	MAGVCN
K11	Lop 1 khoa 1	K1108	11	GV07
K12	Lop 2 khoa 1	K1205	12	GV09
K13	Lop 3 khoa 1	K1305	12	GV14

MONHOC

MAMH	TENMH	TCLT	TCTH	MAKHOA
THDC	Tin hoc dai cuong	4	1	KHMT
CTR	Cau truc roi rac	5	0	KHMT
CSDL	Co so du lieu	3	1	HTTT
CTDLGT	Cau truc du lieu va giao thuat	3	1	KHMT
PTTKTT	Phan tich thiet ke thuật toán	3	0	KHMT
DHMT	Do hoa may tinh	3	1	KHMT
KTMT	Kien truc may tinh	3	0	KTMT
TKCSDL	Thiet ke co so du lieu	3	1	HTTT
PTTKHTTT	Phan tich thiet ke he thong thong tin	4	1	HTTT
HDH	He dieu hanh	4	0	KTMT
NMCNPM	Nhap mon cong nghe phan mem	3	0	CNPM
LTCFW	Lap trinh C for win	3	1	CNPM
LTHDT	Lap trinh huong doi tuong	3	1	CNPM

GIANGDAY

MALOP	MAMH	MAGV	HOCKY	NAM	TUNGAY	DENNGAY
K11	THDC	GV07	1	2006	2/1/2006	12/5/2006
K12	THDC	GV06	1	2006	2/1/2006	12/5/2006
K13	THDC	GV15	1	2006	2/1/2006	12/5/2006
K11	CTR	GV02	1	2006	9/1/2006	17/5/2006
K12	CTR	GV02	1	2006	9/1/2006	17/5/2006
K13	CTR	GV08	1	2006	9/1/2006	17/5/2006
K11	CSDL	GV05	2	2006	1/6/2006	15/7/2006
K12	CSDL	GV09	2	2006	1/6/2006	15/7/2006
K13	CTDLGT	GV15	2	2006	1/6/2006	15/7/2006
K13	CSDL	GV05	3	2006	1/8/2006	15/12/2006
K13	DHMT	GV07	3	2006	1/8/2006	15/12/2006
K11	CTDLGT	GV15	3	2006	1/8/2006	15/12/2006

Giáo trình Cơ sở dữ liệu

K12	CTDLGT	GV15	3	2006	1/8/2006	15/12/2006
K11	HDH	GV04	1	2007	2/1/2007	18/2/2007
K12	HDH	GV04	1	2007	2/1/2007	20/3/2007
K11	DHMT	GV07	1	2007	18/2/2007	20/3/2007

GIAOVIEN

MAGV	HOTEN	HOCVI	HOCHAM	GIOITINH	NGSINH	NGVL	HESO	MUCLUONG	MAKHOA
GV01	Ho Thanh Son	PTS	GS	Nam	2/5/1950	11/1/2004	5.00	2,250,000	KHMT
GV02	Tran Tam Thanh	TS	PGS	Nam	17/12/1965	20/4/2004	4.50	2,025,000	HTTT
GV03	Do Nghiem Phung	TS	GS	Nu	1/8/1950	23/9/2004	4.00	1,800,000	CNPM
GV04	Tran Nam Son	TS	PGS	Nam	22/2/1961	12/1/2005	4.50	2,025,000	KTMT
GV05	Mai Thanh Danh	ThS	GV	Nam	12/3/1958	12/1/2005	3.00	1,350,000	HTTT
GV06	Tran Doan Hung	TS	GV	Nam	11/3/1953	12/1/2005	4.50	2,025,000	KHMT
GV07	Nguyen Minh Tien	ThS	GV	Nam	23/11/1971	1/3/2005	4.00	1,800,000	KHMT
GV08	Le Thi Tran	KS	Null	Nu	26/3/1974	1/3/2005	1.69	760,500	KHMT
GV09	Nguyen To Lan	ThS	GV	Nu	31/12/1966	1/3/2005	4.00	1,800,000	HTTT
GV10	Le Tran Anh Loan	KS	Null	Nu	17/7/1972	1/3/2005	1.86	837,000	CNPM
GV11	Ho Thanh Tung	CN	GV	Nam	12/1/1980	15/5/2005	2.67	1,201,500	MTT
GV12	Tran Van Anh	CN	Null	Nu	29/3/1981	15/5/2005	1.69	760,500	CNPM
GV13	Nguyen Linh Dan	CN	Null	Nu	23/5/1980	15/5/2005	1.69	760,500	KTMT
GV14	Truong Minh Chau	ThS	GV	Nu	30/11/1976	15/5/2005	3.00	1,350,000	MTT
GV15	Le Ha Thanh	ThS	GV	Nam	4/5/1978	15/5/2005	3.00	1,350,000	KHMT

KETQUATHI

MAHV	MAMH	LT	NGTHI	DIEM	KQUA
K1101	CSDL	1	20/7/2006	10.00	Dat
K1101	CTDLGT	1	28/12/2006	9.00	Dat
K1101	THDC	1	20/5/2006	9.00	Dat
K1101	CTR	1	13/5/2006	9.50	Dat
K1102	CSDL	1	20/7/2006	4.00	K Dat
K1102	CSDL	2	27/7/2006	4.25	K Dat
K1102	CSDL	3	10/8/2006	4.50	K Dat
K1102	CTDLGT	1	28/12/2006	4.50	K Dat
K1102	CTDLGT	2	5/1/2007	4.00	K Dat
K1102	CTDLGT	3	15/1/2007	6.00	Dat
K1102	THDC	1	20/5/2006	5.00	Dat
K1102	CTR	1	13/5/2006	7.00	Dat
K1103	CSDL	1	20/7/2006	3.50	K Dat
K1103	CSDL	2	27/7/2006	8.25	Dat
K1103	CTDLGT	1	28/12/2006	7.00	Dat
K1103	THDC	1	20/5/2006	8.00	Dat
K1103	CTR	1	13/5/2006	6.50	Dat

Giáo trình Cơ sở dữ liệu

K1104	CSDL	1	20/7/2006	3.75	K Dat
K1104	CTDLGT	1	28/12/2006	4.00	K Dat
K1104	THDC	1	20/5/2006	4.00	K Dat
K1104	CTR	1	13/5/2006	4.00	K Dat
K1104	CTR	2	20/5/2006	3.50	K Dat
K1104	CTR	3	30/6/2006	4.00	K Dat
K1201	CSDL	1	20/7/2006	6.00	Dat
K1201	CTDLGT	1	28/12/2006	5.00	Dat
K1201	THDC	1	20/5/2006	8.50	Dat
K1201	CTR	1	13/5/2006	9.00	Dat
K1202	CSDL	1	20/7/2006	8.00	Dat
K1202	CTDLGT	1	28/12/2006	4.00	K Dat
K1202	CTDLGT	2	5/1/2007	5.00	Dat
K1202	THDC	1	20/5/2006	4.00	K Dat
K1202	THDC	2	27/5/2006	4.00	K Dat
K1202	CTR	1	13/5/2006	3.00	K Dat

K1202	CTR	2	20/5/2006	4.00	K Dat
K1202	CTR	3	30/6/2006	6.25	Dat
K1203	CSDL	1	20/7/2006	9.25	Dat
K1203	CTDLGT	1	28/12/2006	9.50	Dat
K1203	THDC	1	20/5/2006	10.00	Dat
K1203	CTR	1	13/5/2006	10.00	Dat
K1204	CSDL	1	20/7/2006	8.50	Dat
K1204	CTDLGT	1	28/12/2006	6.75	Dat
K1204	THDC	1	20/5/2006	4.00	K Dat
K1204	CTR	1	13/5/2006	6.00	Dat
K1301	CSDL	1	20/12/2006	4.25	K Dat
K1301	CTDLGT	1	25/7/2006	8.00	Dat
K1301	THDC	1	20/5/2006	7.75	Dat
K1301	CTR	1	13/5/2006	8.00	Dat
K1302	CSDL	1	20/12/2006	6.75	Dat
K1302	CTDLGT	1	25/7/2006	5.00	Dat
K1302	THDC	1	20/5/2006	8.00	Dat
K1302	CTR	1	13/5/2006	8.50	Dat
K1303	CSDL	1	20/12/2006	4.00	K Dat
K1303	CTDLGT	1	25/7/2006	4.50	K Dat
K1303	CTDLGT	2	7/8/2006	4.00	K Dat
K1303	CTDLGT	3	15/8/2006	4.25	K Dat
K1303	THDC	1	20/5/2006	4.50	K Dat
K1303	CTR	1	13/5/2006	3.25	K Dat

Giáo trình Cơ sở dữ liệu

K1303	CTRR	2	20/5/2006	5.00	Dat
K1304	CSDL	1	20/12/2006	7.75	Dat
K1304	CTDLGT	1	25/7/2006	9.75	Dat
K1304	THDC	1	20/5/2006	5.50	Dat
K1304	CTRR	1	13/5/2006	5.00	Dat
K1305	CSDL	1	20/12/2006	9.25	Dat
K1305	CTDLGT	1	25/7/2006	10.00	Dat
K1305	THDC	1	20/5/2006	8.00	Dat
K1305	CTRR	1	13/5/2006	10.00	Dat

HỌC VIÊN

MAHV	HO	TEN	NGSINH	GIOITINH	NOISINH	MALOP
K1101	Nguyen Van	A	27/1/1986	Nam	TpHCM	K11
K1102	Tran Ngoc	Han	14/3/1986	Nu	Kien Giang	K11
K1103	Ha Duy	Lap	18/4/1986	Nam	Nghe An	K11
K1104	Tran Ngoc	Linh	30/3/1986	Nu	Tay Ninh	K11
K1105	Tran Minh	Long	27/2/1986	Nam	TpHCM	K11
K1106	Le Nhat	Minh	24/1/1986	Nam	TpHCM	K11
K1107	Nguyen Nhu	Nhut	27/1/1986	Nam	Ha Noi	K11
K1108	Nguyen Manh	Tam	27/2/1986	Nam	Kien Giang	K11
K1109	Phan Thi Thanh	Tam	27/1/1986	Nu	Vinh Long	K11
K1110	Le Hoai	Thuong	5/2/1986	Nu	Can Tho	K11
K1111	Le Ha	Vinh	25/12/1986	Nam	Vinh Long	K11
K1201	Nguyen Van	B	11/2/1986	Nam	TpHCM	K12
K1202	Nguyen Thi Kim	Duyen	18/1/1986	Nu	TpHCM	K12
K1203	Tran Thi Kim	Duyen	17/9/1986	Nu	TpHCM	K12
K1204	Truong My	Hanh	19/5/1986	Nu	Dong Nai	K12
K1205	Nguyen Thanh	Nam	17/4/1986	Nam	TpHCM	K12
K1206	Nguyen Thi Truc	Thanh	4/3/1986	Nu	Kien Giang	K12
K1207	Tran Thi Bich	Thuy	8/2/1986	Nu	Nghe An	K12
K1208	Huynh Thi Kim	trieu	8/4/1986	Nu	Tay Ninh	K12
K1209	Pham Thanh	trieu	23/2/1986	Nam	TpHCM	K12
K1210	Ngo Thanh	Tuan	14/2/1986	Nam	TpHCM	K12
K1211	Do Thi	Xuan	9/3/1986	Nu	Ha Noi	K12
K1212	Le Thi Phi	Yen	12/3/1986	Nu	TpHCM	K12
K1301	Nguyen Thi Kim	Cuc	9/6/1986	Nu	Kien Giang	K13
K1302	Truong Thi My	Hien	18/3/1986	Nu	Nghe An	K13
K1303	Le Duc	Hien	21/3/1986	Nam	Tay Ninh	K13
K1304	Le Quang	Hien	18/4/1986	Nam	TpHCM	K13
K1305	Le Thi	Huong	27/3/1986	Nu	TpHCM	K13
K1306	Nguyen Thai	Huu	30/3/1986	Nam	Ha Noi	K13

K1307	Tran Minh	Man	28/5/1986	Nam	TpHCM	K13
K1308	Nguyen Hieu	Nghia	8/4/1986	Nam	Kien Giang	K13
K1309	Nguyen Trung	Nghia	18/1/1987	Nam	Nghe An	K13
K1310	Tran Thi Hong	Tham	22/4/1986	Nu	Tay Ninh	K13
K1311	Tran Minh	Thuc	4/4/1986	Nam	TpHCM	K13
K1312	Nguyen Thi Kim	Yen	7/9/1986	Nu	TpHCM	K13

1. Tăng hệ số lương thêm 0.2 cho những giáo viên là trưởng khoa.
2. Cập nhật giá trị điểm trung bình tất cả các môn học (DIEMTB) của mỗi học viên (tất cả các môn học đều có hệ số 1 và nếu học viên thi một môn nhiều lần, chỉ lấy điểm của lần thi sau cùng).
3. Cập nhật giá trị cho cột GHICHU là ‘Cấm thi’ đối với trường hợp: học viên có một môn bất kỳ thi lần thứ 3 dưới 5 điểm.
4. Cập nhật giá trị cho cột XEPLOAI trong quan hệ HOCVIEN như sau:
 - o Nếu DIEMTB ≥ 9 thì XEPLOAI = ‘XS’
 - o Nếu $8 \leq \text{DIEMTB} < 9$ thì XEPLOAI = ‘G’
 - o Nếu $6.5 \leq \text{DIEMTB} < 8$ thì XEPLOAI = ‘K’
 - o Nếu $5 \leq \text{DIEMTB} < 6.5$ thì XEPLOAI = ‘TB’
 - o Nếu $\text{DIEMTB} < 5$ thì XEPLOAI = ‘Y’

PHẦN III TRUY VẤN DỮ LIỆU

STRUCTURED QUERY LANGUAGE - SQL

I. NỘI DUNG HƯỚNG DẪN THỰC HÀNH

1. Sử dụng các câu lệnh SQL có liên quan đến bảng dữ liệu
2. Các hàm thông dụng trong SQL
3. Các lỗi thường gặp trong SQL

II. LÝ THUYẾT

```
SELECT [TOP <Số mẫu tin> ] [Distinct]
      <danh sách biểu thức cột cần tổng hợp> | *
      FROM <danh sách bảng>
      WHERE <điều kiện>
      [
          GROUP BY <danh sách cột cần nhóm>
          HAVING <điều kiện trên nhóm>
      ]
      [ ORDER BY <Cột 1 cần sắp xếp> [Asc|Desc]
      [ [, ...], <cột n cần sắp xếp> [Asc |Desc] ]
```

III THỰC HÀNH

Bài 1: Sử dụng CSDL quản lý hàng hóa để trả lời câu hỏi sau:

1. In ra danh sách các sản phẩm (MASP,TENSP) do ‘Trung Quốc’ sản xuất.
2. In ra danh sách các sản phẩm (MASP, TENSP) có đơn vị tính là ‘Cây’, ‘Quyển’.
3. In ra danh sách các sản phẩm (MASP,TENSP) có mã sản phẩm bắt đầu là ‘B’ và kết thúc là ‘01’.
4. In ra danh sách các sản phẩm (MASP,TENSP) do ‘Trung Quốc’ sản xuất có giá từ 30.000 đến 40.000.
5. In ra danh sách các sản phẩm (MASP,TENSP) do ‘Trung Quốc’ hoặc ‘Thái Lan’ sản xuất có giá từ 30.000 đến 40.000.
6. **In ra các số hóa đơn, trị giá hóa đơn bán ra trong ngày 1/1/2007 và ngày 2/1/2007.**
7. In ra các số hóa đơn, trị giá hóa đơn trong tháng 1/2007, sắp xếp theo ngày (tăng dần) và trị giá của hóa đơn (giảm dần).
8. In ra danh sách các khách hàng (MAKH, HOTEN) đã mua hàng trong ngày 1/1/2007.
9. In ra số hóa đơn, trị giá các hóa đơn do nhân viên có tên ‘Nguyen Van B’ lập trong ngày 28/10/2006.
10. **In ra danh sách các sản phẩm (MASP,TENSP) được khách hàng có tên ‘Nguyen Van A’ mua trong tháng 10/2006.**
11. Tìm các số hóa đơn đã mua sản phẩm có mã số ‘BB01’ hoặc ‘BB02’.
12. Tìm các số hóa đơn đã mua sản phẩm có mã số ‘BB01’ hoặc ‘BB02’, mỗi sản phẩm mua với số lượng từ 10 đến 20.

13. Tìm các số hóa đơn mua cùng lúc 2 sản phẩm có mã số ‘BB01’ và ‘BB02’, mỗi sản phẩm mua với số lượng từ 10 đến 20.
14. In ra danh sách các sản phẩm (MASP,TENSP) do ‘Trung Quốc’ sản xuất hoặc các sản phẩm được bán ra trong ngày 1/1/2007.
15. In ra danh sách các sản phẩm (MASP,TENSP) không bán được.
16. In ra danh sách các sản phẩm (MASP,TENSP) không bán được trong năm 2006.
17. In ra danh sách các sản phẩm (MASP,TENSP) do ‘Trung Quốc’ sản xuất không bán được trong năm 2006.
18. Tìm số hóa đơn đã mua tất cả các sản phẩm do Singapore sản xuất.
19. Tìm số hóa đơn trong năm 2006 đã mua ít nhất tất cả các sản phẩm do Singapore sản xuất.
20. Có bao nhiêu hóa đơn không phải của khách hàng đăng ký thành viên mua?
21. Có bao nhiêu sản phẩm khác nhau được bán ra trong năm 2006.
22. Cho biết trị giá hóa đơn cao nhất, thấp nhất là bao nhiêu ?
23. Trị giá trung bình của tất cả các hóa đơn được bán ra trong năm 2006 là bao nhiêu?
24. Tính doanh thu bán hàng trong năm 2006.
25. Tìm số hóa đơn có trị giá cao nhất trong năm 2006.
26. Tìm họ tên khách hàng đã mua hóa đơn có trị giá cao nhất trong năm 2006.
27. In ra danh sách 3 khách hàng (MAKH, HOTEN) có doanh số cao nhất.
28. In ra danh sách các sản phẩm (MASP, TENSP) có giá bán bằng 1 trong 3 mức giá cao nhất.
29. In ra danh sách các sản phẩm (MASP, TENSP) do ‘Thai Lan’ sản xuất có giá bằng 1 trong 3 mức giá cao nhất (của tất cả các sản phẩm).
30. In ra danh sách các sản phẩm (MASP, TENSP) do ‘Trung Quốc’ sản xuất có giá bằng 1 trong 3 mức giá cao nhất (của sản phẩm do ‘Trung Quốc’ sản xuất).
31. * In ra danh sách 3 khách hàng có doanh số cao nhất (sắp xếp theo kiểu xếp hạng).
32. Tính tổng số sản phẩm do ‘Trung Quốc’ sản xuất.
33. Tính tổng số sản phẩm của từng nước sản xuất.
34. Với từng nước sản xuất, tìm giá bán cao nhất, thấp nhất, trung bình của các sản phẩm.
35. Tính doanh thu bán hàng mỗi ngày.
36. Tính tổng số lượng của từng sản phẩm bán ra trong tháng 10/2006.
37. Tính doanh thu bán hàng của từng tháng trong năm 2006.
38. Tìm hóa đơn có mua ít nhất 4 sản phẩm khác nhau.
39. Tìm hóa đơn có mua 3 sản phẩm do ‘Viet Nam’ sản xuất (3 sản phẩm khác nhau).
40. Tìm khách hàng (MAKH, HOTEN) có số lần mua hàng nhiều nhất.

41. Tháng mấy trong năm 2006, doanh số bán hàng cao nhất ?
42. Tìm sản phẩm (MASP, TENSP) có tổng số lượng bán ra thấp nhất trong năm 2006.
43. Mỗi nước sản xuất, tìm sản phẩm (MASP,TENSP) có giá bán cao nhất.
44. Tìm nước sản xuất sản xuất ít nhất 3 sản phẩm có giá bán khác nhau.
45. Trong 10 khách hàng có doanh số cao nhất, tìm khách hàng có số lần mua hàng nhiều nhất.

Bài 2: Sử dụng CSDL quản lý học viên để trả lời câu hỏi sau:

1. In ra danh sách (mã học viên, họ tên, ngày sinh, mã lớp) lớp trưởng của các lớp.
2. In ra bảng điểm khi thi (mã học viên, họ tên , lần thi, điểm số) môn CTRR của lớp ‘K12’, sắp xếp theo tên, họ học viên.
3. In ra danh sách những học viên (mã học viên, họ tên) và những môn học mà học viên đó thi lần thứ nhất đã đạt.
4. In ra danh sách học viên (mã học viên, họ tên) của lớp ‘K11’ thi môn CTRR không đạt (ở lần thi 1).
5. Danh sách học viên (mã học viên, họ tên) của lớp ‘K11’ thi môn CTRR không đạt (ở tất cả các lần thi).
6. Tìm tên những môn học mà giáo viên có tên ‘Tran Tam Thanh’ dạy trong học kỳ 1 năm 2006.
7. Tìm những môn học (mã môn học, tên môn học) mà giáo viên chủ nhiệm lớp ‘K11’ dạy trong học kỳ 1 năm 2006.
8. Tìm họ tên lớp trưởng của các lớp mà giáo viên có tên ‘Nguyen To Lan’ dạy môn ‘Co So Du Lieu’.
9. In ra danh sách những môn học (mã môn học, tên môn học) phải học liền trước môn ‘Co So Du Lieu’.
10. Môn ‘Cau Truc Roi Rac’ là môn bắt buộc phải học liền trước những môn học (mã môn học, tên môn học) nào.
11. Tìm họ tên giáo viên dạy môn CTRR cho cả hai lớp ‘K11’ và ‘K12’ trong cùng học kỳ 1 năm 2006.
12. Tìm những học viên (mã học viên, họ tên) thi không đạt môn CSDL ở lần thi thứ 1 nhưng chưa thi lại môn này.
13. Tìm giáo viên (mã giáo viên, họ tên) không được phân công giảng dạy bất kỳ môn học nào.
14. Tìm giáo viên (mã giáo viên, họ tên) không được phân công giảng dạy bất kỳ môn học nào thuộc khoa giáo viên đó phụ trách.
15. Tìm họ tên các học viên thuộc lớp ‘K11’ thi một môn bất kỳ quá 3 lần vẫn ‘K Dat’ hoặc thi lần thứ 2 môn CTRR được 5 điểm.
16. Tìm họ tên giáo viên dạy môn CTRR cho ít nhất hai lớp trong cùng một học kỳ của một năm học.

17. Danh sách học viên và điểm thi môn CSDL (chỉ lấy điểm của lần thi sau cùng).
18. Danh sách học viên và điểm thi môn ‘Co So Du Lieu’ (chỉ lấy điểm cao nhất của các lần thi).
19. Khoa nào (mã khoa, tên khoa) được thành lập sớm nhất.
20. Có bao nhiêu giáo viên có học hàm là ‘GS’ hoặc ‘PGS’.
21. Thống kê có bao nhiêu giáo viên có học vị là ‘CN’, ‘KS’, ‘Ths’, ‘TS’, ‘PTS’ trong mỗi khoa.
22. Mỗi môn học thống kê số lượng học viên theo kết quả (đạt và không đạt).
23. Tìm giáo viên (mã giáo viên, họ tên) là giáo viên chủ nhiệm của một lớp, đồng thời dạy cho lớp đó ít nhất một môn học.
24. Tìm họ tên lớp trưởng của lớp có số lượng cao nhất.
25. * Tìm họ tên những LOPTRG thi không đạt quá 3 môn (mỗi môn đều thi không đạt ở tất cả các lần thi).
26. Tìm học viên (mã học viên, họ tên) có số môn đạt điểm 9,10 nhiều nhất.
27. Trong từng lớp, tìm học viên (mã học viên, họ tên) có số môn đạt điểm 9,10 nhiều nhất.
28. Trong từng học kỳ của từng năm, mỗi giáo viên phân công dạy bao nhiêu môn học, bao nhiêu lớp.
29. Trong từng học kỳ của từng năm, tìm giáo viên (mã giáo viên, họ tên) giảng dạy nhiều nhất.
30. Tìm môn học (mã môn học, tên môn học) có nhiều học viên thi không đạt (ở lần thi thứ 1) nhất.
31. Tìm học viên (mã học viên, họ tên) thi môn nào cũng đạt (chỉ xét lần thi thứ 1).
32. Tìm học viên (mã học viên, họ tên) thi môn nào cũng đạt (chỉ xét lần thi sau cùng).
33. Tìm học viên (mã học viên, họ tên) đã thi tất cả các môn đều đạt (chỉ xét lần thi thứ 1).
34. Tìm học viên (mã học viên, họ tên) đã thi tất cả các môn đều đạt (chỉ xét lần thi sau cùng).
35. Tìm học viên (mã học viên, họ tên) có điểm thi cao nhất trong từng môn (lấy điểm ở lần thi sau cùng).

Bài 3: cho CSDL về quản lý chuyến bay như sau:

KHACHHANG(MAKH, TEN, DCHI, DTHOAI)

Mỗi khách hàng có một mã để phân biệt với những người khác, có tên, địa chỉ và số điện thoại để liên lạc.

NHANVIEN(MANV,TEN, DCHI, DTHOAI, LUONG, LOAINV)

Mỗi nhân viên của hãng hàng không có một tên, lương, địa chỉ, điện thoại và có một mã nhân viên duy nhất để phân biệt với các nhân viên khác. Nếu nhân viên là phi

công thì giá trị thuộc tính LOAINV bằng 1, nếu nhân viên là tiếp viên thì giá trị LOAINV bằng 0.

LOAIMB(MALOAI, HANGSX)

Mỗi loại máy bay có một mã loại để phân biệt với các loại máy bay khác và do một hãng sản xuất.

MAYBAY(SOHIEU, MALOAI)

Mỗi máy bay có một số hiệu để phân biệt với các máy bay khác trong cùng một loại.

CHUYENBAY(MACB, SBDI, SBDEN, GIODI, GIODEN)

Một chuyến bay có một mã chuyến bay để phân biệt với các chuyến bay khác. Mỗi chuyến bay xuất phát từ một sân bay (SBDI) và hạ cánh tại một sân bay khác (SBDEN); khởi hành vào một giờ (GIDI) và hạ cánh vào một giờ khác (GODEN).

LICHBAY(NGAYDI, MACB, SOHIEU, MALOAI)

Mỗi chuyến bay có thể có nhiều lịch bay. Mỗi lịch bay được xác định bởi một ngày khởi hành (NGAYDI) cụ thể và mã chuyến bay. Một lịch bay cụ thể chỉ sử dụng một máy bay. Giả sử mỗi chuyến bay chỉ được bố trí tối đa một lần cho một ngày.

DATCHO(MAKH, NGAYDI, MACB)

Mỗi khách hàng có thể đặt chỗ theo lịch bay của hãng hàng không đưa ra. Giả sử mỗi khách hàng chỉ được phép đặt tối đa một chỗ trên một chuyến bay vào một ngày cụ thể.

KHANANG(MANV, MALOAI)

Khả năng có thể lái loại máy bay của một phi công được biểu diễn bởi quan hệ KHANANG. Một phi công có thể biết lái nhiều loại máy bay khác nhau, và ngược lại, mỗi loại máy bay có thể có nhiều phi công có khả năng lái.

PHANCONG(MANV, NGAYDI, MACB)

Các nhân viên được phân công vào một hay nhiều lịch bay. Đối với phi công, họ phải được phân công lái những loại máy bay phù hợp với khả năng.

Bảng mô tả thuộc tính

Tên tắt	Điễn giải	Miền giá trị
TEN	Tên nhân viên hoặc khách hàng	Chuỗi (15)
DCHI	Địa chỉ nhân viên hoặc khách hàng	Chuỗi (50)
DTHOAI	Điện thoại nhân viên hoặc khách hàng	Chuỗi (12)
LUONG	Lương nhân viên	Số thực (10,2)
MANV	Mã nhân viên	Chuỗi (15)
MAKH	Mã khách hàng	Chuỗi (15)
LOAINV	Nhân viên là phi công = 1, là tiếp viên = 0	Bit (1/0)
HANGSX	Hãng sản xuất máy bay	Chuỗi (15)
MALOAI	Mã loại máy bay	Chuỗi (15)
SOHIEU	Số hiệu máy bay	Số nguyên
MACB	Mã chuyến bay	Chuỗi (4)
SBDI	Sân bay xuất phát của chuyến bay	Chuỗi (3)
SBDEN	Sân bay đến của chuyến bay	Chuỗi (3)
GIODI	Giờ máy bay cất cánh	Giờ (HH:MM)
GIODEN	Giờ máy bay hạ cánh	Giờ (HH:MM)
NGAYDI	Ngày đi của chuyến bay	Ngày (MM/DD/YYYY)

TRẢ LỜI TRUY VẤN

- Cho biết mã số, tên phi công, địa chỉ, điện thoại của các phi công đã từng lái máy bay loại B747.
- Cho biết mã số và ngày đi của các chuyến bay xuất phát từ sân bay DCA trong khoảng thời gian từ 14 giờ đến 18 giờ.
- Cho biết tên những nhân viên được phân công trên chuyến bay có mã số 100 xuất phát tại sân bay SLC. Các dòng dữ liệu xuất ra không được phép trùng lặp.
- Cho biết mã loại và số hiệu máy bay đã từng xuất phát tại sân bay MIA. Các dòng dữ liệu xuất ra không được phép trùng lặp.
- Cho biết mã chuyến bay, ngày đi, cùng với tên, địa chỉ, điện thoại của tất cả các hành khách đi trên chuyến bay đó. Sắp xếp theo thứ tự tăng dần của mã chuyến bay và theo ngày đi giảm dần.
- Cho biết mã chuyến bay, ngày đi, cùng với tên, địa chỉ, điện thoại của tất cả những nhân viên được phân công trong chuyến bay đó. Sắp xếp theo thứ tự tăng dần của mã chuyến bay và theo ngày đi giảm dần.
- Cho biết mã chuyến bay, ngày đi, mã số và tên của những phi công được phân công vào chuyến bay hạ cánh xuống sân bay ORD.
- Cho biết các chuyến bay (mã số chuyến bay, ngày đi và tên của phi công) trong đó phi công có mã 1001 được phân công lái.
- Cho biết thông tin (mã chuyến bay, sân bay đi, giờ đi, giờ đến, ngày đi) của những chuyến bay hạ cánh xuống DEN. Các chuyến bay được liệt kê theo ngày đi giảm dần và sân bay xuất phát (SBDI) tăng dần.
- Với mỗi phi công, cho biết hãng sản xuất và mã loại máy bay mà phi công này có khả năng bay được. Xuất ra tên phi công, hãng sản xuất và mã loại máy bay.

11. Cho biết mã phi công, tên phi công đã lái máy bay trong chuyến bay mã số 100 vào ngày 11/01/2000.
12. Cho biết mã chuyến bay, mã nhân viên, tên nhân viên được phân công vào chuyến bay xuất phát ngày 10/31/2000 tại sân bay MIA vào lúc 20:30
13. Cho biết thông tin về chuyến bay (mã chuyến bay, số hiệu, mã loại, hãng sản xuất) mà phi công ‘Quang’ đã lái.
14. Cho biết tên của những phi công chưa được phân công lái chuyến bay nào.
15. Cho biết tên khách hàng đã đi chuyến bay trên máy bay của hãng ‘Boeing’
16. Cho biết mã các chuyến bay chỉ bay với máy bay số hiệu 10 và mã loại B747.
17. Với mỗi sân bay (SBDEN), cho biết số lượng chuyến bay hạ cánh xuống sân bay đó. Kết quả được sắp xếp theo thứ tự tăng dần của sân bay đến.
18. Với mỗi sân bay (SBDI), cho biết số lượng chuyến bay xuất phát từ sân bay đó, sắp xếp theo thứ tự tăng dần của sân bay xuất phát.
19. Với mỗi sân bay (SBDI), cho biết số lượng chuyến bay xuất phát theo từng ngày. Xuất ra mã sân bay đi, ngày và số lượng.
20. Với mỗi sân bay (SBDEN), cho biết số lượng chuyến bay hạ cánh theo từng ngày. Xuất ra mã sân bay đến, ngày và số lượng.
21. Với mỗi lịch bay, cho biết mã chuyến bay, ngày đi cùng với số lượng nhân viên không phải là phi công của chuyến bay đó.
22. Số lượng chuyến bay xuất phát từ sân bay MIA vào ngày 11/01/2000.
23. Với mỗi chuyến bay, cho biết mã chuyến bay, ngày đi, số lượng nhân viên được phân công trên chuyến bay đó, sắp theo thứ tự giảm dần của số lượng.
24. Với mỗi chuyến bay, cho biết mã chuyến bay, ngày đi, cùng với số lượng hành khách đã đặt chỗ của chuyến bay đó, sắp theo thứ tự giảm dần của số lượng.
25. Với mỗi chuyến bay, cho biết mã chuyến bay, ngày đi, tổng lương của phi hành đoàn (các nhân viên được phân công trong chuyến bay), sắp xếp theo thứ tự tăng dần của tổng lương.
26. Cho biết lương trung bình của các nhân viên không phải là phi công.
27. Cho biết mức lương trung bình của các phi công.
28. Với mỗi loại máy bay, cho biết số lượng chuyến bay đã bay trên loại máy bay đó hạ cánh xuống sân bay ORD. Xuất ra mã loại máy bay, số lượng chuyến bay.
29. Cho biết sân bay (SBDI) và số lượng chuyến bay có nhiều hơn 2 chuyến bay xuất phát trong khoảng 10 giờ đến 22 giờ.
30. Cho biết tên phi công được phân công vào ít nhất 2 chuyến bay trong cùng một ngày.
31. Cho biết mã chuyến bay và ngày đi của những chuyến bay có ít hơn 3 hành khách đặt chỗ.
32. Cho biết số hiệu máy bay và loại máy bay mà phi công có mã 1001 được phân công lái trên 2 lần.

33. Với mỗi hãng sản xuất, cho biết số lượng loại máy bay mà hãng đó đã sản xuất. Xuất ra hãng sản xuất và số lượng.
34. Cho biết hãng sản xuất, mã loại và số hiệu của máy bay đã được sử dụng nhiều nhất.
35. Cho biết tên nhân viên được phân công đi nhiều chuyến bay nhất.
36. Cho biết thông tin của phi công (tên, địa chỉ, điện thoại) lái nhiều chuyến bay nhất.
37. Cho biết sân bay (SBDEN) và số lượng chuyến bay của sân bay có ít chuyến bay đáp xuống nhất.
38. Cho biết sân bay (SBDI) và số lượng chuyến bay của sân bay có nhiều chuyến bay xuất phát nhất.
39. Cho biết tên, địa chỉ, và điện thoại của khách hàng đã đi trên nhiều chuyến bay nhất.
40. Cho biết mã số, tên và lương của các phi công có khả năng lái nhiều loại máy bay nhất.
41. Cho biết thông tin (mã nhân viên, tên, lương) của nhân viên có mức lương cao nhất.
42. Cho biết tên, địa chỉ của các nhân viên có lương cao nhất trong phi hành đoàn (các nhân viên được phân công trong một chuyến bay) mà người đó tham gia.
43. Cho biết mã chuyến bay, giờ đi và giờ đến của chuyến bay bay sớm nhất trong ngày.
44. Cho biết mã chuyến bay có thời gian bay dài nhất. Xuất ra mã chuyến bay và thời gian bay (tính bằng phút).
45. Cho biết mã chuyến bay có thời gian bay ít nhất. Xuất ra mã chuyến bay và thời gian bay.
46. Cho biết mã chuyến bay và ngày đi của những chuyến bay bay trên loại máy bay B747 nhiều nhất.
47. Với mỗi chuyến bay có trên 3 hành khách, cho biết mã chuyến bay và số lượng nhân viên trên chuyến bay đó. Xuất ra mã chuyến bay và số lượng nhân viên.
48. Với mỗi loại nhân viên có tổng lương trên 600000, cho biết số lượng nhân viên trong từng loại nhân viên đó. Xuất ra loại nhân viên, và số lượng nhân viên tương ứng.
49. Với mỗi chuyến bay có trên 3 nhân viên, cho biết mã chuyến bay và số lượng khách hàng đã đặt chỗ trên chuyến bay đó.
50. Với mỗi loại máy bay có nhiều hơn một chiếc, cho biết số lượng chuyến bay đã được bố trí bay bằng loại máy bay đó. Xuất ra mã loại và số lượng.
51. Cho biết mã những chuyến bay đã bay tất cả các máy bay của hãng ‘Boeing’.
52. Cho biết mã và tên phi công có khả năng lái tất cả các máy bay của hãng ‘Airbus’.
53. Cho biết tên nhân viên (không phải là phi công) được phân công bay vào tất cả các chuyến bay có mã 100.
54. Cho biết ngày đi nào mà có tất cả các loại máy bay của hãng ‘Boeing’ tham gia.
55. Cho biết loại máy bay của hãng ‘Boeing’ nào có tham gia vào tất cả các ngày đi.
56. Cho biết mã và tên các khách hàng có đặt chỗ trong tất cả các ngày từ 31/10/2000 đến 1/1/2000.

Bài 4: cho cơ sở dữ liệu về quản lý đề án như sau:

NHANVIEN (MANV, HONV, TENLOT, TENNV, NGSINH, DCHI, PHAI, MLUONG, MA_NQL, PHONG)

Mỗi nhân viên có Mã nhân viên (MANV) duy nhất để phân biệt với các nhân viên khác, có họ tên (HONV, TENLOT, TENNV), ngày sinh (NGAYSINH), địa chỉ (DIACHI), phái Nam hoặc Nữ (PHAI), mức lương (LUONG), người quản lý trực tiếp (MANQL) và thuộc về một phòng ban (PHONG).

PHONGBAN (MAPHG, TENPHG, TRGPHG, NGNC)

Mỗi một phòng ban có một mã phòng duy nhất (MAPH) để phân biệt với các phòng ban khác, có tên phòng (TENPHG), người trưởng phòng (TRGPHG), và ngày nhận chức của trưởng phòng (NGNC).

DIADIEMPHG (MAPHG, DIADIEM)

Mỗi một phòng ban (MAPHG) có thể có nhiều địa điểm làm việc khác nhau (DIADIEM).

DEAN (MADA, TENDA, DDIEM_DA, PHONG, NGBD_DA, NGKT_DA)

Mỗi một đề án có một mã đề án duy nhất (MADA) để phân biệt với các đề án khác, có tên đề án (TENDA), địa điểm thực hiện (DDIEMDA), và do một phòng ban chủ trì đề án đó (PHONG), đồng thời mỗi đề án đều có ngày bắt đầu (NGBD_DA), ngày kết thúc đề án (NGKT_DA).

PHANCONG (MANV, MADA, THOIGIAN)

Mỗi một nhân viên (MANV) được phân công tham gia đề án (MADA) dưới dạng tham gia số giờ trên 1 tuần (THOIGIAN)

THANNHAN(MANV, MATN, HOTN, TENTN, PHAI, NGAYSINH, QUANHE)

Mỗi thân nhân có Mã thân nhân (MATN) duy nhất để phân biệt với các thân nhân khác, có họ tên (HOTN, TENTN), phái (PHAI) ngày sinh (NGAYSINH) và mối quan hệ (QUANHE) với nhân viên.

TRẢ LỜI CÁC TRUY VẤN SAU:

1. Tìm những nhân viên (MANV, HONV, TENLOT, TENNV) làm việc ở phòng NC'
2. Tìm những nhân viên (MANV, HONV, TENLOT, TENNV, PHAI) có mức lương trên 2.000.000.
3. Tìm họ tên nhân viên và tên phòng ban nhân viên đó trực thuộc có mức lương từ 2.000.000 đến 3.000.000.
4. Tìm các nhân viên có mức lương trên 2.500.000 ở phòng NC hoặc các nhân viên có mức lương trên 3.000.000 ở phòng DH.
5. Cho biết họ tên đầy đủ của các nhân viên ở 'Tp HCM'.
6. Cho biết họ tên đầy đủ của các nhân viên có họ bắt đầu bằng ký tự 'N'.
7. Cho biết ngày sinh và địa chỉ của nhân viên 'Đinh Bá Tiên'.
8. Cho biết các nhân viên có năm sinh trong khoảng 1960 đến 1965.

9. Cho biết các nhân viên và tuổi của nhân viên.
10. Lập danh sách các nhân viên dưới 18 tuổi của các nhân viên có mã là ‘001’.
11. Danh sách tất cả các nhân viên nữ trên 30 tuổi.
12. Với mỗi phòng ban, cho biết tên phòng ban và địa điểm phòng.
13. Tìm tên những người trưởng phòng của từng phòng ban.
14. Tìm TENDA, MADA, DDIEM_DA, TENPHG, MAPHG, TRPHG, NGNC.
15. Tìm tên và địa chỉ của tất cả các nhân viên của phòng ‘Nghiên cứu’.
16. Tìm tên những nữ nhân viên và tên người thân của họ.
17. Tìm mã số và tên những nhân viên thuộc phòng ‘Nghiên Cứu’ tham gia đề án ‘Tin Học Hóa’ với thời gian làm việc 20 giờ/tuần.
18. Với mọi đề án ở ‘Hà Nội’, liệt kê các mã số đề án (MADA), tên phòng ban chủ trì đề án.
19. Với mỗi đề án ở ‘Hà Nội’, cho biết tên đề án, tên phòng ban, họ tên và ngày nhận chức của trưởng phòng của phòng ban chủ trì đề án đó.
20. Với mỗi đề án ở ‘Hà Nội’, cho biết tên phòng (PHONG), họ tên trưởng phòng (HONV, TENLOT, TENNV) cũng như địa chỉ (DCHI) và ngày sinh (NGSINH) của người ấy.
21. Với mỗi nhân viên, cho biết họ tên nhân viên và họ tên người quản lý trực tiếp của nhân viên đó.
22. Với mỗi nhân viên, cho biết họ tên nhân viên và họ tên trưởng phòng của phòng ban mà nhân viên đó làm việc.
23. Cho biết họ tên nhân viên (HONV, TENLOT, TENNV) và tên các đề án mà nhân viên ấy tham gia nếu có.
24. Cho biết tên các đề án mà nhân viên ‘Đinh Bá Tiến’ đã tham gia.
25. Cho biết số lượng đề án của công ty.
26. Cho biết số lượng đề án do phòng ‘Nghiên Cứu’ chủ trì.
27. Cho biết lương trung bình của tất cả các nữ nhân viên.
28. Cho biết số tiền lương của nhân viên ‘Đinh Bá Tiến’.
29. Với mỗi đề án, liệt kê tên đề án (TENDA) và tổng số giờ làm việc một tuần của tất cả các nhân viên tham dự đề án đó.
30. Với mỗi đề án, cho biết có bao nhiêu nhân viên tham gia đề án đó
31. Với mỗi đề án, cho biết tên đề án và số lượng công việc của đề án này.
32. Với mỗi công việc trong đề án có mã đề án là ‘Đào tạo’, cho biết số lượng nhân viên được phân công.
33. Với mỗi nhân viên, cho biết họ và tên nhân viên và nhân viên đó có bao nhiêu thành phần.
34. Với mỗi nhân viên, cho biết số lượng nhân viên mà nhân viên đó quản lý trực tiếp.
35. Với mỗi phòng ban, liệt kê tên phòng ban (TENPHG) và lương trung bình của những nhân viên làm việc cho phòng ban đó.

36. Với các phòng ban có mức lương trung bình trên 30,000, liệt kê tên phòng ban và số lượng nhân viên của phòng ban đó.
37. Với mỗi phòng ban, cho biết tên phòng ban và số lượng đề án mà phòng ban đó chủ trì.
38. Với mỗi phòng ban, cho biết tên phòng ban, họ tên người trưởng phòng và số lượng đề án mà phòng ban đó chủ trì.
39. Cho biết số đề án diễn ra tại từng địa điểm.
40. Cho biết danh sách các đề án (MADA) có: nhân công với họ (HONV) là ‘Dinh’ hoặc có người trưởng phòng chủ trì đề án với họ (HONV) là ‘Dinh’.
41. Danh sách những nhân viên (HONV, TENLOT, TENNV) có trên 2 thân nhân.
42. Danh sách những nhân viên (HONV, TENLOT, TENNV) không có thân nhân nào.
43. Danh sách những trưởng phòng (HONV, TENLOT, TENNV) có tối thiểu một thân nhân.
44. Tìm họ (HONV) của những trưởng phòng chưa có gia đình.
45. Cho biết họ tên nhân viên (HONV, TENLOT, TENNV) có mức lương trên mức lương trung bình của phòng ‘Nghiên cứu’
46. Cho biết tên phòng ban và họ tên trưởng phòng của phòng ban có đông nhân viên nhất.
47. Tìm họ tên (HONV, TENLOT, TENNV) và địa chỉ (DCHI) của những nhân viên làm việc cho một đề án ở ‘TP HCM’ nhưng phòng ban mà họ trực thuộc lại không toạ lạc ở thành phố ‘TP HCM’.
48. Tổng quát câu 47, tìm họ tên và địa chỉ của các nhân viên làm việc cho một đề án ở một thành phố nhưng phòng ban mà họ trực thuộc lại không toạ lạc ở thành phố đó.
49. Danh sách những nhân viên (HONV, TENLOT, TENNV) làm việc trong mọi đề án của công ty
50. Danh sách những nhân viên (HONV, TENLOT, TENNV) được phân công tất cả đề án do phòng số ‘Nghiên cứu’ chủ trì.
51. Tìm họ tên nhân viên tham gia tất cả các đề án do phòng ban của nhân viên đó trực thuộc chủ trì.
52. Tìm những nhân viên (HONV, TENLOT, TENNV) được phân công tất cả đề án mà nhân viên ‘Đinh Bá Tiến’ làm việc.
53. Cho biết những nhân được phân công cho tất cả các công việc trong đề án ‘Sản phẩm X’.
54. Cho biết danh sách nhân viên tham gia vào tất cả các đề án ở TP HCM.
55. Cho biết phòng ban chủ trì tất cả các đề án ở TP HCM.

IV. MỘT SỐ HÀM THÔNG DỤNG

1. Xử lý chuỗi

Lưu ý: SQL Server không phân biệt chữ hoa và chữ thường. Vì vậy, điều kiện: $lower(ten_ctr)='khach san QUOC TE' \Rightarrow$ vẫn cho ra kết quả

Một số hàm xử lý chuỗi thông dụng:

LEN(<chuỗi>): Trả về chiều dài chuỗi

LOWER(<chuỗi>): Đổi <chuỗi> sang dạng chữ thường

Ví dụ: Select LOWER('Tran PHI phap') \Rightarrow 'tran phi phap'

UPPER(<chuỗi>): Đổi <chuỗi> sang dạng chữ hoa

Ví dụ: Select UPPER('Tran PHI phap') \Rightarrow 'TRAN PHI PHAP'

LEFT(<chuỗi>, <số n>): Trả về chuỗi mới gồm n ký tự bên trái của <chuỗi>

Ví dụ: Select LEFT('Tin Hoc', 3) \Rightarrow 'Tin'

RIGHT(<chuỗi>, <số n>): Trả về chuỗi mới gồm n ký tự bên phải của <chuỗi>

SUBSTRING(<chuỗi>, <số n>, <số m>): Trả về chuỗi con bằng cách lấy từ <chuỗi> m ký tự từ vị trí n

Ví dụ: Select SUBSTRING('Tin Mới Học', 5, 3) \Rightarrow 'Mới'

STUFF(<chuỗi 1>, <số n>, <số m>, <chuỗi 2>): Thay m ký tự trong <chuỗi 1> từ vị trí n bằng <chuỗi 2>

Ví dụ: Select STUFF ('Tin Học', 5, 0, 'Mới ') \Rightarrow 'Tin Mới Học' -- xen vào

Select STUFF('Tin Hãy Học', 5, 3, 'Mới') \Rightarrow 'Tin Mới Học' -- thay thế

Select STUFF('Tin Hãy Học', 5, 4, '') \Rightarrow 'Tin Hoc' -- xóa

CHARINDEX (<chuỗi 1>, <chuỗi 2> [,<số n>]): Trả về vị trí xuất hiện lần đầu tiên của chuỗi 1 trong chuỗi 2, bắt đầu tìm từ vị trí thứ n của chuỗi 2.

Ví dụ: Select CHARINDEX ('qua', 'noi qua ma hong qua') \Rightarrow 5

Select CHARINDEX ('qua', 'noi qua ma hong qua', 6) \Rightarrow 17

PATINDEX (<%mẫu tìm kiếm%>, <chuỗi>): Trả về vị trí bắt đầu xuất hiện mẫu lần đầu tiên trong chuỗi, nếu không tìm thấy mẫu thì trả về 0.

Ví dụ: Select PATINDEX('%om%', 'lom khom duoi nui tieu vai chu') \Rightarrow 2

Select PATINDEX('%__om %', 'lom khom duoi nui tieu vai chu') \Rightarrow 5

2.Xử lý ngày

Lưu ý: - Kiểu DATETIME của SQL Server bao gồm nhiều thành phần:

ngày / tháng / năm / giờ / phút / giây / phần triệu của giây

- Để tránh nhầm lẫn, nên mô tả năm bằng 4 chữ số, mô tả tháng bằng tên tắt của tháng (Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec).

GETDATE(): Cho kết quả là ngày hiện hành.

DATEPART(<thành phần>, <ngày>): Trả về giá trị của <thành phần> trong <ngày>

Các thành phần thông dụng của ngày gồm có:

TT	Thành phần	Viết tắt	Ý nghĩa
1	Year	yy, yyyy	Năm
2	Quarter	qq, q	Quý
3	Month	mm, m	Tháng
4	Dayofyear	dy, y	Ngày của năm (1 → 366)
5	Day	dd, d	Ngày của tháng (1 → 31)
6	Week	wk, ww	Tuần trong năm (1 → 52)
7	WeekDay	wd	Ngày trong năm (1 → 7)
8	Hour	hh	Giờ (1→ 24)

Ví dụ: Select DATEPART (year, GETDATE()) ⇒ 2006

Select DATEPART(week, '20 Aug, 2006') ⇒ 34

DATEADD (<thành phần>, <số n>, <ngày>): Cộng vào <thành phần> của <ngày> một số n.

Ví dụ: Select DATEADD (week, 5, '20 Aug, 2006') ⇒ 2006-09-24

Cho biết 10 ngày sau kể từ ngày 05/08/2006 là ngày thứ mấy:

Select DATEPART (weekday, DATEADD (day, 10, '5 Aug, 2006')) ⇒ 3 (thứ ba)

DATEDIFF (<thành phần> , <ngày bắt đầu>, <ngày kết thúc>): Trả về số khác biệt giữa 2 thành phần của ngày bắt đầu và ngày kết thúc.

Ví dụ: Cho biết từ ngày 05/08/2006 đến ngày 11/9/2006 có bao nhiêu ngày:

Select DATEDIFF (day, '5 Aug, 2006', '11 Sep, 2006') ⇒ 37 ngày

Select DATEDIFF (hour, '5 Aug, 2006', '11 Sep, 2006') ⇒ 888 giờ

DATENAME(<thành phần>, <ngày>): Trả về tên của <thành phần> trong <ngày>.

Ví dụ: Select DATENAME (month, '5 Aug, 2006') ⇒ August

Select DATENAME (weekday, '5 Aug, 2006') ⇒ Saturday

3. Các hàm xử lý số

FLOOR(<số>): Cho kết quả là số nguyên gần nhất nhỉn hơn hoặc bằng <số>

Ví dụ: SELECT FLOOR(123.45) ⇒ 123

FLOOR(-123.45) ⇒ -124

FLOOR(\$123.45) ⇒ 123.0000

CEILING(<số>): Cho kết quả là số nguyên gần nhất lớn hơn hoặc bằng <số>

Ví dụ: SELECT CEILING(123.45) ⇒ 124

CEILING(-123.45)	⇒	-123
CEILING(\$123.45)	⇒	124.0000

4. Hàm chuyển đổi kiểu

CAST (<dữ liệu> AS <kiểu>): Trả về <dữ liệu> với <kiểu> mới.

Ví dụ: SELECT 10 + cast ('34.5' as float) ⇒ 44.5

SELECT right(CAST(124 AS char(4)) , 2) ⇒ 4

5. Cấu trúc phân nhánh CASE

Đánh giá danh sách các điều kiện và trả về biểu thức phù hợp. CASE có hai dạng:

Dạng đơn giản: So sánh hai biểu thức để xác định kết quả trả về.

CASE <Tên cột>

WHEN <Giá trị 1> THEN <Kết quả 1>

[**WHEN <Giá trị 2> THEN <Kết quả 2> ...]**]

[

ELSE <Kết quả ngoại lệ>

]

END

Dạng tìm kiếm: Đánh giá các điều kiện để xác định kết quả trả về.

CASE

WHEN <Biểu thức điều kiện 1> THEN <Kết quả trả về 1>

[**WHEN <Biểu thức điều kiện 2> THEN <Kết quả trả về 2> ...]**]

[

ELSE <Kết quả trả về ngoại lệ>

]

END

Ví dụ: Select CASE

When 10 > 2 THEN 'DUNG'

When 10 = 3 THEN 'SAI'

END

⇒ DUNG

V. CÁC LỖI THƯỜNG GẶP TRONG SQL

3.1. Tên cột không tìm thấy trong bảng

Thông báo lỗi: Invalid column name <tên cột>

Khi gặp lỗi này hãy xem lại tên cột (còn gọi là tên thuộc tính/ tên trường) trong câu lệnh đã đúng chưa.

3.2. Tên bảng không tìm thấy

Thông báo lỗi: Invalid object name <tên bảng>

Khi gặp lỗi này hãy xem lại tên bảng trong câu lệnh đã đúng chưa.

3.3. Lỗi cú pháp

Thông báo lỗi: *Incorrect Syntax near <tù gân lỗi nhất>*

Khi gặp lỗi này hãy xem lại cú pháp trong câu lệnh đã đúng chưa, có thể sai từ khóa hoặc thiếu dấu ngoặc của hàm hoặc của select con, ...

3.4. Tên cột có mặt ở nhiều bảng

Thông báo lỗi: *Ambiguous column name <tên cột>*

Khi gặp lỗi này hãy xem lại <tên cột> gây lỗi trong câu lệnh đã có đặt tên bảng hoặc bí danh đúng trước chua.

Cú pháp: <tên bảng hoặc bí danh>. <tên cột>

3.5. Phép so sánh không tương thích kiểu

Thông báo lỗi: *Error converting data type ...*

hoặc *Syntax error converting ...*

Khi gặp lỗi này hãy xem lại trong câu lệnh đã có phép toán so sánh nào đó không tương thích kiểu dữ liệu.

Chương 4: RÀNG BUỘC TOÀN VẸN và CÁC LOẠI PHỤ THUỘC DỮ LIỆU

4.1. Ràng buộc toàn vẹn (RBTV)

4.1.1. Khái niệm RBTV

Trong CSDL luôn tồn tại nhiều mối liên hệ giữa các thuộc tính, các bộ, các bảng với nhau. Các mối liên hệ này là những điều kiện bất biến mà tất cả các bộ của những quan hệ có liên quan đều phải thỏa mãn tại mọi thời điểm. Những điều kiện bất biến đó được gọi là RBTV.

RBTV (Integrity constraint) viết tắt tiếng việt là RBTV, là một điều kiện được định nghĩa liên quan đến một hay nhiều quan hệ khác nhau. Các mối liên hệ ràng buộc là những điều kiện bất biến mà mọi thể hiện của quan hệ đều phải thỏa mãn trong mọi thời điểm.

Trong thực tế RBTV là các quy tắc quản lý được áp đặt trên các đối tượng của thế giới thực. Mục đích của RBTV là bảo đảm tính nhất quán của dữ liệu, bảo đảm rằng dữ liệu luôn biểu diễn đúng ngữ nghĩa trong thực tế.

Nhiệm vụ của người phân tích thiết kế là phải phát hiện càng đầy đủ các RBTV càng tốt và mô tả chúng một cách chính xác trong hồ sơ phân tích thiết kế. RBTV được xem như là một công cụ để diễn đạt ngữ nghĩa của CSDL. Một CSDL được thiết kế cồng kềnh nhưng nó thể hiện được đầy đủ ngữ nghĩa của thực tế vẫn có giá trị cao hơn rất nhiều so với một cách thiết kế gọn nhẹ nhưng nghèo nàn về ngữ nghĩa vì thiếu các RBTV của CSDL.

Công việc kiểm tra RBTV thường được tiến hành vào thời điểm cập nhật dữ liệu. Những RBTV phát sinh cần phải được ghi nhận và xử lý một cách tường minh.

RBTV và kiểm tra sự vi phạm RBTV là hai trong số những vấn đề quan trọng trong quá trình phân tích thiết kế cơ sở dữ liệu, nếu không quan tâm đúng mức đến những vấn đề trên, thì có thể dẫn đến những hậu quả nghiêm trọng về tính an toàn và toàn vẹn dữ liệu, đặc biệt là đối với những cơ sở dữ liệu lớn.

4.1.2. Các yếu tố của RBTV

Mỗi RBTV có bốn yếu tố: bối cảnh, điều kiện, bảng tam ảnh hướng và hành động phải cần thực hiện khi phát hiện có RBTV bị vi phạm:

a. Bối cảnh

Bối cảnh của RBTV 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 RBTV có thể là một hoặc nhiều quan hệ. Chẳng hạn với RBTV R trên thì bối cảnh của nó là quan hệ Sinhvien.

b. Biểu diễn

Biểu diễn RBTV là sự mô tả, và biểu diễn hình thức nội dung của nó.

Biểu diễn một RBTV 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 RBTV cũng có thể được biểu diễn bằng phụ thuộc hàm.

Ví dụ: Một số RBTV 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.

c. **Bảng tầm ảnh hưởng**

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

Thời điểm cần phải kiểm tra RBTV chính là thời điểm cập nhật dữ liệu.

Một bảng tầm ảnh hưởng của một RBTV có dạng sau:

Tên RBTV	Thêm	Xóa	Sửa
r ₁	+	-	+(A ₁)
r ₂	+	-	+(A ₂)
r ₃	+	-	+(A ₃)
...
r _n	+	-	+(A _n)

Trong đó, dấu + thể hiện thao tác có thể gây ra vi phạm RBTV. Trong trường hợp +(A_i) cho biết thao tác sửa có thể gây vi phạm trên thuộc tính A_i. Dấu - thể hiện thao tác không vi phạm RBTV.

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

Khi một RBTV 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?

4.2. Phân loại RBTV

Trong quá trình phân tích thiết kế CSDL, yêu cầu cần thiết là phải tìm được những RBTV tiềm ẩn trong CSDL. Việc phân loại RBTV cho phép người phân tích tìm kiếm đầy đủ, tránh bỏ sót những RBTV. Các loại RBTV được phân thành hai dạng chính:

- RBTV có bối cảnh là một quan hệ
- RBTV có bối cảnh là nhiều quan hệ

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

RBTV có bối cảnh là một quan hệ được chia thành ba loại: RBTV miền giá trị, RBTV liên bộ và RBTV liên thuộc tính.

a. Ràng buộc toàn vẹn miền giá trị

Quy định rõ về miền giá trị của một thuộc tính.

Ví dụ: Thời gian phân công tham gia đề án của một nhân viên không quá 40h/tuần

- Bối cảnh: Quan hệ PHANCONG
- Biểu diễn: $\forall pc \in PHANCONG (pc.ThoiGian \leq 40)$
- Bảng tầm ảnh hưởng:

RB1	Thêm	Xóa	Sửa
PHANCONG	+	-	+ (ThoiGian)

Ví dụ: Điểm của môn học phải là thang điểm 10

- Bối cảnh: Quan hệ KETQUA (MaMH, MaLop, MaSV, Diem)
- Biểu diễn: $\forall kq \in KETQUA (kq.Diem \geq 0 \wedge kq.Diem \leq 10)$
- Bảng tầm ảnh hưởng:

RB2	Thêm	Xóa	Sửa
KETQUA	+	-	+ (Diem)

b. Ràng buộc toàn vẹn liên thuộc tính

Quy định các ràng buộc giữa các thuộc tính khác nhau trong cùng một quan hệ.

Ví dụ: Ngày trả sách phải bằng hoặc sau ngày mượn sách.

- Bối cảnh: Quan hệ MUONSACH (MaSach, MaDocGia, NgayMuon, NgayHenTra, NgayThucTra)
- Biểu diễn: $\forall ms \in MUONSACH (ms.NgayMuon \leq ms.NgayHenTra \wedge ms.NgayMuon \leq ms.NgayThucTra)$
- Bảng tầm ảnh hưởng:

RB3	Thêm	Xóa	Sửa

MUONSACH	+	-	+ (NgayMuon, NgayHenTra, NgayThucTra)
----------	---	---	---------------------------------------

c. Ràng buộc toàn vẹn liên bộ

Quy định sự tồn tại của một hoặc nhiều bộ phụ thuộc vào sự tồn tại của một hoặc nhiều bộ khác trong cùng quan hệ.

- ♦ RBTV khóa chính là RBTV liên bộ

Ví dụ: Mỗi đề án trong công ty có một mã duy nhất để phân biệt với các đề án khác.

- Bối cảnh: Quan hệ DEAN
- Điều kiện: $\forall da1, da2 \in DEAN : da1 \neq da2 \Rightarrow (da1.MaDA \neq da2.MaDA)$
- Bảng tầm ảnh hưởng:

RB4	Thêm	Xóa	Sửa
DEAN	+	-	+ (MaDA)

- ♦ RBTV về số lượng các bộ trong một quan hệ

Ví dụ: Mỗi sinh viên trong một học kỳ được đăng ký không quá 8 môn học.

- Bối cảnh: Quan hệ DANGKY(MaSV, MaMH)
- Biểu diễn: $\forall dk1 \in DANGKY :$

$\text{count}(dk2 \in DANGKY | dk2.MaSV = dk1.MaSV) \leq 8,$

$\text{count}()$ là hàm đếm số bộ của một quan hệ thỏa điều kiện trong ngoặc () .

- Bảng tầm ảnh hưởng:

RB5	Thêm	Xóa	Sửa
DANGKY	+	-	+ (MaSV)

4.2.2. Ràng buộc toàn vẹn có bối cảnh là nhiều quan hệ

RBTV có bối cảnh là nhiều quan hệ được chia thành năm loại:

- RBTV tham chiếu
- RBTV liên bộ - liên quan hệ
- RBTV liên thuộc tính - liên quan hệ
- RBTV do thuộc tính tổng hợp
- RBTV do chu trình

a. RBTV tham chiếu

Quy định giá trị xuất hiện của một tập thuộc tính trong một quan hệ phải xuất hiện trong một tập thuộc tính trong một quan hệ khác. RBTV này còn được gọi là RBTV tham chiếu, RBTV phụ thuộc tồn tại hay RBTV khóa ngoại.

Ví dụ: RBTV trên 2 quan hệ

Một nhân viên phải thuộc về một phòng trong công ty, nghĩa là trong quan hệ NHANVIEN, nếu một mã phòng (Phong) mà nhân viên trực thuộc xuất hiện, thì mã phòng này phải xuất hiện trong quan hệ PHONGBAN, cụ thể là thuộc tính (MaPhong). Như vậy:

- Bối cảnh: NHANVIEN, PHONGBAN
- Biểu diễn: $\forall nv \in NHANVIEN ((nv.Phong = NULL) \vee (\exists pb \in PHONGBAN (nv.Phong = pb.MaPhong)))$
- Bảng tầm ảnh hưởng:

RB6	Thêm	Xóa	Sửa
NHANVIEN	+	-	+(Phong)
PHONGBAN	-	+	+(MaPhong)

Ví dụ: RBTB trên 1 quan hệ

Người quản lý (MaNQL) của một nhân viên cũng phải là một nhân viên trong công ty

- Bối cảnh: NHANVIEN
- Điều kiện: $\forall nv \in NHANVIEN ((nv.MaNQL = NULL) \vee (\exists nv1 \in NHANVIEN (nv.MaNQL = nv1.MaNV)))$
- Bảng tầm ảnh hưởng:

RB7	Thêm	Xóa	Sửa
NHANVIEN	+	-	+(MaNV, MaNQL)

b. RBTB liên bộ - liên quan hệ

Quy định về từng từng nhóm các bộ của nhiều quan hệ bối cảnh khác nhau.

Ví dụ: Một hóa đơn bán hàng phải có ít nhất một mặt hàng, nghĩa là một chi tiết hóa đơn bán hàng phải có ít nhất một mặt hàng.

- Bối cảnh: HOADON, CTIETHD
- Biểu diễn: $\forall hd \in HOADON (\exists cthd \in CTIETHD(hd.MaHD = cthd.MaHD))$
- Bảng tầm ảnh hưởng:

RB8	Thêm	Xóa	Sửa
HOADON	+	-	+(MaHD)
CTIETHD	-	+	+(MaHD)

c. RBTB liên thuộc tính - liên quan hệ

Quy định về mối liên hệ giữa các thuộc tính trên nhiều quan hệ bối cảnh khác nhau.

Ví dụ: Giả sử cho phép thanh toán tiền nhiều lần và thanh toán sau khi mua hàng, khi đó ngày thanh toán tiền theo một hóa đơn mua hàng phải bằng hoặc sau ngày mua hàng.

- Bối cảnh: HOADON (MaHD, MaKH, NgayHD, TriGia)
THANHTOAN (MaHD, NgayTT, LanTT, SoTienTT)
- Biểu diễn: $\forall \text{hd} \in \text{HOADON} (\exists \text{tt} \in \text{THANHTOAN} (\text{hd.MaHD} = \text{tt.MaHD} \Rightarrow \text{hd.NgayHD} \leq \text{tt.NgayTT}))$
- Bảng tầm ảnh hưởng:

RB9	Thêm	Xóa	Sửa
HOADON	+	-	+(MaHD, NgayHD)
THANHTOAN	+	-	+(MaHD, NgayTT)

d. RBT do thuộc tính tổng hợp

Quy định về mối liên hệ giữa các thuộc tính do sự có mặt của thuộc tính tính toán.

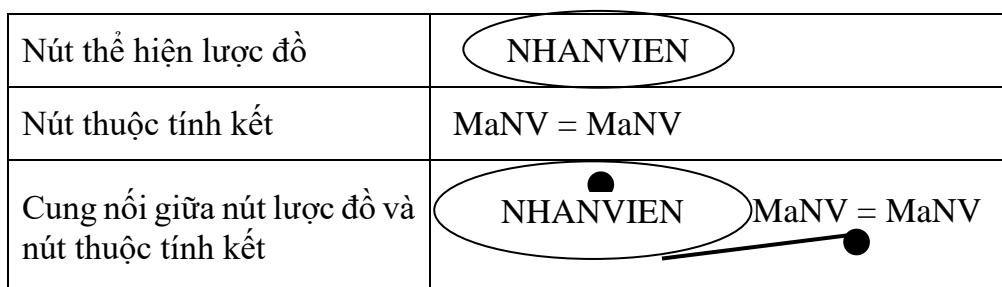
Ví dụ: Điểm trung bình của sinh viên bằng trung bình của các môn mà sinh viên theo học.

- Bối cảnh: SINHVIEN (MaSV, HoSV, TenSV, Khoa, DTB)
KETQUA (MaSV, MaMon, Diem)
- Biểu diễn: $\forall \text{sv} \in \text{SINHVIEN} (\exists \text{kq} \in \text{KETQUA} (\text{sv.MaSV} = \text{kq.MaSV} \Rightarrow \text{sv.DTB} = \text{AVG}(\text{kq.Diem})))$
- Bảng tầm ảnh hưởng:

RB10	Thêm	Xóa	Sửa
SINHVIEN	+	-	+(MaSV, DTB)
KETQUA	+	+	+(MaSV, Diem)

e. RBT do chu trình

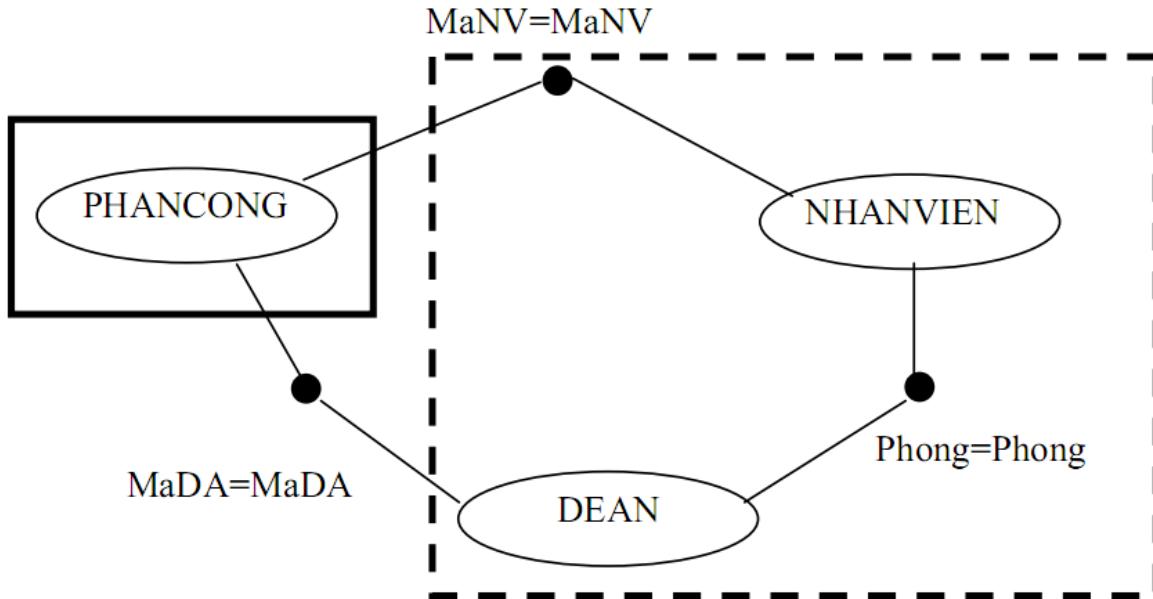
Xảy ra khi có sự hiện diện của chu trình. Để nhận diện chu trình, người ta biểu diễn lược đồ CSDL như sau:



Ví dụ: Một nhân viên chỉ được phân công vào các đề án do phòng mình chủ trì.

- Bối cảnh: NHANVIEN, DEAN, PHANCONG

Đồ thị thể hiện chu trình như sau:



- Biểu diễn: $\forall pc \in PHANCONG (\exists nvda \in NV_DA (nvda.MaNV = pc.MaNV \wedge nvda.MaDA = pc.MaDA))$
với $NV_DA \leftarrow NHANVIEN \bowtie Phong = MaPhong DEAN$
- Bảng tầm ảnh hưởng:

RB11	Thêm	Xóa	Sửa
NHANVIEN	-	+	+(MaNV, Phong)
DEAN	-	+	+(MaDA, Phong)
PHANCONG	+	-	+(MaDA, MaNV)

4.3. Các loại phụ thuộc dữ liệu

4.3.1. Các vấn đề gặp phải khi tổ chức dữ liệu

Trước khi bàn về cách thiết kế một cơ sở dữ liệu tốt, chúng ta hãy phân tích xem tại sao trong một số lược đồ quan hệ lại tồn tại những vấn đề rắc rối. Chẳng hạn cho lược đồ quan hệ Thi (MASV, HOTEN, MONHOC, DIEMTHI), và sau đây là một quan hệ trên lược đồ quan hệ KQThi.

MASV	HOTEN	MONHOC	DIEMTHI
00CDTH189	Nguyễn Văn Thành	Câu Trúc Dữ Liệu	7
00CDTH189	Nguyễn Văn Thành	Cơ Sở Dữ Liệu	9
00CDTH211	Trần Thu Hà	Kỹ Thuật Lập Trình	5
00CDTH189	Nguyễn Văn Thành	Kỹ Thuật Lập Trình	8

Quan hệ này ghi kết quả điểm thi các môn của các sinh viên. Chúng ta có thể nhận thấy một số vấn đề sau:

1. Dư thừa (redundancy): Họ tên của các sinh viên được lặp lại mỗi lần cho mỗi môn thi.
2. Mâu thuẫn tiềm ẩn (potentia inconsistancy) hay bất thường khi cập nhật. Do hậu quả của dư thừa, chúng ta có thể cập nhật họ tên của một sinh viên trong một bộ nào đó nhưng vẫn để lại họ tên cũ trong những bộ khác. Vì vậy, chúng ta có thể không có một họ tên duy nhất đối với mỗi sinh viên như chúng ta mong muốn.
3. Bất thường khi chèn (insertion anomaly). Chúng ta không thể biết họ tên của một sinh viên nếu hiện tại sinh viên đó không dự thi môn nào.
4. Bất thường khi xoá (deletion anomaly). Chúng ta có thể xoá tất cả các môn thi của một sinh viên, vô ý làm mất dấu vết để tìm ra họ tên của sinh viên này.

Những vấn đề nêu trên sẽ được giải quyết nếu chúng ta phân rã lược đồ quan hệ KQThi thành hai lược đồ quan hệ:

Sinhvien (MASV, HOTEN)

Ketqua (MASV, MONHOC, DIEMTHI)

Lúc này lược đồ quan hệ Sinhvien cho biết họ tên của mỗi sinh viên chỉ xuất hiện đúng một lần, do vậy không có dư thừa. Ngoài ra, chúng ta cũng có thể nhập họ tên của một sinh viên dù hiện tại sinh viên đó chưa có kết quả thi môn nào. Tuy nhiên lúc này ta nhận thấy rằng để tìm danh sách họ tên của các sinh viên ứng với môn thi cơ sở dữ liệu thì chúng ta phải thực hiện một phép kết nối, còn với một quan hệ duy nhất Thi chúng ta có thể dễ dàng trả lời bằng cách thực hiện một phép chọn và một phép chiếu. Làm sao để đưa được một lược đồ cơ sở dữ liệu chưa tốt về một lược đồ cơ sở dữ liệu tốt hơn? chương này và chương sau nhằm giải quyết vấn đề này.

4.3.2. Phụ thuộc hàm

Phụ thuộc hàm (functional dependency) dùng để biểu diễn một cách hình thức các ràng buộc toàn vẹn (RBTV). Phụ thuộc hàm có tầm quan trọng rất lớn trong việc giải quyết các bài toán tìm khóa, phủ tối thiểu và chuẩn hóa cơ sở dữ liệu.

a. Định nghĩa phụ thuộc hàm

Xét quan hệ DEAN(MaDA, TenDA, DdiemDA, Phong). Nhận thấy rằng quan hệ DEAN có MaDA là khóa, nghĩa là từ MaDA có thể xác định được tất cả các thông tin về tên đề án, địa điểm thực hiện đề án và phòng ban chủ trì đề án. Có thể phát biểu lại như sau:

- MaDA xác định TenDA hay TenDA phụ thuộc hàm vào MaDA
- MaDA xác định DdiemDA hay DdiemDA phụ thuộc hàm vào MaDA
- MaDA xác định Phong hay Phong phụ thuộc hàm vào MaDA được ký hiệu:
 - MaDA → TenDA
 - MaDA → DdiemDA
 - MaDA → Phong

Một cách tổng quát, ta có:

$Q(A_1, A_2, \dots, A_n)$ là lược đồ quan hệ

$Q^+ = \{A1, A2, \dots, An\}$

X, Y là hai tập con của Q^+

t_1, t_2 là hai bộ bất kỳ của Q . Khi đó: $X \rightarrow Y \Leftrightarrow (t_1.X = t_2.X \Rightarrow t_1.Y = t_2.Y)$

Ta nói X xác định Y hay Y phụ thuộc hàm vào X . X được gọi là vé trái phụ thuộc hàm, Y được gọi là vé phải phụ thuộc hàm.

Phụ thuộc hàm hiển nhiên

$X \rightarrow Y$ gọi là phụ thuộc hàm hiển nhiên nếu $Y \subseteq X$

Phụ thuộc hàm nguyên tố

$X \rightarrow Y$ được gọi là phụ thuộc hàm nguyên tố (hoặc nói cách khác Y được gọi là phụ thuộc đầy đủ vào X) nếu $\forall X' \subsetneq X$ đều không có phụ thuộc hàm $X' \rightarrow Y$.

Như vậy, các phụ thuộc hàm MaDA \rightarrow TenDA, MaDA \rightarrow DdiemDA, MaDA \rightarrow Phong là phụ thuộc hàm nguyên tố.

Xét quan hệ CTIETHOADON (SoHD, MaHang, SoLuong, DonGia, ThanhTien) và các phụ thuộc hàm như sau:

- $\{SoHD, MaHang\} \rightarrow SoLuong$
- $\{SoHD, MaHang\} \rightarrow DonGia$
- $\{SoHD, MaHang\} \rightarrow ThanhTien$

Nhận thấy rằng SoLuong phụ thuộc đầy đủ vào $\{SoHD, MaHang\}$ nhưng DonGia chỉ phụ thuộc vào MaHang (là một thuộc tính khóa) chứ không phụ thuộc đầy đủ vào khóa $\{SoHD, MaHang\}$. Như vậy, trên một lược đồ quan hệ có thể tồn tại nhiều phụ thuộc hàm.

Tập các phụ thuộc hàm được ký hiệu F.

Ví dụ: Cho lược đồ quan hệ Q(ABCDE), r là quan hệ xác định trên Q như sau:

A	B	C	D	E
a1	b1	c1	d1	e1
a1	b2	c2	d2	e1
a2	b1	c3	d3	e1
a2	b1	c4	d3	e1
a3	b2	c3	d1	e1

Trong các phụ thuộc hàm $A \rightarrow D$; $AB \rightarrow D$; $E \rightarrow A$; $A \rightarrow E$ thì phụ thuộc hàm $AB \rightarrow D$; $A \rightarrow E$ thoả r.

b. 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 dùng hệ tiên đề Armstrong (1974), gồm các luật sau:

Với $X, Y, Z, W \subseteq Q^+$. Phụ thuộc hàm có các tính chất sau:

1. Luật phản xạ (reflexivity)

$$X \supseteq Y \Rightarrow X \rightarrow Y$$

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 kết hợp (the union rule)

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

5. Luật tựa bắc cầu (the pseudotransitivity rule)

$$\text{Cho } X \rightarrow Y, WY \rightarrow Z \Rightarrow XW \rightarrow Z$$

6. Luật phân rã (the decomposition rule)

$$\text{Cho } X \rightarrow Y, Z \subseteq Y \Rightarrow X \rightarrow Z$$

Ví dụ: Cho lược đồ quan hệ R(A, B, C, D, E, G, H) và tập phụ thuộc hàm

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

Chứng tỏ $AB \rightarrow E$ được suy diễn từ F.

(1) $AB \rightarrow C$

(2) $AB \rightarrow AB$ (tính phản xạ)

(3) $AB \rightarrow B$ (tính phân rã)

(4) $B \rightarrow D$

(5) $AB \rightarrow D$ (tính bắc cầu 3+4)

(6) $AB \rightarrow CD$ (tính hợp 1+5)

(7) $CD \rightarrow E$

(8) $AB \rightarrow E$ (tính bắc cầu 6+7)

c. **Bao đóng của tập phụ thuộc hàm và bao đóng của tập thuộc tính**

- Bao đóng của tập phụ thuộc hàm F

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 lược đồ quan hệ Q (A, B, C, D, E, G, H) và F được cho như sau:

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

$$\text{Khi đó } F^+ = \{B \rightarrow A; DA \rightarrow CE; D \rightarrow H; GH \rightarrow C; AC \rightarrow D;$$

$$BC \rightarrow AC; BC \rightarrow D; DA \rightarrow AH; DG \rightarrow C; BC \rightarrow AD; \dots\}$$

- Bao đóng của tập thuộc tính X

Cho lược đồ quan hệ Q. Giả sử F là tập các phụ thuộc hàm trong Q, $X \subseteq Q^+$.

Bao đóng của tập thuộc tính X đối với F ký hiệu là X^+ (hoặc X_F^+) là tập tất cả các thuộc tính $A \in Q^+$ được suy ra từ X dựa vào các phụ thuộc hàm trong F và hệ tiên đề Armstrong, nghĩa là: $X^+ = \{A : A \in Q^+ \text{ và } X \rightarrow A \in F^+\}$.

Thuật toán tìm bao đóng của tập thuộc tính

Input: (Q, F) , $X \subseteq Q^+$

Bước 1: đặt $X_0 = X$

Bước 2 : Lặp lại việc tìm X_i

- $X_{i+1} = X_i \cup Z$, $\exists(Y \rightarrow Z) \in F (Y \subseteq X_i)$, loại $(Y \rightarrow Z)$ ra khỏi F
- Nếu $X_{i+1} = X_i$ hoặc $X_i = Q^+$ chuyển sang bước 3

Bước 3: Dừng và kết luận $X^+ = X_i$

Output: X^+

Ví dụ 1: Cho F là tập các phụ thuộc hàm bao gồm:

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

Giải

Bước 1: khởi tạo $X_0 = BD$

Bước 2: Lặp lại việc tính các X_i

X_1 , ta tìm các phụ thuộc hàm có phía bên trái là B, D hoặc BD, ta thấy chỉ có một phụ thuộc hàm $D \rightarrow EG$ nên ta thêm E và G vào X_0 để có: $X_1 = BDEG$

X_2 , ta tìm các phía bên trái nằm trong X_1 . Đó là $D \rightarrow EG$ và $BE \rightarrow C$

Như vậy: $X_2 = BCDEG$

X_3 , ta tìm thấy các phụ thuộc hàm $C \rightarrow A, BC \rightarrow D, CE \rightarrow AG$

Nên $X_3 = ABCDEG = Q^+$

Bước 3: Kết luận $BD^+ = ABCDEG$

Ví dụ 2: Cho lược đồ quan hệ R(A, B, C, D, E, G, H) và tập phụ thuộc hàm

$F = \{f_1: B \rightarrow A, f_2: DA \rightarrow CE, f_3: D \rightarrow H, f_4: GH \rightarrow C, f_5: AC \rightarrow D\}$

Tìm AC^+

Giải

Bước 1: $X_0 = AC$

Bước 2: Từ f_1 đến f_4 không thoả, f_5 thoả nên $X_1 = AC \cup D = ACD$

Lặp lại bước 2:

f_1 không thoả, f_2 thoả nên $X_2 = ACD \cup CE = ACDE$

f_3 thoả nên $X_3 = ACDE \cup H = ACDEH$

f_4 không thỏa, f_5 đã thỏa.

Lặp lại bước 2:

f_2, f_3 và f_5 đã thỏa, f_1 và f_4 không thỏa. Nên $X_4 = X_3 = ACDEH$

Vậy: $AC^+ = ACDEH$

Ví dụ 3: Cho lược đồ quan hệ $Q(A, B, C, D, E, G, H)$ và tập phụ thuộc hàm F

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

Hãy tính: B^+ ; H^+ ; BC^+

Giải

Ta có: $B^+ = BA$ (do có phụ thuộc hàm $B \rightarrow A$)

$H^+ = H$ (do có phụ thuộc hàm $H \rightarrow H$)

$BC^+ = BCADEH$ (do có các phụ thuộc hàm $B \rightarrow A$; $AC \rightarrow D$;

$DA \rightarrow CE$; $D \rightarrow H$)

4.3.3. Phủ tối thiểu (minimal cover)

a. Hai tập phụ thuộc hàm tương đương

Cho F và G là hai tập phụ thuộc hàm, F và G tương đương nếu $F^+ = G^+$ (mỗi phụ thuộc hàm thuộc F đều thuộc G^+ và mỗi phụ thuộc hàm thuộc G đều thuộc F^+).

Ký hiệu: $F \equiv G$ (F phủ G hoặc G phủ F)

F được gọi là phủ G nếu $F^+ \supseteq G^+$

♦ Thuật toán xác định F và G có tương đương

Bước 1:

Với mỗi phụ thuộc hàm $X \rightarrow Y \in F$, kiểm tra $X \rightarrow Y$ có là thành viên của G không.

Bước 2:

Với mỗi phụ thuộc hàm $X \rightarrow Y \in G$, kiểm tra $X \rightarrow Y$ có là thành viên của F không?

Bước 3:

Nếu cả bước 1 và 2 đều đúng thì kết luận $F \equiv G$

Ví dụ: Cho lược đồ quan hệ $R(A, B, C, D, E)$ và tập phụ thuộc hàm

$$F = \{A \rightarrow BC; A \rightarrow D; CD \rightarrow E\}$$

$$G = \{A \rightarrow BCE; A \rightarrow ABD; CD \rightarrow E\}$$

Cho biết F và G có tương đương không?

Giải

Bước 1:

$$A^+ = ABCDE \Rightarrow (A \rightarrow BC \in G^+ \wedge A \rightarrow D \in G^+)$$

$$\text{Hơn nữa ta có } (CD \rightarrow E \in F \wedge CD \rightarrow E \in G)$$

Vậy mọi phụ thuộc hàm trong F đều là thành viên của G

Bước 2:

$$A^+ = ABCDE \Rightarrow (A \rightarrow BCE \in F^+ \wedge A \rightarrow ABD \in F^+)$$

Hơn nữa ta có $(CD \rightarrow E \in G \wedge CD \rightarrow E \in F)$

Vậy mọi phụ thuộc hàm trong G đều là thành viên của F

Bước 3: Kết luận $F \equiv G$

b. Phủ tối thiểu của một tập phụ thuộc hàm

Bổ đề: Mỗi tập các phụ thuộc hàm F đều được phủ bởi tập các phụ thuộc hàm G mà về phái của các phụ thuộc hàm G chỉ gồm một thuộc tính.

Định nghĩa

F được gọi là một tập phụ thuộc hàm tối thiểu nếu F thoả đồng thời ba điều kiện sau:

Điều kiện a : Vé phái của F chỉ có một thuộc tính.

Điều kiện b : Không $\exists f: X \rightarrow A \in F$ và $Z \subset X$ mà:

$$F^+ = (F - (X \rightarrow A)) \cup (Z \rightarrow A)^+$$

Điều kiện c : Không $\exists X \rightarrow A \in F$ mà:

$$F^+ = (F - (X \rightarrow A))^+$$

Trong đó,

Vé phái của mỗi phụ thuộc hàm ở điều kiện a chỉ có một thuộc tính, nên bảo đảm không có thuộc tính nào ở vé phái là dư thừa.

Điều kiện b bảo đảm không có một thuộc tính nào tham gia vé trái của phụ thuộc hàm là dư thừa.

Điều kiện c bảo đảm cho tập F không có một phụ thuộc hàm nào là dư thừa.

Chú ý rằng, một tập phụ thuộc hàm luôn tìm ra ít nhất một phủ tối thiểu và nếu thứ tự các phụ thuộc hàm trong tập F là khác nhau thì có thể sẽ thu được những phủ tối thiểu khác nhau.

♦ **Thuật toán tìm phủ tối thiểu của tập phụ thuộc hàm**

Bước 1: Loại các thuộc tính có vé trái dư thừa của mọi phụ thuộc hàm.

Bước 2: Phân rã các phụ thuộc hàm có vé phái nhiều thuộc tính thành các phụ thuộc hàm có vé phái một thuộc tính.

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

Ví dụ: Cho lược đồ quan hệ R(A, B, C, D, E, G, H) và tập phụ thuộc hàm

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

Tìm phủ tối thiểu của F.

Giải

Bước 1:

- Với $DA \rightarrow CE$:

Giả sử D thừa thì $A \rightarrow CE \in F^+$: ta có $A^+ = A$, $CE \notin A^+$ nên D không thừa

Giả sử A thừa thì $D \rightarrow CE \in F^+$: ta có $D^+ = DH$, $CE \notin D^+$ nên A không thừa

• Với $GH \rightarrow C$:

Giả sử G thừa thì $H \rightarrow C \in F^+$: ta có $H^+ = H$, $C \notin H^+$ nên G không thừa

Giả sử H thừa thì $G \rightarrow C \in F^+$: ta có $G^+ = G$, $C \notin G^+$ nên H không thừa

• Với $AC \rightarrow D$:

Giả sử A thừa thì $C \rightarrow D \in F^+$: ta có $C^+ = C$, $D \notin C^+$ nên A không thừa

Giả sử C thừa thì $A \rightarrow D \in F^+$: ta có $A^+ = A$, $D \notin A^+$ nên C không thừa

Vậy mọi phụ thuộc hàm đều là đầy đủ.

Bước 2: Phân rã vé phải ta có $F = \{B \rightarrow A, DA \rightarrow C, DA \rightarrow E, D \rightarrow H, GH \rightarrow C, AC \rightarrow D\}$

Bước 3: Loại các phụ thuộc hàm dư thừa

• Với $B \rightarrow A$:

Ta có: $B^+ \setminus \{B \rightarrow A\} = B$, $A \notin B^+ \setminus \{B \rightarrow A\}$ nên $B \rightarrow A$ là không thừa

• Với $DA \rightarrow C$:

Ta có: $DA^+ \setminus \{DA \rightarrow C\} = DAEH$, $C \notin DA^+ \setminus \{DA \rightarrow C\}$ nên $DA \rightarrow C$ là không thừa

• Với $DA \rightarrow E$:

Ta có: $DA^+ \setminus \{DA \rightarrow E\} = DACH$, $E \notin DA^+ \setminus \{DA \rightarrow E\}$ nên $DA \rightarrow E$ là không thừa

• Với $D \rightarrow H$:

Ta có: $D^+ \setminus \{D \rightarrow H\} = D$, $H \notin D^+ \setminus \{D \rightarrow H\}$ nên $D \rightarrow H$ là không thừa

• Với $GH \rightarrow C$:

Ta có: $GH^+ \setminus \{GH \rightarrow C\} = GH$, $C \notin GH^+ \setminus \{GH \rightarrow C\}$ nên $GH \rightarrow C$ là không thừa

• Với $AC \rightarrow D$:

Ta có: $AC^+ \setminus \{AC \rightarrow D\} = AC$, $D \notin AC^+ \setminus \{AC \rightarrow D\}$ nên $AC \rightarrow D$ là không thừa

Vậy phủ tối thiểu của F: $PTT(F) = \{B \rightarrow A; DA \rightarrow C; DA \rightarrow E; D \rightarrow H; GH \rightarrow C; AC \rightarrow D\}$

4.4. Phép tách các sơ đồ quan hệ

Phép tách một sơ đồ quan hệ $R(A_1, A_2, \dots, A_n)$ là sự thay thế R bởi các quan hệ R_1, R_2, \dots, R_k . Trong đó mỗi R_j ($1 \leq j \leq k$) có tập thuộc tính A_p, \dots, A_q (A_p, \dots, A_q : Các thuộc tính của R).

Mục tiêu của việc tách sơ đồ quan hệ là khắc phục nhược điểm của cơ sở dữ liệu. Công việc này thuộc về những thành viên trong nhóm thiết kế CSDL.

4.5. Kết nối không mất thông tin

Nếu R là sơ đồ quan hệ được tách thành các quan hệ R_1, R_2, \dots, R_k và F là một tập phụ thuộc hàm, ta nói phép tách có kết nối không mất thông tin nếu mọi quan hệ r của R thỏa mãn điều kiện $r \in F$

$$r = \Pi_{R_1}(r) * \Pi_{R_2}(r) * \dots * = \Pi_{R_k}(r)$$

Bố đề 4:

Cho R là một sơ đồ quan hệ, $p=\{ R_1, R_2, \dots, R_k \}$ là một phép tách của R , r là một quan hệ trên R , và $r_i = \Pi_{R_i}(r)$, gọi mp là ánh xạ định nghĩa bởi:

$$mp(r) = \bigtriangleup_{i=1}^k \Pi_{R_i}(r).$$

Ta có:

- (i) $r \subseteq mp(r)$
- (ii) Nếu $s = mp(r)$ thì $\Pi_{R_i}(s) = r_i$.
- (iii) $mp(mp(r)) = mp(r)$

4.6. Kiểm tra kết nối không mất thông tin

Giải thuật:

- Input: Một sơ đồ quan hệ $R(A_1, A_2, \dots, A_n)$, một tập phụ thuộc hàm F , và một phép tách $p=\{ R_1, R_2, \dots, R_k \}$
- Output: Quyết định xem có một phép tách với một kết nối không mất thông tin.
- Phương pháp:

Ta xây dựng một bảng k dòng và n cột, cột j tương ứng thuộc tính thứ j và dòng i ứng với sơ đồ quan hệ R_i . Tại vị trí giao của dòng i và cột j , đặt ký hiệu a_{ij} nếu a_j có trong R_i ; ký hiệu b_{ij} nếu ngược lại.

Lặp lại việc xét từng phụ thuộc hàm $X \rightarrow Y$ trong F cho đến khi không còn thay đổi được bảng. Mỗi lần xét $X \rightarrow Y$, ta tìm các dòng mang trị bằng nhau trên tập thuộc tính X . Nếu tìm được hai dòng như vậy, làm cho các ký hiệu trên hai dòng đó giống nhau tại các thuộc tính của Y : Nếu một ký hiệu là a_j , ta cho ký hiệu thứ hai cũng là a_j ; nếu chúng là b_{ij} , ta cho chúng cùng là b_{ij} hoặc cùng

Nếu sau khi sửa bảng theo cách như trên, ta thấy có dòng nào đó có dạng a_1, a_2, \dots, a_k thì kết nối không mất thông tin, ngược lại là mất thông tin.

Ví dụ 1:

Cho quan hệ $R(A, B, C, D, E, F)$.

Xét phép tách – kết nối $p = \{R_1, R_2, R_3\}$ trong đó:

$R_1(A, B, D, E)$,

$R_2(A, C, D, F)$,

$R_3(B, C, E, F)$,

Và phụ thuộc hàm $F = \{A \rightarrow B, F \rightarrow E\}$.

Bước 1: Thành lập bảng ban đầu gồm 4 hàng và 7 cột:

	A	B	C	D	E	F
R_1	a_1	a_2	b_{13}	a_4	a_5	b_{16}
R_2	a_1	b_{22}	a_3	a_4	b_{25}	a_6
R_3	b_{31}	a_2	a_3	b_{34}	a_5	a_6

Bước 2: Áp dụng $A \rightarrow B$ suy ra $b_{22} = a_2$

	A	B	C	D	E	F
R_1	a_1	a_2	b_{13}	a_4	a_5	b_{16}
R_2	a_1	a_2	a_3	a_4	b_{25}	a_6
R_3	b_{31}	a_2	a_3	b_{34}	a_5	a_6

Bước 3: Áp dụng $F \rightarrow E$ suy ra $b_{25} = a_5$

	A	B	C	D	E	F
R_1	a_1	a_2	b_{13}	a_4	a_5	b_{16}
R_2	a_1	a_2	a_3	a_4	a_5	a_6
R_3	b_{31}	a_2	a_3	b_{34}	a_5	a_6

Như vậy tồn tại hàng thứ 2 R_2 chứa các ký tự $\{a_1, a_2, a_3, a_4, a_5, a_6\}$. Suy ra phép tách p có kết nối không mất thông tin.

Ví dụ 2: Cho $R(A, B, C, D, E)$ và các phép tách.

$S_1(A, D)$

$S_2(A, B)$

$S_3(B, E)$

$S_4(C, D, E)$

Giáo trình Cơ sở dữ liệu

$S_5(A, E)$

và tập phụ thuộc hàm

$F(A \rightarrow C, B \rightarrow C, C \rightarrow D, E \rightarrow C, CE \rightarrow D)$

Bước 1: Bảng ban đầu gồm 6 hàng và 6 cột:

	A	B	C	D	E
S_1	a_1	b_{12}	b_{13}	a_4	b_{15}
S_2	a_1	a_2	b_{23}	b_{24}	b_{25}
S_3	b_{31}	a_2	b_{33}	b_{34}	a_5
S_4	b_{41}	b_{42}	a_3	a_4	a_5
S_5	a_1	b_{52}	b_{53}	b_{54}	a_5

Bước 2: Áp dụng $A \rightarrow C$ và $B \rightarrow C$ suy ra $b_{13} = b_{23} = b_{53} = b_{33}$

	A	B	C	D	E
R_1	a_1	b_{12}	b_{13}	a_4	b_{15}
R_2	a_1	a_2	b_{23}	b_{24}	b_{25}
R_3	b_{31}	a_2	b_{33}	b_{34}	a_5
R_4	b_{41}	b_{42}	a_3	a_4	a_5
R_5	a_1	b_{52}	b_{53}	b_{54}	a_5

Bước 3: Áp dụng $C \rightarrow D$ suy ra $b_{24} = b_{34} = b_{54} = a_4$

	A	B	C	D	E
R_1	a_1	b_{12}	b_{13}	a_4	b_{15}
R_2	a_1	a_2	b_{13}	a_4	b_{25}
R_3	b_{31}	a_2	b_{13}	a_4	a_5
R_4	b_{41}	b_{42}	a_3	a_4	a_5
R_5	a_1	b_{52}	b_{13}	a_4	a_5

Bước 4: Áp dụng $DE \rightarrow C$ suy ra $b_{13} = a_3$

	A	B	C	D	E
--	---	---	---	---	---

R_1	a_1	b_{12}	a_3	a_4	b_{15}
R_2	a_1	a_2	a_3	a_4	b_{25}
R_3	b_{31}	a_2	a_3	a_4	a_5
R_4	b_{41}	b_{42}	a_3	a_4	a_5
R_5	a_1	b_{52}	a_3	a_4	a_5

Bước 5: Áp dụng CE → A suy ra $b_{41} = b_{31} = a_1$

	A	B	C	D	E
R_1	a_1	b_{12}	a_3	a_4	b_{15}
R_2	a_1	a_2	a_3	a_4	b_{25}
R_3	a_1	a_2	a_3	a_4	a_5
R_4	a_1	b_{42}	a_3	a_4	a_5
R_5	a_1	b_{52}	a_3	a_4	a_5

Như vậy, tồn tại hàng thứ 3 R_3 có chứa các ký tự (a_1, a_2, a_3, a_4, a_5). Suy ra các phép tách kết nối không mất thông tin

4.7. Phép tách bảo toàn phụ thuộc

a. Định nghĩa phép chiếu của F trên tập thuộc tính:

Phép chiếu của F trên tập thuộc tính Z, ký hiệu $\pi_Z(F)$, là các tập phụ thuộc hàm $X \rightarrow Y$ trong F^+ sao cho $XY \subset Z$, chú ý $X \rightarrow Y$ không cần trong F.

b. Định nghĩa phép tách bảo toàn phụ thuộc hàm:

Ta nói phép tách bảo tồn một tập phụ thuộc hàm F nếu hội của tất cả các phụ thuộc hàm trong $\pi_{F_i}(F)$ ($i = 1.. k$) suy diễn ra mọi phụ thuộc hàm trong F.

c. Ứng dụng của phép tách bảo tồn phụ thuộc hàm:

Các phụ thuộc hàm trong F có thể được xem như các ràng buộc trọn vẹn đối với quan hệ R. Nếu chiếu của các phụ thuộc hàm không suy diễn được F, thì khi biểu diễn R dưới dạng $p = (R_1, R_2, \dots, R_k)$, ta có thể nhận thấy các giá trị hiện hành của các R_i biểu diễn một quan hệ không thỏa F, ngay cả khi không làm mất kết nối và thỏa F. Một giải pháp đưa ra là mỗi khi cập nhật một R_i , nên xét xem có vi phạm ràng buộc không.

Ví dụ: cho lược đồ quan hệ R(C, S, Z) có tập phụ thuộc hàm

$CS \rightarrow Z$

$Z \rightarrow C$

Phép tách R thành SZ và CZ có một kết nối không mất thông tin vì
 $(SZ \cap CZ) \rightarrow (CZ - SZ)$

Tuy nhiên, phép chiếu của $F\{CS \rightarrow Z, Z \rightarrow C\}$ nên SZ chỉ cho các phụ thuộc hàm tầm thường về tính phản xạ. Trong khi phép chiếu của F trên CZ cho phụ thuộc hàm $Z \rightarrow C$; các phụ thuộc hàm tầm thường và chúng không suy ra được $CS \rightarrow Z$. Do đó, phép tách không bảo toàn phụ thuộc hàm.

4.8. Kiểm tra sự bảo toàn phụ thuộc hàm

$R \{R_1, R_2, \dots, R_k\}$

$F \{\pi_{R_1}(F) \dots \pi_{R_n}(F)\}$

Ta chứng minh: $G = \bigcup_{i=1}^n \pi_{R_i}(F)$

$X \rightarrow Y \in F \Rightarrow X \rightarrow Y \in G^+$

$Y \in X^+ \{ \text{tính } X^+ \text{ là ta tính } X_0, X_1, \dots \}$

Nói chung, có vẽ dễ dàng kiểm tra xem một phép tách có bảo toàn phụ thuộc hàm. Chỉ cần tính F^+ và chiếu của F lên các R_i . Sau đó, lấy bội của chúng và kiểm tra xem có là phủ của F hay không.

Tuy nhiên, trên thực tế việc tính phủ của F là công việc không hề đơn giản. Do đó, có một cách khác kiểm tra tính bảo tồn mà không cần tính phủ của F và lại tốn ít thời gian.

a. Giải thuật: Kiểm tra sự bảo toàn phụ thuộc hàm

Input: Lược đồ quan hệ $R = \{A_1, A_2, \dots, A_n\}$, tập phụ thuộc hàm F và phép tách:

$p = \{R_1, \dots, R_k\}$

Output: Kết luận về tính bảo toàn phụ thuộc hàm của phân tách p .

Phương pháp:

Ký hiệu: $G = \bigcup_{i=1}^n R_i(F)$

Ta không tính G mà chỉ kiểm tra xem nó có phủ F hay không. Cho $X \rightarrow Y \in F$, kiểm tra xem bao đóng của X theo G , ký hiệu X_G^+ , có chứa Y hay không.

Thủ thuật tính X_G^+ mà không cần có G là xét tác dụng của quá trình tính bao đóng của X đối với các chiếu của F lên R_i .

Xuất phát từ X , ta tính X_G^+ như sau:

(1): Đặt $X_0 = X$, $t = 1$;

(2): Tính X_t trên cơ sở X_{t-1} :

(i): Đặt $X_t = X_{t-1}$

(ii): Với mỗi lược đồ con $R_i \in F$, $R_i \notin X_t$, ta thực hiện phép toán:

$$X_t = X_t \cup ((X_t \cap R_i)^+ \cap R_i)$$

Trong quá trình tính toán nếu $X_t = R$ thì kết thúc, và kết luận $X_G^+ = R$. Ngược lại sang bước (3).

(3): Nếu $X_t = X_{t-1}$ thì kết thúc, và kết luận $X_G^+ = X_t$.

Ngược lại tăng t lên 1 đơn vị và quay về bước (2).

Nếu Y là tập con của X_G^+ có được từ việc thực hiện các bước trên, thì khi đó $X \rightarrow Y \in G^+$. Nếu mỗi $X \rightarrow Y \in F$ được chứng minh thuộc G^+ bằng cách đó thì phép tách bảo toàn phụ thuộc hàm, ngược lại là không.

Ví dụ:

Xét tập thuộc tính A, B, C, D với phép tách

$\{AB, BC, CD\}$

Và tập phụ thuộc hàm F

{

$A \rightarrow B$

$B \rightarrow C$

$C \rightarrow D$

$D \rightarrow A$

}

Mỗi thuộc tính xác định mọi thuộc tính khác. Ta có thể lầm tưởng rằng khi chiếu F lên AB, BC, CD . Ta sẽ không có được $D \rightarrow A$, nhưng cảm giác đó là sai vì khi chiếu F thực ra là chiếu F^+ lên các sơ đồ quan hệ. Như vậy, khi chiếu lên AB không những ta nhận được $A \rightarrow B$ mà còn $B \rightarrow A$. Tương tự, ta nhận được $C \rightarrow B$ trong $\pi_{BC}(F)$ và $D \rightarrow C$ trong $\pi_{DC}(F)$ và ba phụ thuộc hàm còn suy diễn logic được $D \rightarrow A$.

$$G = \pi_{AB}(F) \cup \pi_{BC}(F) \cup \pi_{CD}(F)$$

Bắt đầu từ $Z = \{D\}$. Áp dụng giải thuật cho $R_1 = \{A, B\}$

Ta có: $Z = \{D\} \cup ((\{D\} \cap \{A, B\}^+) \cap \{A, B\}) = \{D\}$

Z : không đổi

Tính Z trên $R_2 = \{B, C\}$. Z cũng không đổi, áp dụng trên R_3 thì $Z = \{C, D\}$ vì $Z = \{D\} \cup ((\{D\} \cap \{C, D\}^+) \cap \{C, D\}) = \{C, D\}$

Trong bước kế tiếp, áp dụng với BC cho $Z = \{C, D\}$ thì $Z = \{B, C, D\}$.

Bước 3, $Z = \{A, B, C, D\} \Rightarrow Z$. Không còn thay đổi được nữa.

Vậy, $\{D\}$ chứa A và tôn trọng G nên kết luận $G = D \rightarrow A$. Các phần tử khác của F ở trong G^+ , ta kết luận ‘Phép tách bảo toàn phụ thuộc hàm’.

BÀI TẬP CHƯƠNG 4

Bài 1.

a. Cho lược đồ quan hệ $Q(A, B, C, D)$, r là một quan hệ trên Q như sau:

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

Giáo trình Cơ sở dữ liệu

Những phụ thuộc hàm nào sau đây không thoả r?

$$D \rightarrow A; \quad A, C \rightarrow D; \quad CD \rightarrow A; \quad D \rightarrow B;$$

b. Cho lược đồ quan hệ Q(A, B, C, D), r là quan hệ 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 thoả r?

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

c. Cho lược đồ quan hệ Q(A, B, C, D), 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 thoả r?

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

Bài 2.

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

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

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

b. Cho lược đồ quan hệ Q(A, B, C, D, E, G, H) 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^+ ?

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

c. Cho lược đồ quan hệ Q = (A, B, C, D) và tập phụ thuộc hàm $F = \{A \rightarrow B; A \rightarrow C\}$

Những phụ thuộc hàm nào được suy ra từ F?

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

Bài 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 $BD \rightarrow C \in F^+$

Bài 4. Hãy tìm phủ tối thiểu của F.

- a. Q(A, B, C, D, E, G) và F = {AB → C; C → A; BC → D; ACD → B; D → EG; BE → C; CG → BD; CE → AG}
- b. Q(A, B, C, D, E, G) và F = {AB → C; C → A; BC → D; ACD → B; D → EG; BE → C; CG → BD; CE → AG}
- c. Q(A, B, C) và F = {A → B; A → C; B → A; C → A; B → C}
- d. Q(A, B, C, D, E, G, H) và F = {A → H; AB → C; BC → D; G → B}
- e. Q(A, B, C, S, X, Y, Z) và F = {S → A; AX → B; S → B; BY → C; CZ → X}
- f. Q(A, B, C, D, E, G, H, I, J) và F = {BG → D; G → J; AI → C; CE → H; BD → G; JH → A; D → I}
- g. Q(A, B, C, D, E, G, H, I, J) và F = {BH → I; GC → A; I → J; AE → G; D → B; I → H}
- h. Q{A, B, C, D, E, F, G, H, K, L, M} và F = {A → B, C → D, E → F, G → AHK, AH → G, GLC → M }
- i. cho Q(A, B, C, D, E, I) và F={A → C, AB → C, C → DI, EC → AB, EI → C}
- j. cho Q(TENTAU, LOAITAU, MACHUYEN, LUONGHANG, BENCANG, NGAY) và F = {TENTAU → LOAITAU

MACHUYEN → TENTAU, LUONGHANG

TENTAU, NGAY → BENCANG, MACHUYEN}

- k. Cho lược đồ quan hệ Q(MSCD, MSSV, CD, HG) và tập phụ thuộc F như sau:

F = {MSCD → CD; CD → MSCD;
CD, MSSV → HG; MSCD, HG → MSSV;
CD, HG → MSSV; MSCD, MSSV → HG}

- l. cho Q(P, Q, X, Y, Z, W) và F={ X → YZ, ZW → P, P → Z, W → XPQ, XYQ → YW, WQ → YZ}

Bài 5.

Cho lược đồ CSDL

KEHOACH (NGAY, GIO, PHONG, MONHOC, GIAOVIEN)

F={NGAY, GIO, PHONG → MONHOC

MONHOC, NGAY → GIAOVIEN

NGAY, GIO, PHONG → GIAOVIEN

MONHOC → GIAOVIEN

}

- a. Tính $\{NGAY, GIO, PHONG\}^+$; $\{MONHOC\}^+$

- b. Tìm phủ tối thiểu của F

Bài 6. Tìm bao đóng

- a. Cho lược đồ quan hệ R(A, B, C, D, E, G) và tập phụ thuộc hàm

$$F = \{A \rightarrow C, A \rightarrow EG, B \rightarrow D, G \rightarrow E\}$$

Tìm AB^+ , CGD^+ , A^+

b. Cho lược đồ quan hệ $R(A, B, C, D, E, G)$ và tập phụ thuộc hàm

$$F = \{B \rightarrow C, A \rightarrow EG, B \rightarrow A, G \rightarrow E\}$$

Tìm AB^+ , CGD^+ , A^+

c. Cho lược đồ quan hệ $R(A, B, C, D, E)$ và tập phụ thuộc hàm

$$F = \{A \rightarrow C, BC \rightarrow D, D \rightarrow E, E \rightarrow A\}$$

Tìm AB^+ , BD^+ , D^+

d. Cho lược đồ quan hệ $\alpha = (U, F)$ với $U = ABCDEG$ và $F = \{A \rightarrow C, BC \rightarrow D, D \rightarrow E, E \rightarrow A\}$.

Hãy tính : $(AB)^+$; $((DE)^+ A)^+$

e. Cho lược đồ quan hệ $\alpha = (U, F)$ với $U = ABCDEG$ và $F = \{B \rightarrow C, AC \rightarrow D, D \rightarrow G, AG \rightarrow E\}$

Hãy tính : AB^+ ; BD^+

Bài 7.

Cho lược đồ quan hệ $R = (U, F)$ với

$U = ABCDEGH$ và tập phụ thuộc hàm

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

$f = AB \rightarrow E$, chứng minh rằng với mọi quan hệ R trên U nếu R thỏa F thì R cũng thỏa f .

Bài 8.

Cho lược đồ quan hệ R (U, F) với

$U = ABCDEGHIJ$ và tập phụ thuộc hàm

$$F = \{AB \rightarrow E, AG \rightarrow J, BE \rightarrow I, E \rightarrow G, GI \rightarrow H\}$$

$$f = AB \rightarrow GH$$

Chứng minh rằng f suy dẫn được từ F

Bài 9.

Cho lược đồ quan hệ $R(U, F)$

$U = ABCDEGH$ và tập phụ thuộc hàm

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

Hãy chứng minh

$$AB \rightarrow E, \quad BG \rightarrow C, \quad AB \rightarrow G$$

Bài 10.

Cho lược đồ quan hệ và tập phụ thuộc hàm

$$F = \{AB \rightarrow E, AG \rightarrow I, BE \rightarrow I, E \rightarrow G, GI \rightarrow H\}$$

Chứng minh rằng $AB \rightarrow GH$ suy dẫn được từ F

Bài 11.

Cho lược đồ quan hệ $(=(u, F))$ và tập phụ thuộc hàm

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

Chứng minh rằng $AB \rightarrow E$ và $AB \rightarrow G$ suy dẫn được từ F

Bài 12.

Cho lược đồ quan hệ R(ABCD) và $F=\{A \rightarrow B, BC \rightarrow D\}$

hãy cho biết các phụ thuộc hàm nào dưới đây có thể suy dẫn được từ F

$$AC \rightarrow D, \quad B \rightarrow D, \quad AD \rightarrow B$$

Bài 13.

Loại bỏ các phụ thuộc hàm dư thừa trong tập phụ thuộc hàm sau:

$$F=\{X \rightarrow Y, Y \rightarrow X, Y \rightarrow Z, Z \rightarrow Y, X \rightarrow Z, Z \rightarrow X\}$$

Bài 14.

Nếu $X \rightarrow Y \in F$, $A \in X$, thuộc tính A được gọi là dư thừa nếu

$$\{X - A\} \rightarrow Y \in F^+$$

hãy loại bỏ các thuộc tính dư thừa trong các tập sau:

a. $F=\{X \rightarrow YW, XW \rightarrow Z, Z \rightarrow Y, XY \rightarrow Z\}$

b. $F=\{A \rightarrow BC, E \rightarrow C, D \rightarrow AEF, ABF \rightarrow BD\}$

Bài 15.

Sử dụng các luật của hệ tiên đề Armstrong chứng minh các tính chất sau:

a. Tính tựa bắc cầu: Nếu $X \rightarrow Y$ và $YZ \rightarrow W$ thì $XZ \rightarrow W$

b. Tính phản xạ chặt $X \rightarrow X$

c. Tính cộng tính : Nếu $X \rightarrow Y$ và $Z \rightarrow W$ thì $XZ \rightarrow YW$

d. Tính chất hợp : Nếu $X \rightarrow Y$ và $X \rightarrow Z$ thì $X \rightarrow YZ$

e. Tính tách : Nếu $X \rightarrow YZ$ thì $X \rightarrow Y$ và $X \rightarrow Z$

f. Tính tích luỹ: Nếu $X \rightarrow YZ$, $Z \rightarrow VW$ thì $X \rightarrow YVW$

Bài 16.

Cho tập các thuộc tính quản lý tổng đài - cước thuê bao:

MaTD : Mã tổng đài

TênTD: Tên tổng đài

DL: Dung lượng

MC: Mức cấp

SDT: Số điện thoại thuê bao

NSX: Nước sản xuất

NG: Ngày lắp đặt tổng đài

GTR: Giá trị tổng đài

VT: Nơi đặt tổng đài

TenTB: Tên thuê bao

TTB: Mã thuê bao

DC: Địa chỉ thuê bao

SDTD: Số điện thoại đến/đi

NGG: Ngày thực hiện đàm thoại

GB: Giờ bắt đầu đàm thoại

KT: Giờ kết thúc đàm thoại

DD: Cho biết cuộc gọi đang thực hiện là cuộc gọi đi hoặc cuộc gọi đến

a) Hãy xác định một phép tách có bảo toàn phụ thuộc.

b) Hãy xác định một phép tách – kết nối không mất thông tin.

Bài 17. Cho tập các thuộc tính quản lý phát hành báo chí QLBC :

MaK: Mã khách

TenK: Tên khách

DC: Địa chỉ

MaB: Mã báo

TenB: Tên báo

Gia: Giá báo

SL: số lượng

KY: kỳ phát hành

a) Hãy xác định một phép tách – kết nối không mất thông tin và bảo toàn phụ thuộc hàm.

b) Hãy xác định một phép tách – kết nối không mất thông tin.

Bài 18. Kiểm tra phép tách sau có mất thông tin không.

a) $\alpha = (U, F)$ với $U = ABCD$, $F = \{A \rightarrow B, AC \rightarrow D\}$, $\delta = \{AB, ACD\}$

b) $\alpha = (U, F)$ với $U = ABCDE$, $F = \{A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A\}$, $\delta = \{AD, AB, BE, CDE\}$

c) Xác định và giải thích dạng chuẩn cao nhất của lược đồ quan hệ $\alpha = (U, F)$ với $U = ABCD$, $F = \{A \rightarrow C, D \rightarrow B, C \rightarrow ABD\}$

Bài 19. Cho lược đồ quan hệ $\alpha = (U, F)$ với

$U = ABCDEGH$

$F = \{CD \rightarrow H, E \rightarrow B, D \rightarrow G, BH \rightarrow E, CH \rightarrow DG, C \rightarrow A\}$

Hỏi phép tách $p = (ABCDE, BCH, CDEGH)$ có kết nối mất thông tin không?

Bài 20.

Cho $\alpha=(u, F)$ với

$U=ABCDE$ và

$F=\{A \rightarrow C, B \rightarrow C, A \rightarrow D, DE \rightarrow C, CE \rightarrow A\}$

Phép tách : $\delta=\{AD, AB, BE, CDE, AE\}$

Kiểm tra tính kết nối của phép tách có không mất thông tin ?

Bài 21.

Cho $\alpha=(u, F)$ với

$U=ABCDEF$ và

$F=\{AB \rightarrow C, C \rightarrow B, ABD \rightarrow E, F \rightarrow A\}$

Phép tách: $\delta=\{BC, AC, ABDE, ABDF\}$

Kiểm tra tính kết nối của phép tách có không mất thông tin ?

Bài 22.

Cho $\alpha=(u, F)$ với

$U=ABCDEG$

$F=\{D \rightarrow G, C \rightarrow A, CD \rightarrow E, A \rightarrow B\}$

Phép tách: $\delta=\{DG, AC, SCE, AB\}$

Kiểm tra tính kết nối của phép tách có không mất thông tin ?

Bài 23.

Cho $\alpha=(u, F)$ với

$U=ABCDE$ và

$F=\{A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A\}$

Phép tách : $\delta=\{AC, CD, BE, BC, AE\}$

Kiểm tra tính kết nối của phép tách có không mất thông tin ?

Chương 5: CHUẨN HÓA SƠ ĐỒ QUAN HỆ

Chương này sẽ giới thiệu sự cần thiết phải chuẩn hóa cơ sở dữ liệu, các dạng chuẩn phổ biến, các đặc trưng của từng dạng chuẩn. Kỹ thuật chuyển đổi một quan hệ chưa chuẩn hóa về một nhóm các quan hệ ở dạng chuẩn 3NF không bị mất thông tin, trong một số trường hợp các phép tách vẫn bảo toàn các phụ thuộc. Nội dung của chương bao gồm:

Sự cần thiết chuẩn hóa sơ đồ quan hệ.

- Dạng chuẩn 1 - 1NF.
- Dạng chuẩn 2 - 2NF.
- Dạng chuẩn 3 - 3NF.
- Dạng chuẩn 4 – 4NF.
- Dạng chuẩn BOYCE-CODD.

Chuyển cơ sở dữ liệu từ dạng chuẩn thấp sang dạng chuẩn 3NF hoặc dạng chuẩn Boyce – Codd.

5.1. Khóa của sơ đồ quan hệ

Trong một lược đồ quan hệ luôn luôn tồn tại một thuộc tính hay một tập các thuộc tính có khả năng biểu diễn duy nhất các thuộc tính còn lại. Nói cách khác giá trị của một hay nhiều thuộc tính có thể xác định duy nhất giá trị các thuộc tính khác. Số chứng minh thư sẽ xác định được tất cả thông tin về người mang chứng minh thư đó. Tập thuộc tính có tính chất như trên được gọi là khóa của lược đồ quan hệ. Thông thường trong một lược đồ quan hệ có thể tồn tại nhiều khóa. Trong số đó sẽ chọn một khóa làm khóa chính sao cho đơn giản và không nhập nhằng thông tin.

5.1.1. Khóa (Key)

Cho lược đồ quan hệ $R(A_1, A_2, \dots, A_n)$, A là tập thuộc tính của quan hệ R , F là tập phụ thuộc hàm trên A , K là tập con của A . Khi đó, K gọi là một khóa của R nếu thỏa hai điều kiện sau:

- $K^+ = A$
- Không tồn tại $K' \in K$ sao cho $K'^+ = A$

Ví dụ 1: Cho quan hệ MONHOC (MSMH, TenMon)

Trong đó:

MSMH: Mã số môn học.

TenMon: Tên môn học.

Ý nghĩa của quan hệ MonHoc. Mỗi môn học có một mã số và một tên.

Ta thấy cả hai thuộc tính MSMH và TenMon đều thỏa hai điều kiện đã nói trên nên quan hệ MonHoc có đến hai khóa $\{MaMon\}$ và $\{TenMon\}$. Nhưng khi thiết kế CSDL ta chỉ chọn một trong hai thuộc tính trên, thường sẽ chọn MSMH là khóa chính.

Ví dụ 2: Đối với quan hệ NHA (SoNha, Duong, DienTich):

Trong đó:

SoNha: Số nhà.

Duong: Đường.

DienTich: Diện tích.

Ý nghĩa của quan hệ NHA. Mỗi nhà được đặc trưng bởi số nhà, đường như là địa chỉ để xác định vị trí nhà và mỗi nhà có một diện tích.

Ta thấy {SoNha, Duong} thỏa hai điều kiện của khóa.

Có thể định nghĩa khóa trong một quan hệ như sau: Nếu 2 bộ bất kỳ trùng nhau trên các thành phần của khóa thì cũng trùng nhau trên các thành phần không khóa. Nếu trùng nhau trên một số thành phần của khóa thì không thể trùng nhau trên các thành phần không khóa. Giá trị các thành phần của khóa không thể nhận giá trị null hay các giá trị không xác định.

Ví dụ: Cho quan hệ NHANVIEN (MSNV, HoTen, NgaySinh, DiaChi).

Trong đó:

MSNV: Mã số nhân viên.

HoTen: Họ và tên nhân viên.

NgaySinh: Ngày sinh của nhân viên.

DiaChi: Địa chỉ của nhân viên.

Tùy trường hợp mà ta có các khóa khác nhau mà quan hệ NHANVIEN có khóa khác nhau:

- 1). {MSNV} là khóa vì không có hai nhân viên nào cùng mã số..
- 2). Nếu không có nhân viên nào trùng họ tên nhau, ta có {HoTen} là một khóa.
- 3). Nếu không có hai nhân viên nào trùng địa chỉ, ta có {DiaChi} là một khóa.
- 4). Nếu trong cơ quan có nhân viên trùng họ tên, hoặc trùng địa chỉ nhưng không có trường hợp hai nhân viên cùng họ tên cùng địa chỉ thì {HoTen, DiaChi} là một khóa.
- 5). Nếu không có nhân viên nào trùng họ tên và trùng cả ngày sinh thì {HoTen, NgaySinh} cũng là một khóa.

5.1.2. Siêu khóa

Cho lược đồ quan hệ R(A₁, A₂, ..., A_n), A là tập thuộc tính của quan hệ R, F là tập phụ thuộc hàm trên A, K là khóa của quan hệ R. Khi đó K' được gọi là một siêu khóa của R nếu K ⊂ K'.

Ví dụ: Cho quan hệ NHANVIEN (MSNV, HoTen, NgaySinh, DiaChi) có MSNV là khóa thì ta có các siêu khóa sau:

{MSNV}

{MSNV, HoTen }

{MSNV, NgaySinh }

{MSNV, DiaChi }

...

5.1.3. Thuộc tính khóa và thuộc tính không khóa

Cho lược đồ quan hệ $R(A_1, A_2, \dots, A_n)$, A là tập thuộc tính của quan hệ R , F là tập phụ thuộc hàm trên A , K là một khóa của R .

A_i là một thuộc tính khóa nếu $A_i \in K$, ngược lại A_i là thuộc tính không khóa.

Ví dụ: Cho lược đồ quan hệ KETQUA (MSSV, MSMH, LanThi, DiemThi)

Trong đó:

MSSV: Mã số sinh viên.

MSMH: Mã số môn học.

LanThi: Lần thi.

DiemThi: Điểm thi.

Ý nghĩa: Quan hệ KETQUA ghi nhận điểm thi, lần thi của các môn học cho các sinh viên.

Trong đó khóa của quan hệ KETQUA là {MSSV, MMH, LanThi} vậy ta có ba thuộc tính khóa MSSV, MMH và LanThi (vì thế không được gọi MSSV là khóa hoặc MMH là khóa)

5.1.4. Thuật toán tìm khóa

Đầu tiên, chúng ta cần hiểu một vài khái niệm :

Ta gọi :

- Q là tập thuộc tính của quan hệ
- F là tập phụ thuộc hàm
- $L(\text{left})$: là các thuộc tính xuất hiện bên trái
- $R(\text{right})$: là các thuộc tính xuất hiện ở về phải
- $S(\text{superkey})$: là tập các siêu khóa
- $K(\text{key})$: là tập các khóa
- Tập thuộc tính nguồn (TN) : bao gồm các thuộc tính chỉ xuất hiện ở về trái, không xuất hiện ở về phải của F và các thuộc tính không xuất hiện ở cả về trái và về phải của F .

Vậy $TN = Q - R$

Nghĩa là ta lấy Q trừ cho R để tìm thuộc tính chỉ xuất hiện ở L và các thuộc tính không xuất hiện ở cả L và R

Ví dụ : Cho tập cơ sở dữ liệu $Q = \{A, B, C, D, E\}$ $L = \{A, B\}$ $R = \{B, C, E\}$

$TN = Q - R = \{A, D\}$

- Tập thuộc tính đích (TD) : Bao gồm các thuộc tính chỉ xuất hiện ở R , không xuất hiện ở L .

Vậy $TD = R - L$

Ví dụ : Cho $L = \{A, B, C, D, E\}$ $R = \{E, F, G, H\}$

$TD = \{F, G, H\}$

Giáo trình Cơ sở dữ liệu

- Tập thuộc tính trung gian (TG) : Chứa các thuộc tính xuất hiện ở cả L và R

Vậy TG = L Giao R (Giao của 2 tập hợp để lấy thuộc tính chung của 2 Tập hợp đó)

Ví dụ : Cho L = {A,B,C,D,E} R = {D,E,F,G}

Vậy TG = L /cap R = {D,E}

Thuật toán :

Bước 1 :

Tìm tập thuộc tính nguồn TN và Tập thuộc tính trung gian TG, bằng cách ví dụ ở trên thì các bạn có thể dễ dàng tìm thấy 2 tập thuộc tính này.

Bước 2 :

- Nếu TG = \emptyset
- Thì K(Key) = TN, và kết thúc thuật toán, xuất ra K của tập cơ sở dữ liệu $\langle Q, F \rangle$
- Ngược lại, nếu TG $\neq \emptyset$
- Thì qua bước 3

Bước 3 :

Tìm tất cả các tập con X_i của TG

Bước 4 :

Tìm Siêu khóa (S_i) bằng cách với mọi X_i

Nếu $(TN \cup X_i)^+ = Q$ thì khi đó $S_i = TN \cup X_i$

Bước 5 :

- Tìm Khóa(K_i) bằng cách loại bỏ các siêu khóa không tối thiểu
- Với mọi $S_i S_j$ thuộc S
- Nếu S_i chứa trong S_j thì loại bỏ S_j ra khỏi tập siêu khóa. Khi đó, tập S còn lại chính là tập khóa cần tìm

Ví dụ :

Ta có S = {AB, ABC, ED, EDF}

Ta thấy AB chứa trong ABC, ED chứa trong EDF vậy chúng ta cần phải loại bỏ ABC và EDF.

Vậy S = {AB,ED} chính là tập khóa cần tìm

Chúng ta có một ví dụ mẫu như sau :

Ví dụ : Cho một tập cơ sở dữ liệu R = $\langle Q, F \rangle$

Với Q = {ABC} F = {AB \rightarrow C, C \rightarrow A}. Tìm tất cả các khóa thuộc tập cơ sở dữ liệu trên.

Bài Làm :

L = {ABC} R = {CA}

TN = {B} TG = {AC} $\neq \emptyset$ nên ta làm tiếp bước 3

Ta có tập con Xi của tập TG = { ϕ , A,C,AC}

Ta lấy từng thuộc tính thuộc tập con Xi của tập TG hợp với TN ta có các thuộc tính sau:

$S_1 = TN \cup 0 = B$ Ta có $B^+ = B \neq Q$ nên $S_1 = A$ không là siêu khóa

$S_2 = TN \cup A = AB$ Ta có $AB^+ = ABC = Q$ nên $S_2 = AB$ là siêu khóa

$S_3 = TN \cup C = BC$ Ta có $BC^+ = ABC = Q$ nên $S_3 = BC$ là siêu khóa

$S_4 = TN \cup AC = ABC$ Ta có $ABC^+ = ABC = Q$ nên $S_4 = ABC$ là siêu khóa

Vậy ta có tập siêu khóa $S = \{AB, BC, ABC\}$.

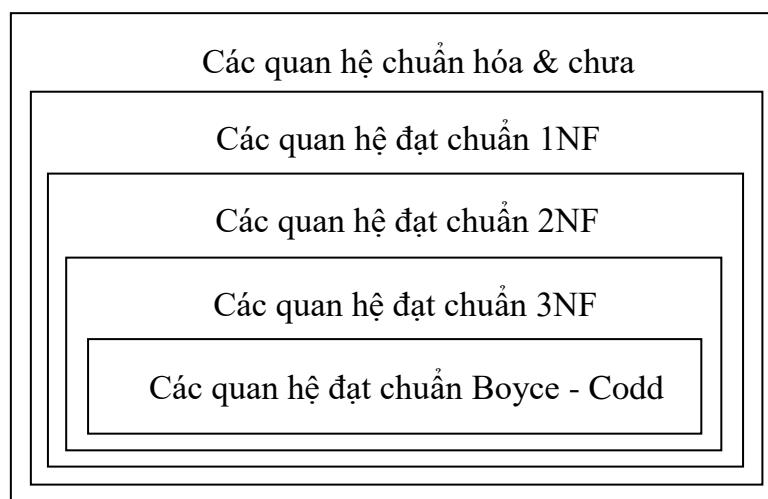
Tuy nhiên, vì AB và BC chứa trong ABC nên loại bỏ siêu khóa ABC ra khỏi tập siêu khóa

Vậy ta có, tập khóa K = {AB, BC} là khóa của lược đồ quan hệ

5.2. Các dạng chuẩn của sơ đồ quan hệ

Phần này sẽ tìm hiểu và nghiên cứu khái niệm chuẩn hóa dữ liệu quan hệ. Đã có nhiều kết quả trong lĩnh vực chuẩn hóa, song vẫn còn nhiều vấn đề bất lợi khi tìm kiếm thông tin. Vì vậy người ta cũng chỉ mới đề xuất các dạng sao cho tiện lợi, đơn giản hơn trong quá trình thao tác các ngôn ngữ con dữ liệu.

Trong ‘ Further Normalization of the DataBase relational Model’ lần đầu tiên E.F. Codd đã đưa ra các khái niệm về chuẩn hóa quan hệ. Ông đã chia các quan hệ thành 3 lớp và gọi chúng là các quan hệ dạng chuẩn thứ nhất 1NF, dạng chuẩn thứ hai 2NF và dạng chuẩn thứ ba 3NF. Trong ‘ Multivalued Dependencies and a New Normal Form Relational Databases’ R. Fagin đã đưa ra khái niệm dạng chuẩn thứ tư, theo ông đó là dạng tiện lợi hơn so với quan hệ dạng chuẩn 3. Mục đích của chương này là minh họa những ưu điểm của các quan hệ dạng chuẩn 4 và trình bày phương pháp chuyển đổi một quan hệ chưa là dạng chuẩn (dạng chuẩn thấp) về một nhóm tương đương các quan hệ dạng chuẩn 3NF hoặc Boyce - Codd.



Hình 5.1: Sơ đồ thể hiện mối liên hệ giữa các dạng chuẩn

Sự cần thiết chuẩn hóa dữ liệu: Khi thực hiện các phép lưu trữ trên các quan hệ chưa được chuẩn hóa thường xuất hiện các dị thường thông tin. Nghĩa là trong dữ liệu lưu trữ, sự dư thừa, mất dữ liệu, mâu thuẫn hay không nhất quán dữ liệu có thể xảy ra khi cập nhật, bổ sung hay sửa đổi dữ liệu. Dị thường thông tin là nguyên nhân gây cản trở cho việc tìm kiếm, hỏi đáp thông tin. Mục tiêu của chuẩn hóa dữ liệu là triệt tiêu mức cao nhất khả năng

xuất hiện các dị thường thông tin khi thực hiện các phép lưu trữ. Có như vậy mục tiêu của các hệ cơ sở mới được bảo đảm. Dữ liệu lưu trữ phản ánh thế giới hiện thực khách quan, đầy đủ hơn và sinh động hơn.

Khi thiết kế và cài đặt các hệ CSDL, chuẩn hóa là quá trình khảo sát danh sách các thuộc tính và áp dụng tập các quy tắc phân tích vào danh sách đó, biến đổi chúng thành nhiều tập nhỏ hơn sao cho:

- Tối thiểu việc lặp lại.
- Tránh dị thường thông tin.
- Xác định và giải quyết được sự không rõ ràng, nhập nhằng trong suy diễn.

Quá trình chuẩn hóa sơ đồ quan hệ dựa trên các khái niệm về phụ thuộc hàm, phụ thuộc đầy đủ, khóa, các thuộc tính không khóa... Một mô hình được xem là mô hình chuẩn hóa tốt, lý tưởng là mô hình ở đó mỗi một thuộc tính không khóa phụ thuộc hàm vào khóa.

Quan hệ chuẩn hóa là những quan hệ mà mỗi giá trị thuộc tính trong bộ là những thuộc tính nguyên tố, không phân chia ra được. Nói cách khác, mỗi toạ độ của hàng và cột trong quan hệ chỉ có đúng một giá trị chứ không phải là một tập các giá trị.

Ví dụ: Lược đồ quan hệ QLCAP(TC, GTR, MC, MNSX, TNSX)

Trong đó:

TC : mã tuyến cáp.

GTR : giá trị của cáp.

MC : mã số cáp.

MNSX : mã nước sản xuất.

TNSX : tên nước sản xuất.

Ý nghĩa dữ liệu như sau:

- Trong một tuyến cáp, giá trị của một loại cáp được xác định duy nhất.
- Mỗi mã cáp xác định mã nước sản xuất cáp.
- Mã nước xác định tên nước sản xuất.

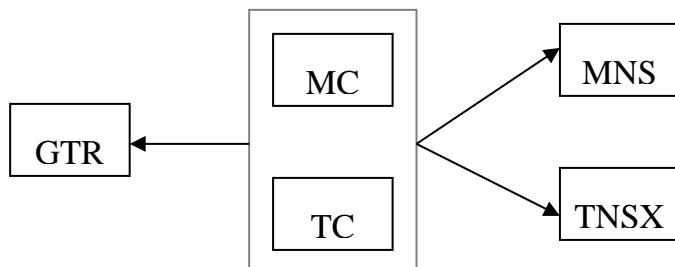
Sơ đồ quan hệ có các tập phụ thuộc hàm sau F:

$(TC, MC) \rightarrow GTR$

$MC \rightarrow MNSX$

$MNSX \rightarrow TNSX$

Các thuộc tính khóa: TC, MC và các thuộc tính không khóa: GTR, TNSX, MNSX.



Sơ đồ các phụ thuộc hàm trong lược đồ quan hệ QLCAP

TC	GTR	MC	N	NSX
T01	200	C01	HAQ	Hàn Quốc
T01	250	C02	HAQ	Hàn Quốc
T01	220	C03	VTN	Việt Nam
T02	500	C01	HAQ	Hàn Quốc
T02	400	C04	JAN	Nhật Bản
T03	100	C05	RUS	Nga
T04	400	C06	CHN	Trung Quốc
T04	450	C03	VTN	Việt Nam

5.2.1. Thuộc tính nguyên tố, không nguyên tố.

Cho sơ đồ quan hệ R(A1, A2, ..., An), A là tập thuộc tính của quan hệ R, F là tập phụ thuộc hàm trên A, K là một khóa của R. Nếu B là thành phần của một khóa thì B là nguyên tố, ngược lại B không nguyên tố.

Ví dụ: Cho sơ đồ quan hệ R(A, B, C, D) có các phụ thuộc hàm

{

AB → C

B → D

BC → A

}

Với quan hệ R thì ta có thể tìm được hai khóa hoặc AB hoặc BC

Vậy ta có A, B, C là thuộc tính khóa và D là thuộc tính không khóa.

5.2.2. Các dạng phụ thuộc hàm**a. Phụ thuộc hàm từng phần:**

Cho phụ thuộc hàm X → A được gọi là phụ thuộc hàm từng phần nếu X là con thực sự của khóa.

b. Phụ thuộc hàm đầy đủ:

Cho phụ thuộc hàm X → A được gọi là phụ thuộc hàm đầy đủ nếu không tồn tại Y → X để cho Y → A.

=> Phụ thuộc hàm đầy đủ vào khóa ngược với phụ thuộc hàm từng phần.

c. Phụ thuộc truyền:

Cho phụ thuộc hàm $X \rightarrow A$ được gọi là phụ hàm truyền nếu X không là con thực sự của khóa nào trong quan hệ.

d. Phụ thuộc trực tiếp:

Cho phụ thuộc hàm $X \rightarrow A$ được gọi là phụ thuộc hàm trực tiếp nếu không tồn tại tập thuộc tính Y , Y khác X và Y khác A thỏa: $X \rightarrow Y$ và $Y \rightarrow A$

Vậy, phụ thuộc trực tiếp vào khóa ngược với phụ thuộc truyền.

Ví dụ: cho quan hệ SACH (MS, TuaSach, Loai)

Trong đó:

MS: Mã số sách, bao gồm số thứ tự của sách ghép với tên loại của sách.

TuaSach: Tựa sách.

Loai: Tên loại của sách.

Vậy ta có các phụ thuộc hàm:

$MS \rightarrow Loai$ là phụ thuộc từng phần.

5.2.3. Sự chuẩn hóa một quan hệ

Các quan hệ trong cơ sở dữ liệu cần đưa về một dạng chuẩn nhất định, nhằm tránh xảy ra các bất lợi như đã đề cập. Một sơ đồ quan hệ muốn thỏa dạng chuẩn cao thì phải thỏa các điều kiện của dạng chuẩn thấp trước. Thông thường một cơ sở dữ liệu thì các quan hệ phải thỏa ít nhất được chuẩn 3NF (quan hệ đó đã thỏa chuẩn 1NF và 2NF).

5.2.4. Dạng chuẩn thứ nhất - 1NF

Cho quan hệ $R(A_1, A_2, \dots, A_n)$, A là tập thuộc tính của quan hệ R , F là tập phụ thuộc hàm được gọi là dạng chuẩn 1 – 1NF, khi và chỉ khi các thuộc tính chỉ chứa các giá trị nguyên tố. Thường các quan hệ chuẩn hóa là các quan hệ dạng chuẩn 1NF. Tuy nhiên, cấu trúc biểu diễn dữ liệu trong các quan hệ dạng 1NF còn nhiều điều bất tiện. Vì vậy khi thao tác thực hiện các phép chèn thêm, sửa đổi hay bổ sung cập nhật dữ liệu thường xuất hiện dị thường thông tin, không thể chấp nhận được trong quá trình tìm kiếm.

Theo định nghĩa, quan hệ QLCAP(TC, GTR, MC, NSNSX, NSX) là một quan hệ dạng chuẩn thứ nhất 1NF.

Ví dụ: cho quan hệ SACH (MS, TuaSach, Loai)

Trong đó:

MS: mã số sách, bao gồm số thứ tự của sách ghép với mã loại của sách.

TuaSach: tựa sách.

Loai: tên loại của sách.

=> Quan hệ SACH không thỏa dạng chuẩn thứ nhất (1NF).

5.2.5. Dạng chuẩn thứ hai – 2NF

Lược đồ quan hệ $R(A_1, A_2, \dots, A_n)$ được gọi là dạng chuẩn 2NF, khi và chỉ khi nó là dạng chuẩn 1NF và các thuộc tính không khóa phụ thuộc hàm đầy đủ vào khóa. Nói cách khác,

quan hệ R là dạng chuẩn thứ nhất và không tồn tại các phụ thuộc hàm $X \rightarrow Y \in F^+$ sao cho X là tập con thực sự của khóa và Y là thuộc tính không khóa.

Lược đồ quan hệ QLCAP(TC, GTR, MC, NNSX, TNSX) không là dạng chuẩn 2NF, vì các thuộc tính không khóa NNSX và TNSX phụ thuộc hàm vào MC $\subset \{TC, MC\}$, tức là MC phụ thuộc không đầy đủ vào khóa {TC, MC}. Với cấu trúc lưu trữ của QLCAP, khi thực hiện các phép chèn thêm, loại bỏ hay sửa đổi dữ liệu sẽ xuất hiện dị thường thông tin. Không thể chèn thêm thông tin về một loại cáp mới khi chưa xác định giá trị của tuyếen cáp, vì TC không thể nhận giá trị không xác định. Tương tự, không thể xoá một tuyếen duy nhất có một loại cáp duy nhất, vì như vậy sẽ mất thông tin về tuyếen cáp và loại cáp đó. Có thể tách QLCAP thành 2 lược đồ quan hệ TUYEN và CAP_NSX về dạng chuẩn 2NF không mất thông tin. Cấu trúc lưu trữ của TUYEN và CAP_NSX phản ánh thế giới dữ liệu khách quan hơn, trung thực hơn so với cấu trúc của quan hệ QLCAP.

Ví dụ: Cho quan hệ R(A, B, C, D, E, G}

Và tập phụ thuộc hàm F

$$\begin{aligned} & \{ \\ & AB \rightarrow C \\ & D \rightarrow EG \\ & C \rightarrow A \\ & BE \rightarrow C \\ & BC \rightarrow D \\ & CG \rightarrow BD \\ & ACD \rightarrow B \\ & CE \rightarrow AG \\ & \} \end{aligned}$$

Các khóa của lược đồ quan hệ gồm:

$$K_1 = \{AB\}$$

$$K_2 = \{BE\}$$

$$K_3 = \{CG\}$$

$$K_4 = \{CE\}$$

$$K_5 = \{CD\}$$

$$K_6 = \{BC\}.$$

Như vậy không tồn tại các thuộc tính không khóa, vì vậy lược đồ quan hệ trên là dạng chuẩn 2NF.

5.2.6. Dạng chuẩn thứ ba 3NF

Cho lược đồ quan hệ được $R(A_1, A_2, \dots, A_n)$ có tập thuộc tính là A gọi là dạng chuẩn 3NF, khi và chỉ khi không tồn tại phụ thuộc hàm $X \rightarrow Y \in F^+$ sao cho $X^+ \neq A$, $Y \not\subseteq X$ và Y là thuộc tính không khóa. Nói cách khác nếu $X \rightarrow Y \in F^+$, $Y \not\subseteq X$ thì khi đó hoặc X là khóa của lược đồ quan hệ hoặc Y là một thuộc tính của khóa.

Trong lược đồ 2NF, cấm tất cả các thuộc tính không khóa phụ thuộc vào các tập con thực sự của khóa. Trong dạng chuẩn 3NF, cấm các thuộc tính không khóa phụ thuộc hàm vào tất cả các tập thuộc tính có bao đóng khác A.

Các khẳng định sau là tương đương:

- a) Không tồn tại phụ thuộc hàm $X \rightarrow Y \in F^+$ sao cho $X^+ \neq A$, $Y \not\subseteq X$ và Y là thuộc tính không khóa.
- b) Nếu $X \rightarrow Y \in F^+$, $Y \not\subseteq X$ khi đó hoặc X là khóa của lược đồ quan hệ hoặc Y là một thuộc tính của khóa.
- c) Không tồn tại thuộc tính không khóa phụ thuộc bắc cầu vào khóa.
- d) Tập các thuộc tính không khóa bằng rỗng.

Như vậy một lược đồ quan hệ dạng chuẩn 3NF thì cũng là dạng chuẩn 2NF.

Ví dụ: Cho các quan hệ NHASX (nhà sản xuất) và TP (thành phố) như sau:

NHASX(MSNSX, Von, ThanhPho)

MSNSX: Mã số nhà sản xuất.

Von: Vốn của nhà sản xuất.

ThanhPho: Thành phố nhà sản xuất.

TP(ThanhPho, QuocGia)

ThanhPho: Thành phố.

QuocGia: Quốc gia của thành phố.

Hai quan hệ trên có các phụ thuộc hàm sau:

MSNSV → Von, ThanhPho

ThanhPho → QuocGia

Hai quan hệ NSX và TP đạt chuẩn 3NF.

Từ một lược đồ quan hệ 2NF nhưng không phải 3NF được biến đổi thành một họ tương đương các lược đồ quan hệ dạng 3NF. Quá trình chuyển đổi là quá trình khả nghịch, do đó không làm mất thông tin trong quá trình chuyển đổi. Các lược đồ quan hệ kết quả sau biến đổi được gọi là các lược đồ quan hệ chiếu. Nhóm các lược đồ quan hệ chiếu tương đương với lược đồ quan hệ ban đầu theo nghĩa là lược đồ quan hệ ban đầu có thể nhận được bằng cách thực hiện các phép kết nối các lược đồ quan hệ chiếu tương ứng. Như vậy quá trình biến đổi sẽ không làm mất thông tin. Thông tin của lược đồ quan hệ ban đầu có thể nhận được từ các lược đồ quan hệ chiếu.

Ví dụ: cho quan hệ

SINHVIEN (MSSV, HoTen, Nganh, Khoa)

Quan hệ SINHVIEN không thỏa chuẩn 3NF vì Nganh → Khoa

Nếu quan hệ SINHVIEN tách thành hai quan hệ

SV (MSSV, HoTen, Nganh)

NGANH (Nganh, Khoa)

Quan hệ SV và NGANH thỏa dạng chuẩn 3NF.

5.2.7. Dạng chuẩn BOYCE-CODD - BCNF

Khái niệm lược đồ quan hệ dạng chuẩn 1NF và dạng chuẩn 2NF là những khái niệm trung gian trên chặng đường đi tới chuẩn 3NF. Tuy nhiên với khái niệm chuẩn 3NF chưa thỏa mãn cho các trường hợp lược đồ quan hệ có hơn một khóa, đặc biệt các trường hợp khóa giao nhau. Heath là người đầu tiên, sau đó Boyce Codd đã đưa ra định nghĩa dạng chuẩn Boyce - Codd có thể thỏa mãn các yêu cầu trên. Khái niệm dạng chuẩn Boyce Codd không dựa vào các khái niệm dạng chuẩn 1NF và 2NF, tức là không dựa vào các khái niệm về phụ thuộc hàm đầy đủ và phụ thuộc bắc cầu.

Lược đồ quan hệ $R(A_1, A_2, \dots, A_n)$ có tập thuộc tính A, được gọi là lược đồ dạng chuẩn Boyce - Codd (BCNF), nếu với mọi phụ thuộc $X \rightarrow Y \in F^+$, thì khi đó hoặc $Y \subseteq X$ (phụ thuộc tầm thường), hoặc X là một khóa của lược đồ quan hệ. Tức là nếu $X \rightarrow Y \in F^+, Y \notin X^+ \text{ thì } X^+ = A$. Từ định nghĩa trên có thể suy ra rằng:

- a) Các thuộc tính không khóa phụ thuộc hoàn toàn vào khóa.
- b) Các thuộc tính khóa phụ thuộc hoàn toàn vào tất cả khóa khác.

Có thể định nghĩa cách khác như sau: $R(A_1, A_2, \dots, A_n)$ được gọi là dạng chuẩn Boyce - Codd, khi và chỉ khi mọi định thuộc là khóa của lược đồ quan hệ.

Các khẳng định sau là tương đương

- a) Nếu $X \rightarrow Y \in F^+, Y \notin X$ khi đó X là khóa của lược đồ quan hệ.
- b) Nếu $X \rightarrow Y \in F^+, Y \notin X$ khi đó $X^+ = R$.
- c) Mọi định thuộc là khóa của lược đồ quan hệ. (X là định thuộc khi và chỉ khi tồn tại $Y \subseteq R$ sao cho $X \rightarrow Y \in F^+$ là phụ thuộc đầy đủ).

Định nghĩa dạng chuẩn 3NF và BCNF giống nhau trừ mệnh đề ‘hoặc Y là một thuộc tính của khóa’. Điều này có nghĩa là nếu lược đồ quan hệ dạng chuẩn BCNF thì cũng là 3NF. Trong định nghĩa 3NF, loại trừ các thuộc tính không khóa phụ thuộc vào các thuộc tính có bao đóng khác A, còn trong định nghĩa BCNF thì loại trừ tất cả các thuộc tính phụ thuộc vào các thuộc tính có bao đóng khác A.

Các dạng chuẩn lược đồ quan hệ lồng nhau, nghĩa là nếu lược đồ quan hệ dạng chuẩn BCNF thì cũng là dạng chuẩn 3NF, nếu đã là 3NF cũng là dạng chuẩn 2NF và nếu 2NF cũng là dạng chuẩn 1NF. Ký hiệu $BCNF \subset 3NF \subset 2NF \subset 1NF$.

Ví dụ 1: Cho Q (A, B, C, D, E, I)

$$F = \{ACD \rightarrow EBI, CE \rightarrow AD\}$$

- Xác định khóa của Q có hai khóa là {ACD} và {CE}
- Chia nhỏ về phải của các phụ thuộc hàm:

$$ACD \rightarrow E$$

$$ACD \rightarrow B$$

$$ACD \rightarrow I$$

$$CE \rightarrow A$$

$$CE \rightarrow D$$

Giáo trình Cơ sở dữ liệu

- Mọi phụ thuộc hàm trong F đều có vé trái là một siêu khóa

Vậy Q đạt dạng chuẩn Boyce - Codd.

Ví dụ 2: Cho sơ đồ quan hệ

LAIXE(Xe, NguoiLai)

TAINAN(Xe, NanNhan, DiaDiem, ThoiDiem)

Có các phụ thuộc hàm:

Xe, NanNhan → DiaDiem

Xe, NanNhan → ThoiDiem

Xe → NguoiLai

Là hai quan hệ thỏa dạng chuẩn Boyce – Codd.

5.2.8. Dạng chuẩn thứ tư (4NF)

Dạng chuẩn thứ tư là dạng mở rộng của chuẩn Boyce-Codd, áp dụng cho các sơ đồ quan hệ có các phụ thuộc hàm đa trị.

Cho quan hệ R và D là các tập phụ trên R. Ta nói R là dạng chuẩn thứ tư nếu ta có phụ thuộc hàm đa trị $X \rightarrow Y$ bất kỳ, trong đó Y khác rỗng hoặc không là tập con của X và X hợp Y không chứa hết thuộc tính của R, thì X phải là một siêu khóa của R.

Như vậy, ta nhận xét rằng khi D chỉ chứa các phụ thuộc hàm, nếu R ở dạng chuẩn thứ tư thì nó cũng ở dạng chuẩn Boyce-Codd. Từ đó, ta cũng suy ra khi D chỉ chứa các phụ thuộc hàm, nếu R không ở dạng chuẩn Boyce-Codd thì cũng không ở dạng chuẩn thứ tư.

Ví dụ: Cho quan hệ TKBieu(ThoiGian, Phong, Mon) với phụ thuộc hàm đa trị

Mon → ThoiGian, Phong.

Phụ thuộc đa trị này không vi phạm 4NF vì hội các thuộc tính hiện diện trên đã chứa hết tất cả các thuộc tính của TKBieu.

BÀI TẬP CHƯƠNG 5

Bài 1.

Cho lược đồ quan hệ $\alpha = (U, F)$ với

$U = ABCD$, $F = \{D \rightarrow B, C \rightarrow A, B \rightarrow ACD\}$

Xác định dạng chuẩn cao nhất của lược đồ quan hệ trên

Bài 2.

Cho lược đồ quan hệ $\alpha = (U, F)$ với

$U = ABCD$, $F = \{CD \rightarrow B, A \rightarrow C, B \rightarrow ACD\}$

Xác định dạng chuẩn cao nhất của lược đồ quan hệ trên

Bài 3.

Cho (U, F) với

$U = XYZW$ và tập

$F = \{Y \rightarrow W, W \rightarrow Y, XY \rightarrow Z\}$

Dạng chuẩn cao nhất của lược đồ là gì?

Bài 4.

Cho (U, F) với

$U = ABCDEG$ và tập phụ thuộc hàm

$F = \{AB \rightarrow C, AC \rightarrow E, EG \rightarrow D, AB \rightarrow G\}$

$\delta = \{DEG, ABDEG\}$

Phép tách trên có mất thông tin không?

Hãy chứng minh mọi quan hệ chỉ có 2 thuộc tính đè ở dạng chuẩn BCNF?

Bài 5.

Xét quan hệ $R(ABCDE)$ và tập phụ thuộc hàm

$F = \{AB \rightarrow CE, E \rightarrow AB, C \rightarrow D\}$

Hãy tìm dạng chuẩn cao nhất của lược đồ?

Bài 6.

Xét quan hệ $R(ABCDEF)$ và tập phụ thuộc hàm

$F = \{A \rightarrow B, C \rightarrow DG, AC \rightarrow E, D \rightarrow G\}$

Hãy tìm khóa của lược đồ

Hãy tìm dạng chuẩn cao nhất của lược đồ

Bài 7.

Xét quan hệ $R(ABCD)$ và tập phụ thuộc hàm

$F = \{AB \rightarrow D, AC \rightarrow BD, B \rightarrow C\}$

Hãy tìm dạng chuẩn cao nhất của lược đồ

Bài 8.

Cho $\alpha = (U, F)$ với

$U = ABCDEF$

$F = \{AB \rightarrow C, C \rightarrow B, ABD \rightarrow E, F \rightarrow A\}$

Lược đồ có ở dạng BCNF không

Bài 9.

Cho tập các thuộc tính quản lý tổng đài- cước thuê bao:

MaTD : Mã tổng đài

TenTD: Tên tổng đài

DL: Dung lượng

MC: Màu cáp

Giáo trình Cơ sở dữ liệu

SDT: Số điện thoại thuê bao

NSX: Nước sản xuất

NG: Ngày lắp đặt tổng đài

GTR: Giá trị tổng đài

VT: Nơi đặt tổng đài

TenTB: Tên thuê bao

MaTB: Mã thuê bao

DC: Địa chỉ thuê bao

SDTD: Số điện thoại đến/đi

NGG: Ngày thực hiện đàm thoại

GioB: Giờ bắt đầu đàm thoại

GioKT: Giờ kết thúc đàm thoại

DD: Cho biết trang thại cuộc gọi đang thực hiện là cuộc gọi đi hay cuộc gọi đến

Hãy xác định một phép tách – kết nối không mất thông tin về BCNF.

Bài 10.

Cho tập các thuộc tính quản lý phát hành báo chí QLBC gồm các thuộc tính:

MaK: Mã khách

TenK: Tên khách

DC: Địa chỉ

MaB: Mã báo

TenB: Tên báo

Gia: Giá báo

SL: số lượng

KY: kỳ phát hành

a) Hãy xác định một phép tách – kết nối không mất thông tin về 3NF có bảo toàn phụ thuộc hàm.

b) Hãy xác định một phép tách – kết nối không mất thông tin về BCNF.

Bài 11.

Giả sử các thuộc tính (TenBC, MaBC, MaKhach, DC, NgNhan, ST, DCN, BCN, NG, HD)

TenBC: Bưu cục

MaBC: Mã số bưu cục

MaKhach: Mã số của khách gửi

DC: Địa chỉ khách

NgNhan: Người nhận

DCN: Địa chỉ người nhận

ST: Số tiền

BCN: Bưu cục nhận

BCD: Mã bưu cục đến

NG: Ngày gửi

HD: Hóa đơn.

a) Hãy xác định một phép tách – kết nối không mất thông tin có bảo toàn phụ thuộc hàm ở dạng chuẩn 3NF.

b) Hãy xác định một phép tách – kết nối không mất thông tin ở dạng chuẩn 3NF.

TÀI LIỆU THAM KHẢO

- [1] PGS.TSKH. Trần Quốc Chiến - Giáo trình Cơ sở dữ liệu - NXB Đà Nẵng
- [2] Huỳnh Văn Đức - Cơ sở dữ liệu - NXB lao động.
- [3] Nguyễn Xuân Huy - Hệ CSDL quan hệ - ĐH Tổng Hợp.
- [4] Phạm Thị Xuân Lộc - Bài giảng Cơ sở dữ liệu
Khoa Công nghệ thông tin, Trường Đại học Cần Thơ.
- [5] Phạm Đức Nhiệm - Giáo trình Cơ sở dữ liệu quan hệ - NXB Hà Nội 2005
- [6] Trần Phuoc Tuân - Bài giảng Đại số quan hệ - Khoa Toán – Tin
Trường ĐH.SP.TPHCM
- [7] Đỗ Trung Tuân - Cơ sở dữ liệu - NXB Đại học quốc gia Hà Nội
- [8] Lê Tiến Vương - Nhập môn Cơ sở dữ liệu quan hệ
NXB Khoa học và Kỹ thuật, 1995.

MỤC LỤC

Chương 1: TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU	1
1.1. Mở đầu	1
1.1.1. Định nghĩa cơ sở dữ liệu	1
1.1.2. Định nghĩa môn cơ sở dữ liệu	1
1.1.3. Hệ quản trị cơ sở dữ liệu (DataBase Management System - DBMS)	1
1.1.4. Ưu điểm của cơ sở dữ liệu	2
1.1.5. Các mức biểu diễn một CSDL	2
1.1.6. Tính độc lập của dữ liệu.....	3
1.2. Các mô hình của cơ sở dữ liệu	3
1.2.1. Mô hình phân cấp (Hierachical Model):	4
1.2.2. Mô hình mạng (Network Model)	4
1.2.3. Mô hình quan hệ (Relational Model)	4
1.2.4. Mô hình thực thể quan hệ (Entity Relationship Model):	11
1.2.5. Mô hình dữ liệu hướng đối tượng (Object Oriented Data Model).....	12
1.3. Ngôn ngữ trên dữ liệu.....	13
1.3.1. Ngôn ngữ tự nhiên và ngôn ngữ hình thức	13
1.3.2. Ngôn ngữ mô tả dữ liệu (Data Definition Language – DDL).....	13
1.3.3. Ngôn ngữ thao tác dữ liệu (Data Manipulation Language – DML)	14
1.3.4. Các dạng thể hiện của ngôn ngữ	14
1.3.6. Quản lý giao tác	15
BÀI TẬP CHƯƠNG 1	16
Chương 2: ĐẠI SỐ QUAN HỆ	18
2.1 Giới thiệu	18
2.2. Các phép toán cơ bản	18
2.2.1 Phép chọn (Selection)	18
2.2.2 Phép chiếu (Projection)	20
2.2.3 Quan hệ tương thích.....	21
2.2.4 Phép trừ (Minus)	22
2.2.5 Phép hợp (Union)	23

2.2.6 Phép tích Decartes	24
2.3 Các phép toán khác	25
2.3.1 Phép giao (Intersection).....	25
2.3.2 Phép chia (Division)	26
2.3.3 Phép kết nối (Join).....	27
2.3.4 Phép kết nối tự nhiên (Natural Join).....	28
2.3.5 Phép kết nối mở rộng (Outer Join)	29
2.3.6 Phép đặt tên lại cho biểu thức và thuộc tính.....	32
2.3.7 Các hàm kết tập	33
2.4 Một số khái niệm bổ sung.....	40
2.4.1 Đại số hệ	40
2.4.2 Đại số quan hệ	41
2.4.3 Biểu thức quan hệ	41
BÀI TẬP CHƯƠNG 2.....	41
Chương 3: NGÔN NGỮ SQL	45
3.1 Giới thiệu ngôn ngữ SQL.....	45
3.1.1 Sự xuất hiện của SQL.....	45
3.1.2 Các tiện ích SQL	46
3.2 Ngôn ngữ mô tả dữ liệu.....	46
3.2.1 Tạo cơ sở dữ liệu	48
3.2.2 Tạo cấu trúc cho bảng.....	49
3.2.3 Thay đổi cấu trúc bảng	55
3.2.4 . Xóa các thành phần trong cơ sở dữ liệu	57
3.3 Ngôn ngữ thao tác.....	58
3.3.1 Thêm dữ liệu vào bảng	59
3.3.2 Xóa dữ liệu ra khỏi bảng	60
3.3.3 Cập nhật dữ liệu trong bảng	62
3.4 Truy vấn dữ liệu	63
3.4.1 Tìm kiếm đơn giản	63
3.4.2 Tìm kiếm phần tử duy nhất (Loại bỏ các phần tử trùng nhau).....	67
3.4.3 Tìm kiếm tất cả các phần tử	67

3.4.4 Tìm kiếm có sắp xếp	69
3.4.5 Hạn chế số lượng mẫu tin	70
3.5 Các hàm tổng hợp	71
3.5.1 Hàm SUM	71
3.5.2 Hàm AVG	72
3.5.3 Hàm COUNT	73
3.5.4 Hàm MAX.....	73
3.5.5 Hàm MIN	74
3.5.6 Tính toán theo nhóm	75
3.5.7 Phát biểu SELECT lồng nhau	79
3.5.8 Truy vấn trên tập hợp	81
3.6 Khung nhìn (View).....	83
3.6.1 Tạo khung nhìn	83
3.6.2 Xóa khung nhìn	84
3.7 Các lệnh về an toàn dữ liệu	84
3.7.1 Tạo quyền truy xuất	84
3.7.2 Hủy bỏ quyền truy xuất.....	85
3.7.3 Các hàm thông dụng	85
BÀI TẬP CHƯƠNG 3	86
PHẦN 1: NGÔN NGỮ MÔ TẢ	86
I. NỘI DUNG THỰC HIỆN.....	86
II. LÝ THUYẾT	86
1. Giới thiệu sơ lược hệ quản trị Cơ sở dữ liệu SQL Server 2008.....	86
2. Các kiểu dữ liệu (data type) thông dụng	88
3. Tạo và sử dụng CSDL.....	89
4. Thao tác trên bảng.....	93
5. Hướng dẫn thực hành	95
III. THỰC HÀNH	96
Bài 1. Cơ sở dữ liệu về Quản lý giải bóng đá (<i>QLBongDa</i>)	96
Bài 2. Cơ sở dữ liệu QLCongTrinh (Quản lý công trình).....	99
Bài 3. Cơ sở dữ liệu QLGiangDay như sau:	100

BÀI 4 : CSDL quản lý hàng hóa như sau :	100
Bài 5 : Mô tả CSDL quản lý học viên như sau.....	102
PHẦN II NGÔN NGỮ THAO TÁC	105
I. NỘI DUNG THỰC HIỆN	105
II LÝ THUYẾT	105
1. Thêm dữ liệu.....	105
2. Xóa dữ liệu	105
3 Cập nhật dữ liệu trong bảng	105
III THỰC HÀNH	105
Bài 1: Sử dụng CSDL QLBongDa (ở phần I) thực hiện các yêu cầu sau:	105
BÀI 2 : Sử dụng lại CSDL quản lý hàng hóa để thực hiện :	108
KHACHHANG	108
SANPHAM.....	108
HOADON	109
SOHD.....	110
BÀI 3 : Sử dụng CSDL quản lý học viên (ở phần I)	110
PHẦN III TRUY VẤN DỮ LIỆU	115
I. NỘI DUNG HƯỚNG DẪN THỰC HÀNH	115
II. LÝ THUYẾT	116
III THỰC HÀNH	116
Bài 1: Sử dụng CSDL quản lý hàng hóa để trả lời câu hỏi sau:	116
Bài 2: Sử dụng CSDL quản lý học viên để trả lời câu hỏi sau:.....	118
Bài 3: cho CSDL về quản lý chuyến bay như sau:	119
Bài 4: cho cơ sở dữ liệu về quản lý đề án như sau:	124
IV. MỘT SỐ HÀM THÔNG DỤNG	126
1. Xử lý chuỗi	126
2.Xử lý ngày	127
3. Các hàm xử lý số	128
4. Hàm chuyển đổi kiểu.....	129
5. Cấu trúc phân nhánh CASE.....	129
V. CÁC LỖI THƯỜNG GẶP TRONG SQL.....	129

Chương 4: RÀNG BUỘC TOÀN VẸN và CÁC LOẠI PHỤ THUỘC DỮ LIỆU.....	131
 4.1 Ràng buộc toàn vẹn (RBTV).....	131
4.1.1 Khái niệm RBTV	131
4.1.2 Các yếu tố của RBTV	131
4.1.3 Hành động cần phải có khi phát hiện có RBTV bị vi phạm	132
 4.2 Phân loại RBTV	133
4.2.1 Ràng buộc toàn vẹn có bối cảnh là một quan hệ.....	133
4.2.2 Ràng buộc toàn vẹn có bối cảnh là nhiều quan hệ	134
 4.3. Các loại phụ thuộc dữ liệu.....	137
4.3.1 Các vấn đề gặp phải khi tổ chức dữ liệu	137
4.3.2 Phụ thuộc hàm.....	138
4.3.3 Phủ tối thiểu (minimal cover)	142
 4.4 Phép tách các sơ đồ quan hệ	144
 4.5 Kết nối không mất thông tin	144
 4.6 Kiểm tra kết nối không mất thông tin.....	145
 4.7 Phép tách bảo toàn phụ thuộc	148
 4.8 Kiểm tra sự bảo toàn phụ thuộc hàm	149
BÀI TẬP CHƯƠNG 4	150
Chương 5: CHUẨN HÓA SƠ ĐỒ QUAN HỆ	157
 5.1 Khóa của sơ đồ quan hệ	157
5.1.1 Khóa (Key).....	157
5.1.2 Siêu khóa	158
5.1.3 Thuộc tính khóa và thuộc tính không khóa.....	159
5.1.4 Thuật toán tìm khóa	159
 5.2 Các dạng chuẩn của sơ đồ quan hệ	161
5.2.1 Thuộc tính nguyên tố, không nguyên tố	163
5.2.2 Các dạng phụ thuộc hàm	163
5.2.3 Sự chuẩn hóa một quan hệ	164
5.2.4 Dạng chuẩn thứ nhất - 1NF	164
5.2.5 Dạng chuẩn thứ hai – 2NF	164

5.2.6 Dạng chuẩn thứ ba 3NF.....	165
5.2.7 Dạng chuẩn BOYCE-CODD - BCNF.....	167
5.2.8 Dạng chuẩn thứ tư (4NF).....	168
BÀI TẬP CHƯƠNG 5.....	168
TÀI LIỆU THAM KHẢO	172