BỘ CÔNG THƯƠNG

TRƯỜNG CAO ĐẲNG THƯƠNG MẠI VÀ DU LỊCH

GIÁO TRÌNH

MÔN HỌC: CƠ SỞ DỮ LIỆU

NGÀNH: CÔNG NGHỆ THÔNG TIN (ƯDPM)

TRÌNH ĐỘ: TRUNG CẤP

(Ban hành kèm theo Quyết định số: 405/QĐ-CĐKT ngày 05 tháng 07 năm 2022 của Trường Cao đẳng Thương mại và Du lịch)

Thành phố Thái Nguyên, năm 2022

(Lưu hành nội bộ)

TUYÊN BỐ BẢN QUYỀN

Tài liệu này thuộc loại sách giáo trình nên các nguồn thông tin có thể được phép dùng nguyên bản hoặc trích dùng cho các mục đích về đào tạo và tham khảo.

Mọi mục đích khác mang tính lệch lạc hoặc sử dụng với mục đích kinh doanh thiếu lành mạnh sẽ bị nghiêm cấm.

LỜI GIỚI THIỆU

"Môn học 'Cơ sở dữ liệu' là một môn học quan trọng và thiết yếu trong lĩnh vực công nghệ thông tin và quản lý thông tin. Trong môn học này, chúng ta sẽ khám phá cách tổ chức, truy xuất và quản lý dữ liệu trong một hệ thống thông tin hiện đại.

Chúng ta sẽ bắt đầu với một lời giới thiệu sâu rộng về cơ sở dữ liệu và lý do tại sao chúng đóng vai trò quan trọng trong thế giới số hóa ngày nay. Chúng ta sẽ tìm hiểu về mô hình dữ liệu, cách biểu diễn thông tin, và tạo quan hệ giữa các dữ liệu.

Qua môn học, bạn sẽ nắm vững ngôn ngữ truy vấn dữ liệu, SQL, và làm quen với các phương pháp thiết kế cơ sở dữ liệu. Chúng ta sẽ xem xét cách tối ưu hóa cơ sở dữ liệu để đảm bảo hiệu suất tốt nhất.

Cuối cùng, bạn sẽ hiểu về quản trị cơ sở dữ liệu, bao gồm bảo mật, sao lưu và phục hồi dữ liệu. Khóa học này sẽ là nền tảng quan trọng cho sự phát triển trong các ngành công nghệ, quản trị doanh nghiệp, và khoa học dữ liệu. Chúng ta sẽ bắt đầu hành trình khám phá thế giới của cơ sở dữ liệu cùng với bạn."

Tài liệu học tập được biên soạn theo đúng chương trình đào tạo và các quy định về cách trình bày của Nhà trường. Nội dung của tài liệu học tập bao gồm các chương, trong mỗi chương bao gồm các phần nội dung chủ yếu như sau:

- Mục tiêu của chương.
- Nội dung bài giảng lý thuyết.
- Bài tập vận dụng.

Nhằm tạo điều kiện cho người học có một bộ tài liệu tham khảo mang tính tổng hợp, thống nhất và mang tính thực tiễn sâu hơn. Nhóm người dạy chúng tôi đề xuất và biên soạn ***Giáo trình Tin học văn phòng*** dành riêng cho người học trình độ Trung cấp.

Nội dung của giáo trình bao gồm các chương sau:

Chương 1: Đại cương về các hệ cơ sở dữ liệu

Chương 2: Các mô hình quan hệ

Chương 3: Ngôn ngữ thao tác dữ liệu

Chương 4: Hệ quản trị cơ sở dữ liệu SQL

Trong quá trình biên soạn, chúng tôi đã tham khảo và trích dẫn từ nhiều tài liệu được liệt kê tại mục Danh mục tài liệu tham khảo. Chúng tôi chân thành cảm ơn các tác giả của các tài liệu mà chúng tôi đã tham khảo.

Bên cạnh đó, giáo trình cũng không thể tránh khỏi những sai sót nhất định. Nhóm tác giả rất mong nhận được những ý kiến đóng góp, phản hồi từ quý đồng nghiệp, các bạn người học và bạn đọc.

Trân trọng cảm ơn./.

Thành phố Thái Nguyên, ngày 20 tháng 08 năm 2022

MỤC LỤC

LỜI GIỚI THIỆU 2
CHƯƠNG 1. ĐẠI CƯƠNG VỀ CÁC HỆ CƠ SỞ DỮ LIỆU 13
1.1. Các hệ thống xử lý tệp truyền thống 14
1.2. Hệ quản trị cơ sở dữ liệu 16
1.2.1. Các khái niệm cơ bản 16
1.2.2. Các khả năng của một hệ quản trị cơ sở dữ liệu 17
1.2.3. Kiến trúc của một hệ quản trị cơ sở dữ liệu 18
1.2.4. Người quản trị cơ sở dữ liệu 20
1.2.5. Người sử dụng cơ sở dữ liệu 20
1.3. Phân loại các hệ cơ sở dữ liệu 20
1.3.1. Hệ cơ sở dữ liệu tập trung 20
1.3.2. Hệ cơ sở dữ liệu phân tán 22
CHƯƠNG 2. CÁC MÔ HÌNH QUAN HỆ 25
2.1. Mô hình thực thể liên kết (mô hình ER) 26
2.1.1. Các khái niệm cơ bản 26
2.1.2. Sơ đồ thực thể liên kết 31
2.1.3. Tính năng của các liên kết 31
2.2. Mô hình dữ liệu liên kết 32
2.2.1. Các khái niệm cơ bản 32
2.2.2. Chuyển sơ đồ ER sang mô hình quan hệ 33
2.3. Mô hình dữ liệu mạng 36
2.3.1. Các khái niệm cơ bản 36
2.3.2. Chuyển sơ đồ ER sang mô hình mạng 37
2.4. Mô hình dữ liệu phân cấp 38
2.4.1. Các khái niệm cơ bản 38
2.5. Mô hình dữ liệu hướng đối tượng 43
2.5.1. Các khái niệm cơ bản 43
2.5.2. Chuyển sơ đồ ER sang mô hình dữ liệu hướng đối tượng 43
CHƯƠNG 3. NGÔN NGỮ THAO TÁC DỮ LIỆU 47
3.1. Đại số quan hệ 48
3.1.1. Phép hợp 48
3.1.2. Phép giao 49

3.1.3. Phép trừ 50
3.1.4. Tích Đề các 50
3.1.5. Phép chiếu 51
3.1.6. Phép chọn 52
3.1.7. Phép kết nối 54
3.1.8. Phép chia 57
3.2. Khóa của sơ đồ quan hệ 58
3.2.1. Siêu khóa (Super Key) 58
3.2.2. Khóa (Key) 58
3.2.3. Khóa dự tuyển (Candidate Key) – Khóa chính (Primary Key) 60
3.2.4. Khóa ngoại (Foreign Key) 61
3.3. Bài tập ứng dụng 61
3.3.1. Phép toán tập hợp và phép toán quan hệ 61
3.3.2. Mô hình thực thể kết hợp 62
CHƯƠNG 4. HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU SQL 65
4.1. Giới thiệu 66
4.2. Các lệnh truy vấn SELECT 67
4.2.1. Câu lệnh tìm kiếm tổng quát 67
4.2.2. Tìm thông tin từ các cột của bảng - Mệnh đề SELECT 68
4.2.3. Chọn các dòng của bảng - Mệnh đề WHERE 70
4.2.4. Kết nối nhiều bảng dữ liệu 72
4.2.5. Sắp xếp các dòng của bảng - Mệnh đề ORDER BY: 72
4.2.6. Các lệnh truy vấn lồng nhau 73
4.2.7. Gom nhóm dữ liệu - Mệnh đề GROUP BY 74
4.2.8. Hàm tính toán theo nhóm 75
4.3. Các lệnh truy vấn UPDATE 75
4.4. Các lệnh truy vấn INSERT 77

GIÁO TRÌNH MÔN HỌC

1. Tên môn học: CƠ SỞ DỮ LIỆU

2. Mã môn học: MH13

3. Vị trí, tính chất, ý nghĩa và vai trò của môn học:

3.1. Vị trí: Giáo trình dành cho người học trình độ Trung cấp tại trường Cao đẳng Thương mại và Du lịch.

3.2. Tính chất: Giáo trình cung cấp kiến thức, kỹ năng và năng lực tự chủ và trách nhiệm cho người học. Môn học "Cơ sở dữ liệu" tập trung vào việc thiết kế, truy vấn và quản lý dữ liệu, kết hợp lý thuyết với thực hành và sử dụng ngôn ngữ SQL. Nó giúp phát triển kiến thức và kỹ năng liên quan đến quản lý dữ liệu, thiết kế cơ sở dữ liệu, và tối ưu hóa hiệu suất hệ thống, có tính ứng dụng cao trong nhiều lĩnh vực: (1) có bộ giáo trình phù hợp với chương trình đào tạo của trường; (2) dễ dàng tiếp thu cũng như vận dụng các kiến thức và kỹ năng được học vào môi trường học tập và thực tế thuộc lĩnh vực văn ph.

3.3. Ý nghĩa và vai trò của môn học: Môn học "Cơ sở dữ liệu" có vai trò quan trọng trong việc hiểu cách quản lý và sử dụng dữ liệu một cách hiệu quả. Nó có tính ứng dụng cao trong công nghệ thông tin, quản lý doanh nghiệp, khoa học dữ liệu và nhiều lĩnh vực khác. Đồng thời, môn học này giúp học viên nắm vững kiến thức và kỹ năng cần thiết để tạo, quản lý, và bảo mật cơ sở dữ liệu..

4. Mục tiêu của môn học:

4.1. Về kiến thức:

A1. Hiểu về cơ sở dữ liệu, bao gồm các khái niệm cơ bản như mô hình thực thể-quan hệ, ngôn ngữ truy vấn dữ liệu, và quy trình thiết kế cơ sở dữ liệu.

A2. Hiểu cách sử dụng ngôn ngữ SQL để truy vấn, thêm, cập nhật và xóa dữ liệu từ cơ sở dữ liệu.

A4. Hiểu thiết kế cơ sở dữ liệu từ việc xác định yêu cầu, tạo bảng và quan hệ, đến việc tối ưu hóa cơ sở dữ liệu để đảm bảo hiệu suất tốt nhất

4.2 .Về kỹ năng:

B1. Biết về cơ sở dữ liệu, bao gồm các khái niệm cơ bản như mô hình thực thể-quan hệ, ngôn ngữ truy vấn dữ liệu, và quy trình thiết kế cơ sở dữ liệu.B2. Kỹ năng soạn thảo và biên tập văn bản

B3. Biết cách sử dụng ngôn ngữ SQL để truy vấn, thêm, cập nhật và xóa dữ liệu từ cơ sở dữ liệu.

B4. Biết thiết kế cơ sở dữ liệu từ việc xác định yêu cầu, tạo bảng và quan hệ, đến việc tối ưu hóa cơ sở dữ liệu để đảm bảo hiệu suất tốt nhất B5. Kỹ năng sử dụng email và quản lý tài liệu

4.3 Về năng lực tự chủ và trách nhiệm:

C1. Năng lực quản lý thời gian

C2. Trách nhiệm với công việc

C3. Năng lực học tập và làm việc độc lập

C4. Năng lực quản lý thông tin

C5. Tự chủ trong việc giải quyết vấn đề

5. Nội dung của môn học

5.1. Chương trình khung

Mã MH	**Tên môn học**	**Số tín chỉ**	**Thời gian học tập** *(giờ)*			
			Tổng số	**Trong đó**		
				Lý thuyết	**Thực hành/ thực tập/ bài tập/ thảo luận**	**Thi/ Kiểm tra**
I	**Các môn học chung**	**12**	**255**	**94**	**148**	**13**
MH01	Chính trị	2	30	15	13	2
MH02	Pháp luật	1	15	9	5	1
MH03	Giáo dục thể chất	1	30	4	24	2
MH04	Giáo dục quốc phòng và an ninh	2	45	21	21	3
MH05	Tin học	2	45	15	29	1
MH06	Ngoại ngữ	4	90	30	56	4
II	**Các môn học chuyên môn**	**64**	**1560**	**504**	**1013**	**43**
II.1	***Môn học cơ sở***	***16***	***240***	***224***	***-***	***13***
MH07	Tin học văn phòng	2	30	12	17	1

MH08	Bảng tính Excel	2	30	12	17	1
MH09	Cấu trúc máy tính	2	30	28	-	2
MH10	Mạng máy tính	2	30	15	14	1
MH11	Lập trình cơ bản	2	30	28	-	2
MH12	Cấu trúc dữ liệu và giải thuật	2	30	28	-	2
MH13	Cơ sở dữ liệu	2	30	28	-	2
MH14	Lắp ráp và bảo trì máy tính	2	30	28	-	2
II.2	***Môn học chuyên môn***	***46***	***1290***	***313***	***948***	***28***
MH15	Ngoại ngữ ch.ngành CNTT	4	60	57	-	3
MH16	Hệ điều hành Windows Server	2	30	28	-	2
MH17	Quản trị CSDL với Access 1	3	45	43	-	2
MH18	Quản trị CSDL với SQL Server	3	45	27	17	1
MH19	Lập trình Windows 1 (VB.NET)	3	45	43	-	2
MH20	Thiết kế và quản trị website	3	45	43	-	2
MH21	Đồ họa ứng dụng	2	30	28	-	2
MH22	An toàn và bảo mật thông tin	2	30	28	-	2
MH23	TH xây dựng phần mềm quản lý	4	120	-	114	6
MH24	TH thiết kế và quản trị website	4	120	-	114	6
MH25	Thực tập tốt nghiệp	16	720	-	720	
II.3	***Môn học tự chọn****(chọn 1 trong 2)*	***2***	***30***	***28***	***-***	***2***
MH26	Kỹ năng giao tiếp, phục vụ khách hàng	2	30	28	-	2
MH27	Lập trình mạng	2	30	28	-	2
	Tổng cộng	**76**	**1815**	**598**	**1161**	**56**

5.2. Chương trình chi tiết môn học

Số TT	Tên chương, mục	Thời gian (giờ)			
		Tổng số	Lý thuyết	Thực hành	Kiểm tra
1	Chương 1: Đại cương về các hệ cơ sở dữ liệu	2	2		
2	Chương 2: Các mô hình quan hệ	6	6		
3	Chương 3: Ngôn ngữ thao tác dữ liệu	12	11		1
4	Chương 4: Hệ quản trị cơ sở dữ liệu SQL	10	9		1
	Cộng	**30**	**28**		**2**

6. Điều kiện thực hiện môn học:

6.1. Phòng học Lý thuyết/Thực hành: Đáp ứng phòng học chuẩn

6.2. Trang thiết bị dạy học: Projetor, máy vi tính, bảng, phấn

6.3. Học liệu, dụng cụ, mô hình, phương tiện: Giáo trình, mô hình học tập,…

6.4. Các điều kiện khác: Người học tìm hiểu thực tế về công tác xây dựng phương án khắc phục và phòng ngừa rủi ro tại doanh nghiệp.

7. Nội dung và phương pháp đánh giá:

7.1. Nội dung:

- Kiến thức: Đánh giá tất cả nội dung đã nêu trong mục tiêu kiến thức
- Kỹ năng: Đánh giá tất cả nội dung đã nêu trong mục tiêu kỹ năng.
- Năng lực tự chủ và trách nhiệm: Trong quá trình học tập, người học cần:

+ Nghiên cứu bài trước khi đến lớp.

+ Chuẩn bị đầy đủ tài liệu học tập.

+ Tham gia đầy đủ thời lượng môn học.

+ Nghiêm túc trong quá trình học tập.

7.2. Phương pháp:

Người học được đánh giá tích lũy môn học như sau:

7.2.1. Cách đánh giá

Việc đánh giá kết quả học tập của người học được thực hiện theo quy định tại Thông tư 04/2022/TT-BLĐTBXH ngày 30/3/2022 của Bộ Lao động – Thương binh và Xã hội; Quy chế Tổ chức đào tạo trình độ trung cấp, trình độ cao đẳng theo phương thức tích luỹ mô-đun, tín chỉ của Nhà trường ban hành kèm theo Quyết định số 246/QĐ-CĐTMDL ngày 01/6/2022 của Hiệu trưởng Trường cao đẳng Thương mại và Du lịch và hướng dẫn cụ thể theo từng môn học/mô-đun trong chương trình đào tạo

Điểm đánh giá	**Trọng số**
+ Điểm kiểm tra thường xuyên (Hệ số 1)	40%
+ Điểm kiểm tra định kỳ (Hệ số 2)	
+ Điểm thi kết thúc môn học	60%

7.2.2. Phương pháp đánh giá

Phương pháp đánh giá	Phương pháp tổ chức	Hình thức kiểm tra	Thời điểm kiểm tra
Thường xuyên	Viết/ Thuyết trình	Tự luận/ Trắc nghiệm	Sau 20 giờ.
Định kỳ	Viết/ Thuyết trình	Tự luận/ Trắc nghiệm	Sau 28 giờ
Kết thúc môn học	Viết	Tự luận và trắc nghiệm	Sau 30 giờ

7.2.3. Cách tính điểm

- Điểm đánh giá thành phần và điểm thi kết thúc môn học được chấm theo thang điểm 10 (từ 0 đến 10), làm tròn đến một chữ số thập phân.

- Điểm môn học là tổng điểm của tất cả điểm đánh giá thành phần của môn học nhân với trọng số tương ứng. Điểm môn học theo thang điểm 10 làm tròn đến một chữ số thập phân, sau đó được quy đổi sang điểm chữ và điểm số theo thang điểm 4 theo quy định của Bộ Lao động Thương binh và Xã hội về đào tạo theo tín chỉ.

8. Hướng dẫn thực hiện môn học

8.1. Phạm vi, đối tượng áp dụng: Đối tượng Trung cấp Công nghệ thông tin (ƯDPM)

8.2. Phương pháp giảng dạy, học tập môn học

8.2.1. Đối với người dạy

*** Lý thuyết:** Áp dụng phương pháp dạy học tích cực bao gồm: thuyết trình ngắn, nêu vấn đề, hướng dẫn đọc tài liệu, bài tập tình huống, câu hỏi thảo luận….

*** Bài tập:** Phân chia nhóm nhỏ thực hiện bài tập theo nội dung đề ra.

*** Thảo luận:** Phân chia nhóm nhỏ thảo luận theo nội dung đề ra.

*** Hướng dẫn tự học theo nhóm:** Nhóm trưởng phân công các thành viên trong nhóm tìm hiểu, nghiên cứu theo yêu cầu nội dung trong bài học, cả nhóm thảo luận, trình bày nội dung, ghi chép và viết báo cáo nhóm.

8.2.2. Đối với người học: Người học phải thực hiện các nhiệm vụ như sau:

- Nghiên cứu kỹ bài học tại nhà trước khi đến lớp. Các tài liệu tham khảo sẽ được cung cấp nguồn trước khi người học vào học môn học này (trang web, thư viện, tài liệu...)
- Tham dự tối thiểu 80% các buổi giảng lý thuyết. Nếu người học vắng >20% số tiết lý thuyết phải học lại môn học mới được tham dự kì thi lần sau.
- Tham dự đủ các bài kiểm tra thường xuyên, định kỳ.
- Tham dự thi kết thúc môn học.
- Chủ động tổ chức thực hiện giờ tự học.

9. Tài liệu tham khảo:

[[1].Bùi Thị Hòa. *Bài giảng cơ sở dữ liệu*. Nhà xuất bản Bách Khoa – Hà Nội, 2012

[2].Hồ thuần, Nguyễn Quang Vinh, Nguyễn Xuân Huy, *Nhập môn các hệ cơ sở dữ liệu (sách dịch)*, Nhà xuất bản Thống kê, 1986.

[3].Vũ Đức Thi, *Cơ sở dữ liệu- Kiến thức và thực hành*, Nhà xuất bản Thống kê, 1997.

[4].Nguyễn Bá Tường, *Cơ sở dữ liệu- Lý thuyết và thực hành*, Nhà xuất bản Khoa học và Kỹ thuật, 2001.

[5].Nguyễn Kim Anh, *Nguyên lý các hệ cơ sở dữ liệu*, Nhà xuất bản ĐHQG Hà Nội, 2004.

CHƯƠNG 1. ĐẠI CƯƠNG VỀ CÁC HỆ CƠ SỞ DỮ LIỆU

❖ GIỚI THIỆU CHƯƠNG 1

Chương 1, "Đại cương về hệ cơ sở dữ liệu," giới thiệu về cơ sở dữ liệu và lý do tại sao chúng quan trọng trong quản lý thông tin. Chương này định nghĩa các khái niệm cơ bản và giới thiệu lịch sử và loại cơ sở dữ liệu. Cuối cùng, nó tập trung vào lợi ích của cơ sở dữ liệu trong cuộc sống hàng ngày.

❖ MỤC TIÊU CHƯƠNG 1

Sau khi học xong chương này, người học có khả năng:

➢ ***Về kiến thức:***

- *Hiểu được khái niệm cơ sở dữ liệu và tầm quan trọng của nó trong việc lưu trữ và quản lý thông tin. Kiến thức về chèn hình ảnh và bảng biểu*
- *Hiểu các khái niệm quan trọng như dữ liệu, thông tin, hệ quản trị cơ sở dữ liệu (DBMS), và mô hình dữ liệu*
- *Hiểu hiểu về lịch sử và sự phát triển của cơ sở dữ liệu, cùng với việc tìm hiểu về các hệ quản trị cơ sở dữ liệu phổ biến*

➢ ***Về kỹ năng:***

- *Biết áp dụng các khái niệm cơ bản về cơ sở dữ liệu, bao gồm dữ liệu, thông tin, hệ quản trị cơ sở dữ liệu (DBMS), và mô hình dữ liệu.*
- *Biết đánh giá và so sánh các loại cơ sở dữ liệu khác nhau, như cơ sở dữ liệu quan hệ và cơ sở dữ liệu NoSQL*
- *Biết tư duy về lịch sử và phát triển của cơ sở dữ liệu, cùng với khả năng tìm hiểu và hiểu về các hệ quản trị cơ sở dữ liệu phổ biến*

➢ ***Về năng lực tự chủ và trách nhiệm:***

- *Năng lực về quản lý thời gian, trách nhiệm với công việc*
- *Năng lực học tập và làm việc độc lập*
- *Tự chủ trong việc giải quyết vấn đề*

❖ PHƯƠNG PHÁP GIẢNG DẠY VÀ HỌC TẬP CHƯƠNG 1

- *Đối với người dạy: sử dụng phương pháp giảng giảng dạy tích cực (diễn giảng, vấn đáp, dạy học theo vấn đề); yêu cầu người học thực hiện câu hỏi thảo luận và bài tập chương (cá nhân hoặc nhóm).*

- *Đối với người học: chủ động đọc trước giáo trình trước buổi học; hoàn thành đầy đủ câu hỏi thảo luận và bài tập tình huống theo cá nhân hoặc nhóm và nộp lại cho người dạy đúng thời gian quy định.*

❖ **ĐIỀU KIỆN THỰC HIỆN CHƯƠNG 1**

- ***Phòng học chuyên môn hóa/nhà xưởng:*** Phòng học thực hành
- ***Trang thiết bị máy móc:*** Máy chiếu, máy tính và các thiết bị dạy học khác
- ***Học liệu, dụng cụ, nguyên vật liệu:*** Chương trình môn học, giáo trình, tài liệu tham khảo, giáo án, phim ảnh, và các tài liệu liên quan.
- ***Các điều kiện khác:*** Không có

❖ **KIỂM TRA VÀ ĐÁNH GIÁ CHƯƠNG 1**

- **Nội dung:**

✓ *Kiến thức: Kiểm tra và đánh giá tất cả nội dung đã nêu trong mục tiêu kiến thức*

✓ *Kỹ năng: Đánh giá tất cả nội dung đã nêu trong mục tiêu kĩ năng.*

✓ *Năng lực tự chủ và trách nhiệm: Trong quá trình học tập, người học cần:*

+ Nghiên cứu bài trước khi đến lớp

+ Chuẩn bị đầy đủ tài liệu học tập.

+ Tham gia đầy đủ thời lượng môn học.

+ Nghiêm túc trong quá trình học tập.

- **Phương pháp:**

✓ ***Điểm kiểm tra thường xuyên:*** *1 điểm kiểm tra (hình thức: hỏi miệng)*

✓ ***Kiểm tra định kỳ lý thuyết:*** *không có*

❖ **NỘI DUNG**

1.1. Các hệ thống xử lý tệp truyền thống

Trong những năm gần đây, thuật ngữ "Cơ sở dữ liệu" (CSDL - Database) đã trở nên khá quen thuộc không chỉ riêng với những người làm Tin học mà còn đối với cả những người làm trong nhiều lĩnh vực khác như Thống kê, Kinh tế, Quản lý Doanh nghiệp v.v… Các ứng dụng của Tin học vào công tác quản lý ngày càng nhiều hơn và càng đa dạng hơn. Có thể nói hầu hết các lĩnh vực kinh tế, xã hội, giáo dục, y tế v.v… đều đã ứng dụng các thành tựu mới của Tin học vào phục vụ công tác chuyên môn của mình. Chính vì lẽ đó mà ngày càng nhiều người quan tâm đến lĩnh vực thiết kế và xây dựng các CSDL.

Mục đích của Chương 1 chỉ đơn giản là cung cấp các khái niệm cơ bản về CSDL để các học viên có một cái nhìn ban đầu về một CSDL và một hệ quản trị CSDL. Trước hết chúng ta sẽ tìm hiểu lý do tại sao cần phải có một CSDL.

Hệ thống các tập tin truyền thống (File System):

Cho đến nay vẫn còn một số đơn vị kinh tế, hành chính sự nghiệp v.v… sử dụng mô hình hệ thống các tập tin cổ điển: Chúng được tổ chức riêng rẽ, phục vụ cho một mục đích của một đơn vị hay một đơn vị con trực thuộc cụ thể. Chẳng hạn, ta hãy xét ví dụ sau:

Ví dụ: Tại một công ty người ta trang bị máy vi tính cho tất cả các phòng, ban nghiệp vụ. Bộ phận Văn phòng sử dụng máy vi tính để soạn thảo văn bản bằng Microsoft Word do thủ trưởng yêu cầu về tình hình hoạt động của đơn vị, trong đó có chỉ tiêu về tổng số công nhân viên chức chia theo trình độ chuyên môn được đào tạo. Phòng Kế toán sử dụng máy vi tính để tính lương và in danh sách lương của từng bộ phận trong đơn vị dựa trên danh sách cán bộ viên chức cùng hệ số lương và các hệ số phụ cấp của họ do phòng Tổ chức cung cấp. Thông tin mà phòng Kế toán quản lý và khai thác là: Họ và Tên, Hệ số lương, Hệ số phụ cấp, Phụ cấp khác của các công nhân viên chức (CNVC) xếp theo từng phòng ban và sử dụng công cụ văn phòng là Microsoft Excel. Phòng Tổ chức quản lý thông tin lý lịch của CNVC chi tiết hơn gồm: Họ, Tên (để riêng thành một cột "Tên" để tiện sắp xếp Alphabet), Giới tính, Ngày sinh, Ngày tuyển dụng, Hoàn cảnh gia đình, Quá trình đào tạo, Hệ số lương, Hệ số phụ cấp, Ngày xếp lương trên… nhưng thiếu thông tin về Phụ cấp khác của CNVC. Phần mềm được sử dụng để quản lý là FoxPro for Windows.

Trong khi đó, tại Tổng công ty của họ, các phòng ban nghiệp vụ cũng được trang bị máy vi tính. Phòng Tổ chức cán bộ tại Tổng công ty sử dụng phần mềm Microsoft Access để quản lý CNVC gồm các cán bộ chủ chốt từ trường phó phòng, quản đốc và phó quản đốc xí nghiệp trở lên của các công ty con trực thuộc. Thông tin quản lý tại đây cũng giống như thông tin quản lý tại phòng tổ chức của công ty con.

Nhận xét:

+ Ưu điểm:

- Việc xây dựng hệ thống các tập tin riêng tại từng đơn vị quản lý ít tốn thời gian bởi khối lượng thông tin cần quản lý và khai thác là nhỏ, không đòi hỏi đầu tư vật chất và chất xám nhiều, do đó triển khai ứng dụng nhanh.
- Thông tin được khai thác chỉ phục vụ cho mục đích hẹp nên khả năng đáp ứng nhanh chóng, kịp thời.

+ Nhược điểm:

- Do thông tin được tổ chức ở mỗi phòng ban mỗi khác, cũng như phần mềm công cụ để triển khai mỗi nơi cũng rất khác nhau nên sự phối hợp tổ chức và khai thác ở các phòng ban là khó khăn. Thông tin ở phòng ban này không sử dụng được cho phòng ban khác, tại đơn vị con với đơn vị cấp trên. Cùng một thông tin được nhập vào máy tại nhiều nơi khác nhau gây ra lãng phí công sức nhập tin và không gian lưu trữ trên các vật mang tin. Sự trùng lắp thông tin có thể dẫn đến tình trạng không nhất quán dữ liệu. Chẳng hạn, nhân viên Nguyễn Văn Quang được ghi đầy đủ ở phòng Tổ chức, nhưng tại phòng Kế toán chỉ ghi tắt là Nguyễn V Quang.
- Thông tin được tổ chức ở nhiều nơi nên việc cập nhật cũng dễ làm mất tính nhất quán dữ liệu. Một cán bộ chủ chốt của công ty có thay đổi về hoàn cảnh gia đình (mới cưới vợ / lấy chồng, sinh thêm con…) có thể được cập nhật ngay tại đơn vị nhưng sau một thời gian mới được cập nhật tại Tổng công ty.
- Do hệ thống được tổ chức thành các hệ thống tập tin riêng lẻ nên thiếu sự chia sẻ thông tin giữa các nơi. Việc kết nối các hệ thống này hay việc nâng cấp ứng dụng sẽ là rất khó khăn.

Qua phân tích trên chúng ta nhận thấy việc tổ chức dữ liệu theo hệ thống các tập tin có nhiều nhược điểm. Việc xây dựng một hệ thống tin đảm bảo được tính chất nhất quán dữ liệu, không trùng lặp thông tin mà vẫn đáp ứng được nhu cầu khai thác đồng thời của tất cả các phòng ban ở công ty và tổng công ty là thực sự cần thiết.

1.2. Hệ quản trị cơ sở dữ liệu

1.2.1. Các khái niệm cơ bản

Như chúng ta đã biết, kích thước và độ phức tạp của CSDL rất khác nhau.

Ví dụ:

- Danh bạ điện thoại của một quốc gia, một thành phố… chứa tới hàng triệu số và những thông tin cần thiết về khách hàng.
- Trong trường Đại học có tới hàng ngàn sinh viên. Nhà trường phải quản lý tất cả những thông tin liên quan đến sinh viên như: Tên, ngày sinh, quê quán, địa chỉ, kết quả học tập…

Ta thấy, bộ nhớ là vấn đề cần phải được giải quyết. Tuy nhiên, vấn đề quan trọng hơn ở đây lại là cách thức tổ chức dữ liệu trong một CSDL để phục vụ cho việc truy cập, tìm kiếm, cập nhật… nhanh chóng và an toàn hơn.

Để giải quyết tốt tất cả các vấn đề đặt ra cho một CSDL như đã nêu trên, cần phải có một hệ thống các phần mềm chuyên dụng được gọi là hệ quản trị CSDL (Database

Management System - DBMS). Đó là các công cụ hỗ trợ tích cực cho các nhà phân tích và thiết kế CSDL và những người khai thác CSDL.

Các khái niệm:

Định nghĩa một CSDL: Tạo cấu trúc dữ liệu tương ứng với mô hình dữ liệu được chọn, bao gồm việc đặc tả các kiểu dữ liệu, các cấu trúc và các ràng buộc cho các dữ liệu sẽ được lưu trữ trong CSDL.

Xây dựng một CSDL: Là quá trình lưu trữ các dữ liệu trên các phương tiện lưu trữ được hệ quản trị CSDL kiểm soát.

Thao tác một CSDL: Bao gồm các chức năng như truy vấn CSDL để lấy ra các dữ liệu cụ thể, cập nhật CSDL để phản ánh các thay đổi trong thế giới nhỏ và tạo ra các báo cáo từ các dữ liệu.

Một hệ quản trị CSDL là một tập hợp các chương trình giúp cho người sử dụng quản lý cấu trúc và dữ liệu của CSDL, điều khiển truy xuất dữ liệu trong CSDL, duy trì và khai thác một CSDL.

Cho đến nay có khá nhiều hệ quản trị CSDL mạnh được đưa ra thị trường như: Visual FoxPro, Microsoft Access, SQL Server, DB2, Sybase, Paradox, Informix, Oracle… với các chất lượng khác nhau.

1.2.2. Các khả năng của một hệ quản trị cơ sở dữ liệu

Một hệ quản trị CSDL thực hiện các chức năng sau:

Cung cấp môi trường tạo lập CSDL

Một hệ quản trị CSDL phải cung cấp một môi trường cho người dùng dễ dàng khai báo kiểu dữ liệu, các cấu trúc dữ liệu thể hiện thông tin và các ràng buộc trên dữ liệu. Với các hệ quản trị CSDL hiện nay, người dùng có thể tạo lập CSDL thông qua các giao diện đồ hoạ. Để thực hiện được chức năng này, mỗi hệ quản trị CSDL cung cấp cho người dùng một ngôn ngữ định nghĩa dữ liệu. Ta có thể hiểu ngôn ngữ định nghĩa dữ liệu thực chất là hệ thống các kí hiệu để mô tả CSDL.

Cung cấp môi trường cập nhật và khai thác dữ liệu

Ngôn ngữ để người dùng diễn tả yêu cầu cập nhật hay tìm kiếm, kết xuất thông tin được gọi là ngôn ngữ thao tác dữ liệu.

Thao tác dữ liệu gồm:

- Cập nhật: Nhập thêm, sửa, xoá dữ liệu.
- Khai thác: Tìm kiếm, kết xuất dữ liệu.

Trong thực tế, ngôn ngữ định nghĩa dữ liệu và ngôn ngữ thao tác dữ liệu là 2 thành phần của một ngôn ngữ CSDL duy nhất. Ngôn ngữ CSDL được sử dụng phổ biến hiện nay là SQL (Structured Query Language).

Cung cấp công cụ kiểm soát, điều khiển truy cập vào CSDL

Để góp phần đảm bảo được các yêu cầu đặt ra cho một hệ CSDL, hệ quản trị CSDL phải có các bộ chương trình thực hiện những nhiệm vụ sau:

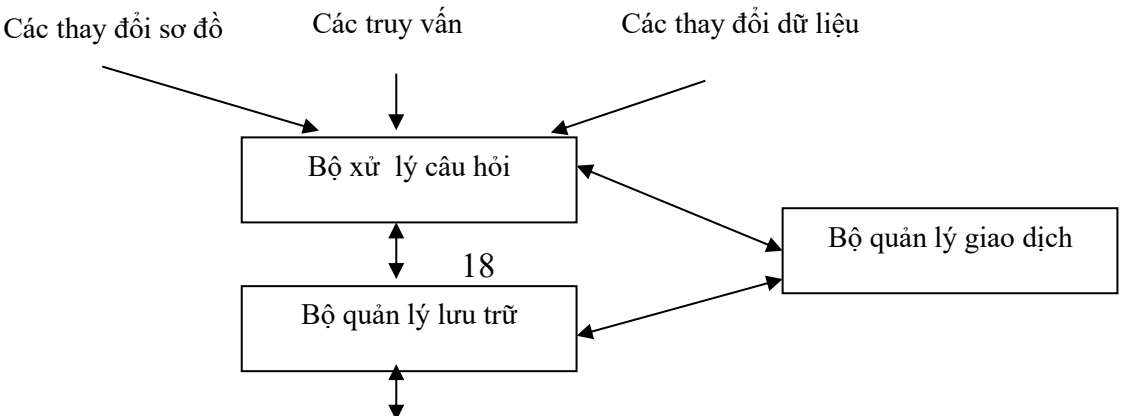
- Phát hiện và ngăn chặn sự truy cập không được phép: Chức năng này góp phần đáp ứng yêu cầu an toàn và bảo mật thông tin.
- Duy trì tính nhất quán của dữ liệu.
- Tổ chức và điều khiển các truy cập đồng thời để bảo vệ các ràng buộc toàn vẹn và tính nhất quán.
- Khôi phục CSDL khi có sự cố ở phần cứng hay phần mềm.
- Quản lí các mô tả dữ liệu.

Nói chung, mọi hệ quản trị CSDL đều có thể cung cấp các chương trình nêu trên, nhưng các hệ quản trị CSDL khác nhau có chất lượng và khả năng khác nhau khi đáp ứng các nhu cầu thực tế. Chẳng hạn, một hệ quản trị CSDL dùng trên một máy tính cá nhân chỉ cung cấp những phương tiện bảo vệ dữ liệu, duy trì tính nhất quán dữ liệu, khôi phục dữ liệu một cách hạn chế. Trong khi đó, những hệ quản trị CSDL lớn, phục vụ cùng lúc cho nhiều người dùng, ngoài các phương tiện kể trên có thể có thêm những chức năng khác như xử lí các truy cập đồng thời…

Các hệ quản trị CSDL luôn phát triển theo hướng đáp ứng các đòi hỏi ngày càng cao hơn của người dùng, bởi vậy các chức năng của hệ quản trị CSDL ngày càng được mở rộng hơn.

1.2.3. Kiến trúc của một hệ quản trị cơ sở dữ liệu

Chúng ta sẽ phác thảo kiến trúc và thấy cách thức của một hệ quản trị CSDL điển hình. Ta có sơ đồ kiến trúc hình 1.1:



Hình 1.1. Các thành phần chính của hệ quản trị CSDL

- Dữ liệu, siêu dữ liệu: Đáy kiết trúc là thiết bị nhớ ngoài lưu trữ dữ liệu và siêu dữ liệu. Trong phần này không chỉ chứa dữ liệu được trữ trong CSDL mà chứa cả các siêu dữ liệu, tức là thông tin cấu trúc của CSDL. Ví dụ: Trong hệ quản trị cơ sở dữ liệu quan hệ, các siêu dữ liệu bao gồm các tên của các quan hệ, tên các thuộc tính của các quan hệ, và các kiểu dữ liệu đối với các thuộc tính này.

Bộ quản lý lưu trữ: Nhiệm vụ của bộ quản lý lưu trữ là lấy ra các thông tin được yêu cầu từ những thiết bị lưu trữ dữ liệu và thay đổi những thông tin này khi được yêu cầu bởi các mức trên nó của hệ thống.

Bộ xử lý câu hỏi: Bộ xử lý câu hỏi điều khiển không chỉ các câu hỏi mà cả các yêu cầu thay đổi dữ liệu hay siêu dữ liệu. Nhiệm vụ của nó là tìm ra cách tốt nhất một thao tác được yêu cầu và phát ra lệnh đối với bộ quản lý lưu trữ và thực thi thao tác đó.

Bộ quản trị giao dịch: Bộ quản trị giao dịch có trách nhiệm đảm bảo tính toàn vẹn của hệ thống. Phải đảm bảo rằng một số thao tác thực hiện đồng thời không cản trở mỗi thao tác khác và hệ thống không mất dữ liệu thậm chí cả khi lỗi hệ thống xảy ra.

+ Tương tác với bộ xử lý câu hỏi, do vậy nó phải biết dữ liệu nào được thao tác bởi các thao tác hiện thời để tránh sự đụng độ giữa các thao tác và cần thiết nó có thể làm trễ một số truy vấn nhất định hay một số thao tác cập nhật để đụng độ không thể xảy ra.

+ Tương tác với bộ quản lý lưu trữ bởi vì các sơ đồ đối với việc bảo vệ dữ liệu thường kéo theo việc lưu trữ một nhật ký các thay đổi đối với dữ liệu. Hơn nữa, việc sắp thứ tự các thao tác một cách thực sự được nhật ký này sẽ chứa trong một bản ghi đối với mỗi thay đổi khi gặp lỗi hệ thống, các thay đổi chưa được ghi vào đĩa có thể được thực hiện lại.

Các kiểu thao tác đối với hệ quản trị CSDL: Tại đỉnh kiến trúc, ta thấy có 3 kiểu thao tác:

+ Các truy vấn: Đây là các thao tác hỏi đáp về dữ liệu được lưu trữ trong CSDL. Chúng được sinh ra theo hai cách sau:

Thông qua giao diện truy vấn chung. Ví dụ: Hệ quản trị CSDL quan hệ cho phép người sử dụng nhập các câu lệnh truy vấn SQL mà nó được chuyển qua bộ xử lý câu hỏi và được trả lời.

Thông qua các giao diện chương trình ứng dụng: Một hệ quản trị CSDL điển hình cho phép người lập trình viết các chương trình ứng dụng gọi đến hệ quản trị CSDL này và truy vấn CSDL.

+ Các cập nhật dữ liệu: Đây là các thao tác thay đổi dữ liệu như xoá, sửa dữ liệu trong CSDL. Giống như các truy vấn, chúng có thể được phát ra thông qua giao diện chung hoặc thông qua giao diện của chương trình.

+ Các thay đổi sơ đồ: Các lệnh này thường được phát bởi một người sử dụng được cấp phép, thường là những người quản trị CSDL mới được phép thay đổi sơ đồ của CSDL hay tạo lập một CSDL mới.

1.2.4. Người quản trị cơ sở dữ liệu

Người quản trị CSDL: Đây là những người hiểu biết về tin học, về các hệ quản trị CSDL và hệ thống máy tính. Họ là người tổ chức CSDL (khai báo cấu trúc CSDL, ghi nhận các yêu cầu bảo mật cho các dữ liệu cần bảo vệ…), do đó họ phải nắm rõ các vấn đề kỹ thuật về CSDL để có thể phục hồi dữ liệu khi có sự cố. Họ là những người cấp quyền hạn khai thác CSDL, do vậy họ có thể giải quyết được các vấn đề tranh chấp dữ liệu, nếu có.

1.2.5. Người sử dụng cơ sở dữ liệu

Người sử dụng CSDL không chuyên về lĩnh vực tin học và CSDL, do đó CSDL cần có các công cụ để cho những người sử dụng không chuyên có thể sử dụng để khai thác CSDL khi cần thiết.

1.3. Phân loại các hệ cơ sở dữ liệu

1.3.1. Hệ cơ sở dữ liệu tập trung

Với hệ CSDL tập trung, toàn bộ dữ liệu được lưu trữ tại một máy hoặc một dàn máy. Những người dùng từ xa có thể truy cập vào CSDL thông qua các phương tiện truyền thông dữ liệu. Nói chung có ba kiểu kiến trúc tập trung:

a. Hệ CSDL cá nhân

Hệ CSDL cá nhân là hệ CSDL có một người dùng, người này vừa thiết kế, vừa tạo lập, vừa cập nhật và bảo trì CSDL, đồng thời cũng là người khai thác thông tin, tự lập và hiển thị các báo cáo.

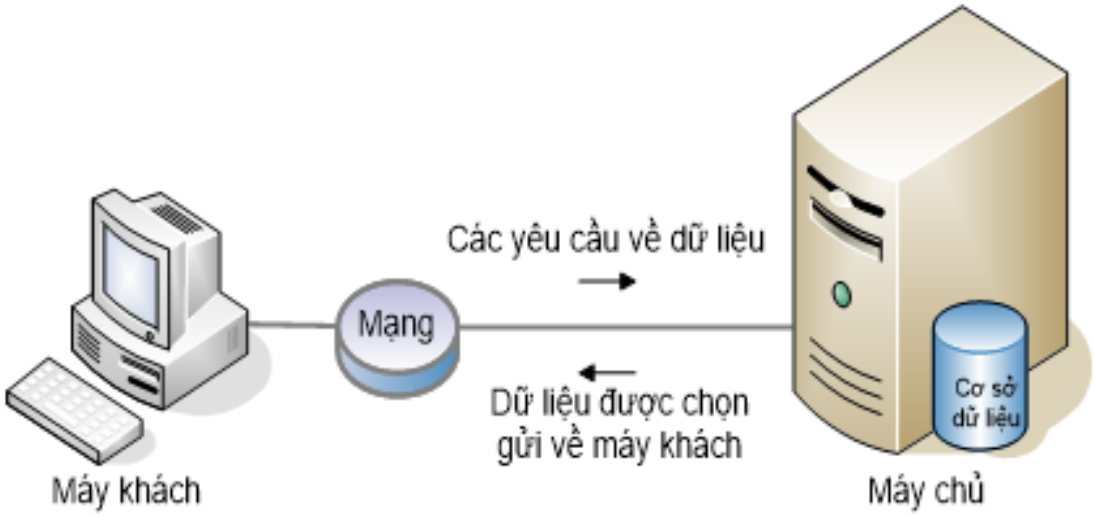
b. Hệ CSDL trung tâm

Hệ CSDL trung tâm là hệ CSDL với dữ liệu được lưu trữ trên máy tính trung tâm, nhiều người sử dụng từ xa có thể truy cập CSDL này thông qua các thiết bị đầu cuối và các phương tiện truyền thông. Tùy thuộc vào quy mô của tổ chức, máy tính trung tâm này là một dàn máy hay một máy. Các hệ CSDL trung tâm thường rất lớn và có nhiều người dùng, ví dụ các hệ thống đăng kí và bán vé máy bay, các hệ thống thông tin của tổ chức tài chính,...

c. Hệ CSDL khách - chủ

Trong kiến trúc khách-chủ, các thành phần (của hệ QTCSDL) tương tác với nhau tạo nên hệ thống gồm thành phần yêu cầu tài nguyên và thành phần cấp tài nguyên. Hai thành phần này không nhất thiết phải cài đặt trên cùng một máy tính.

Thành phần cấp tài nguyên thường được cài đặt tại một máy chủ trên mạng (cục bộ) Còn thành phần yêu cầu tài nguyên có thể cài đặt tại nhiều máy khác trên mạng (ta gọi là các máy khách).



Hình 1.2. Mô hình khách - chủ

Phần mềm CSDL trên máy khách quản lí các giao diện khi thực hiện chương trình. Kiến trúc loại này có một số ưu điểm sau:

Khả năng truy cập rộng rãi đến các CSDL.

Nâng cao khả năng thực hiện: các CPU ở máy chủ và máy khách khác nhau có thể cùng chạy song song, mỗi CPU thực hiện nhiệm vụ của riêng nó.

Chi phí cho phần cứng có thể được giảm do chỉ cần máy chủ có cấu hình đủ mạnh để lưu trữ và quản trị CSDL.

Chi phí cho truyền thông được giảm do một phần các thao tác được giải quyết trên máy khách, chỉ cần: yêu cầu về truy cập CSDL gửi đến máy chủ và dữ liệu kết quả gửi về cho máy khách.

Nâng cao khả năng đảm bảo tính nhất quán của dữ liệu vì các ràng buộc được định nghĩa và kiểm tra chỉ tại máy chủ.

Kiến trúc này phù hợp với việc mở rộng các hệ thống.

1.3.2. Hệ cơ sở dữ liệu phân tán

a. Khái niệm CSDL phân tán

CSDL phân tán là những hệ thống cho phép người dùng truy cập không chỉ dữ liệu đặt tại chỗ mà cả những dữ liệu để ở xa.

CSDL phân tán là một tập hợp dữ liệu có liên quan (về logic) được dùng chung và phân tán về mặt vật lí trên một mạng máy tính. Một hệ QTCSDL phân tán là một hệ thống phần mềm cho phép quản trị CSDL phân tán và làm cho người sử dụng không nhận thấy sự phân tán về lưu trữ dữ liệu. Người dùng truy cập vào CSDL phân tán thông quan chương trình ứng dụng. Các chương trình ứng dụng được chia làm hai loại:

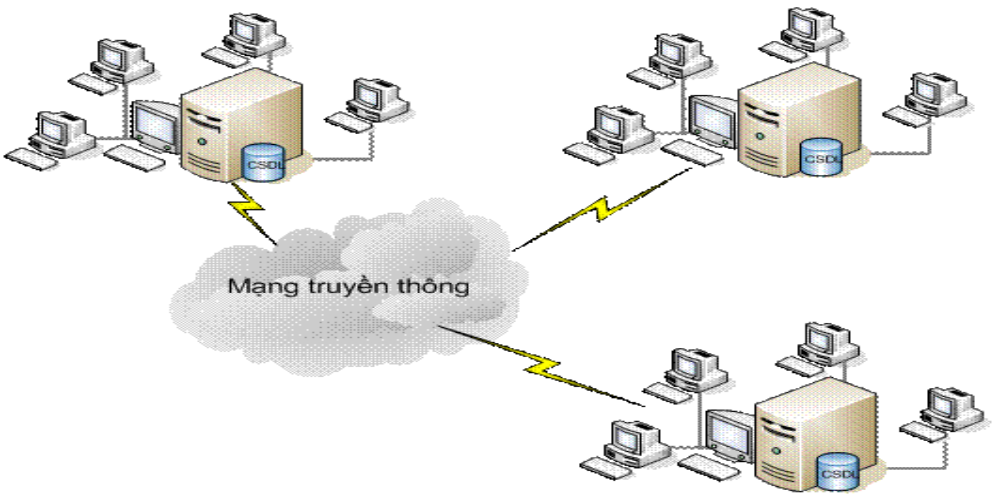
Chương trình không yêu cầu dữ liệu từ nơi khác.

Chương trình có yêu cầu dữ liệu từ nơi khác.

Có thể chia các hệ CSDL phân tán thành 2 loại chính: thuần nhất và hỗn hợp.

Hệ CSDL phân tán thuần nhất: các nút trên mạng đều dùng cùng một hệ QTCSDL.

Hệ CSDL phân tán hỗn hợp: các nút trên mạng có thể dùng các hệ QTCSDL khác nhau.



Hình 1.3. Mô hình CSDL phân tán

b. Một số ưu điểm và hạn chế của các hệ CSDL phân tán

Sự phân tán dữ liệu và các ứng dụng có một số ưu điểm so với các hệ CSDL tập trung:

Cấu trúc phân tán dữ liệu thích hợp cho bản chất phân tán của nhiều người dùng.

Dữ liệu được chia sẻ trên mạng nhưng vẫn cho phép quản trị dữ liệu địa phương (dữ liệu đặt tại mỗi trạm).

Dữ liệu có tính sẵn sàng cao.

Dữ liệu có tính tin cậy cao vì khi một nút gặp sự cố, có thể khôi phục được dữ liệu tại đây do bản sao của nó có thể được lưu trữ tại một nút khác nữa.

Hiệu năng của hệ thống được nâng cao hơn.

Cho phép mở rộng các tổ chức một cách linh hoạt. Có thể thêm nút mới vào mạng máy tính mà không ảnh hưởng đến hoạt động của các nút sẵn có.

So với các hệ CSDL tập trung, hệ CSDL phân tán có một số hạn chế như sau:

Hệ thống phức tạp hơn vì phải làm ẩn đi sự phân tán dữ liệu đối với người dùng.

Chi phí cao hơn.

Đảm bảo an ninh khó khăn hơn.

Đảm bảo tính nhất quán dữ liệu khó hơn.

Việc thiết kế CSDL phân tán phức tạp hơn

❖ TÓM TẮT CHƯƠNG 1

Trong chương này, một số nội dung chính được giới thiệu:

1. **Tầm quan trọng của Cơ sở Dữ liệu**: Chương giới thiệu về tầm quan trọng của cơ sở dữ liệu trong việc lưu trữ, quản lý, và truy xuất thông tin..
2. **Khái niệm cơ bản**: Chương này định nghĩa các khái niệm quan trọng như dữ liệu (data), thông tin (information), hệ quản trị cơ sở dữ liệu (DBMS), và mô hình dữ liệu (data model).
3. **Lịch sử và Phát Triển**: Chương giới thiệu về lịch sử và sự phát triển của cơ sở dữ liệu. Nó cung cấp cái nhìn tổng quan về cách cơ sở dữ liệu đã phát triển qua thời gian, bao gồm các sự kiện quan trọng và tiến bộ trong lĩnh vực này.

❖ CÁC BÀI TẬP HỆ THỐNG KIẾN THỨC

1. **Khái niệm Cơ Sở Dữ liệu**: Hãy viết một bài luận ngắn về tầm quan trọng của cơ sở dữ liệu trong cuộc sống hàng ngày của bạn. Bạn có thể bao gồm ví dụ về cách cơ sở dữ liệu được sử dụng trong ứng dụng hoặc dự án cá nhân của bạn. Đảm bảo rõ ràng về lý do tại sao cơ sở dữ liệu là một yếu tố quan trọng trong việc quản lý thông tin.

2. **So sánh Loại Cơ Sở Dữ liệu**: Chọn hai loại cơ sở dữ liệu khác nhau (ví dụ: cơ sở dữ liệu quan hệ và NoSQL) và so sánh chúng dựa trên những ưu điểm và hạn chế

của từng loại. Tạo một biểu đồ hoặc bảng so sánh để minh họa điểm mạnh và điểm yếu của mỗi loại cơ sở dữ liệu.

3. **Lịch sử Cơ Sở Dữ liệu**: Tìm hiểu về lịch sử phát triển của cơ sở dữ liệu và viết một bài báo tóm tắt về các sự kiện quan trọng trong quá trình này. Bạn có thể bao gồm các sự kiện như sự ra đời của hệ quản trị cơ sở dữ liệu đầu tiên hoặc các cải tiến quan trọng trong lĩnh vực cơ sở dữ liệu.

CHƯƠNG 2. CÁC MÔ HÌNH QUAN HỆ

❖ GIỚI THIỆU CHƯƠNG 2

Chương 2, "Các mô hình quan hệ," tập trung vào nắm vững kiến thức về cơ sở dữ liệu quan hệ, một loại cơ sở dữ liệu phổ biến trong lĩnh vực quản lý thông tin. Chương này sẽ giới thiệu về các yếu tố cơ bản của cơ sở dữ liệu quan hệ, bao gồm các bảng, mối quan hệ giữa chúng, và ngôn ngữ SQL để truy vấn dữ liệu. Học viên sẽ học cách thiết kế bảng và quan hệ giữa chúng để lưu trữ thông tin một cách hiệu quả và tối ưu hóa cơ sở dữ liệu.

MỤC TIÊU CHƯƠNG 2

Sau khi học xong chương này, người học có khả năng:

➢ ***Về kiến thức:***

- *Hiểu về khái niệm cơ sở dữ liệu quan hệ và tại sao nó quan trọng trong việc quản lý thông tin*
- *Nắm vững cấu trúc cơ bản của bảng trong cơ sở dữ liệu quan hệ, bao gồm cách đặt tên, định dạng dữ liệu, và quan hệ giữa các bảng*
- *Hiểu cách sử dụng ngôn ngữ SQL để truy vấn dữ liệu từ cơ sở dữ liệu quan hệ. Điều này bao gồm việc viết câu truy vấn SELECT để trích xuất thông tin cụ thể.*
- *Hiểu cáchthiết kế bảng và quan hệ giữa chúng để lưu trữ thông tin một cách hiệu quả và tối ưu hóa cơ sở dữ liệu.*

➢ ***Về kỹ năng:***

- *Biết cách xác định các yếu tố quan trọng như khóa chính, khóa ngoại, và các ràng buộc dữ liệu để lưu trữ thông tin một cách cấu trúc và hiệu quả.*
- *Biết cách sử dụng ngôn ngữ SQL để thực hiện các truy vấn dữ liệu*
- *Biết cách đánh giá hiệu suất của cơ sở dữ liệu quan hệ và tối ưu hóa cơ sở dữ liệu để đảm bảo hoạt động một cách hiệu quả*

➢ ***Về năng lực tự chủ và trách nhiệm:***

- *Năng lực về quản lý thời gian, trách nhiệm với công việc*
- *Năng lực học tập và làm việc độc lập*
- *Tự chủ trong việc giải quyết vấn đề*

❖ PHƯƠNG PHÁP GIẢNG DẠY VÀ HỌC TẬP CHƯƠNG 2

- *Đối với người dạy: sử dụng phương pháp giảng giảng dạy tích cực (diễn giảng, vấn đáp, dạy học theo vấn đề); yêu cầu người học thực hiện câu hỏi thảo luận và bài tập chương (cá nhân hoặc nhóm).*

- *Đối với người học: chủ động đọc trước giáo trình trước buổi học; hoàn thành đầy đủ câu hỏi thảo luận và bài tập tình huống theo cá nhân hoặc nhóm và nộp lại cho người dạy đúng thời gian quy định.*

❖ **ĐIỀU KIỆN THỰC HIỆN CHƯƠNG 2**

- ***Phòng học chuyên môn hóa/nhà xưởng:*** Phòng học thực hành
- ***Trang thiết bị máy móc:*** Máy chiếu, máy tính và các thiết bị dạy học khác
- ***Học liệu, dụng cụ, nguyên vật liệu:*** Chương trình môn học, giáo trình, tài liệu tham khảo, giáo án, phim ảnh, và các tài liệu liên quan.
- ***Các điều kiện khác:*** Không có

❖ **KIỂM TRA VÀ ĐÁNH GIÁ CHƯƠNG 2**

- **Nội dung:**

✓ *Kiến thức: Kiểm tra và đánh giá tất cả nội dung đã nêu trong mục tiêu kiến thức*

✓ *Kỹ năng: Đánh giá tất cả nội dung đã nêu trong mục tiêu kĩ năng.*

✓ *Năng lực tự chủ và trách nhiệm: Trong quá trình học tập, người học cần:*

+ Nghiên cứu bài trước khi đến lớp

+ Chuẩn bị đầy đủ tài liệu học tập.

+ Tham gia đầy đủ thời lượng môn học.

+ Nghiêm túc trong quá trình học tập.

- **Phương pháp:**

✓ ***Điểm kiểm tra thường xuyên:*** *1 điểm kiểm tra (hình thức: hỏi miệng)*

✓ ***Kiểm tra định kỳ lý thuyết:*** *không có*

NỘI DUNG

2.1. Mô hình thực thể liên kết (mô hình ER)

2.1.1. Các khái niệm cơ bản

Mô hình quan hệ biểu thị CSDL như một tập các quan hệ. Mỗi quan hệ có thể được biểu diễn như một bảng giá trị, mỗi một dòng trong bảng biểu thị một tập hợp các giá trị dữ liệu liên quan với nhau. Tên bảng và tên các cột dùng để giúp giải thích ý nghĩa của các giá trị trong mỗi hàng. Mọi giá trị trong một cột đều cùng một kiểu dữ liệu.

Theo thuật ngữ mô hình quan hệ, mỗi hàng được gọi là một bộ, mỗi đầu cột được gọi là một thuộc tính, và bảng được gọi là một quan hệ. Kiểu dữ liệu mô tả các kiểu của dữ liệu xuất hiện trong mỗi cột gọi là một miền.

a. Thực thể (Entity)

Thực thể là một đối tượng cụ thể hay trừu tượng trong thế giới thực mà nó tồn tại và có thể phân biệt được với các đối tượng khác. Tên thực thể là các danh từ.

Ví dụ:

- Bạn Nguyễn Văn A là một thực thể cụ thể.
- Sinh viên là một thực thể trừu tượng.

b. Thuộc tính (Attribute)

Thuộc tính là các thông tin riêng biệt cần lưu trữ của các đối tượng.

Tập tất cả các thuộc tính của quan hệ R là R+

Chẳng hạn, với bài toán quản lý điểm thi của sinh viên:

- Thực thể Sinh viên có một số thuộc tính: Họ tên, ngày sinh, giới tính, tỉnh thường trú, học bổng, lớp đang theo học… các đặc trưng này gọi là thuộc tính của đối tượng sinh viên.
- Mỗi môn học có một số thuộc tính: Mã môn, Tên môn, Số đơn vị học trình.

Một thuộc tính hay một tập thuộc tính mà giá trị của nó xác định duy nhất mỗi thực thể trong tập các thực thể được gọi là khoá đối với tập thực thể này. (Khái niệm khoá sẽ được trình bày sau)

Mỗi thuộc tính có các thành phần: Tên gọi, kiểu dữ liệu, miền giá trị.

Tên gọi

Thường được đặt một cách ngắn gọn, gợi nhớ.

Ví dụ: Đối tượng SinhVien, và GiangVien đều có thuộc tính tên thì đặt là TenSV, TenGV.

Quy ước: Trong lý thuyết, nếu không cần lưu ý đến ngữ nghĩa, thì tên của các thuộc tính sẽ được ký hiệu bằng các chữ cái in hoa đầu tiên trong bảng chữ cái La-tinh: A, B, C, D…

Trong cùng một đối tượng không thể có 2 thuộc tính cùng tên.

Tên thuộc tính phải được đặt một cách gợi nhớ, không nên đặt tên thuộc tính quá dài (vì như thế sẽ làm cho việc viết các câu lệnh truy vấn trở nên vất vả hơn), nhưng cũng không nên đặt tên thuộc tính quá ngắn (vì nó sẽ không cho thấy ngữ nghĩa của thuộc tính).

Không đặt trùng tên 2 thuộc tính mang ngữ nghĩa khác nhau thuộc 2 đối tượng khác nhau.

Kiểu dữ liệu

Các kiểu dữ liệu thường được sử dụng trong các hệ quản trị CSDL:

- Text (hoặc Character, Char, String): Kiểu văn bản.
- Number (hoặc Numeric, Float): Kiểu số
- Yes/No (hoặc Boolean): Kiểu luận lý
- Date/Time: Kiểu thời gian (Ngày/tháng/năm + Giờ: phút: giây)
- Memo (hoặc VarChar): Kiểu văn bản có độ dài thay đổi.

Chẳng hạn, với sinh viên Nguyễn Văn Thành thì các thuộc tính Họ và tên, Mã số sinh viên thuộc kiểu chuỗi, thuộc tính Ngày sinh thuộc kiểu ngày tháng, Hộ khẩu thường trú kiểu chuỗi, thuộc tính Hình ảnh kiểu hình ảnh…

Mỗi hệ quản trị CSDL có thể gọi tên các kiểu dữ liệu nói trên bằng các tên gọi khác nhau, ngoài ra còn bổ sung thêm một số kiểu dữ liệu riêng của mình. Chẳng hạn, Microsoft Access có kiểu dữ liệu OLE để chứa các đối tượng nhúng như hình ảnh, âm thanh, video… ORACLE có kiểu dữ liệu LONG cho phép chứa dữ liệu có kích thước lớn tới 2 tỷ bytes.

Miền giá trị (Domain)

Tùy thuộc vào quy tắc quản lý, mỗi thuộc tính có thể chỉ chứa một số giá trị nào đó của kiểu dữ liệu. Tập giá trị hợp lệ của một thuộc tính được gọi là miền giá trị của thuộc tính.

Ký hiệu miền giá trị của một thuộc tính A: MGT(A) hay Dom(A)

Một số ví dụ về định nghĩa miền:

- Họ tên: Tập hợp các dãy chữ cái có độ dài <= 30.
- Tuổi: Tập các số nguyên nằm trong khoảng từ 1 đến 80.
- Điểm học tập: Có miền giá trị từ 0 đến 10.
- Giới tính: Tập hợp gồm 2 giá trị “Nam”, “Nữ”.

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

Loại giá trị của thuộc tính

- Đơn trị: Các thuộc tính có giá trị duy nhất cho một thực thể. (VD: Số CMND, Mã SV…)
- Đa trị: Các thuộc tính có một tập giá trị cho cùng một thực thể. (VD: Bằng cấp)
- Suy diễn được: Các thuộc tính có giá trị được suy ra từ các thuộc tính khác. (VD: Thuộc tính Tuổi được suy ra từ thuộc tính Năm sinh…)

c. Lược đồ quan hệ (Relation Schema)

Tập tất cả các thuộc tính cần quản lý của một đối tượng cùng với mối liên hệ giữa chúng được gọi là lược đồ quan hệ. Lược đồ quan hệ Q với tập thuộc tính {A1, A2, …, An} được viết là Q(A1, A2, …, An). Tập các thuộc tính của Q được ký hiệu là Q+.

Thường khi thành lập một lược đồ, người thiết kế luôn gắn cho nó một ý nghĩa nhất định, ý nghĩa đó gọi là tân từ của lược đồ quan hệ đó. Khi phát biểu tân từ cho một lược đồ quan hệ, người thiết kế cần phải mô tả đầy đủ ý nghĩa để người khác tránh hiểu nhầm. Dựa vào tân từ người ta xác định được tập thuộc tính khóa của lược đồ quan hệ. (Khái niệm khoá sẽ được trình bày sau)

Chẳng hạn, đối với lược đồ quan hệ sinh viên (được đặt tên là Sv):

SV(MASV, HOSV, TENSV, NU, NGAYSINH, MALOP, HOCBONG, TINH)

+ Tên quan hệ: Sv

+ Tập thuộc tính: {MASV, HOSV, TENSV, NU, NGAYSINH, MALOP, HOCBONG, TINH}, kí hiệu Sv+

+ Tân từ của lược đồ quan hệ sinh viên là: “Mỗi sinh viên có một mã sinh viên (MASV) duy nhất. Mỗi mã sinh viên xác định tất cả các thuộc tính của sinh viên đó như họ tên (HOTEN), nam/nữ (NU), ngày sinh (NGAYSINH), lớp đang theo học (MALOP), học bổng (HOCBONG), tỉnh cư trú (TINH).

+ Bậc của lược đồ quan hệ sinh viên là: 8

Tập hợp các lược đồ quan hệ cùng nằm trong một hệ thống quản lý được gọi là một lược đồ CSDL. Trong một lược đồ CSDL, các tên lược đồ quan hệ là duy nhất.

Ví dụ: Lược đồ CSDL quản lý điểm sinh viên có thể gồm các lược đồ quan hệ sau:

Sv(MASV, HOSV, TENSV, NU, NGAYSINH, MALOP, TINH, HOCBONG)

Lop(MALOP, TENLOP, SISO, MAKHOA)

Kh(MAKHOA, TENKHOA, SOCBGD)

Mh(MAMH, TENMH, SOTIET)

Kq(MASV, MAMH, DIEMTHI)

d. Quan hệ (Relation)

Sự thể hiện của lược đồ quan hệ Q ở một thời điểm nào đó được gọi là quan hệ. Tại những thời điểm khác nhau thì quan hệ sẽ có những thể hiện khác nhau. Như vậy, trên một lược đồ quan hệ có thể định nghĩa rất nhiều quan hệ.

Thường ta dùng các ký hiệu như R, S, Q để chỉ các lược đồ quan hệ, còn quan hệ được định nghĩa trên nó tương ứng được ký hiệu là là r, s, q.

Một quan hệ có n thuộc tính được gọi là quan hệ n ngôi.
Để chỉ quan hệ r xác định trên lược đồ quan hệ Q ta có thể viết r(Q).
Mỗi quan hệ (hay bảng quan hệ) như là một bảng 2 chiều gồm các dòng và các cột
Mỗi dòng của bảng là một bộ (còn gọi là bản ghi, hoặc mẩu tin).
Mỗi cột của bảng là một thuộc tính (còn gọi là trường, hoặc vùng tin).
Ví dụ: Xét một thể hiện của lược đồ quan hệ SINHVIEN như sau:

MASV	HOTEN	MAMH	MAKHOA	DIEMTHI
99001	TRAN DAN THU	CSDL	CNTT	3.0
99002	NGUYEN HA DA THAO	CSDL	CNTT	8.0
99001	TRAN DAN THU	THVP	CNTT	6.0
99005	LE THANH TRUNG	THVP	AV	5.0

+ Quan hệ SINHVIEN là quan hệ 5 ngôi.
+ Mỗi một cột là một thuộc tính của quan hệ SINHVIEN.
+ Mỗi một dòng là một Sinh viên.
Chú ý:
- Các thuộc tính của một quan hệ đều có tên khác nhau.
- Trong một quan hệ, tất cả các giá trị của một thuộc tính có cùng một miền giá trị.
- Quan hệ là một bảng không chứa các hàng giống hệt nhau. Mỗi hàng là duy nhất, không thể có 2 hàng có cùng các giá trị ở tất cả vùng tin.
- Thứ tự của các hàng là không quan trọng.
- Mỗi quan hệ là một bảng, nhưng không phải mọi bảng đều là quan hệ.
e. Bộ (Tuple)
Bộ là tập mỗi giá trị liên quan của tất cả các thuộc tính của một lược đồ quan hệ.
Thường người ta dùng các chữ cái thường (như t, p, q…) để biểu diễn các bộ. Chẳng hạn để nói bộ t thuộc quan hệ r ta viết: t □ □r.
Chẳng hạn quan hệ SINHVIEN sau có 4 bộ:

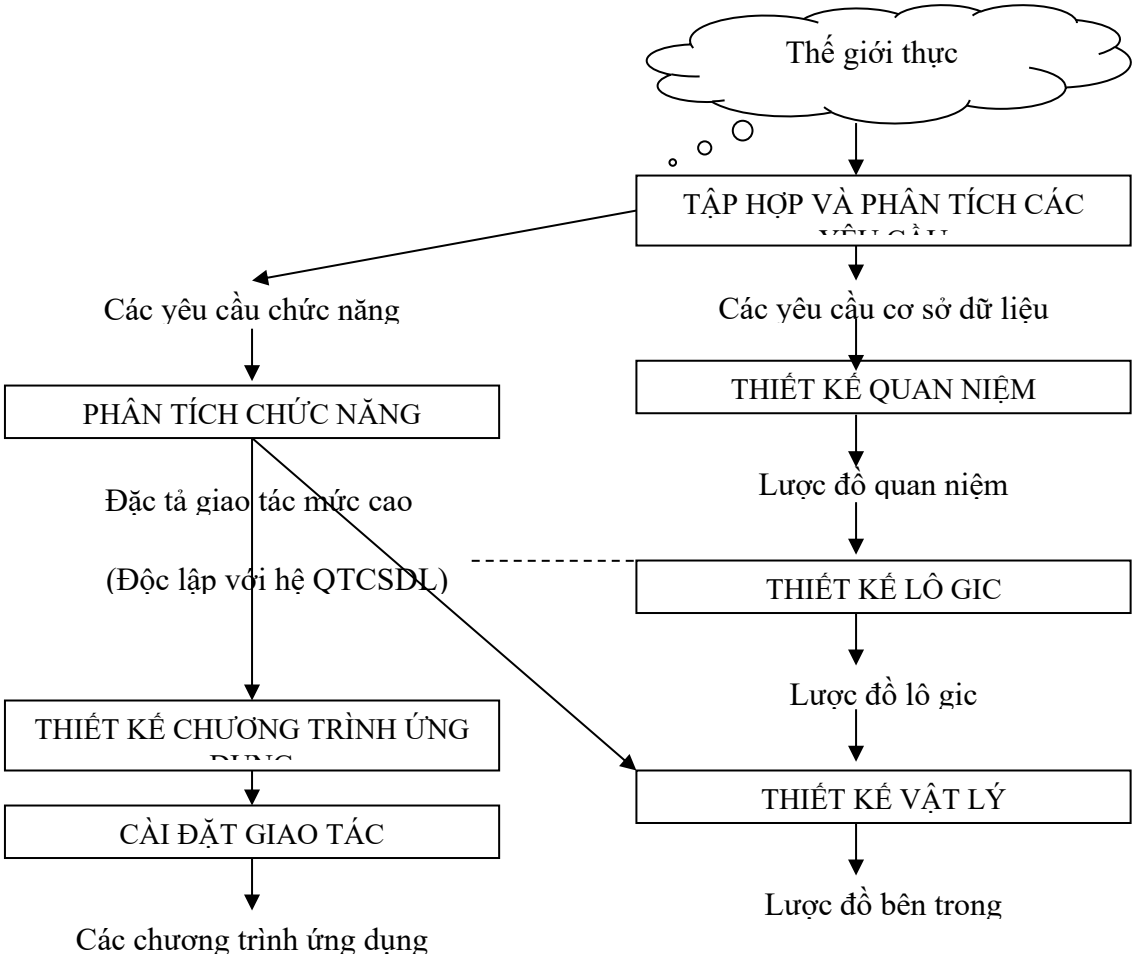
MASV	HOTEN	MAMH	MAKHOA	DIEMTHI
99001	TRAN DAN THU	CSDL	CNTT	3.0
99002	NGUYEN HA DA THAO	CSDL	CNTT	8.0
99001	TRAN DAN THU	THVP	CNTT	6.0

99005	LE THANH TRUNG	THVP	AV	5.0

Ta có: t1 = (99001, TRAN DAN THU, CSDL, CNTT, 3.0) là một bộ của quan hệ SINHVIEN.

2.1.2. Sơ đồ thực thể liên kết

Gọi tắt là mô hình ER (Entity-Relationship Model). Đó là một mô hình dữ liệu mức quan niệm phổ biến, tập trung vào các cấu trúc dữ liệu và các ràng buộc. Mô hình này thường được sử dụng để thiết kế các ứng dụng cơ sở dữ liệu và nhiều công cụ thiết kế cơ sở dữ liệu sử dụng các khái niệm của nó.



Hình 2.1. Sơ đồ thực thể liên kết

2.1.3. Tính năng của các liên kết

Mô hình thực thể kết hợp để biểu diễn mối liên kết giữa các thực thể trong thiết kế mô hình mức khái niệm.

Các thành phần chủ yếu của mô hình thực thể liên kết là thực thể, thuộc tính và mối kết hợp.

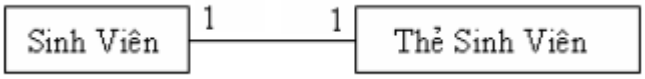
2.2. Mô hình dữ liệu liên kết

2.2.1. Các khái niệm cơ bản

a. Liên kết

Một liên kết là một sự ghép nối giữa hai hay nhiều thực thể. Phản ánh sự liên hệ giữa các thực thể. Có các kiểu liên kết sau.

Quan hệ 1-1 (đọc là liên kết một một): Hai thực thể A và B có mối quan hệ 1-1 nếu một thực thể kiểu A tương ứng với một thực thể kiểu B và ngược lại:



Quan hệ 1-n (đọc là liên kết một nhiều): Hai thực thể A và B có mối Quan hệ 1- n nếu một thực thể kiểu A tương ứng với nhiều thực thể kiểu B và một thực thể của B chỉ tương ứng với một thực thể kiểu A:

Quan hệ n-n (đọc là liên kết nhiều nhiều): Hai thực thể A và B có mối Quan hệ n-n nếu một thực thể kiểu A tương ứng với nhiều thực thể kiểu B và ngược lại".

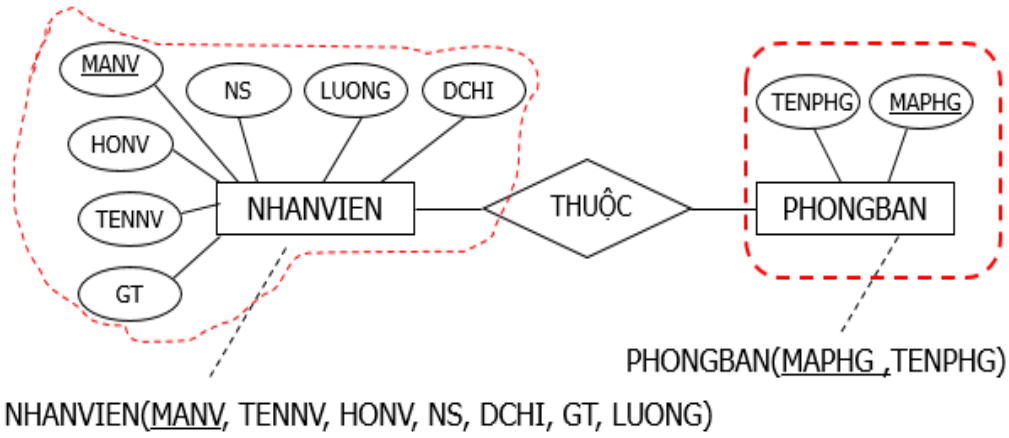
b. Hệ thống kí hiệu

Thực thể	Thực thể được biểu diễn thành hình chữ nhật trong đó có ghi tên thực thể
Thuộc	Thuộc tính được biểu diễn thành hình Elip trong đó có ghi tên thuộc tính
Ouanh	Quan hệ
Thực thể khoá	Khoá của thuộc tính
TTđa trị	TT đa trị

2.2.2. Chuyển sơ đồ ER sang mô hình quan hệ

a. Tập thực thể

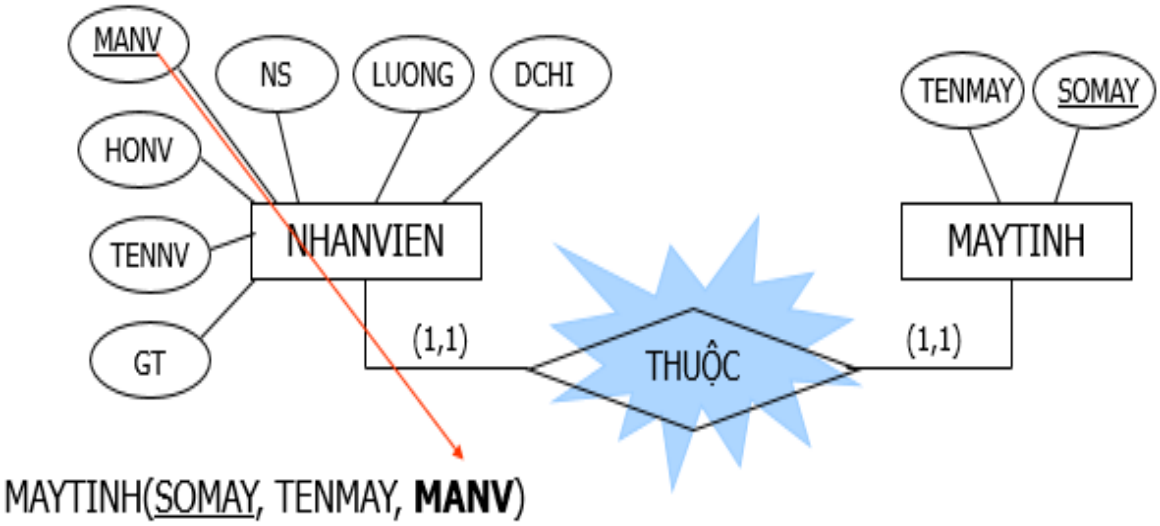
Mỗi thực thể chuyển thành một quan hệ cùng tên và danh sách thuộc tính. Thuộc tính khoá trở thành khoá chính của quan hệ Ví dụ chuyển tập thực thể



Hình 2.2. Tập thực thể

b. Mối kết hợp 1 – 1

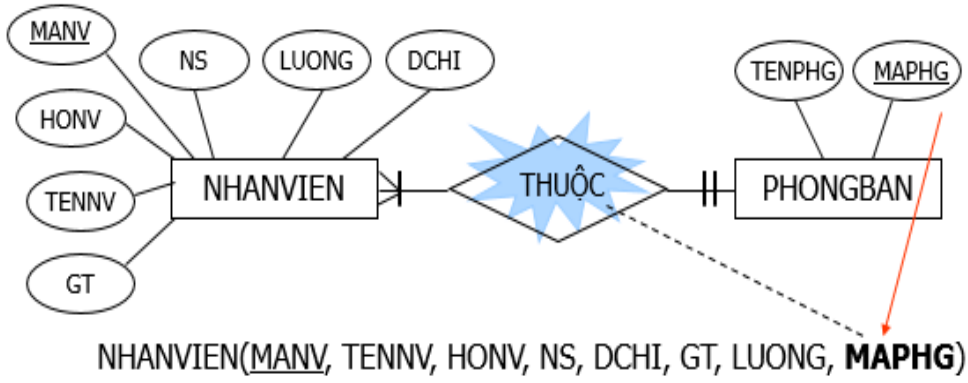
Thuộc tính khoá bên này làm khoá ngoại bên kia hoặc ngược lại. Bên dưới là ví dụ chuyển mối kết hợp 1 – 1



Hình 2.3. Mối kết hợp 1-1

c. Mối kết hợp 1 – N

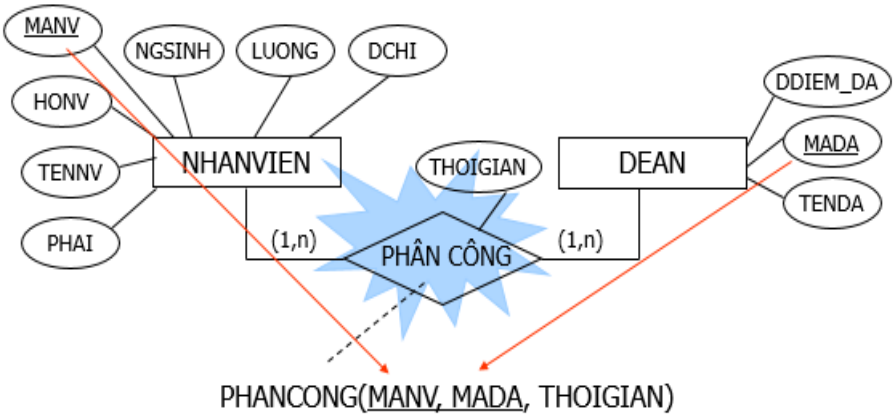
Thuộc tính khoá bên 1 làm khoá ngoại bên nhiều. Ví dụ:



Hình 2.4. Mối kết hợp 1-N

d. Mối kết hợp N – N

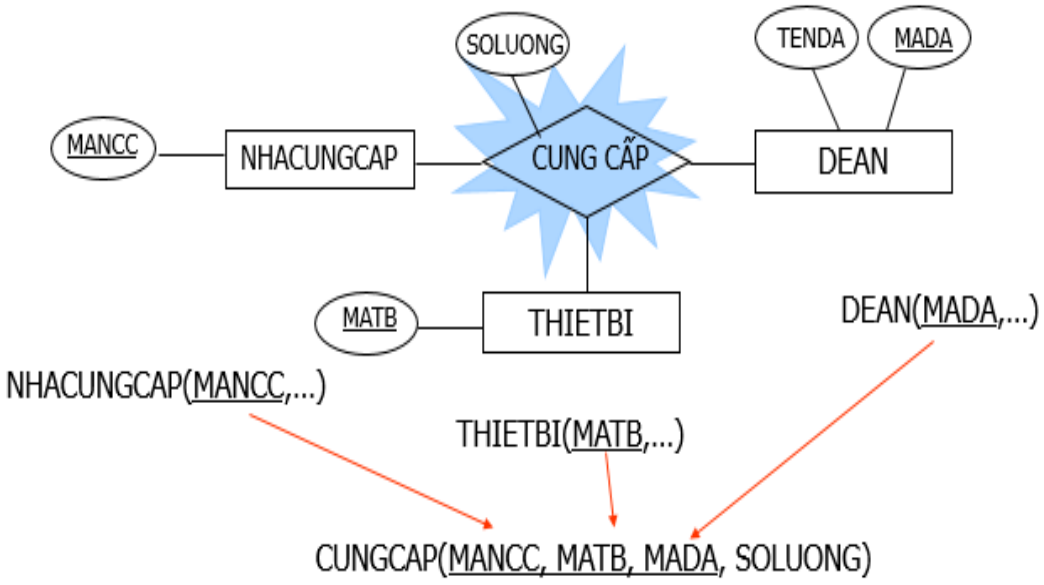
Chuyển thành quan hệ mới có khoá chính gồm 2 thuộc tính khoá của 2 quan hệ; thuộc tính mối kết hợp (nếu có) trở thành thuộc tính của quan hệ mới. Ví dụ



Hình 2.5. Mối kết hợp N-N

e. Mối kết hợp 3 ngôi (Ba thực thể tham gia vào mối kết hợp)

Chuyển thành quan hệ mới, có khoá chính gồm 3 thuộc tính khoá của 3 thực thể tham gia mối kết hợp. Thuộc tính mối kết hợp (nếu có) trở thành thuộc tính của quan hệ mới. Ví dụ

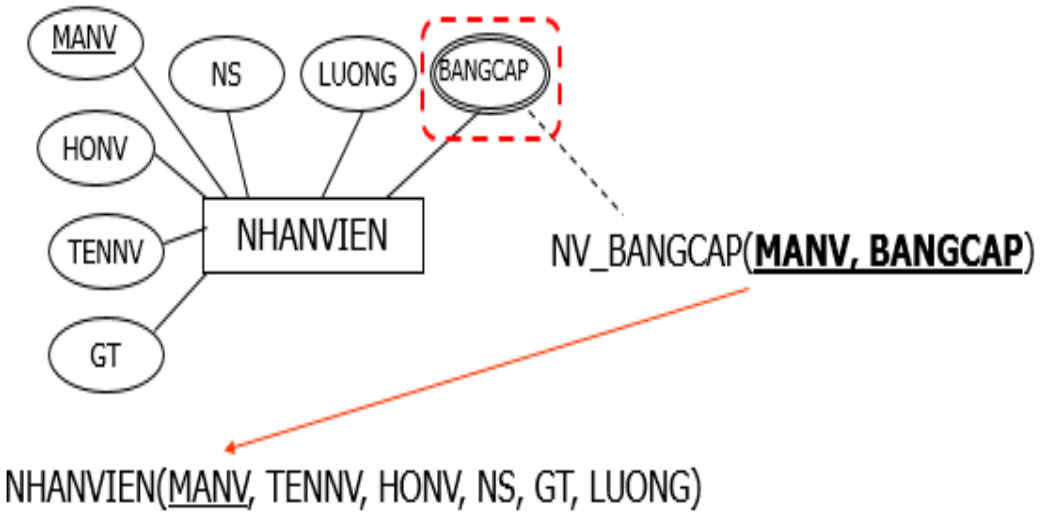


Hình 2.6. Kết hợp 3 ngôi

f. Thuộc tính đa trị (Thuộc tính có nhiều giá trị cho một thể hiện)

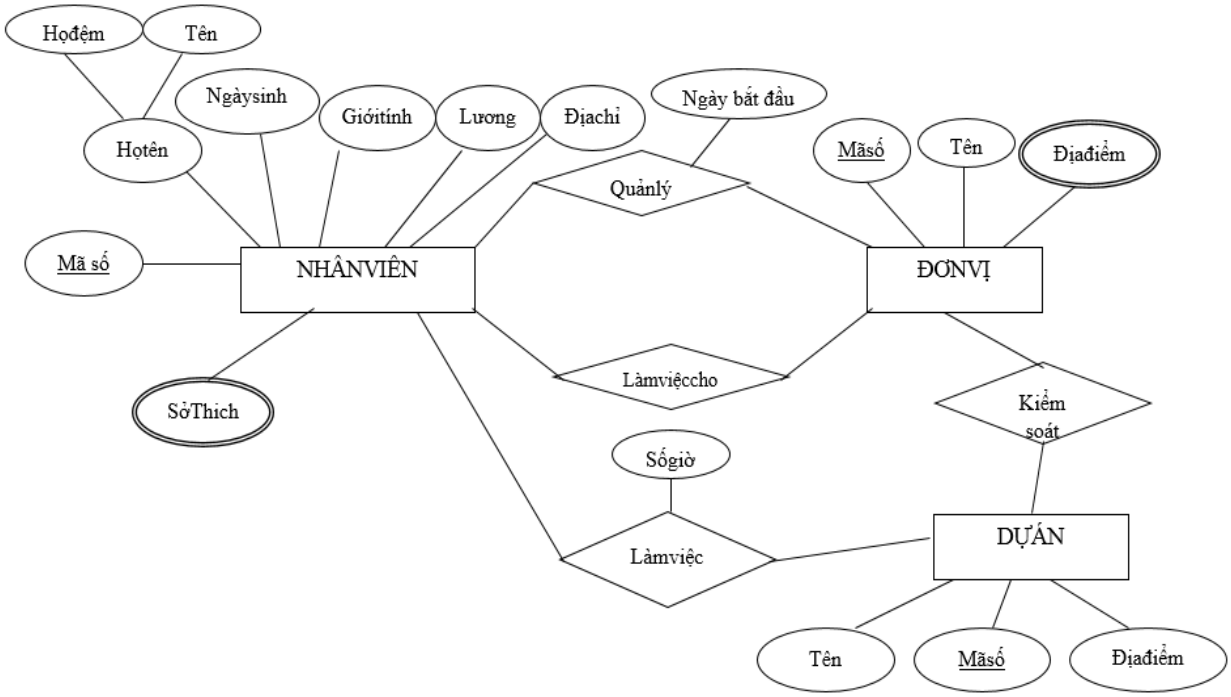
Chuyển thành quan hệ mới có khoá chính gồm thuộc tính đa trị và thuộc tính khoá của thực thể. Sau khi chuyển thành quan hệ mới, thuộc tính đa trị sẽ biến mất khỏi thực thể cũ.

Ví dụ:



Hình 2.7. Thuộc tính đa trị

g. Mô hình dữ liệu quan hệ – Ví dụ

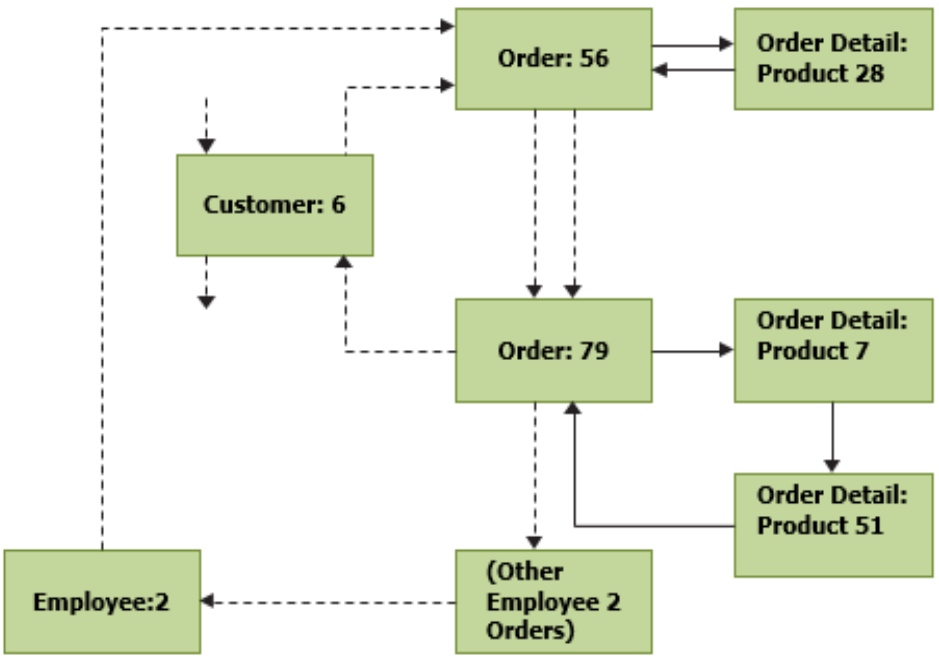


Hình 2.8. Mô hình dữ liệu quan hệ

2.3. Mô hình dữ liệu mạng

2.3.1. Các khái niệm cơ bản

Các file riêng biệt trong hệ thống file phẳng được gọi là các bản ghi. Tập hợp bản ghi cùng kiểu tạo thành một kiểu thực thể dữ liệu. Các kiểu thực thể kết nối với nhau thông qua mối quan hệ cha-con. Mô hình dữ liệu mạng biểu diễn bởi một đồ thị có hướng, và các mũi tên chỉ từ kiểu thực thể cha sang kiểu thực thể con.



Hình 2.9. Mô hình dữ liệu mạng

2.3.2. Chuyển sơ đồ ER sang mô hình mạng

Mô hình dữ liệu mạng (Network Data Model) là mô hình được biểu diễn bởi những mẫu tin (record), loại mẫu tin, loại liên hệ giữa các mẫu tin và bản số của mẫu tin. Ngoài cái tên mô hình dữ liệu mang ra thì nó còn có những cái tên khác như mô hình mạng, mô hình lưới. Bây giờ ta tìm hiểu chi tiết cụ thể:

+ Mẫu tin: Là tin mô tả một đối tượng trong thế giới thực. Ví dụ sinh viên Nguyễn Văn A sẽ có các thông tin tên sinh viên, ngày sinh, tuổi, ... và mỗi sinh viên chính là một mẫu tin.

+ Loại mẫu tin: Chính là tên của đối tượng. Ví dụ NHANVIEN là một loại mẫu tin, PHONG là một loại mẫu tin. Để thể hiện một loại mẫu tin trong sơ đồ thì ta dùng ký hiệu dưới đây để thể hiện.

+Loại liên hệ: Mô tả mối liên hệ giữa các loại mẫu tin, ví dụ mỗi sinh viên sẽ học trong một lớp và một lớp có nhiều sinh viên. Ký hiệu của nó là:

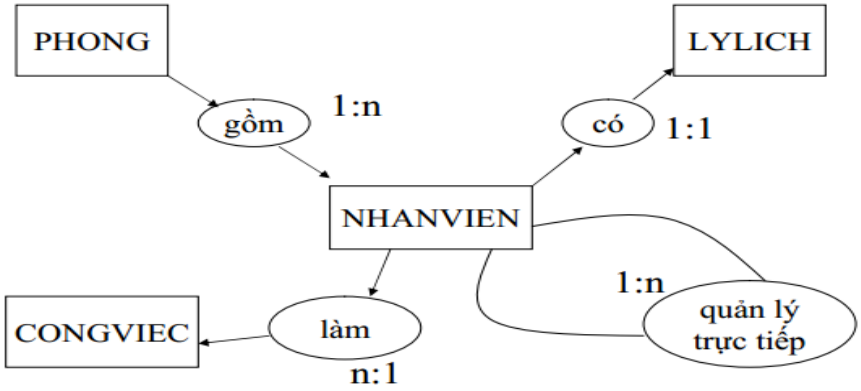
Bản số: chỉ ra số lượng các mẫu tin tham gia trong các mối liên hệ, chúng ta sẽ có những mối liên hệ sau đây:

(1:1) (one-to-one): mỗi mẫu tin của loại mẫu tin chủ kết hợp với đúng 1 mẫu tin của loại mẫu tin thành viên.

(1:n) (one-to-many): mỗi mẫu tin của loại mẫu tin chủ kết hợp với 1 hay nhiều mẫu tin thành viên.

(n:1) (many-to-one): nhiều mẫu tin của loại mẫu tin chủ kết hợp với những 1 mẫu tin của loại mẫu tin thành viên.

(Recursive): một loại mẫu tin chủ cũng có thể đồng thời là loại mẫu tin thành viên với chính



Hình 2.10. Chuyển sơ đồ ER sang mô hình mạng

Giải thích:

Mỗi phòng có nhiều nhân viên (n) và mỗi nhân viên chỉ thuộc một phòng (1), vậy đây là mối quan hệ (1-n)

Mỗi nhân viên có một lý lịch (1) và mỗi lý lịch thuộc về một nhân viên (1), đây là quan hệ (1-1)

Mỗi nhân viên làm nhiều công việc (n) và mỗi công việc có một nhân viên (1) làm, đây là mối quan hệ (n-1)

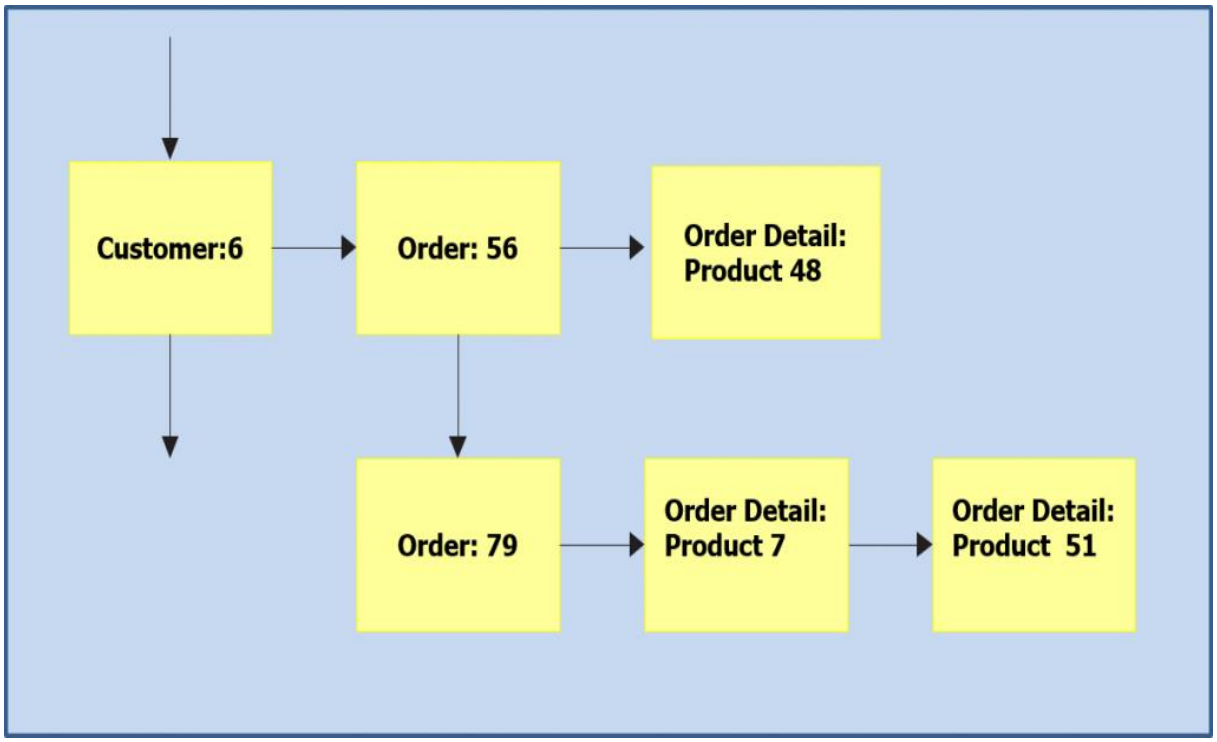
Môi nhân viên sẽ có một người quản lý trực tiếp (trưởng nhóm), và mỗi nhóm có thể có nhiều nhân viên, đây là mối quan hệ (Recursive) vì nó tự chỉ đến nó

Các bạn nhớ để ý luồng đi của mũi tên nhé, nó đi theo hướng cấp cao đến cấp thấp. Ví dụ đi từ PHONG đi đến NHANVIEN vì NHANVIEN là con của PHONG, từ NHANVIEN đi đến LYLICH vì mỗi lý lịch thuộc về NHANVIEN. Từ mũi tên đó ta sẽ tìm được mối quan hệ (1-n) (n-1) (1,1) (Recursive) một cách dễ dàng.

2.4. Mô hình dữ liệu phân cấp

2.4.1. Các khái niệm cơ bản

Tổ chức theo hình cây, mỗi nút biểu diễn một thực thể dữ liệu. Liên hệ dữ liệu thể hiện trên liên hệ giữa nút cha và nút con. Mỗi nút cha có thể có một hoặc nhiều nút con, nhưng mỗi nút con chỉ có thể có một nút cha.

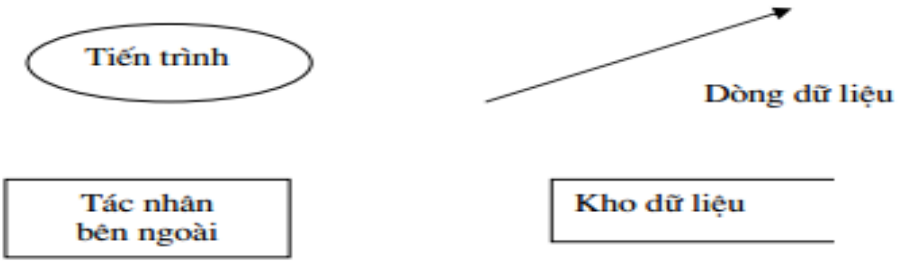


Hình 2.11. Mô hình dữ liệu phân cấp

Sơ đồ phân rã chức năng cho biết các công việc cần làm -> chưa biết mối quan hệ giữa chúng cũng như giữa các chức năng với môi trường bên ngoài -> cần có mô tả bổ sung. => ra đời của mô hình luồng dữ liệu.

a. Định nghĩa: là công cụ mô tả các dòng thông tin liên hệ giữa các chức năng với nhau và giữa các chức năng với môi trường bên ngoài.

b. Khái niệm và ký hiêu.



*. Tác nhân bên ngoài: Là một người hoặc một vật nào đó, tương tác với hệ thống, sử dụng hệ thống.

Tác nhân sẽ gửu thông điệp đến hệ thống hoặc nhận thông điệp xuất phát từ hệ thống hoặc là thay đổi các thông tin cùng với hệ thống.

- Một tác nhân cũng có thể là người hoặc là một hệ thống khác.
- Tác nhân cũng có thể được xếp loại.

+ Một tác nhân chính là tác nhân sử dụng những chức năng chính của hệ thống.

+ Một tác nhân phụ (vd: nhà quản trị hệ thống)

Cả hai tác nhân này đều được mô hình hóa để đảm bảo mô tả đầy đủ những chức năng của hệ thống, mặc dù những chức năng chính mới thực sự nằm trong mối quan tâm chủ yếu của khách hàng. Khi tìm tác nhân, ta thường đặt các câu hỏi ai sẽ là người thực hiện chức năng này

*. Dòng dữ liệu

- Mô tả luồng dữ liệu đi giữa tác nhân và công việc hoặc giữa công việc và kho dữ liệu.
- Không có luồng dữ liệu đi giữa tác nhân với tác nhân hoặc giữa công việc với công việc.

*. Các tiến trình trong hệ thống.
- Là các công việc có tác động hoặc làm biến đổi thông tin.
- Tên ghi trong tiến trình chính là các chức năng trong sơ đồ phân rã chức năng.
- Là cầu nối giữa tác nhân và hệ thống, nó sẽ giúp các tác nhân thực hiện được yêu cầu của họ.
*. Kho dữ liệu
- Là nơi lưu trữ thông tin của hệ thống, một hệ thống sẽ có một hoặc nhiều kho dữ liệu, tùy thuộc vào độ phức tạp của từng bài toán. Song người ta thường cố gắng tổ chức sao cho càng ít khi thì càng tốt (ít nhất là một kho). Nếu có nhiều kho -> khi viết chương trình sẽ phải tạo móc nối tới các kho khác nhâu -> khó khăn cho người coding.
*. Ý nghĩa của sơ đồ luồng dữ liệu
- Trong sơ đồ phân rã chức năng, cho biết công việc và mức phân cấp (sự chi tiết). Sơ đồ luồng dữ liệu bổ sung khiếm khuyết mà sơ đồ phân rã chức năng còn thiếu, chính là quan hệ giữa các công việc.
- Giúp cho việc thấy rõ hoạt động xử lý, cơ sở bước đầu thiết kế xử lý.
- Công cụ giao tiếp giữa nhà thiết kế, phần tích và người sử dụng.
* Một số quy tắc khi vẽ biểu đồ luồng dữ liệu
Vẽ biểu đồ luồng dữ liệu cần tuân theo các quy tắc sau - các "cái vào" của một tiến trình cần khác với các "cái ra" của nó. Nguyên tắc này nhấn mạnh rằng, các dữ liệu qua một tiến trình phải có thay đổi. Ngược lại tiến trình là không cần thiết vì không tác động gì đến các luồng thông tin đi qua nó.
- Mỗi tiến trình phải có tên duy nhất. Tuy nhiên, một số tác nhân ngoài và kho dữ liệu có thể được vẽ lặp lại. - Các luồng dữ liệu đi vào một tiến trình phải đủ để tạo thành các luồng dữ liệu đi ra.
Tiến trình
- Không một tiến trình nào chỉ có cái ra mà không có cái vào. Đối tượng chỉ có cái ra thì chỉ có thể là tác hân.
- Không một tiến trình nào chỉ có cái vào. Một đối tượng chỉ có cái vào chỉ có thể là tác nhân
Kho dữ liệu
- Không có luồng dữ liệu từ một kho đến một kho dữ liệu khác.
- Dữ liệu không thể di chuyển trực tiếp từ một tác nhân đến một kho dữ liệu và ngược lại.

Tác nhân

- Dữ liệu không thể di chuyển trực tiếp từ một tác nhân đến một tác nhân. Luồng dữ liệu

- Một luồng dữ liệu không thể quay lại nơi mà nó vừa đi khỏi.

- Một luồng dữ liệu đi vào kho có nghĩa là kho dữ liệu được cập nhật, một luồng dữ liệu đi ra khỏi kho có nghĩa là kho dữ liệu đựoc đọc.

- Không có luồng dữ liệu đi giữa các tiến trình với nhau.

c. Phân rã biểu đồ luồng dữ liệu.

*. Định nghĩa: Quá trình phân nhỏ mỗi tiến trình của một biểu đồ DFD thành một DFD mới (nếu có thể) gọi là phân ra biểu đồ luồng dữ liệu đã cho.

*. Đảm bảo sự cân bằng giữa hai mức kề nhau.

Khi phân rã một tiến trình của một biểu đồ DFD thanh một biểu đồ ở mức sau thì mọi luồng dữ liệu vào và ra, tác tác nhân ngoài và kho dữ liệu liên quan với nó phải được bảo toàn (giữ nguyên) trong biểu đồ mức sau.

*. Biểu đồ DFD sơ cấp

Quá trình phân rã luồng dữ liệu sẽ dừng lại khi đạt đến biểu đồ mức thấp nhất – gọi là biểu đồ luồng dữ liệu sơ cấp. Những quy tắc sau đây là những gợi ý để dừng quá trình phân rã.

- Khi một tiến trình là một quyết định hay một tính toán đơn giản, một thao tác dữ liệu như đọc, cập nhật, thêm mới, ghi, xoá.

- Khi mỗi luồng dữ liệu không cần chia nhỏ hơn để chỉ ra rằng các dữ liệu khác nhau đều đã được quản lý. Biểu đồ DFD sơ cấp chỉ ra mức chi tiết nhất các thao tác mà hệ thống mới cần thực hiện.

Chú ý: Khi phân rã biểu đồ luồng dữ liệu, nếu đã xây dựng được biểu đồ phân rã chức năng thì có thể lấy mỗi chức năng tương ứng của nó làm một tiến trình ở biểu đồ luồng dữ liệu có mức tương đương.

d. Các biểu đồ DFD của một ứng dụng.

*. Biểu đồ ngữ cảnh của hệ thống

Biểu đồ ngữ cảnh biểu diễn hệ thống ở mức cao nhất. Trong biểu đồ này chỉ gồm ba loại thành phần.

- Một tiến trình duy nhất mô tả toàn hệ thống, trong đó có tên hệ thống và có chỉ số là 0.

- Các tác nhân bên ngoài.

- Các luồng dữ liệu giữa tác nhân và hệ thống mô tả sự tương tác giữa hệ thống với môi trường. Biểu đồ ngữ cảnh cho ta một cái nhìn tổng quát về hệ thống trong môi trường của nó. Các tác nhân của hệ thống cần phải xác định đầy đủ. Sự thiếu các tác nhân sẽ là nguyên nhân là cho hệ thống được xây dựng không có khả năng hoạt động tốt trên thực tế.

Biểu đồ DFD mức 1. Được phân ra dựa trên biểu đồ ngữ cảnh.

Nguyên tắc phân rã biểu đồ luồng dữ liệu ở trên, ta tiến hành phân rã (làm mịn) biểu đồ ngữ cảnh để nhận được biểu đồ DFD mức 1.

*. Các bước tiến hành:

- Thay thế: Thay thế tiến trình duy nhất của biểu đồ ngữ cảnh bằng các tiến trình con (tương ứng với các chức năng mức 1 trong biểu đồ phân rã chức năng)
- Giữ nguyên: Toàn bộ các tác nhân ngoài và các luồng dữ liệu trong biểu đồ ngữ cảnh và chuyển sang biểu đồ luồng dữ liệu mức 0.
- Thêm vào:

+ Các kho dữ liệu tương ứng với các hồ sơ dữ liệu đã chọn trong danh sách.

+ Các luồng dữ liệu từ các tiến trình đến các kho.

* Biểu đồ DFD mức i

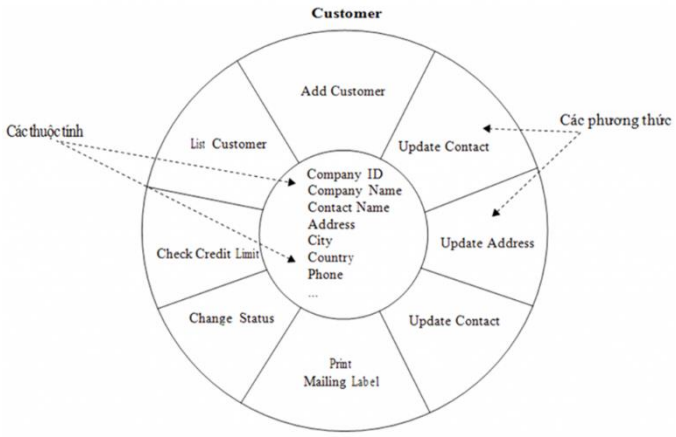
Việc phát triển các biểu đồ luồng dữ liệu chi tiết là quá trình tiếp tục làm mịn các biểu đồ luồng dữ liệu đã nhận trước đó. Một biểu đồ luồng dữ liệu mức i nhận được từ việc phân rã một tiến trình thuộc biểu đồ luồng dữ liệu mức i -1. Quá trình làm mịn biểu đồ luồng dữ liệu mức i -1 được thực hiện như sau.

- Thay thế: Tiến trình được xét của biểu đồ luồng dữ liệu mức i-1 bằng các tiến trình con tương ứng với các chức năng của mức tương ứng trong biểu đồ phân rã chức năng.
- Giữ nguyên: Toàn bộ các tác nhân ngoài, các kho dữ liệu, và các luồng dữ liệu liên quan đến tiến trình được xét.
- Thêm vào: Luồng dữ liệu giữa các tiến trình con với các kho dữ liệu. Việc đánh số hiệu các tiến trình trong một biểu đồ mới nhận được cần tuân theo những nguyên tắc sau: Số hiệu của một tiến trình gồm hai phần: (phần đầu và phần số thứ tự) Phần đầu là số hiệu của tiến trình được sử dụng để phân rã. Phần số thứ tự là thứ tự của tiến trình con. Chú ý: Một biểu đồ chỉ nên phát triển trên một trang giấy.

2.5. Mô hình dữ liệu hướng đối tượng

2.5.1. Các khái niệm cơ bản

Mỗi đối tượng bao gồm các thuộc tính, phương thức (hành vi) của đối tượng. Các đối tượng trao đổi với nhau thông qua các phương thức. Một đối tượng có thể được sinh ra từ việc thừa kế từ đối tượng khác, nạp chồng (hay định nghĩa lại) phương thức của đối tượng khác…



Hình 2.13. Mô hình dữ liệu hướng đối tượng

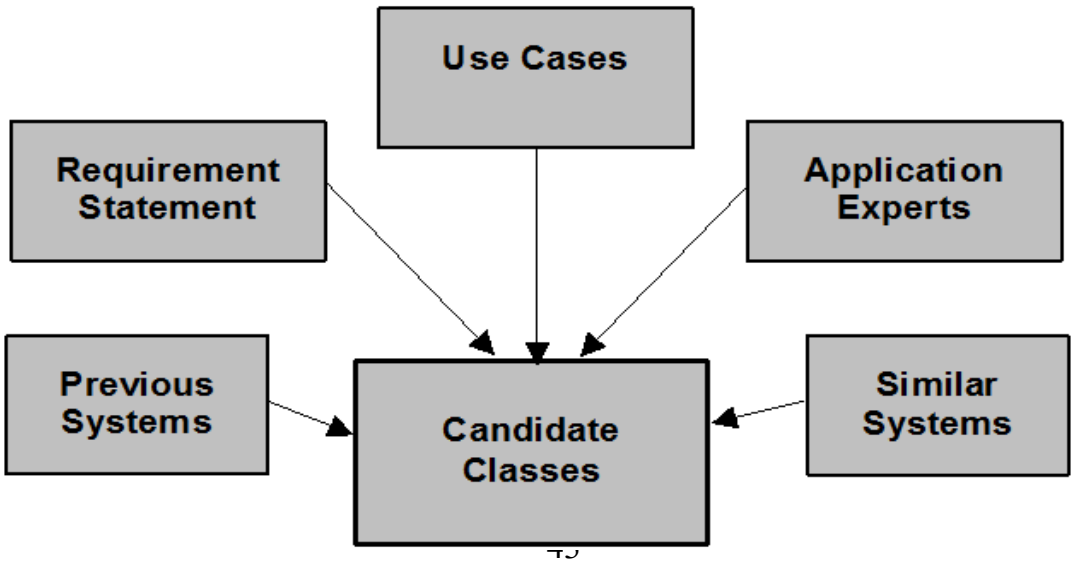
2.5.2. Chuyển sơ đồ ER sang mô hình dữ liệu hướng đối tượng

Mô hình dữ liệu hướng đối tượng là mô hình khó xây dựng nhất so với các mô hình khác. Để xây dựng được mô hình hướng đối tượng, ta phải hiểu được hệ thống một cách rõ ràng và có kinh nghiệm về lập trình hướng đối tượng mới có thể xây dựng thành công mô hình này.

Thực hiện theo các bước sau đây để xây dựng mô hình hướng đối tượng

Bước 1: Tìm các Classes dự kiến

Entity Classes (các lớp thực thể) là các thực thể có thật và hoạt động trong hệ thống, bạn dựa vào các nguồn sau để xác định chúng.



- Requirement statement: Các yêu cầu. Chúng ta phân tích các danh từ trong các yêu cầu để tìm ra các thực thể.
- Use Cases: Phân tích các Use Case sẽ cung cấp thêm các Classes dự kiến.
- Previous và Similar System: có thể sẽ cung cấp thêm cho bạn các lớp dự kiến.
- Application Experts: các chuyên gia ứng dụng cũng có thể giúp bạn.

Xem xét, ví dụ ATM ở trên chúng ta có thể thấy các đối tượng là Entity Class như sau:

- Customers: khách hàng giao dịch là một thực thể có thật và quản lý trong hệ thống.
- Accounts: Tài khoản của khách hàng cũng là một đối tượng thực tế.
- ATM Cards: Thẻ dùng để truy cập ATM cũng được quản lý trong hệ thống.
- ATM Transactions: Các giao dịch được lưu giữ lại, nó cũng là một đối tượng có thật.
- Banks: Thông tin ngân hàng bạn đang giao dịch, nếu có nhiều nhà Bank tham gia vào hệ thống bạn phải quản lý nó. Lúc đó Bank trở thành đối tượng bạn phải quản lý.
- ATM: Thông tin ATM bạn sẽ giao dịch. Nó cũng được quản lý tương tự như Banks.

Lưu ý: Chỉ các thực thể bên trong hệ thống được xem xét, các thực thể bên ngoài hệ thống không được xem xét. Ví dụ Customers là những người khách hàng được quản lý trong hệ thống chứ không phải người dùng máy ATM bên ngoài. Bạn phải lưu ý điều này để phân biệt Class và Actor.

Bước 2: Tìm các thuộc tính và phương thức cho lớp

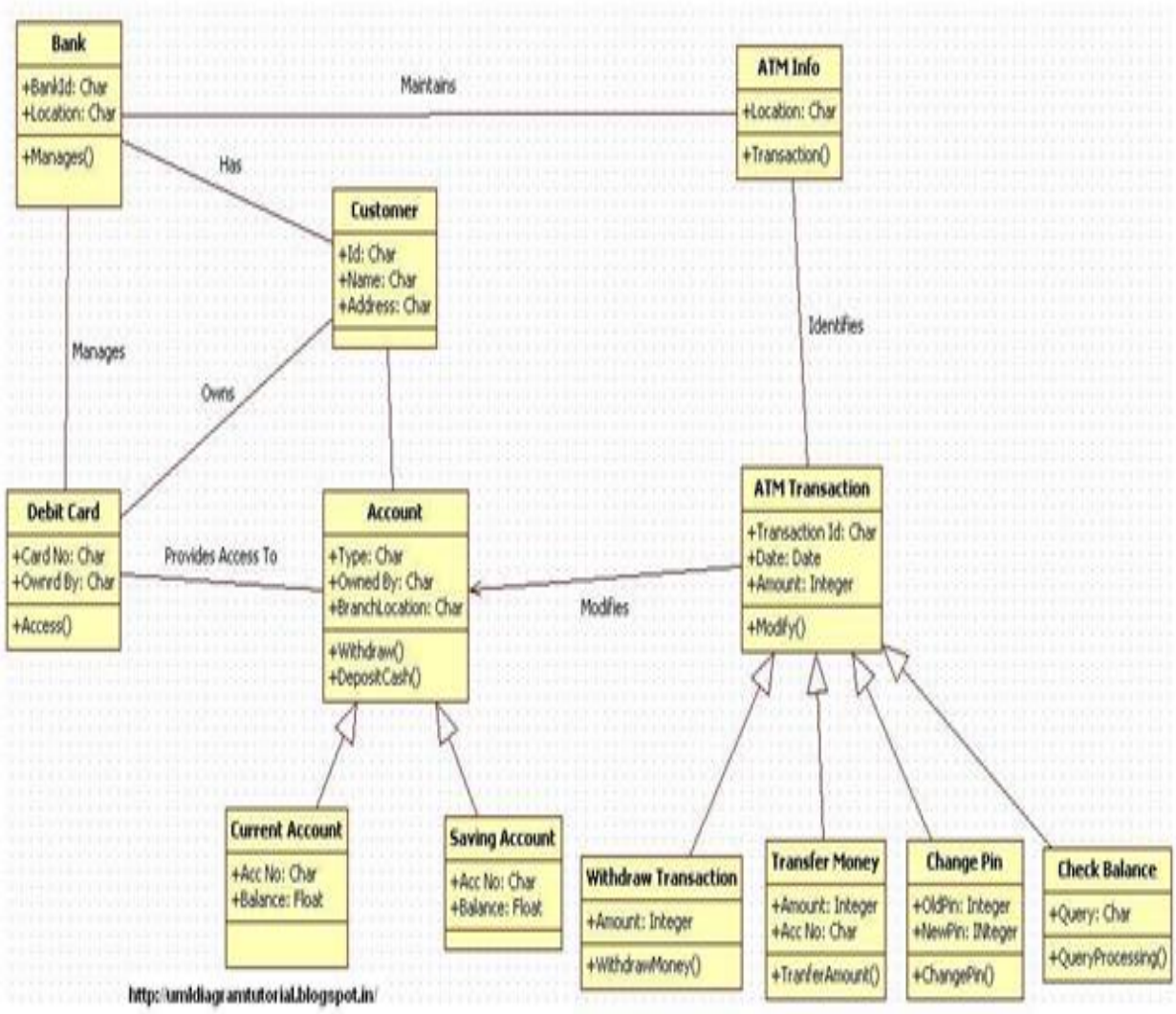
- Tìm thuộc tính: phân tích thông tin từ các form mẫu có sẵn, bạn sẽ tìm ra thuộc tính cho các đối tượng của lớp. Ví dụ các thuộc tính của lớp Customer sẽ thể hiện trên Form đăng ký thông tin khách hàng.
- Tìm phương thức: phương thức là các hoạt động mà các đối tượng của lớp này có thể thực hiện. Chúng ta sẽ bổ sung phương thức đầy đủ cho các lớp khi phân tích Sequence Diagram sau này.

Bước 3: Xây dựng các quan hệ giữa các lớp và phát hiện các lớp phát sinh

- Phân tích các quan hệ giữa các lớp và định nghĩa các lớp phát sinh do các quan hệ sinh ra. Chúng ta phân tích các thực thể ở trên và nhận thấy.

Lớp Accounts có thể chia thành nhiều loại tài khoản như Current Accounts và Saving Accounts và có quan hệ thừa kế với nhau.

Lớp ATM Transactions cũng có thể chia thành nhiều loại giao dịch như Deposit, Withdraw, Transferv.v.. và chúng cũng có quan hệ thừa kế với nhau. Tách chúng ra và vẽ chúng lên bản vẽ chúng ta sẽ có được mô hình hướng đối tượng cho hệ thống ATM như sau:



Hình 2.14. Mô hình hướng đối tượng cho hệ thống ATM

*. Đặc tả Class

Nhìn vào mô hình hướng đối tượng, chúng ta có thể thấy cấu trúc của hệ thống gồm những lớp nào nhưng để cài đặt chúng, chúng ta phải đặc tả chi tiết hơn nữa. Trong đó, cần mô tả:

- Các thuộc tính: Tên, kiểu dữ liệu, kích thước
- Các phương thức:

Tên

Mô tả

Tham số đầu vào: Tên, kiểu dữ liệu, kích thươcs

Kết quả đầu ra: Tên, kiểu dữ liệu, kích thước

Luồng xử lý

Điều kiện bắt đầu

Điều kiện kết thúc

Tuy nhiên, việc này cũng mất khá nhiều thời gian.

❖ **TÓM TẮT CHƯƠNG 2**

Trong chương này, một số nội dung chính được giới thiệu:

1. **Giới thiệu về Cơ sở Dữ liệu Quan hệ**: Chương bắt đầu bằng việc giới thiệu về cơ sở dữ liệu quan hệ và lý do tại sao nó quan trọng trong lĩnh vực quản lý thông tin.
2. **Cấu Trúc Bảng**: Học viên sẽ làm quen với cấu trúc cơ bản của bảng trong cơ sở dữ liệu quan hệ, bao gồm cách đặt tên, định dạng dữ liệu, và khóa chính (primary key).
3. **Quan hệ giữa Bảng**: Chương giải thích về quan hệ giữa các bảng trong cơ sở dữ liệu quan hệ, bao gồm mối quan hệ thông qua khóa ngoại (foreign key) và cách chúng tương tác.
4. **Sử dụng SQL cho Truy vấn Dữ liệu**: Học viên sẽ học cách sử dụng ngôn ngữ SQL để truy vấn dữ liệu từ cơ sở dữ liệu quan hệ. Chương giới thiệu cú pháp cơ bản của câu truy vấn SELECT và cách sử dụng nó để trích xuất thông tin từ bảng.
5. **Thiết kế Bảng và Quan hệ**: Học viên sẽ phát triển khả năng thiết kế bảng và quan hệ giữa chúng để lưu trữ thông tin một cách hiệu quả và tối ưu hóa cơ sở dữ liệu.

❖ **CÁC BÀI TẬP HỆ THỐNG KIẾN THỨC**

1. Thiết kế Bảng và Quan hệ

Hãy tạo một kịch bản cơ sở dữ liệu đơn giản cho một ứng dụng tùy chọn (ví dụ: quản lý thư viện hoặc quản lý nhân viên). Thiết kế các bảng cần thiết và xác định mối quan hệ giữa chúng. Sử dụng SQL để tạo các bảng này và chèn một số dữ liệu mẫu.

2. Truy vấn Dữ liệu

Sử dụng cơ sở dữ liệu mà bạn đã thiết kế trong Bài tập 1, viết một số câu truy vấn SQL để trích xuất thông tin từ bảng. Ví dụ: lấy danh sách tất cả những người dùng có tuổi dưới 30.

3. Tối ưu hóa Cơ sở Dữ liệu

Đánh giá cơ sở dữ liệu mà bạn đã thiết kế trong Bài tập 1 và đề xuất các cách tối ưu hóa. Có thể bao gồm việc thêm chỉ mục (index), loại bỏ các trường dư thừa, hoặc tái thiết kế cấu trúc bảng để cải thiện hiệu suất.

CHƯƠNG 3. NGÔN NGỮ THAO TÁC DỮ LIỆU

❖ GIỚI THIỆU CHƯƠNG 3

Chương 3, "Ngôn ngữ thao tác dữ liệu," là một phần quan trọng trong việc nắm vững kiến thức về cơ sở dữ liệu. Chương này tập trung vào ngôn ngữ SQL (Structured Query Language), là ngôn ngữ tiêu chuẩn sử dụng để truy vấn, cập nhật và quản lý dữ liệu trong cơ sở dữ liệu quan hệ.

❖ MỤC TIÊU CHƯƠNG 3

Sau khi học xong chương này, người học có khả năng:

➢ ***Về kiến thức:***

- *Hiểu các câu lệnh SQL cơ bản, bao gồm SELECT (truy vấn), INSERT (thêm dữ liệu), UPDATE (cập nhật dữ liệu), DELETE (xóa dữ liệu), và CREATE TABLE (tạo bảng).*
- *Hiểu cách sử dụng ràng buộc dữ liệu (constraints) như khóa chính (primary key) và khóa ngoại (foreign key) để bảo vệ tính toàn vẹn và quy tắc dữ liệu trong cơ sở dữ liệu.*
- *Hiểu cách xử lý lỗi và ngoại lệ trong SQL để đảm bảo việc truy vấn và thao tác dữ liệu diễn ra một cách an toàn và hiệu quả.*

➢ ***Về kỹ năng:***

- *Biết sử dụng ngôn ngữ SQL để viết và thực hiện câu truy vấn dữ liệu*
- *Biết cách sử dụng các câu lệnh SQL như INSERT, UPDATE và DELETE để cập nhật và xóa dữ liệu từ cơ sở dữ liệu*

➢ ***Về năng lực tự chủ và trách nhiệm:***

- *Năng lực về quản lý thời gian, trách nhiệm với công việc*
- *Năng lực học tập và làm việc độc lập*
- *Tự chủ trong việc giải quyết vấn đề*

❖ PHƯƠNG PHÁP GIẢNG DẠY VÀ HỌC TẬP CHƯƠNG 3

- *Đối với người dạy: sử dụng phương pháp giảng giảng dạy tích cực (diễn giảng, vấn đáp, dạy học theo vấn đề); yêu cầu người học thực hiện câu hỏi thảo luận và bài tập chương (cá nhân hoặc nhóm).*
- *Đối với người học: chủ động đọc trước giáo trình trước buổi học; hoàn thành đầy đủ câu hỏi thảo luận và bài tập tình huống theo cá nhân hoặc nhóm và nộp lại cho người dạy đúng thời gian quy định.*

❖ ĐIỀU KIỆN THỰC HIỆN CHƯƠNG 3

- ***Phòng học chuyên môn hóa/nhà xưởng:*** Phòng học thực hành

- ***Trang thiết bị máy móc:*** Máy chiếu, máy tính và các thiết bị dạy học khác
- ***Học liệu, dụng cụ, nguyên vật liệu:*** Chương trình môn học, giáo trình, tài liệu tham khảo, giáo án, phim ảnh, và các tài liệu liên quan.
- ***Các điều kiện khác:*** Không có

❖ **KIỂM TRA VÀ ĐÁNH GIÁ CHƯƠNG 3**

- **Nội dung:**

✓ *Kiến thức: Kiểm tra và đánh giá tất cả nội dung đã nêu trong mục tiêu kiến thức*

✓ *Kỹ năng: Đánh giá tất cả nội dung đã nêu trong mục tiêu kĩ năng.*

✓ *Năng lực tự chủ và trách nhiệm: Trong quá trình học tập, người học cần:*

+ Nghiên cứu bài trước khi đến lớp

+ Chuẩn bị đầy đủ tài liệu học tập.

+ Tham gia đầy đủ thời lượng môn học.

+ Nghiêm túc trong quá trình học tập.

- **Phương pháp:**

✓ ***Điểm kiểm tra thường xuyên:*** *1 điểm kiểm tra (hình thức: hỏi miệng)*

✓ ***Kiểm tra định kỳ lý thuyết:*** *không có*

❖ **NỘI DUNG**

3.1. Đại số quan hệ

3.1.1. Phép hợp

Cho hai quan hệ r và s trên cùng một lược đồ quan hệ R.

Phép hợp của hai quan hệ r và s, ký hiệu $r \cup s$, là một quan hệ gồm các bộ của r hay của s.

$r \cup s = \{ u \mid u \in r \text{ hay } u \in s \}$

Phép hợp có tính giao hoán.

$r \cup s = s \cup r$

Ví dụ:

R	A	B	C
	1	2	3
	4	5	6
	1	2	7
	8	4	5

S	A	B	C
	1	2	3
	3	5	7
	6	2	9

R ∪ S	A	B	C
	1	2	3
	4	5	6
	1	2	7
	8	4	5

3	5	7
6	2	9

3.1.2. Phép giao

Cho hai quan hệ r và s trên cùng một lược đồ quan hệ R.

Phép giao của hai quan hệ r và s, ký hiệu r ∩ s, cho kết quả là một quan hệ gồm các bộ xuất hiện đồng thời trong cả r và s.

r ∩ s = { u | u ∈ r và u ∈ s }

Phép giao của hai quan hệ r và s có thể được tính từ phép trừ:

r ∩ s = r – (r – s) = s – (s – r)

Phép giao có tính giao hoán.

r ∩ s = s ∩ r

Ví dụ 1:

R	A	B	C
	1	2	3
	4	5	6
	1	2	7
	8	4	5

S	A	B	C
	1	2	3
	3	5	7
	6	2	9

R ∩ S	A	B	C
	1	2	3

Ví dụ 2: Cho 2 quan hệ NHAHANG và KHACHSAN như sau:

NHAHANG	TEN	DIACHI	PHUONG	QUAN
	Huy Hoàng	504 Điện Biên Phủ	14	3
	Tương Lai	229 Ngô Chí Thanh	6	5
	Sinh Đôi	300 Lý Thái Tổ	6	10
	Đông Nam Á	230 Ngô Gia Tự	10	1

KHACHSAN	TEN	DIACHI	PHUONG	QUAN
	Viễn Đông	2 Lê Lợi	5	1
	Huy Hoàng	504 Điện Biên Phủ	14	3
	Ánh Hồng	151 Đặng Dung	2	PN
	Tương Lai	229 Ngô Chí Thanh	6	5

Tìm các nơi vừa là khách sạn, vừa là nhà hàng:

NH_KS = NHAHANG ∩ KHACHSAN

Kết quả:

NH_KS	TEN	DIACHI	PHUONG	QUAN
	Huy Hoàng	504 Điện Biên Phủ	14	3
	Tương Lai	229 Ngô Chí Thanh	6	5

3.1.3. Phép trừ

Cho hai quan hệ r và s trên cùng một lược đồ quan hệ R.

Phép trừ (hiệu) của quan hệ r cho s, ký hiệu r – s, cho kết quả là một quan hệ gồm các bộ của r không có trong s.

$r - s = \{ u \mid u \in r \text{ và } u \notin s \}$

Phép hiệu không có tính giao hoán.

$r - s \neq s - r$

Ví dụ:

R	A	B	C
	1	2	3
	4	5	6
	1	2	7

S	A	B	C
	1	2	3
	3	5	7
	6	2	9

R – S	A	B	C
	4	5	6
	1	2	7
	8	4	5

3.1.4. Tích Đề các

Tích Descartes (Đề-các) còn gọi là tích hỗn hợp (cross product) hoặc là nối hỗn hợp (cross join), được ký hiệu là ×. Đó cùng là một phép toán hai ngôi nhưng những quan hệ mà nó áp dụng trên đó không phải là tương thích đồng nhất. Phép toán này được sử dụng để nối các bộ của hai quan hệ vào một kiểu kết hợp.

Cho quan hệ r trên lược đồ quan hệ R (A1, A2, …, Am) và s trên lược đồ quan hệ S (B1, B2, …, Bn).

Phép tích Descartes (Cartesian product) của hai quan hệ r và s, ký hiệu là r × s, là một quan hệ trên lược đồ T(A1, A2, …, Am, B1, B2, …, Bn) gồm các bộ u sao cho m thành phần đầu tiên là một bộ của r và n thành phần cuối cùng là một bộ của s.

$r \times s = \{ (u1, \ldots, um, um+1, \ldots, um+n) \mid (u1, \ldots, um) \in r \text{ và } (um+1, \ldots, um+n) \in s\}$

Như vậy, nếu R có m bộ và S có n bộ thì R× S có m*n bộ. Phép toán này nếu áp dụng một mình thì thông thường không có ý nghĩa. Nó chỉ có lợi khi kết hợp tiếp theo bằng một phép chọn các giá trị tương thích của các thuộc tính xuất phát từ các quan hệ thành phần. Tích Đề-các kết hợp với một phép chọn cho ta một phép nối (join).

Phép tích Descartes có tính giao hoán.

$r \times s = s \times r$

Ví dụ: Ta có hai quan hệ R và S như sau:

R	A	B	C
	1	2	3
	4	5	6
	1	2	7
	8	4	5

S	T	D
	1	5
	3	7

Kết quả phép tích Descartes: R × S

R × S	A	B	C	T	D
	1	2	3	1	5
	1	2	3	3	7
	4	5	6	1	5
	4	5	6	3	7
	1	2	7	1	5
	1	2	7	3	7
	8	4	5	1	5
	8	4	5	3	7

3.1.5. Phép chiếu

Mục đích:

Phép chiếu trên một quan hệ thực chất là loại bỏ đi một số thuộc tính của quan hệ đó.

Nếu ta coi một quan hệ như một bảng thì phép chọn thực hiện chọn ra một số hàng của bảng thoả mãn điều kiện chọn và bỏ qua các hàng không thoả mãn điều kiện chọn. Phép chiếu là phép toán chọn ra một số cột của bảng. Nếu chúng ta chỉ quan tâm đến một số thuộc tính của quan hệ, chúng ta dùng phép chiếu để chiếu lên các thuộc tính đó.

Định nghĩa:

Cho lược đồ quan hệ R(A1, A2, …, An) và một tập con các thuộc tính X, với $X \subset \{A1, A2, \ldots, An\}$.

Gọi t là một bộ thuộc R, A là một thuộc tính, t[A] là giá trị của bộ t tại thuộc tính A. Giả sử X = { B1, B2, …, Bm }. Khi đó, t[X] là giá trị của bộ t trên tập thuộc tính X, t[X] = { t[B1], t[B2], …, t[Bm] }

Phép chiếu r(R) trên một tập thuộc tính X, ký hiệu là ΠX(r) được định nghĩa như sau:

ΠX(r) = { t[X] / t ∈ r }

Phép chiếu của quan hệ R trên tập thuộc tính X là một tập các bộ, được xây dựng bằng cách loại bỏ đi từ các bộ t trong quan hệ R những thuộc tính không nằm trong X.

Kết quả trả về là một quan hệ có m thuộc tính nằm trong X và có cùng thứ tự như thứ tự của chúng có trong danh sách.

Số lượng các bộ kết quả luôn nhỏ hơn hoặc bằng số lượng các bộ trong R. Nếu X không bao gồm các thuộc tính khoá của R thì quan hệ kết quả có thể có những bộ trùng nhau. Phép chiếu loại bỏ mọi bộ trùng lặp, và như vậy, kết quả của phép chiếu là một tập hợp các bộ và là một quan hệ đúng đắn.

Nếu X2 chứa tất cả các thuộc tính có trong X1 thì:

Π X1 (Π X2 (r)) = Π X1 (r)

Phép chiếu không có tính giao hoán.

Ví dụ: Cho quan hệ SINHVIEN như sau:

MASV	HOTEN	MAMH	MAKHOA	DIEMTHI
99001	TRAN DAN THU	CSDL	CNTT	3.0
99002	NGUYEN HA DA THAO	CSDL	CNTT	8.0
99001	TRAN DAN THU	THVP	CNTT	6.0
99005	LE THANH TRUNG	THVP	AV	5.0

Yêu cầu: Lập danh sách cho biết MASV, HOTEN, MAKHOA của các sinh viên thuộc khoa CNTT (Công nghệ Thông tin).

σ(MAKHOA = CNTT)(ΠMASV, HOTEN, MAKHOA(SINHVIEN))

Kết quả:

MASV	HOTEN	MAKHOA
99001	TRAN DAN THU	CNTT
99002	NGUYEN HA DA THAO	CNTT

3.1.6. Phép chọn

Mục đích: Phép chọn là phép chọn lọc ra một tập con các bộ của quan hệ R đã cho, thỏa biểu thức điều kiện cho trước.

Định nghĩa:

Cho quan hệ r(R) và điều kiện F, với F là một biểu thức luận lý có giá trị là đúng (TRUE) hoặc sai (FALSE).

Phép chọn trên r theo điều kiện F, ký hiệu là σF(r), cho kết quả là một quan hệ gồm các bộ của r sao cho thỏa mãn điều kiện F là TRUE.

σF(r) = { u | u ∈ r và F(u) = TRUE }

Biểu thức F là điều kiện chọn, được định nghĩa là một tổ hợp các phép toán logic của các biểu thức so sánh, mỗi biểu thức so sánh là một phép toán so sánh giữa 2 toán hạng. Hai toán hạng được so sánh với nhau có thể là 2 thuộc tính hoặc giữa một thuộc tính và một giá trị hằng.

Ví dụ:

+ Điều kiện: DIEM_THI < DIEM_GIUA_HP (thuộc tính điểm thi DIEM_THI < thuộc tính điểm kiểm tra giữa học phần DIEM_GIUA_HP)

+ Điều kiện: DIEM_THI ≥ 5 (thuộc tính điểm thi DIEM_THI ≥ 5)

Các phép toán luận lý (logic) trong biểu thức F: NOT (¬), AND (∧), OR (∨)

Các phép toán so sánh trong biểu thức F: =, ≠, <, ≤, >, ≥

Biểu thức F cho giá trị đúng (TRUE) hoặc sai (FALSE) đối với mỗi bộ đã cho của quan hệ khi kiểm tra riêng bộ đó. F(u) là điều kiện có được bằng cách thay thế các tên thuộc tính trong điều kiện F bởi các giá trị tương ứng trong bộ u.

Kết quả trả về là một quan hệ mới có cùng danh sách thuộc tính với r, và có số bộ luôn ít hơn hoặc bằng số bộ của r.

Phép chọn là phép toán một ngôi, nghĩa là nó được áp dụng cho một quan hệ.

Phép chọn có tính chất giao hoán:

σF1(σF2(r)) = σF2(σF1(r)) = σF1 ∧ F2(r)

Ta có thể kết hợp một loạt các phép chọn thành một phép chọn đơn giản bằng cách sử dụng phép toán AND (∧). Ví dụ:

σF1(σF2(r)) = σF2 AND F1(r)

Ví dụ: Cho quan hệ SINHVIEN như sau:

MASV	HOTEN	MAMH	MAKHOA	DIEMTHI
99001	TRAN DAN THU	CSDL	CNTT	3.0
99002	NGUYEN HA DA THAO	CSDL	CNTT	8.0
99001	TRAN DAN THU	THVP	CNTT	6.0
99005	LE THANH TRUNG	THVP	AV	5.0

Yêu cầu: Tìm các sinh viên học môn CSDL và có điểm thi trên 5

σ(MAMH = CSDL ∧ DIEMTHI > 5)(SINHVIEN)

Kết quả:

MASV	HOTEN	MAMH	MAKHOA	DIEMTHI
99002	NGUYEN HA DA THAO	CSDL	CNTT	8.0

3.1.7. Phép kết nối

+ Phép kết θ (Theta)

Phép kết được dùng để kết hợp các bộ có liên hệ với nhau từ hai quan hệ thành một bộ. Phép toán này rất quan trọng đối với cơ sở dữ liệu quan hệ có nhiều bảng bởi vì nó cho phép ta xử lý các mối liên kết giữa các quan hệ. Dạng tổng quát của phép kết trên hai quan hệ R(A1, A2, …,An) và S(B1, B2, …, Bm) là:

R ⋈<Điều kiện kết> S

Kết quả của phép kết là một quan hệ Q(A1, A2, …,An, B1, B2, …, Bm) có n+m thuộc tính. Mỗi bộ của Q là một sự kết nối giữa một bộ của R và một bộ của S khi chúng thoả mãn điều kiện kết.

Sự khác nhau giữa tích Đề-các và phép kết là ở chỗ trong phép kết, chỉ có các bộ thoả mãn điều kiện kết mới xuất hiện trong kết quả, trong khi đó trong tích Đề-các, mọi tổ hợp của các bộ đều có trong kết quả. Điều kiện kết được chỉ ra trên các thuộc tính của hai quan hệ R và S và được tính toán cho mỗi tổ hợp các bộ. Mọi tổ hợp bộ mà điều kiện kết là đúng được chứa trong quan hệ kết quả Q như là một bộ đơn.

Một điều kiện kết tổng quát có dạng:

<Điều kiện> AND <Điều kiện> AND … AND <Điều kiện>

Trong đó, mỗi điều kiện có dạng Ai θ Bj, Ai là một thuộc tính của R, Bj là một thuộc tính của S, Ai và Bj có cùng miền và θ là một trong các dấu phép toán so sánh {<, <=, =, >=, >, ≠}. Một phép toán kết với điều kiện tổng quát như vậy gọi là một phép kết θ (Tê-ta). Các bộ có các thuộc tính kết là null không xuất hiện trong kết quả.

Nói một cách ngắn gọn, phép kết θ là phép tích Đề-các có áp dụng điều kiện.

Ví dụ, giả sử ta có hai quan hệ R và S như sau:

R	A	B	C
	1	2	3
	4	5	6
	1	2	7
	8	4	5

T	A	D
	1	5
	3	7

Kết quả phép kết θ: R ⋈R.A > T.A T

R.A	B	C	T.A	D

4	5	6	1	5
4	5	6	3	7
8	4	5	1	5
8	4	5	3	7

+ Phép kết trong (Inner Join/Equijoin)

Phép kết trong là phép kết θ với điều kiện là bằng nhau (=) giữa khóa chính và khóa ngoại.

R ⋈ <R.Khóa chính = S.Khóa ngoại> S

Ví dụ: Cho 2 quan hệ SINHVIEN và KHOA như sau:

SINHVIEN	MASV	HOTEN	MAKHOA
	99001	TRAN DAN THU	CNTT
	99002	NGUYEN HA DA THAO	CNTT
	99001	TRAN DAN THU	CNTT
	99005	LE THANH TRUNG	AV

KHOA	MAKHOA	TENKHOA
	CNTT	CONG NGHE THONG TIN
	AV	ANH VAN

Kết quả phép kết trong: SINHVIEN ⋈ SINHVIEN.MAKHOA = KHOA.MAKHOA KHOA

MASV	HOTEN	MAKHOA	MAKHOA	TENKHOA
99001	TRAN DAN THU	CNTT	CNTT	CONG NGHE THONG TIN
99002	NGUYEN HA DA THAO	CNTT	CNTT	CONG NGHE THONG TIN
99001	TRAN DAN THU	CNTT	CNTT	CONG NGHE THONG TIN
99005	LE THANH TRUNG	AV	AV	ANH VAN

+ Phép kết tự nhiên (Natural Join)

Trong kết quả của phép kết trong (Inner Join), chúng ta thấy luôn luôn có một hoặc nhiều cặp thuộc tính có các giá trị như nhau trong mỗi bộ. Việc có các cặp thuộc tính

có giá trị như nhau là thừa, vì vậy người ta đề nghị một phép kết mới gọi là kết tự nhiên. Phép kết tự nhiên nhằm loại bỏ thuộc tính thứ hai (thuộc tính thừa) trong điều kiện nối bằng.

Phép kết tự nhiên hai quan hệ R và S, ký hiệu là R ⋈ S, cho kết quả là quan hệ chứa các bộ là sự phối hợp của bộ u trong quan hệ R với tất cả các bộ v trong quan hệ S sao cho các trị trên miền thuộc tính chung của hai bộ này là như nhau.

Phép kết tự nhiên không đòi hỏi hai tập thuộc tính R và S giao nhau khác rỗng.

Nếu R ∩ S = ∅ thì R ⋈ S là phép tích Descartes của R và S.

Phép kết tự nhiên của hai quan hệ R và S có thể được tính từ phép tích Đề-các, phép chọn và phép chiếu:

R ⋈ S = ΠT σF (R × S)

Trong đó: + F là biểu thức R.A1 = S.A1 ∧ R.A2 = S.A2 ∧ … ∧ R.Am = S.Am.

+ R ∩ S = {A1, A2, …, Am}.

+ T = R ∩ S

Phép kết tự nhiên có tính giao hoán: R ⋈ S = S ⋈ R

Định nghĩa chuẩn của phép kết tự nhiên đòi hỏi hai thuộc tính nối (hoặc mỗi cặp thuộc tính nối) phải có tên như nhau trong cả hai quan hệ. Nếu các thuộc tính đó không cùng tên thì trước khi nối phải áp dụng phép toán đặt lại tên.

Ví dụ 1: Sử dụng phép kết tự nhiên hai quan hệ SINHVIEN và KHOA trong ví dụ trên:

SINHVIEN ⋈ KHOA

MASV	HOTEN	MAKHOA	TENKHOA
99001	TRAN DAN THU	CNTT	CONG NGHE THONG TIN
99002	NGUYEN HA DA THAO	CNTT	CONG NGHE THONG TIN
99001	TRAN DAN THU	CNTT	CONG NGHE THONG TIN
99005	LE THANH TRUNG	AV	ANH VAN

Ví dụ 2: Cho bảng Lịch thi và bảng Môn học, hãy tổng hợp để in ra bảng kết quả gồm Mã môn, Tên môn thi, Ngày, Phòng.

LICHTHI	MaM	Ngay
	01	12/07/1994
	02	14/07/1994

MONHOC	MaM	TenM
	01	Phân tích hệ thống
	02	Cơ sở dữ liệu

03	16/07/1994
04	17/07/1994

03	Lập trình C
04	Lập trình Web

Kết quả phép kết tự nhiên MONHOC ⋈ LICHTHI:

KQ	MaM	TenM	Ngay
	01	Phân tích hệ thống	12/07/1994
	02	Cơ sở dữ liệu	14/07/1994
	03	Lập trình C	16/07/1994
	04	Lập trình Web	17/07/1994

3.1.8. Phép chia

Cho quan hệ r trên lược đồ quan hệ R(A1, A2, …, Ak, Ak+1, …, Am) và quan hệ s trên lược đồ quan hệ S (A1, A2, …, Ak).

Phép chia quan hệ r cho s với s ≠ ∅, ký hiệu là r ÷ s, cho kết quả là một quan hệ trên lược đồ quan hệ T(Ak+1, …, Am) gồm các bộ (uk+1, …, um) sao cho đối với tất cả các bộ (u1, …, uk) thuộc s thì bộ (u1, …, uk, uk+1, …, um) thuộc r.

Điều đó có nghĩa là để một bộ u xuất hiện trong kết quả T của phép chia, các giá trị trong u phải xuất hiện trong R trong sự kết nối với mọi bộ của S.

Phép chia quan hệ r cho s có thể được tính từ các phép chiếu, phép tích Đề-các, phép hiệu:

r ÷ s = ΠAk+1, …, Am (r) − ΠAk+1, …, Am ((ΠAk+1,…,Am (r) × s) – r)

Phép chia không có tính giao hoán.

r ÷ s ≠ s ÷ r

Ví dụ: Cho các quan hệ A, B1, B2 và B3 như sau:

A	SNO	PNO
	s1	p1
	s1	p2
	s1	p3
	s1	p4
	s2	p1
	s2	p2
	s3	p2
	s4	p2
	s4	p4

B1	PNO
	p2

B2	PNO
	p2
	p4

B3	PNO
	p1
	p2
	p4

Kết quả các phép chia: A÷B1, A÷B2, A÷B3

A÷B1	SNO
	s1
	s2
	s3
	s4

A÷B2	SNO
	s1
	s4

A÷B3	SNO
	s1

3.2. Khóa của sơ đồ quan hệ

3.2.1. Siêu khóa (Super Key)

Một quan hệ được định nghĩa như một tập hợp các bộ. Theo định nghĩa tập hợp, các phần tử của một tập hợp là khác nhau, vì vậy mọi bộ trong quan hệ phải khác nhau. Điều đó có nghĩa là: Không thể có 2 bộ có cùng một tổ hợp giá trị cho tất cả các thuộc tính của chúng.

Thông thường, có tồn tại các tập con của các thuộc tính của một lược đồ quan hệ có tính chất là không có 2 bộ nào ở trong mọi trạng thái quan hệ r của R có cùng một tổ hợp giá trị cho các thuộc tính của nó.

Giả sử chúng ta ký hiệu một tập con như vậy là S, khi đó với 2 bộ khác nhau bất kỳ t1 và t2 trong một trạng thái quan hệ r của lược đồ quan hệ R chúng ta có ràng buộc $\forall r, \forall t1, \forall t2 \in r, t1 \neq t2 \Rightarrow t1[S] \neq t2[S]$

Tập hợp S có các thuộc tính như vậy được gọi là một siêu khoá của lược đồ quan hệ R. Một siêu khoá S xác định rõ một ràng buộc về tính duy nhất, phát biểu rằng không có 2 bộ khác nhau trong một trạng thái r của R có cùng một giá trị cho S.

Một quan hệ có thể có nhiều siêu khoá.

Mỗi quan hệ có ít nhất một siêu khoá mặc định, đó là tập hợp tất cả các thuộc tính của nó.

3.2.2. Khóa (Key)

Một khoá K của một lược đồ quan hệ R là một siêu khoá của R với tính chất là nếu bỏ đi bất kỳ thuộc tính A nào ra khỏi K thì sẽ còn lại một tập K không phải là siêu khoá của R. Như vậy, một khoá là một siêu khoá tối thiểu (siêu khoá chứa ít thuộc tính nhất), nghĩa là đó là một siêu khoá mà ta không thể vứt bỏ thuộc tính nào ra khỏi nó mà vẫn giữ được ràng buộc về tính duy nhất.

Ví dụ: Xét quan hệ SINHVIEN như sau:

Mãsố	Họtên	Ngàysinh	Giớitính	Địachỉ
4515202	Lê Vân	12/09/84	Nữ	Hà Nội
4516802	Hoàng Tùng	21/03/84	Nam	Bắc Ninh

4620503	Trương Định	15/05/85	Nam	Hà Nam
4612203	Phạm An	16/04/85	Nam	Nam Định
4521402	Đỗ Cung	20/01/84	Nam	Nghệ An

+ Thuộc tính {Mã số} là một khoá của SINHVIEN bởi vì không có 2 bộ sinh viên có cùng một giá trị cho Mã số.

+ Mọi tập hợp thuộc tính có chứa Mã số, ví dụ: {Mã số, Họ tên, Ngày sinh}, đều là một siêu khoá. Tuy nhiên, siêu khoá {Mã số, Họ tên, Ngày sinh} không phải là khoá bởi vì nếu bỏ đi thuộc tính Họ tên hoặc Ngày sinh hoặc cả 2 thì nó vẫn còn là một siêu khoá.

Giá trị của một thuộc tính khoá có thể được sử dụng để xác định một cách duy nhất mỗi bộ trong một quan hệ. Ví dụ: Giá trị 4515202 của Mã số xác định một cách duy nhất bộ giá trị tương ứng với sinh viên Lê Vân trong quan hệ SINHVIEN.

Một khoá được xác định từ ý nghĩa của các thuộc tính và tính chất là bất biến, tính chất đó phải thỏa mãn khi chúng ta chèn thêm các bộ mới vào quan hệ. Ví dụ: Ta không thể chỉ định thuộc tính Họ tên của quan hệ SINHVIEN là khoá bởi vì có thể tồn tại 2 sinh viên có cùng Họ tên khi chúng ta chèn thêm thông tin của sinh viên mới vào.

Một lược đồ quan hệ R phải có ít nhất một khóa và có thể có nhiều khóa.

Các thuộc tính thuộc một khóa được gọi là thuộc tính khóa (prime attribute), các thuộc tính còn lại trong lược đồ quan hệ được gọi là các thuộc tính không khóa (nonprime attribute).

Quy ước:

- Trong một bộ bất kỳ của quan hệ, các thuộc tính khoá không chứa giá trị rỗng.
- Các thuộc tính tạo nên khoá của một lược đồ quan hệ được gạch dưới.

Ví dụ: Xét quan hệ Sv(**MASV**, **HOSV,TENSV, NU, NGAYSINH, MALOP**)

+ Thuộc tính khóa: MASV

+ Thuộc tính không khóa: HOSV,TENSV, NU, NGAYSINH, MALOP

Trong thực tế, đối với các loại thực thể tồn tại khách quan (Ví dụ: Sinh viên, Giảng viên, Nhân viên, Hàng hoá…) người thiết kế CSDL thường gán thêm cho các lược đồ quan hệ này một thuộc tính giả gọi là Mã số để làm khoá (Ví dụ: Mã số Sinh viên, Mã số Giảng viên, Mã số Nhân viên, Mã số Hàng hoá…). Trong khi đó, các lược đồ quan hệ biểu diễn cho sự trừu tượng hoá thường có khoá là một tổ hợp của 2 hay nhiều thuộc tính của nó.

Ví dụ: Xét lược đồ quan hệ: Kq(MASV, MAMH, DIEMTHI) quản lý điểm thi của sinh viên ứng (MASV) với từng môn học (MAMH) cụ thể. Một sinh viên có thể thi nhiều môn, và nhiều sinh viên có thể cùng thi chung một môn. Như vậy, ta chọn khóa là tập hợp 2 thuộc tính: {MASV, MAMH}.

3.2.3. Khóa dự tuyển (Candidate Key) – Khóa chính (Primary Key)

Một lược đồ quan hệ có thể có nhiều hơn một khoá. Trong trường hợp đó, mỗi một khoá được gọi là một khoá dự tuyển.

Một trong các khóa dự tuyển được chọn làm khóa tiêu biểu, khóa này được gọi là khóa chính. Ta phải chỉ định một trong các khoá dự tuyển làm khoá chính của quan hệ.

Một quan hệ chỉ có một khóa chính và có thể có nhiều khóa dự tuyển. Chú ý rằng khi một lược đồ quan hệ có nhiều khoá dự tuyển, việc lựa chọn một khoá dự tuyển để làm khoá chính là tuỳ ý, tuy nhiên tốt nhất là chọn khoá chính gồm một thuộc tính hoặc có số lượng các thuộc tính ít nhất.

Khóa chính được chọn để cài đặt trong một hệ quản trị CSDL.

Khi chọn khóa chính ta phải chú ý các tính chất sau:

- Khóa có tính áp dụng khi nó không bỏ sót bất kỳ trường hợp nào của vấn đề.
- Khóa phải có tính duy nhất dùng để phân biệt bộ này với bộ kia trong quan hệ.
- Khóa có tính nhỏ nhất: Khi ta bỏ qua bất kỳ thuộc tính nào của nó thì nó không còn tính duy nhất nữa.
- Khóa có tính ổn định khi giá trị của khóa không thay đổi.

Ví dụ: Xét quan hệ SINHVIEN sau:

SINHVIEN(MaSV, SoCMND, TenSV, Ngaysinh, Gioitinh, QueQuan)

Ta xét các tập thuộc tính sau:

K1={MaSV, TenSV} $\rightarrow$ Là siêu khóa

K2={Ngaysinh, QueQuan} $\rightarrow$ Không là khóa

K3={NgaySinh, TenSV, QueQuan} $\rightarrow$ Không là khóa

K4={MaSV} $\rightarrow$ Là siêu khóa (khóa)

K5={MaSV, TenSV, Ngaysinh, Gioitinh, QueQuan} $\rightarrow$ Là siêu khóa

K6={MaSV, Ngaysinh} $\rightarrow$ Là siêu khóa

K7={SoCMND} $\rightarrow$ Là siêu khóa (khóa)

K4={MaSV}, K7={SoCMND} $\rightarrow$ Là các khóa dự tuyển

K4={MaSV} hoặc K7={SoCMND} $\rightarrow$ Là khóa chính (tùy theo sự lựa chọn trong hệ quản trị CSDL)

3.2.4. Khóa ngoại (Foreign Key)

Một thuộc tính được gọi là khóa ngoại nếu nó là thuộc tính của một lược đồ quan hệ này nhưng lại là khóa chính của lược đồ quan hệ khác.

Xét 2 lược đồ quan hệ R và S, gọi F là tập thuộc tính khóa ngoại của R tham chiếu đến khóa chính của S. Các thuộc tính khóa ngoại của R phải có cùng miền giá trị với thuộc tính khóa chính của S.

Quy ước: Các thuộc tính khoá ngoại được gạch dưới theo kiểu không liền nét.

Ví dụ: Xét lược đồ CSDL gồm các lược đồ quan hệ sau:

Sv(MASV, HOSV,TENSV, NU, NGAYSINH, MALOP)

Lop(MALOP, TENLOP, SISO, MAKHOA)

+ MASV là khoá chính của lược đồ quan hệ Sv.

+ MALOP là khoá chính của lược đồ quan hệ Lop.

+ MALOP là khoá ngoại của lược đồ quan hệ Sv.

Chú ý:

- Trong một lược đồ quan hệ, một thuộc tính có thể vừa tham gia vào khóa chính, vừa tham gia vào khóa ngoại.
- Khóa ngoại có thể tham chiếu đến khóa chính trong cùng một lược đồ quan hệ.
- Có thể có nhiều khóa ngoại tham chiếu đến cùng một khóa chính.

3.3. Bài tập ứng dụng

3.3.1. Phép toán tập hợp và phép toán quan hệ

Cho lược đồ CSDL dùng để quản lý điểm sinh viên được mô tả như sau:

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

Lop(MALOP, TENLOP, MAKHOA)

Khoa(MAKHOA, TENKHOA)

Monhoc(MAMH, TENMH, DONVIHT)

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

Ketqua(MASV, MAMH, LANTHI, DIEMTHI)

Phancong(MALOP, MAMH, MAGV)

Viết biểu thức ngôn ngữ Đại số quan hệ để thực hiện các yêu cầu:

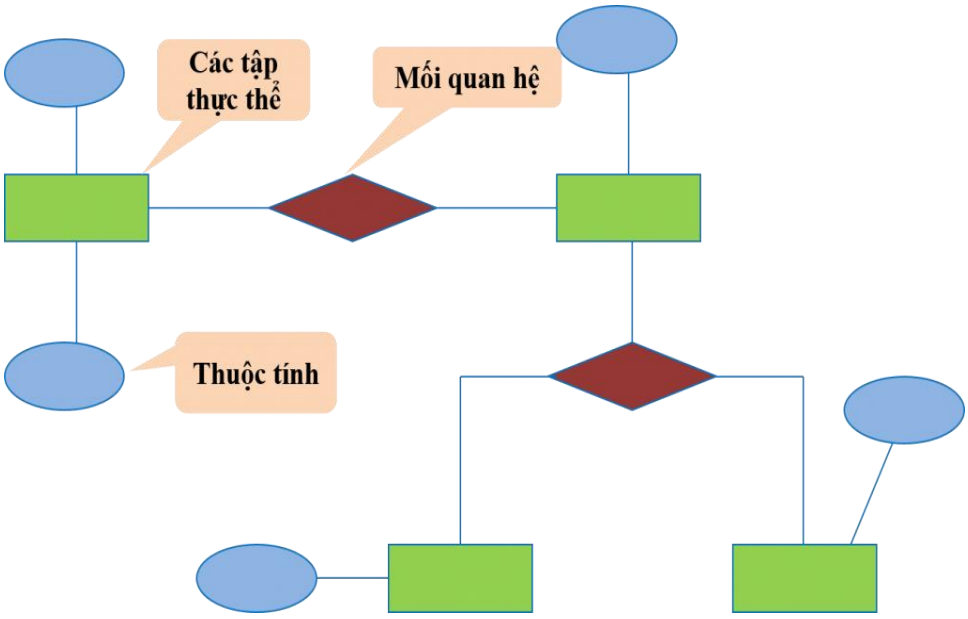
Lập danh sách các sinh viên lớp có mã lớp là CDTH2A, danh sách cần các thông tin: MASV, HOTENSV

Lập danh sách sinh viên nữ và có mã khoa là CNTT, danh sách cần các thông tin: MASV, HOTENSV.

Lập bảng điểm thi lần 1 của tất cả các môn cho sinh viên lớp CDTH2A, danh sách cần các thông tin: MASV, HOTENSV, TENMH, DIEMTHI.
Lập phiếu điểm thi lần 1 các môn cho sinh viên có MASV là 00CDTH189, danh sách cần các thông tin: MAMH,TENMH, DONVIHT, DIEMTHI.
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
Lập danh sách các sinh viên của lớp có MALOP là CDTH2A, danh sách cần các thông tin: MASV, HOTENSV, NGAYSINH, TINH.
Lập danh sách các giảng viên có cấp học vị là THAC SY của khoa có MAKHOA là CNTT, danh sách cần các thông tin: MAGV, HOTENGV, CHUYENNGANH.
Lập bảng điểm thi lần 1 môn học 869 cho tất cả sinh viên thuộc hai lớp có MALOP là CĐTH2A và CĐTH2B, danh sách cần: MASV, HOTENSV, DIEMTHI.
Lập danh sách các giảng viên đã dạy lớp CĐTH2A, danh sách cần các thông tin: MAGV, HOTENGV, TENKHOA, HOCVI, TENMH.
Lập danh sách các môn mà lớp CDTH2A đã học, danh sách cần các thông tin: MAMH, TENMH, DONVIHT, HOTENGV.
Lập danh sách những giảng viên đã dạy sinh viên có MASV là 00CDTH189, danh sách cần MAGV, HOTENGV, HOCVI, CHUYENNGANH, TENKHOA, TENMH
Lập danh sách các sinh viên có mã khoa CNTT có điểm thi lần 1 môn học 869 lớn hơn hoặc bằng 8, danh sách cần MASV, HOTENSV, DIEMTHI, TENLOP.

3.3.2. Mô hình thực thể kết hợp

Mô hình thực thể kết hợp được sử dụng để biểu diễn cơ sở dữ liệu ở mức khái niệm. Mô hình thực thể kết hợp bao gồm có các thực thể, danh sách thuộc tính và những mối kết hợp.
Biểu diễn mô hình thực thể kết hợp dưới dạng sơ đồ thực thể liên kết (Entity Relationship Diagram – ERD).
Mô hình thực thể kết hợp để biểu diễn mối liên kết giữa các thực thể trong thiết kế mô hình mức khái niệm. Các thành phần chủ yếu của mô hình thực thể liên kết là thực thể, thuộc tính và mối kết hợp.
a. Mô hình thực thể kết hợp – Sơ đồ ERD



Hình 3.1. Mô hình thực thể kết hợp

b. Để tạo ERD, chúng ta thực hiện 2 bước sau:

Bước 1. Xác định thực thể và thuộc tính bao gồm thuộc tính khoá và thuộc tính đa trị

Bước 2. Xác định mối quan hệ có thể có giữa các thực thể, thuộc tính mối kết hợp

Bài tập vận dụng: Thiết kế cơ sở dữ liệu quản lý forum

Các thành viên tham gia diễn đàn có mã số là duy nhất, họ tên thành viên và mật khẩu. Các bài viết có mã số bài viết, tiêu đề của bài viết và nội dung bài viết. Các chủ đề của các bài viết có mã số chủ đề (thuộc tính khóa) và tên của chủ đề.

Mỗi bài viết được đăng bởi một thành viên và mỗi thành viên có thể đăng nhiều bài viết. Một bài viết phải thuộc một chủ đề và một chủ đề có thể có nhiều bài viết.

❖ TÓM TẮT CHƯƠNG 3

Trong chương này, một số nội dung chính được giới thiệu:

1. **Ngôn ngữ SQL**: Giới thiệu về SQL và tầm quan trọng của nó trong việc truy vấn, cập nhật và quản lý dữ liệu trong cơ sở dữ liệu quan hệ.
2. **Truy vấn Dữ liệu**: Học sinh sẽ học cách sử dụng câu lệnh SQL SELECT để truy vấn dữ liệu từ cơ sở dữ liệu. Chương bao gồm cách lựa chọn dữ liệu từ bảng, sử dụng điều kiện để lọc dữ liệu, và sắp xếp kết quả.
3. **Cập nhật Dữ liệu**: Giới thiệu về việc cập nhật dữ liệu bằng cách sử dụng câu lệnh SQL INSERT, UPDATE và DELETE
4. **Ràng buộc Dữ liệu**: Học sinh sẽ hiểu cách sử dụng các ràng buộc dữ liệu như khóa chính (primary key) và khóa ngoại (foreign key) để đảm bảo tính toàn vẹn và quy tắc dữ liệu trong cơ sở dữ liệu.

❖ CÁC BÀI TẬP HỆ THỐNG KIẾN THỨC

1. Truy vấn Dữ liệu

Sử dụng một cơ sở dữ liệu đơn giản (ví dụ: danh sách nhân viên hoặc sản phẩm), hãy viết một số câu truy vấn SQL để trích xuất thông tin từ cơ sở dữ liệu. Ví dụ: lấy danh sách tất cả những sản phẩm có giá trên 100 đơn vị tiền.

2. Cập nhật Dữ liệu

Sử dụng cơ sở dữ liệu từ Bài tập 1, hãy viết câu lệnh SQL để cập nhật thông tin về một số bản ghi. Ví dụ: cập nhật giá của sản phẩm có ID là 1 thành 150 đơn vị tiền.

3. Ràng buộc Dữ liệu

Thiết kế một cơ sở dữ liệu đơn giản với mối quan hệ giữa hai bảng (ví dụ: bảng người dùng và bảng sản phẩm) và sử dụng ràng buộc dữ liệu như khóa chính và khóa ngoại. Sau đó, thử thêm một bản ghi mới vào bảng sản phẩm và đảm bảo rằng nó tuân theo các ràng buộc dữ liệu đã định nghĩa.

CHƯƠNG 4. HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU SQL

❖ GIỚI THIỆU CHƯƠNG 4

Chương 4, "Hệ quản trị cơ sở dữ liệu SQL," tập trung vào hệ quản trị cơ sở dữ liệu (DBMS) SQL, là một phần quan trọng trong việc quản lý cơ sở dữ liệu và tương tác với dữ liệu

❖ MỤC TIÊU CHƯƠNG 4

Sau khi học xong chương này, người học có khả năng:

➢ ***Về kiến thức:***

- *hiểu về vai trò của DBMS SQL trong việc quản lý cơ sở dữ liệu.*
- *hiểu cách cài đặt và quản lý một hệ thống cơ sở dữ liệu sử dụng DBMS SQL cụ thể*
- *Hiểu hiểu về lịch sử và sự phát triển của cơ sở dữ liệu, cùng với việc tìm hiểu về các hệ quản trị cơ sở dữ liệu phổ biến*

➢ ***Về kỹ năng:***

- *Biết quản lý cơ sở dữ liệu bằng cách thực hiện các tác vụ như tạo bảng, thêm người dùng, sao lưu và khôi phục dữ liệu, và thiết lập bảo mật.*
- *Biết cách sử dụng các công cụ quản trị DBMS SQL để thực hiện các nhiệm vụ quản lý cơ bản.*

➢ ***Về năng lực tự chủ và trách nhiệm:***

- *Năng lực về quản lý thời gian, trách nhiệm với công việc*
- *Năng lực học tập và làm việc độc lập*
- *Tự chủ trong việc giải quyết vấn đề*

❖ PHƯƠNG PHÁP GIẢNG DẠY VÀ HỌC TẬP CHƯƠNG 4

- *Đối với người dạy: sử dụng phương pháp giảng giảng dạy tích cực (diễn giảng, vấn đáp, dạy học theo vấn đề); yêu cầu người học thực hiện câu hỏi thảo luận và bài tập chương (cá nhân hoặc nhóm).*
- *Đối với người học: chủ động đọc trước giáo trình trước buổi học; hoàn thành đầy đủ câu hỏi thảo luận và bài tập tình huống theo cá nhân hoặc nhóm và nộp lại cho người dạy đúng thời gian quy định.*

❖ ĐIỀU KIỆN THỰC HIỆN CHƯƠNG 4

- ***Phòng học chuyên môn hóa/nhà xưởng:*** Phòng học thực hành
- ***Trang thiết bị máy móc:*** Máy chiếu, máy tính và các thiết bị dạy học khác
- ***Học liệu, dụng cụ, nguyên vật liệu:*** Chương trình môn học, giáo trình, tài liệu tham khảo, giáo án, phim ảnh, và các tài liệu liên quan.

- ***Các điều kiện khác:*** Không có

❖ **KIỂM TRA VÀ ĐÁNH GIÁ CHƯƠNG 4**

- **Nội dung:**

✓ *Kiến thức: Kiểm tra và đánh giá tất cả nội dung đã nêu trong mục tiêu kiến thức*

✓ *Kỹ năng: Đánh giá tất cả nội dung đã nêu trong mục tiêu kĩ năng.*

✓ *Năng lực tự chủ và trách nhiệm: Trong quá trình học tập, người học cần:*

+ Nghiên cứu bài trước khi đến lớp

+ Chuẩn bị đầy đủ tài liệu học tập.

+ Tham gia đầy đủ thời lượng môn học.

+ Nghiêm túc trong quá trình học tập.

- **Phương pháp:**

✓ ***Điểm kiểm tra thường xuyên:*** *1 điểm kiểm tra (hình thức: hỏi miệng)*

✓ ***Kiểm tra định kỳ lý thuyết:*** *không có*

❖ **NỘI DUNG**

4.1. Giới thiệu

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 Manipullation 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 kiểu 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). 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 hoá ngôn ngữ SQL và sau đó tổ chức tiêu chuẩn thế giới (International Standards Organization - ISO) cũng đã công nhân ngôn ngữ này.
Đó là chuẩn SQL-86. tới này SQL đã qua 3 lần chuẩn hoá (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.

4.2. Các lệnh truy vấn SELECT

4.2.1. Câu lệnh tìm kiếm tổng quát

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 phần này, chúng ta chỉ tìm hiểu câu lệnh quan trọng nhất của SQL đó là câu lệnh hỏi, tìm kiếm dữ liệu SELECT. Kết quả của lệnh SELECT là một quan hệ, quan hệ kết quả này có thể kết xuất ra màn hình, máy in, hoặc là trên các thiết bị lưu trữ thông tin khác. Để đơn giản trong cách trình bày, ta xem quan hệ để thực hiện câu truy vấn là quan hệ nguồn và quan hệ kết quả của truy vấn là quan hệ đích.
Mỗi câu lệnh SQL có thể được viết trên nhiều dòng và kết thúc lệnh bởi dấu chấm phẩy (;), tuy nhiên từ khoá, tên hàm, tên thuộc tính, tên bảng, tên đối tượng thì không được phép viết tách xuống hàng. Trong vận dụng thực tế, từ khoá, tên thuộc tính, tên bảng, tên đối tượng được viết in hoa hay chữ thường là như nhau.
Cú pháp tổng quát của câu lệnh SELECT như sau:
SELECT Tên cột<Danh sách thuộc tính>
FROM Tên bảng <Danh sách các quan hệ>
SELECT DISTINCT <Danh sách thuộc tính> <Biểu thức>...
FROM <Danh sách các quan hệ>
WHERE <Biểu thức>
GROUP BY <Danh sách thuộc tính gom nhóm>
HAVING <Điều kiện nhóm>
ORDER BY <Danh sách các thuộc tính> [ASC | DESC]
Trong đó:
<Biểu thức> là sự kết hợp một cách hợp lệ giữa các thuộc tính, các toán tử và các hàm.
Sau đây sẽ là các toán tử và hàm thông dụng nhất:

a. Các toán tử số học:

^ (luỹ thừa), * (nhân), / (chia), MOD (phần dư), + (cộng), - (trừ)

b. Các toán tử luận lý:

NOT (phủ định), AND (phép hội), OR (phép tuyển)

c. Các toán tử tập hợp:

IN (danh sách các giá trị), LIKE, BETWEEN … AND, NOT LIKE, UNION (phép hợp), INTERSECT (phép giao), MINUS (phép trừ)

d. Các toán tử so sánh:

=, <>, >, <, >=, <=

Lưu ý: Trong một số phiên bản của SQL, phép toán <> có thể được viết dạng !=

e. Các hàm xử lý ngày tháng:

DATE(): Trả về ngày tháng năm của hệ thống

TIME(): Trả về giờ phút giây của hệ thống

DAY(Biểu thức ngày): Trả về một trị số từ 1 đến 31 của biểu thức ngày.

MONTH(Biểu thức ngày): Trả về một số từ 1 đến 12 - là tháng của biểu thức ngày

YEAR(Biểu thức ngày): Trả về năm của biểu thức ngày

LEN(Biểu thức chuỗi): Trả về chiều dài của chuỗi

Sau đây ta sẽ lần lượt tìm hiểu kỹ hơn các mệnh đề của câu lệnh.

Ví dụ lược đồ CSDL sau minh họa cho các truy vấn dữ liệu:

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

Lop(MALOP,TENLOP,MAKHOA)

Khoa(MAKHOA,TENKHOA)

Monhoc(MAMH,TENMH,DONVIHT)

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

Ketqua(MASV, MAMH, LANTHI,DIEMTHI)

Phancong(MALOP,MAMH,MAGV)

4.2.2. Tìm thông tin từ các cột của bảng - Mệnh đề SELECT

SELECT DISTINCT * <Danh sách thuộc tính> <Biểu thức>…

FROM <Danh sách các quan hệ>

Những thuộc tính được liệt kê trong mệnh đề SELECT sẽ là các thuộc tính có trong quan hệ đích.

Ký hiệu * theo sau từ khóa SELECT dùng để chỉ tất cả các thuộc tính của quan hệ nguồn sẽ là thuộc tính của quan hệ đích. Danh sách các thuộc tính cách nhau bởi dấu phẩy và thứ tự này cũng là thứ tự của các thuộc tính trong quan hệ đích.

Mệnh đề FROM:

Những quan hệ liên quan đến câu truy vấn được liệt kê sau mệnh đề FROM, các quan hệ này cách nhau bởi dấu phẩy, thứ tự của các quan hệ được chỉ ra ở đây là không quan trọng.

Cần chú ý rằng khi mệnh đề FROM chỉ ra từ 2 quan hệ trở lên, nếu có một thuộc tính ở mệnh đề SELECT là thuộc tính của nhiều hơn một quan hệ thì cần phải chỉ rõ thuộc tính đó thuộc về quan hệ nào theo cú pháp <Tên quan hệ>.<Tên thuộc tính>. Đây là lỗi thường gặp khi thực hành với câu lệnh truy vấn SQL.

Ví dụ 1: Lập danh sách sinh viên gồm MASV, HOTENSV, NU, NGAYSINH, TINH

Cách 1:

```
SELECT    MASV, HOTENSV, NU, NGAYSINH, TINH
FROM      Sinhvien;
```

Cách 2:

```
SELECT    *
FROM      Sinhvien;
```

Khi cần lấy thông tin về tất cả các cột của bảng, chúng ta có thể sử dụng dấu sao (*) thay cho việc liệt kê các tên cột của bảng. Nếu áp dụng cách viết này thì câu lệnh trên tương đương với câu lệnh sau:

```
SELECT * FROM Sinhvien;
```

Cú pháp này chỉ được sử dụng khi câu truy vấn chỉ liên quan đến một quan hệ.

Ví dụ 2: Lập danh sách bao gồm các thông tin về giảng viên như mã số giảng viên, họ và tên giảng viên, học vị, chuyên ngành.

```
SELECT    MAGV, HOTENGV, HOCVI, CHUYENNGANH
FROM      Giangvien;
```

Câu lệnh tìm kiếm thông tin từ các cột của bảng ở trên là cài đặt của phép chiếu trên 4 thuộc tính MAGV, HOTENGV, HOCVI, CHUYENNGANH của quan hệ Giangvien.

Nếu chúng ta muốn đặt tên khác cho tên của các cột của bảng (còn gọi là bí danh - Alias), việc này được thực hiện bằng cách thêm từ khóa AS và theo sau là một tên mới. Nếu tên có chứa các ký tự đặc biệt và/hoặc khoảng trắng thì viết tên đó trong cặp dấu ngoặc vuông ([]).

Chẳng hạn Ví dụ 2 có thể viết lại là:

```
SELECT MAGV AS [Mã Giảng viên], HOTENGV AS [Họ tên] HOCVI [Trình độ]
CHUYENNGANH AS [Chuyên ngành]
```

FROM Giangvien;

Câu lệnh SELECT không chỉ thực hiện việc trích thông tin từ các cột đơn lẻ của bảng mà còn có thể thực hiện các tính toán theo công thức hay biểu thức bất kỳ dựa trên giá trị của các cột trên từng bản ghi của bảng.

Từ khóa DISTINCT nhằm loại bỏ bớt các bộ trùng nhau trong bảng kết quả của lệnh truy vấn (chỉ giữ lại một bộ đại diện cho các bộ giống nhau).

Ví dụ 3: Hãy cho biết các giảng viên của trường thuộc những chuyên ngành nào?

```
SELECT    DISTINCT CHUYENNGANH
FROM      Giangvien;
```

Kết quả của câu lệnh này là tất cả những chuyên ngành mà các giảng trong trường có thể đảm nhận, mỗi chuyên ngành chỉ xuất hiện một lần trong kết quả truy vấn được.

4.2.3. Chọn các dòng của bảng - Mệnh đề WHERE

SELECT DISTINCT * <Danh sách thuộc tính> <Biểu thức>…

FROM <Danh sách các quan hệ>

WHERE <Biểu thức điều kiện>

Trong đó <Biểu thức điều kiện> có giá trị là hoặc đúng (TRUE) hoặc sai (FALSE). Đây là sự cài đặt của phép chọn trong ngôn ngữ đại số quan hệ.

Nếu điều kiện này chỉ liên quan đến một quan hệ thì gọi là điều kiện chọn, nếu điều kiện liên quan đến từ hai quan hệ trở lên thì gọi là điều kiện kết. Các điều kiện chọn và điều kiện kết có thể phối hợp với nhau bởi các toán tử logic (AND, OR, NOT) để tạo nên những biểu thức logic phức tạp hơn. Cần chú ý rằng thứ tự của các điều kiện ở đây là quan trọng: Nếu có thể thì nên thực hiện điều kiện chọn trước khi thực hiện điều kiện kết.

Sau đây là một số ví dụ cho phép chọn.

Ví dụ 4:

Lập danh sách những môn học có số đơn vị học trình >=4. Danh sách cần MAMH, TENMH, DONVIHT.

```
SELECT    MAMH, TENMH, DONVIHT
FROM      Monhoc
WHERE     DONVIHT >=4;
```

Ví dụ 5: Lập danh sách các sinh viên có mã lớp là CDTH2A, CDTH2B, CDTH2C.

```
SELECT * FROM Sinhvien
```

WHERE MALOP="CDTH2A" OR MALOP="CDTH2B" OR MALOP="CDTH2C";

Cũng có thể viết cách khác như sau:

SELECT * FROM Sinhvien

WHERE MALOP IN ("CDTH2A","CDTH2B","CDTH2C");

Ví dụ 6: Lập danh sách những sinh viên lớp CDTH2A có điểm thi lần 1 môn CSDL từ 6 đến 8

SELECT MASV, DIEMTHI

FROM Ketqua

WHERE LANTHI=1 AND DIEMTHI>=6 AND DIEMTHI<=8 AND MAMH="CSDL";

Hoặc có thể viết cách khác

SELECT MASV, DIEMTHI

FROM Ketqua

WHERE LANTHI=1 AND DIEMTHI BETWEEN 6 AND 8 AND MAMH="CSDL"

a. Toán tử so sánh tương đối: LIKE

Mẫu so sánh trong phép toán LIKE là một giá trị kiểu text, đó là một dãy ký tự bất kỳ, trong đó có hai ký tự có ý nghĩa đặc biệt sau đây:

? đại diện cho một ký tự bất kỳ tại vị trí có dấu chấm hỏi.

% đại diện cho một nhóm ký tự bất kỳ tại vị trí đó.

Ví dụ 7: Lập danh sách các sinh viên có họ là Nguyễn đang theo học tại lớp có mã lớp là CDTH2A

SELECT * FROM Sinhvien

WHERE MALOP="CDTH2A" AND HOTENSV LIKE "Nguyễn%";

Lưu ý: Một số hệ quản trị CSDL cho phép sử dụng dấu nháy đơn (') hoặc nháy kép (") bao quanh các giá trị ở dạng chuỗi văn bản.

b. Toán tử BETWEEN … AND

Toán tử BETWEEN … AND lấy ra một miền dữ liệu nằm giữa 2 giá trị kiểu số, chuỗi văn bản hoặc ngày tháng.

SELECT <Thuộc tính>

FROM <Quan hệ>

WHERE <Thuộc tính> BETWEEN <Giá trị 1> AND <Giá trị 2>;

4.2.4. Kết nối nhiều bảng dữ liệu

Để truy vấn thông tin từ nhiều bảng dữ liệu, ta sử dụng phép kết INNER JOIN giữa các bảng trong mệnh đề FROM.

```
SELECT <Thuộc tính>
FROM   <Danh sách các quan hệ>
WHERE <Điều kiện kết>;
```

Hoặc có thể viết cách khác:

```
SELECT <Thuộc tính>
FROM   <Quan hệ 1> INNER JOIN <Quan hệ 2> ON <Điều kiện kết>;
```

Ví dụ kết nối 2 bảng Sinhvien và Lop:

```
SELECT   *
FROM     Sinhvien, Lop
WHERE    Sinhvien.MALOP=Lop.MALOP;
```

Hoặc:

```
SELECT   *
FROM     Sinhvien INNER JOIN Lop ON Sinhvien.MALOP=Lop.MALOP;
```

Ví dụ 7: Lập danh sách các sinh viên lớp CDTH2A có điểm thi môn học có mã môn học là "869" lớn hơn hay bằng 8.0

```
SELECT   Sinhvien.MASV, HOTENSV, NU, NGAYSINH, DIEMTHI
FROM     Sinhvien, Ketqua
WHERE    Sinhvien.MASV=Ketqua.MASV AND MALOP="CDTH2A"
         AND MAMH="869" AND DIEMTHI>=8.0;
```

Cần chú ý rằng do thuộc tính MASV xuất hiện ở cả hai quan hệ Sinhvien và Ketqua, nên khi liệt kê nó ở mệnh đề SELECT cần chỉ rõ ra nó thuộc quan hệ nào.

4.2.5. Sắp xếp các dòng của bảng - Mệnh đề ORDER BY:

Quan hệ đích có thể được sắp xếp tăng/giảm theo một (hoặc nhiều) thuộc tính nào đó bằng cách sử dụng mệnh đề ORDER BY (độ ưu tiên giảm dần tư trái sang phải), từ khóa DESC (Descending) được dùng nếu muốn sắp xếp giảm dần, nếu không có DESC,mặc định CSDL sẽ được sắp xếp tăng dần ASC (Ascending) theo các thuộc tính đã chỉ ra.

Nghĩa là danh sách các lớp được sắp xếp theo cột Mã khoa, nếu cột Mã khoa trùng nhau thì sắp xếp theo cột số học viên. Sau đây là vấn đề truy vấn thông tin từ nhiều bảng dữ liệu.

Ví dụ 8: Lập danh sách cần các thông tin MALOP, TENLOP, MAKHOA sắp xếp theo MAKHOA

```
SELECT    MALOP, TENLOP, MAKHOA
FROM      Lop
ORDER BY MAKHOA
```

4.2.6. Các lệnh truy vấn lồng nhau

Là những câu lệnh mà trong thành phần WHERE có chứa thêm một câu lệnh SELECT khác nữa. Câu lệnh này thường gặp khi dữ liệu cần thiết phải duyệt qua nhiều lần. Đây là một trong những vấn đề khó khăn nhất khi truy vấn dữ liệu.

Ví dụ 9: Lập danh sách những sinh viên lớp CDTH2A có điểm thi lần 1 môn học CSDL cao nhất.

Với câu lệnh này, nếu dùng các ngôn ngữ lập trình không có ngôn ngữ hỏi cấu trúc sẽ rất dài dòng. Đầu tiên phải tìm cho ra số điểm lớn nhất thỏa mãn điều kiện trên, sau đó phải duyệt dữ liệu thêm một lần nữa để chọn ra những bộ thỏa đề bài.

```
SELECT    Sinhvien.MASV, HOTENSV, NU, NGAYSINH, DIEMTHI
FROM      Sinhvien, Ketqua
WHERE     MAMH='CSDL' AND Lanthi=1 AND
          Sinhvien.MASV=Ketqua.MASV AND
          DIEMTHI>=ALL
          (SELECT   DIEMTHI
          FROM            Ketqua, Sinhvien
          WHERE           MAMH='CSDL' AND Lanthi=1
                    AND Sinhvien.MASV=Ketqua.MASV);
```

Ví dụ 10: Lập danh sách những giảng viên cùng khoa với giảng viên NGUYEN VAN THANH

```
SELECT    *
FROM      Giangvien
WHERE     Makhoa IN
          (SELECT   MAKHOA
          FROM            Giangvien
          WHERE           HOTENGV="NGUYEN VAN THANH");
```

Kết quả của câu hỏi con được sử dụng trong phép so sánh với một giá trị khác trong biểu thức điều kiện của câu hỏi bao nó. Các phép so sánh có dạng:

<Phép so sánh>[<Lượng từ>](SELECT Câu hỏi con)

Trong đó phép so sánh có thể là phép so sánh số học hoặc phép so sánh trên tập hợp. <Lượng từ> có thể là ALL, ANY (hoặc SOME). Phép so sánh bằng ANY có thể được thay tương đương bằng phép toán IN, phép so sánh <>ALL có thể thay tương đương bằng phép toán NOT IN.

4.2.7. Gom nhóm dữ liệu - Mệnh đề GROUP BY

Khi cần tính toán trên các bộ theo một nhóm, một thuộc tính nào đó, ta dùng mệnh đề GROUP BY, chẳng hạn cần tính điểm trung bình chung tất cả các môn học cho tất cả các sinh viên, hay là cần tính số lượng sinh viên cho mỗi lớp, mỗi khoa, đếm số lượng sinh viên nữ của mỗi khoa, đếm số lượng sinh viên của mỗi tỉnh…

Mệnh đề GROUP BY <Thuộc tính gom nhóm> dùng để phân nhóm dữ liệu. Những bộ của bảng có cùng giá trị trên các thuộc tính này sẽ tạo thành một nhóm.

Ví dụ 11: Lập bảng Điểm TB lần 1 các môn học của các sinh viên có mã lớp là CDTH2A. Danh sách cần: MASV, HOTENSV, DIEMTB. Trong đó DIEMTB là thuộc tính tự đặt.

```
SELECT    KETQUA.MASV, HOTENSV, AVG(DIEMTHI) AS DIEMTB
FROM      Sinhvien, Ketqua
WHERE     Sinhvien.MASV=Ketqua.MASV
          AND MALOP="CDTH2A" AND LANTHI=1
GROUP BY Ketqua.MASV, HOTENSV
```

Mệnh đề HAVING <Điều kiện trên nhóm>

Nếu cần kiểm tra điều kiện của một nhóm thì dùng mệnh đề HAVING, chẳng hạn như cho biết những sinh viên nào có điểm trung bình các môn ≥ 8, những khoa nào có nhiều hơn 100 sinh viên nữ…

Lưu ý những thuộc tính có tham gia vào mệnh đề GROUP BY để phân nhóm phải được liệt kê trong danh sách thuộc tính theo sau từ khóa SELECT.

Mệnh đề HAVING <điều kiện trên nhóm> được sử dụng như là phép chọn phối hợp với việc phân nhóm dữ liệu.

Ví dụ 12: Lập bảng Điểm TB lần 1 lớn hơn hoặc bằng 8.0 các môn đã thi của các sinh viên có mã lớp là CDTH2A. Danh sách cần: MASV, HOTENSV, DIEMTB. Trong đó DIEMTB là thuộc tính tự đặt.

Giống như ở Ví dụ 12 nhưng có thêm điều kiện là điểm trung bình các môn đã thi lớn hơn hoặc bằng 8.0.

```
SELECT    KETQUA.MASV, HOTENSV, AVG(DIEMTHI) AS DIEMTB
FROM      SINHVIEN, KETQUA, LOP
WHERE     MALOP="CDTH2A" AND LANTHI=1
          AND SINHVIEN.MASV=KETQUA.MASV
GROUP BY KETQUA.MASV, HOTENSV
HAVING    AVG(DIEMTHI)>=8.0;
```

4.2.8. Hàm tính toán theo nhóm

Các hàm tính toán theo nhóm (đầu vào là một tập giá trị và trả về một giá trị đơn):

SUM(Thuộc tính): tính tổng giá trị của các bộ theo thuộc tính đã chỉ ra.

MAX(Thuộc tính): cho biết giá trị lớn nhất của các bộ theo thuộc tính đã chỉ ra.

MIN(Thuộc tính): cho biết giá trị nhỏ nhất của các bộ theo thuộc tính đã chỉ ra.

AVG(Thuộc tính): cho biết giá trị trung bình của các bộ theo thuộc tính đã chỉ ra.

COUNT *: Đếm tất cả các bộ.

COUNT(Thuộc tính): chỉ đếm những bộ mà giá trị của thuộc tính là khác.

COUNT DISTINCT (Thuộc tính): Chỉ đếm những bộ mà giá trị của thuộc tính là khác NULL, hơn nữa, những bộ mà giá trị trùng nhau trên thuộc tính chỉ được đếm là một (đại diện cho cả nhóm).

Chú ý: Cách sử dụng các toán tử và các hàm này còn tuỳ thuộc vào câu lệnh SELECT của ngôn ngữ được sử dụng.

Ví dụ: Cho biết số lượng sinh viên trong toàn trường

```
SELECT    COUNT(MASV)
FROM      Sinhvien;
```

Ví dụ: Cho biết số lượng sinh viên theo từng lớp (mã lớp)

```
SELECT    MALOP, COUNT(MASV)
FROM      Sinhvien
GROUP BY MALOP;
```

Thứ tự dịch một lệnh truy vấn tổng hợp là như sau:

FROM → WHERE → GROUP BY → HAVING → SELECT → ORDER BY

4.3. Các lệnh truy vấn UPDATE

Truy vấn UPDATE trong SQL được sử dụng để sửa đổi các bản ghi đang tồn tại trong một bảng. Bạn có thể sử dụng mệnh đề WHERE với truy vấn UPDATE để cập nhật các hàng đã được lựa chọn, nếu không, tất cả các hàng sẽ bị tác động.

Cú pháp cơ bản của truy vấn UPDATE với mệnh đề WHERE như sau:

UPDATE ten_bang
SET cot1 = giatri1, cot2 = giatri2...., cotN = giatriN
WHERE [dieu_kien];

Bảng SINHVIEN có các bản ghi sau:

```
+----+---------+-----+-----------+--------+
| ID |  TEN    | TUOI|  KHOAHOC  | HOCPHI |
+----+---------+-----+-----------+---------+
|  1 | Hoang   |  21 | CNTT      | 4       |
|  2 | Viet    |  19 | DTVT      | 3.0     |
|  3 | Thanh   |  18 | KTDN      |  4      |
|  4 | Nhan    |  19 | CK        |  4.5    |
|  5 | Huong   |  20 | TCNH      |  5      |
+----+---------+-----+-----------+---------+
```

Ví dụ sau sẽ cập nhật KHOAHOC cho một sinh viên có ID là 4.

UPDATE SINHVIEN
SET KHOAHOC = 'TTCK'
WHERE ID = 4;

Bây giờ, bảng SINHVIEN sẽ có các bản ghi sau:

```
+----+---------+-----+-----------+---------+
| ID |  TEN    | TUOI|  KHOAHOC  | HOCPHI  |
+----+---------+-----+-----------+---------+
|  1 | Hoang   |  21 | CNTT      | 4       |
|  2 | Viet    |  19 | DTVT      | 3.0     |
|  3 | Thanh   |  18 | KTDN      |  4      |
|  4 | Nhan    |  19 | TTCK      |  4.5    |
|  5 | Huong   |  20 | TCNH      |  5      |
+----+---------+-----+-----------+---------+
```

Nếu bạn muốn sửa đổi tất cả giá trị các cột là KHOAHOC và HOCPHI trong bảng này, bạn không cần sử dụng mệnh đề WHERE và khi đó truy vấn UPDATE sẽ như sau:

UPDATE SINHVIEN
SET KHOAHOC = 'CNTT', HOCPHI = 4;

Bây giờ, bảng SINHVIEN sẽ có các bản ghi sau:

```
+----+---------+-----+-----------+------------------+
| ID |  TEN    | TUOI|  KHOAHOC  | HOCPHI  |
+----+---------+-----+-----------+---------+
|  1 | Hoang   |  21 | CNTT      | 4       |
|  2 | Viet    |  19 | CNTT      | 4       |
|  3 | Thanh   |  18 | CNTT      |  4      |
|  4 | Nhan    |  19 | CNTT      |  4      |
|  5 | Huong   |  20 | CNTT      |  4      |
+----+---------+-----+-----------+---------+
```

4.4. Các lệnh truy vấn INSERT

Lệnh INSERT INTO trong SQL được sử dụng để thêm các hàng dữ liệu mới vào một bảng trong Database.

Có hai dạng cú pháp cho lệnh INSERT INTO trong SQL như sau:

INSERT INTO TABLE_TEN (cot1, cot2, cot3,...cotN)]
VALUES (giatri1, giatri2, giatri3,...giatriN);

Ở đây, cot1, cot2,...cotN là tên các cột trong bảng mà bạn muốn chèn dữ liệu.

Bạn có thể không cần xác định tên các cột trong truy vấn SQL nếu bạn đang thêm các giá trị cho tất cả cot của bảng đó. Nhưng bạn nên đảm bảo thứ tự của các giá trị là giống như thứ tự các cột trong bảng.

Cú pháp của lệnh INSERT INTO này như sau:

INSERT INTO TABLE_TEN VALUES (giatri1,giatri2,giatri3,...giatriN);

Ví dụ

Các lệnh dưới đây sẽ tạo 4 bản ghi trong bảng SINHVIEN:

```
INSERT INTO SINHVIEN (ID,TEN,TUOI,KHOAHOC,HOCPHI)
VALUES (1, 'Hoang', 21, 'CNTT', 4.0 );

INSERT INTO SINHVIEN (ID,TEN,TUOI,KHOAHOC,HOCPHI)
VALUES (2, 'Viet', 19, 'DTVT', 3.0 );

INSERT INTO SINHVIEN (ID,TEN,TUOI,KHOAHOC,HOCPHI)
VALUES (3, 'Thanh', 18, 'KTDN', 4.0 );

INSERT INTO SINHVIEN (ID,TEN,TUOI,KHOAHOC,HOCPHI)
VALUES (4, 'Nhan', 19, 'CK', 4.5 );
```

Bạn có thể tạo thêm một bản ghi trong bảng SINHVIEN bởi sử dụng cú pháp thứ hai như sau:

```
INSERT INTO SINHVIEN
VALUES (5, 'Huong', 20, 'TCNH', 5.0 );
```

Tất cả lệnh trên sẽ tạo các bản ghi sau trong bảng SINHVIEN:

```
+----+---------+-----+-----------+---------+
| ID |  TEN    | TUOI|  KHOAHOC  | HOCPHI  |
+----+---------+-----+-----------+---------+
|  1 | Hoang   |  21 | CNTT      | 4.0     |
|  2 | Viet    |  19 | DTVT      | 3.0     |
|  3 | Thanh   |  18 | KTDN      |  4.0    |
|  4 | Nhan    |  19 | CK        |   4.5   |
|  5 | Huong   |  20 | TCNH      |   5.0   |
+----+---------+-----+-----------+---------+
```

TÓM TẮT CHƯƠNG 4

Trong chương này, một số nội dung chính được giới thiệu:

1. **DBMS SQL là gì?**: Chương giới thiệu về khái niệm DBMS SQL và tầm quan trọng của nó trong việc quản lý cơ sở dữ liệu.
2. **Cài đặt DBMS SQL**: Chương bao gồm hướng dẫn về cách cài đặt hệ quản trị cơ sở dữ liệu SQL trên máy tính hoặc máy chủ.
3. **Quản lý Cơ sở Dữ liệu**: Học viên sẽ học cách quản lý cơ sở dữ liệu bằng cách thực hiện các nhiệm vụ quản lý cơ bản.
4. **Sử dụng Công cụ Quản trị**: Chương này giới thiệu về việc sử dụng các công cụ quản trị DBMS SQL để thực hiện các thao tác quản lý cơ bản.

❖ CÁC BÀI TẬP HỆ THỐNG KIẾN THỨC

1. Tại sao DBMS SQL quan trọng trong việc quản lý cơ sở dữ liệu? Hãy nêu ví dụ về vai trò của nó trong lĩnh vực quản lý thông tin.
2. Cuốn sách chỉ ra rằng cài đặt DBMS SQL là một bước quan trọng. Vậy, quy trình cài đặt hệ quản trị cơ sở dữ liệu SQL thông thường bao gồm những bước gì và tại sao chúng quan trọng?
3. Khi quản lý cơ sở dữ liệu, bạn cần thực hiện các nhiệm vụ quản lý cơ bản như tạo bảng, thêm người dùng, và sao lưu dữ liệu. Hãy đưa ra ví dụ cụ thể về một tình huống thực tế mà bạn cần thực hiện một trong những nhiệm vụ này và giải thích quá trình thực hiện