

LỜI GIỚI THIỆU

Cơ sở dữ liệu là môn học cơ sở chuyên ngành quan trọng của chuyên ngành Công nghệ thông tin. Mục đích của giáo trình Cơ sở dữ liệu này là nhằm chuẩn hóa tài liệu học tập cho sinh viên các hệ cao đẳng nghề chuyên ngành CNTT, đồng thời cũng là tài liệu tham khảo đối với các chuyên ngành khác trong lĩnh vực Tin học.

Giáo trình giúp các học viên tiếp cận vấn đề cốt lõi nhất về mặt lý thuyết: các định nghĩa, khái niệm, hệ quả, định lý, giải thuật,... từ đó có thể áp dụng vào bài toán thực tế thiết kế chuẩn hóa cơ sở dữ liệu nói riêng cũng như thiết kế hệ thống thông tin nói chung.

Giáo trình không đi sâu vào việc chứng minh các định lý mà chú trọng đến việc giải thích ý nghĩa thực tế của các công thức lý thuyết để từ đó hướng dẫn học viên cách tiếp cận tư duy logic, nắm vững kỹ thuật tính toán cũng như các bước triển khai giải quyết các bài toán thực tế trên khía cạnh công nghệ.

Nội dung giáo trình được chia làm 5 chương:

Chương I: Giới thiệu những khái niệm cơ bản nhất về mô hình cơ sở dữ liệu. Tìm hiểu về mô hình thực thể kết hợp.

Chương II: Giới thiệu về mô hình dữ liệu quan hệ, các quy tắc chuyển đổi từ mô hình ER sang mô hình dữ liệu quan hệ. Ngoài ra chương 2 còn trình bày các quy tắc, phép toán của ngôn ngữ đại số quan hệ.

Chương III: Trình bày về ngôn ngữ truy vấn dữ liệu quan hệ (SQL), chủ yếu là câu lệnh truy vấn Select và các mệnh đề kết hợp với câu lệnh.

Chương IV: Khái lược về ràng buộc toàn vẹn.

Chương V: Đi sâu vào một số khái niệm như: phụ thuộc hàm, khóa, bao đóng, các dạng chuẩn,... Tìm hiểu giới thiệu một số thuật toán liên quan đến tìm khóa hay cách xác định các dạng chuẩn cho lược đồ quan hệ. Đây là chương đóng vai trò quan trọng trong việc tư duy logic, giúp quá trình thiết kế chuẩn hóa cơ sở dữ liệu được rõ ràng và chính xác hơn.

Ngoài ra, giáo trình còn trình bày thêm vấn đề thiết kế cơ sở dữ liệu bằng các kỹ thuật phân rã theo chuẩn 3 và chuẩn BC.

Sau mỗi chương đều có bài tập hướng dẫn và bài tập tự làm.

Hy vọng cuốn sách còn là tài liệu tham khảo hữu ích cho các đồng nghiệp trong giảng dạy, nghiên cứu trên các lĩnh vực có liên quan.

Rất mong nhận được nhiều ý kiến đóng góp, trao đổi của bạn đọc.

Xin chân thành cảm ơn!

Hà Nam, ngày ... tháng ... năm 2020

“Ký và ghi rõ họ tên”

Biên soạn: Bùi Thị Thu Hằng

MỤC LỤC

LỜI GIỚI THIỆU	3
MỤC LỤC	4
GIÁO TRÌNH MÔN HỌC	8
Mục tiêu của môn học:.....	Error! Bookmark not defined.
Giới thiệu:	10
Mục tiêu:	10
Nội dung chính:	10
1. Một số khái niệm cơ bản.	10
1.1. Định nghĩa cơ sở dữ liệu	10
1.2. Ưu điểm của cơ sở dữ liệu.....	10
1.3. Các đặc trưng của phương pháp cơ sở dữ liệu	11
1.4. Các đối tượng sử dụng CSDL	11
1.5. Hệ Quản Trị Cơ Sở Dữ Liệu (Data Base Management System).....	11
1.6. Các Ứng Dụng Của Cơ Sở Dữ Liệu	12
2. Các mô hình dữ liệu	12
3. Mô hình thực thể kết hợp	13
3.1. Thực Thể (entity).....	13
3.2. Thuộc tính (attribute).....	13
3.3. Loại thực thể (entity type)	13
3.4. Khoá (key)	14
Ví dụ 1.1:	14
Sinhvien	15
3.5. Mối kết hợp (relationship).....	15
Bản số của mối kết hợp:.....	15
BÀI TẬP THỰC HÀNH CỦA HỌC VIÊN.....	16
GỢI Ý BÀI TẬP	17
BÀI TẬP THAM KHẢO.....	17
Mục tiêu:	20
Nội dung chính:	20
1. Một số khái niệm cơ bản.	20
1.1. Thuộc tính.....	20
1.2. Lược Đồ Quan Hệ (relation schema).....	21
1.3. Quan Hệ (relation)	21
1.4. Bộ (Tuple)	21
2. Chuyển mô hình thực thể kết hợp sang mô hình dữ liệu quan hệ.....	22

Ví dụ 2.1:	23
3. Ngôn ngữ đại số quan hệ.....	23
3.1. Phép Hợp 2 quan hệ(Union).....	23
3.2. Phép Giao 2 quan hệ (Intersection)	24
3.3. Phép Trừ 2 quan hệ (Minus).....	24
3.4. Tích Decac của 2 quan hệ (Cartesian Product)	24
3.5. Phép chia 2 quan hệ:	25
3.6. Phép Chiếu (projection).....	26
3.7. Phép Chọn (Selection).....	26
3.8. Phép 0 - kết.....	27
Ai e Bj.....	27
3.9. Phép Kết Tự Nhiên (natural join).....	27
BÀI TẬP THỰC HÀNH CỦA HỌC VIÊN.....	28
GỢI Ý BÀI TẬP	29
Giới thiệu:	31
Mục tiêu:	31
Nội dung chính:	31
1. Mở đầu	31
Mục đích của chuân SQL là.....	31
Một ngôn ngữ quan hệ chuân như SQL sẽ mang lại các lợi ích sau đây:.....	32
Kiểu dữ liệu	32
2. Tìm thông tin từ các cột của bảng - Mệnh đề Select.....	34
3. Chọn các dòng của bảng - Mệnh đề Where	35
4. Sắp xếp các dòng của bảng - Mệnh đề Order by.....	37
5. Câu lệnh truy vấn lồng nhau	38
6. Gom nhóm dữ liệu - mệnh đề Group by	42
BÀI TẬP THỰC HÀNH CỦA HỌC VIÊN.....	43
BÀI TẬP THAM KHẢO.....	46
Giới thiệu:	48
Mục tiêu:	48
Nội dung chính:	48
1. Ràng buộc toàn vẹn.....	48
1.1. Khái niệm ràng buộc toàn vẹn.....	48
1.2. Các yếu tố của ràng buộc toàn vẹn:.....	48
1.2.1. Điều kiện:	48
1.2.2. Bối cảnh:	49

1.2.3.	Bảng tầm ảnh hưởng	49
Quy ước:		49
1.2.4.	Hành động cần phải có khi phát hiện có RBTV bị vi phạm:	50
2.	Phân loại các ràng buộc toàn vẹn:.....	50
Q1:	Khach (MAKH, TENKH, DIACHIKH, DIENTHOAI).....	50
Q2:	Hang (MAHANG,TENHANG,QUYCACH, DVTINH)	50
Q3:	Dathang (SODH,MAHANG, SLDAT, NGAYDH, MAKH).....	51
Q4:	Hoadon (SOHD, NGAYLAP, SODH, TRIGIAHD, NGAYXUAT).....	51
Q5:	Chitiethd (SOHD, MAHANG, GIABAN, SLBAN)	51
Q6:	Phieuthu(SOPT, NGAYTHU, MAKH, SOTIEN)	52
2.1.	Ràng buộc toàn vẹn có bối cảnh là một quan hệ	52
2.1.1.	Ràng buộc toàn vẹn liên bộ:.....	52
2.1.2.	Ràng buộc toàn vẹn về miền giá trị:.....	52
2.1.3.	Ràng buộc toàn vẹn liên thuộc tính:.....	55
2.2.	Ràng buộc toàn vẹn có bối cảnh là nhiều quan hệ:	55
2.2.1.	Ràng buộc toàn vẹn về khóa ngoại:	55
2.2.2.	Ràng buộc toàn vẹn liên thuộc tính liên quan hệ:	55
2.2.3.	Ràng buộc toàn vẹn liên bộ liên quan hệ:	56
	BÀI TẬP THỰC HÀNH CỦA HỌC VIÊN:.....	57
	BÀI TẬP THAM KHẢO:	58
	Giới thiệu:	58
	Mục tiêu:	58
	Nội dung chính:	59
1.	Các vấn đề gặp phải khi tổ chức dữ liệu:	59
2.	Phụ thuộc hàm.....	59
2.1.	Định nghĩa phụ thuộc hàm	59
2.2.	Cách xác định phụ thuộc hàm cho lược đồ quan hệ.....	60
2.3.	Một số tính chất của phụ thuộc hàm - hệ luật dẫn Armstrong	60
3.	Bao đóng của tập phụ thuộc hàm và bao đóng của tập thuộc tính.....	61
3.1.	Bao đóng của tập phụ thuộc hàm F	61
3.2.	Bao đóng của tập thuộc tính X	62
3.3.	Bài toán thành viên.....	62
3.4.	Thuật toán tìm bao đóng của một tập thuộc tính	63
4.	Khóa của lược đồ quan hệ - một số thuật toán tìm khóa.....	64
4.1.	Định nghĩa khóa của quan hệ	64
4.2.	Thuật toán tìm một khóa của một lược đồ quan hệ	64

4.3.	Thuật toán tìm tất cả các khóa của một lược đồ quan hệ	65
5.	Phủ tối thiểu	66
5.1.	Tập phụ thuộc hàm tương đương	66
5.2.	Phủ tối thiểu.....	67
5.2.1.	Phụ thuộc hàm có vế trái dư thừa:.....	67
5.2.2.	Tập phụ thuộc hàm có vế phải một thuộc tính:.....	67
5.2.3.	Tập phụ thuộc hàm không dư thừa:	68
5.3.	Thuật toán tìm phủ tối thiểu	68
6.	Dạng chuẩn của lược đồ quan hệ	68
6.1.	Một số khái niệm liên quan đến các dạng chuẩn.....	68
6.2.	Dạng chuẩn 1 (First Normal Form).....	69
6.3.	Dạng chuẩn 2 (Second Normal Form).....	69
6.4.	Dạng chuẩn 3 (Third Normal Form)	70
6.5.	Dạng chuẩn BCNF (Boyce Codd Normal Form).....	72
BÀI TẬP THỰC HÀNH CỦA HỌC VIÊN:.....		75
BÀI TẬP THAM KHẢO:		76
TÀI LIỆU THAM KHẢO		78

GIÁO TRÌNH MÔN HỌC

Tên môn học: Cơ sở dữ liệu

Mã số của môn học: MH 10;

Thời gian của môn học: 60 giờ; (lý thuyết: 20 giờ; thực hành, thí nghiệm, thảo luận, bài tập: 37 giờ; kiểm tra: 3 giờ)

I. Vị trí, tính chất của môn học

- Vị trí : môn học Cơ sở dữ liệu được bố trí sau khi học xong các môn Tin học, lập trình căn bản, toán ứng dụng.

- Tính chất : là môn học kỹ thuật cơ sở, thuộc các môn học, mô đun đào tạo cơ sở.

II. Mục tiêu môn học

- Về kiến thức:

- + Hiểu được các khái niệm, các thuật ngữ trong cơ sở dữ liệu;
- + Sử dụng ngôn ngữ truy vấn dữ liệu quan hệ thành thạo;
- + Biết được các dạng chuẩn và chuẩn hóa được bài toán cơ sở dữ liệu trước khi cài đặt;

- Về kỹ năng:

- + Thực hiện được các dạng chuẩn và chuẩn hóa được bài toán cơ sở dữ liệu trước khi cài đặt;
- + Vận dụng được các mô hình dữ liệu của cơ sở dữ liệu quan hệ vào việc thiết kế cơ sở dữ liệu áp dụng cho bài toán cụ thể trong thực tế;

- Về năng lực tự chủ và trách nhiệm:

- + Có ý thức làm việc độc lập hoặc làm việc theo nhóm, giải quyết công việc, vấn đề phức tạp khi đưa ra một bài toán truy vấn dữ liệu;
- + Đánh giá chất lượng sản phẩm và kết quả thực hiện của các thành viên trong nhóm sau khi hoàn thành bài tập.
- + Có trách nhiệm trong việc ứng dụng kiến thức nền tảng của môn học nhằm phục vụ cho các môn học cơ sở và chuyên ngành.

III. Nội dung môn học

1 Nội dung tổng quát và phân phối thời gian :

Số TT	Tên chương, mục	Thời gian			
		Tổng số	Lý thuyết	Thực hành, thí nghiệm, thảo luận, bài tập	Kiểm tra* (LT hoặc TH)
1	Chương 1: Tổng quan về cơ sở dữ liệu 1. Một số khái niệm cơ bản	5	2	3	

Số TT	Tên chương, mục	Thời gian			
		Tổng số	Lý thuyết	Thực hành, thí nghiệm, thảo luận, bài tập	Kiểm tra* (LT hoặc TH)
	2. Các mô hình dữ liệu 3. Mô hình thực thể kết hợp				
2	Chương 2: Mô hình dữ liệu quan hệ Các khái niệm cơ bản Chuyển mô hình thực thể kết hợp sang mô hình dữ liệu quan hệ Ngôn ngữ đại số quan hệ	15	6	8	1
3	Chương 3: Ngôn ngữ truy vấn dữ liệu 1. Mở đầu 2. Tìm thông tin từ các cột của bản – Mệnh đề Select 3. Chọn các dòng của bản – Mệnh đề Where 4. Sắp xếp các dòng của bảng – Mệnh đề Order by 5. Câu lệnh truy vấn lồng nhau 6. Gom nhóm dữ liệu – mệnh đề Group by	15	3	11	1
4	Chương 4: Ràng buộc toàn vẹn 1. Ràng buộc toàn vẹn 2. Phân loại ràng buộc toàn vẹn	5	1	4	
5	Chương 5: Lý thuyết thiết kế cơ sở dữ liệu 1. Các vấn đề gặp phải khi tổ chức dữ liệu 2. Các phụ thuộc hàm 3. Bao đóng của tập phụ thuộc hàm và bao đóng của tập thuộc tính 4. Khóa của lược đồ quan hệ - Một	20	8	11	1

Số TT	Tên chương, mục	Thời gian			
		Tổng số	Lý thuyết	Thực hành, thí nghiệm, thảo luận, bài tập	Kiểm tra* (LT hoặc TH)
	số thuật toán tìm khóa 5. Phủ tối thiểu 6. Dạng chuẩn của lược đồ quan hệ				
	Cộng:	60	20	37	3

CHƯƠNG 1. TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU

Mã chương: MHCNTT 10.01

Giới thiệu:

Bài học này giới thiệu khái quát về các mô hình dữ liệu cơ bản, các thuật ngữ, khái niệm liên quan trong cơ sở dữ liệu. Thông qua bài học này người đọc sẽ hình dung được những vấn đề cần tiếp cận, khai thác trong môn học cơ sở dữ liệu.

Mục tiêu:

- Trình bày sơ lược các khái niệm về cơ sở dữ liệu, các mô hình dữ liệu.
- Trình bày chi tiết mô hình thực thể kết hợp (ERD), có thể phân tích dữ liệu và thiết kế được mô hình thực thể kết hợp.
- Thực hiện thao tác an toàn với máy tính.

Nội dung chính:

1. Một số khái niệm cơ bản.

Mục tiêu: Trình bày sơ lược các khái niệm về cơ sở dữ liệu.

1.1. Định nghĩa cơ sở dữ liệu

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

1.2. Ưu điểm của cơ sở dữ liệu

- Giảm sự trùng lặp thông tin xuống mức thấp nhất và do đó bảo đảm được tính nhất quán và toàn vẹn dữ liệu.
- Đảm bảo dữ liệu có thể truy xuất theo nhiều cách khác nhau.
- Khả năng chia sẻ thông tin cho nhiều người sử dụng.

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

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

1.4. Các đối tượng sử dụng CSDL

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

Người quản trị CSDL: Là người hàng ngày chịu trách nhiệm quản lý và bảo trì CSDL như:

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

Người phát triển và lập trình ứng dụng: là những người chuyên nghiệp về lĩnh vực tin học có trách nhiệm thiết kế, tạo dựng và bảo trì thông tin cuối cùng cho người dùng.

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

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

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

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

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

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

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

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

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

1.6. Các Ứng Dụng Của Cơ Sở Dữ Liệu

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

2. Các mô hình dữ liệu

Mục tiêu: Trình bày giới thiệu sơ lược các mô hình dữ liệu.

Mô hình dữ liệu là sự trừu tượng hoá môi trường thực. Mỗi loại mô hình dữ liệu đặc trưng cho một cách tiếp cận dữ liệu khác nhau của những nhà phân tích thiết kế CSDL. Mỗi loại mô hình dữ liệu đều có những ưu điểm và những mặt hạn chế của nó, nhưng vẫn có những mô hình dữ liệu nổi trội và được nhiều người quan tâm nghiên cứu.

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

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

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

- Sang thập kỷ tám mươi, mô hình CSDL thứ ba ra đời, đó là mô hình cơ sở dữ liệu hướng

đối tượng, mô hình cơ sở dữ liệu phân tán, mô hình cơ sở dữ liệu suy diễn,...

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

3. Mô hình thực thể kết hợp

Mục tiêu: Trình bày chi tiết mô hình thực thể kết hợp (ERD), có thể phân tích dữ liệu và thiết kế được mô hình thực thể kết hợp.

Hiện nay mô hình dữ liệu quan hệ thường được dùng trong các hệ quản trị CSDL, đây là mô hình dữ liệu ở mức vật lý. Để thành lập được mô hình này, thường là phải dùng mô hình dữ liệu ở mức quan niệm để đặc tả, một trong những mô hình ở dạng đó là mô hình thực thể kết hợp (sau đó mới dùng một số quy tắc để chuyển hệ thống từ mô hình này về mô hình dữ liệu quan hệ - các quy tắc này sẽ được nói đến trong mục 2.2).

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

3.1. Thực Thể (entity)

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

3.2. Thuộc tính (attribute)

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

Chẳng hạn các thuộc tính của sinh viên Nguyễn Văn Thành là: mã số, sinh viên, giới tính, ngày sinh, hộ khẩu thường trú, lớp đang theo học,.

3.3. Loại thực thể (entity type)

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

Ví dụ các sinh viên có mã sinh viên là “02CĐTH019”, “02CĐTH519”, “02TCTH465”, nhóm lại thành một loại thực thể, được đặt tên là Sinhvien chẳng hạn.

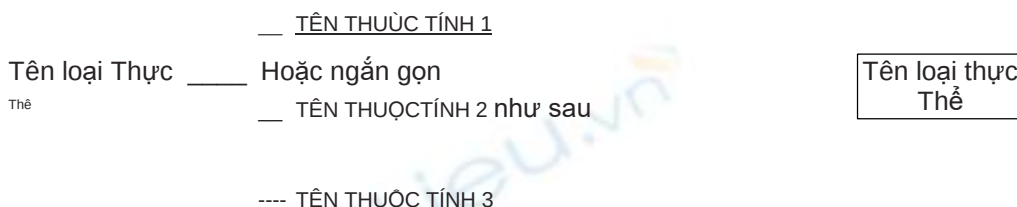
Tương tự trong ứng dụng quản lý điểm của sinh viên (sẽ được trình bày ngay sau đây) ta có các loại thực thể như Monhoc, Lop, Khoa.

3.4. Khoá (key)

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

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

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



Ví dụ 1.1:

Bài toán quản lý điểm của sinh viên được phát biểu sơ bộ như sau: Mỗi sinh viên cần quản lý các thông tin như: họ và tên (HOTENSV), ngày tháng năm sinh (NGAYSINH), giới tính (NU), nơi sinh (NƠISINH), hộ khẩu thường trú (TINH). Mỗi sinh viên được cấp một mã số sinh viên duy nhất (MASV) để phân biệt với mọi sinh viên khác của trường, mỗi sinh viên chỉ thuộc về một lớp nào đó. Mỗi lớp học có một mã số lớp (MALOP) duy nhất để phân biệt với tất cả các lớp học khác trong trường: có một tên gọi (TENLOP) của lớp, mỗi lớp chỉ thuộc về một khoa. Mỗi khoa có một tên gọi (TENKHOA) và một mã số duy nhất (MAKHOA) phân biệt với các khoa khác. Mỗi môn học có một tên gọi (TENMH) cụ thể, được học trong một số đơn vị học trình (DONVIHT) và ứng với môn học là một mã số duy nhất (MAMH) để phân biệt với các môn học khác. Mỗi giảng viên cần quản lý các thông tin: họ và tên (HOTENGV), cấp học vị (HOCVI), thuộc một chuyên ngành (CHUYENNGANH) và được gán cho một mã số duy nhất gọi là mã giảng viên (MAGV) để phân biệt với các giảng viên khác. Mỗi giảng viên có thể dạy nhiều môn ở nhiều khoa, nhưng chỉ thuộc về sự quản lý hành chính của một khoa. Mỗi sinh viên với một môn học được phép thi tối đa 3 lần, mỗi lần thi (LANTHI), điểm thi (DIEMTHI). Mỗi môn học ở mỗi lớp học chỉ phân công cho một giảng viên dạy (tất nhiên là một giảng viên thì có thể dạy nhiều môn ở một lớp).

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

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

3.5. Môi kết hợp (relationship)

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

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

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

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



sau:

Bản số của môi kết hợp:

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

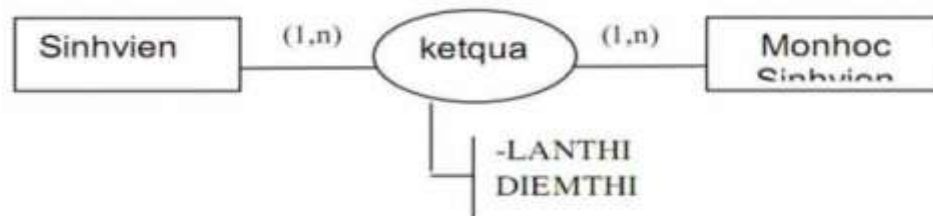
Mỗi bản số là một cặp số (min,max), chỉ số lượng tối thiểu và số lượng tối đa của thực thể khi tham gia vào môi kết hợp đó.

Ví dụ:



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

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



Khoá của môi kết hợp: là hợp của các khoá của các loại thực thể liên quan. Chẳng hạn như thuộc tính MAGV là khoá của loại thực thể Giảngviên, MALOP là thuộc tính khoá của

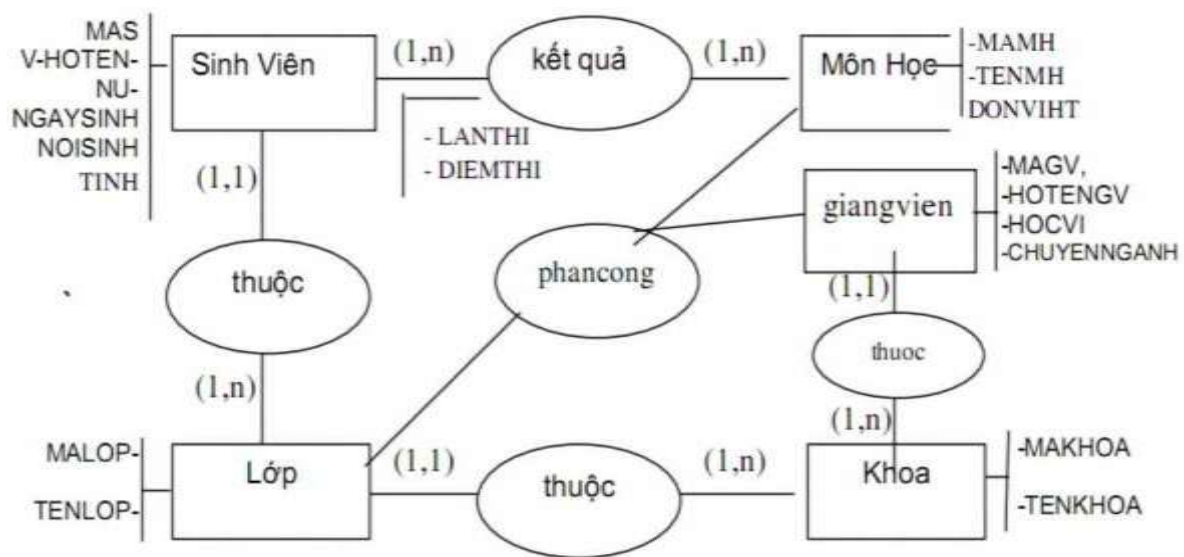
loại thực thể Lop, MAMH là thuộc tính khoá của loại thực thể Monhoc, do đó mỗi kết hợp phancong (giữa các loại thực thể Giangvien,Lop,Monhoc) có khoá là {MAGV,MAMH,MALOP} - phancong là mối kết hợp 3 ngôi.

Việc thành lập mô hình thực thể kết hợp cho một ứng dụng tin học có thể tiến hành theo các bước sau:

- b1. Xác định danh sách các loại thực thể
- b2. Xác định các mối kết hợp giữa các loại thực thể để phác thảo mô hình.
- b3. Lập bản số của các mối kết hợp.

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

Ví dụ 1.2:



BÀI TẬP THỰC HÀNH CỦA HỌC VIÊN

Dựa vào các phân tích sơ bộ dưới đây, hãy lập mô hình thực thể kết hợp (gồm loại thực thể, mối kết hợp, bản số, thuộc tính của loại thực thể, khoá của loại thực thể) cho bài toán quản lý sau:

Bài 1.1. QUẢN LÝ ĐẶT BÁO CỦA KHÁCH HÀNG

Tại một nơi phát hành báo cần phải quản lý thông tin đặt báo của khách hàng. Có nhiều thể loại báo (nhật báo, nguyệt sang, tạp chí, tham luận,...), mỗi thể loại cần lưu lại mã thể loại, tên thể loại. Mỗi tờ báo cần lưu lại mã báo là gì, tên báo (kiến thức ngày nay, tuổi trẻ ngày, tuổi trẻ tuần, Sài Gòn giải phóng,...) hình thức (quý, ngày, tuần), đơn giá kỳ. Mỗi một thể loại báo có nhiều tờ báo. Khách hàng đến đặt báo cần lưu lại thông tin: mã khách hàng, tên khách hàng, ngày đặt. Một khách hàng có thể đặt nhiều tờ báo khác nhau. Cần lưu lại: tờ báo mà khách muốn đặt, số kỳ, thành tiền.

GỢI Ý BÀI TẬP

Bài 1.1:

b1.Xác định danh sách các loại thực thể

-Thực thể **BẢO**: có các thuộc tính: mã bảo (mabao) là thuộc tính định danh; tên bảo (tenbao), hình thức (hinhthuc), đơn giá kỳ (dongiaký).

- Thực thể **THẺ LOẠI**: có các thuộc tính: mã thẻ loại (matheloai) là thuộc tính định danh, tên thẻ loại (tentheloai).

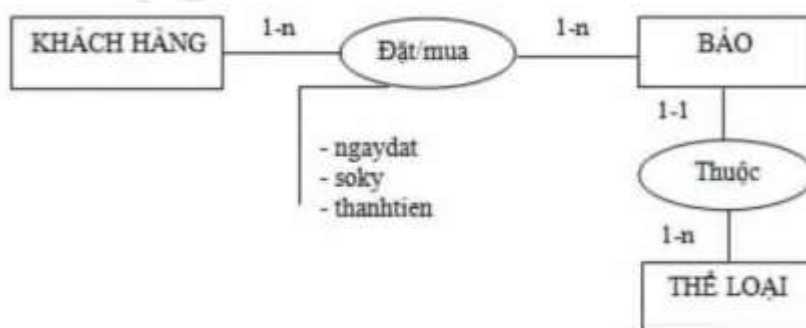
- Thực thể **KHÁCH HÀNG**: có các thuộc tính: mã khách hàng (makhachhang), tên khách hàng (tenkhachhang).

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

- **BẢO** thuộc **THẺ LOẠI**.

- **KHÁCH HÀNG** đặt/mua **BẢO**: có các thuộc tính riêng phát sinh trong quá trình đặt.mua là: ngày đặt (ngaydat), số kỳ (soky), thành tiền (thanhtien).

Phác thảo mô hình:



b3.Lập bản số của các mối kết hợp: gắn các quan hệ sau lên mô hình phác thảo, ta có được mô hình thực thể mối quan hệ sơ bộ cuối cùng.

- Quan hệ Đặt/mua giữa KHÁCH HÀNG và BẢO là quan hệ (1-n)-(1-n)

- Quan hệ Thuộc giữa BẢO và THẺ LOẠI là quan hệ (1-1)-(1-n).

(Các bài tập còn lại được phân tích và giải quyết tương tự bài 1).

BÀI TẬP THAM KHẢO

Bài 1. 2. QUẢN LÝ VIỆC MƯỢN/TRẢ SÁCH Ở MỘT THƯ VIỆN

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

độc giả (MADG), cùng với các thông tin khác như: họ tên (HOTEN), ngày sinh (NGAYSINH), địa chỉ (ĐIACHI), nghề nghiệp (NGHENGHIEP).

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

Bài 1.3. QUẢN LÝ LỊCH DẠY CỦA GIÁO VIÊN

Để quản lý lịch dạy của các giáo viên và lịch học của các lớp, một trường tổ chức như sau: Mỗi giáo viên có một mã số giáo viên (MAGV) duy nhất, mỗi MAGV xác định các thông tin như: họ và tên giáo viên (HOTEN), số điện thoại (DTGV). Mỗi giáo viên có thể dạy nhiều môn cho nhiều khoa nhưng chỉ thuộc sự quản lý hành chính của một khoa nào đó. Mỗi môn học có một mã số môn học (MAMH) duy nhất, mỗi môn học xác định tên môn học (TENMH). Ứng với mỗi lớp thì mỗi môn học chỉ được phân cho một giáo viên. Mỗi phòng học có một số phòng học (PHONG) duy nhất, mỗi phòng có một chức năng (CHUCNANG); chẳng hạn như phòng lý thuyết, phòng thực hành máy tính, phòng nghe nhìn, xưởng thực tập cơ khí,... Mỗi khoa có một mã khoa (MAKHOA) duy nhất, mỗi khoa xác định các thông tin như: tên khoa (TENKHOA), điện thoại khoa (DTKHOA). Mỗi lớp có một mã lớp (MALOP) duy nhất, mỗi lớp có một tên lớp (TENLOP), sĩ số lớp (SISO). Mỗi lớp có thể học nhiều môn của nhiều khoa nhưng chỉ thuộc sự quản lý hành chính của một khoa nào đó. Hàng tuần, mỗi giáo viên phải lập lịch báo giảng cho biết giáo viên đó sẽ dạy những lớp nào, ngày nào (NGAYDAY), môn gì?, tại phòng nào, từ tiết nào (TUTIET) đến tiết nào (ĐENTIET), tựa đề bài dạy (BAIDAY), những ghi chú (GHICHU) về các tiết dạy này, đây là giờ dạy lý thuyết (LYTHUYET) hay thực hành - giả sử nếu LYTHUYET=1 thì đó là giờ dạy thực hành và nếu LYTHUYET=2 thì đó là giờ lý thuyết, một ngày có 16 tiết, sáng từ tiết 1 đến tiết 6, chiều từ tiết 7 đến tiết 12, tối từ tiết 13 đến 16. Một số yêu cầu của hệ thống này như:: Lập lịch dạy trong tuần của các giáo viên. Tổng số dạy của các giáo viên theo từng môn cho từng lớp, ..

Bài 1.4. QUẢN LÝ HỌC VIÊN Ở MỘT TRUNG TÂM TIN HỌC

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

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

Trung tâm có nhiều môn học, mỗi môn học có mã môn học (MAMH) duy nhất, mỗi môn

học xác định tên môn học(TENMH), số tiết lý thuyết (SOTIETLT), số tiết thực hành (SOTIETHH). Mỗi học viên ứng với mỗi môn học có một điểm thi(DIEMTHI) duy nhất. Mỗi lần đóng học phí, học viên sẽ được trung tâm giao cho một phiếu biên lai thu tiền, mỗi biên lai có một số biên lai duy nhất để quản lý. Một số yêu cầu của hệ thống này như: Lập danh sách những học viên khai giảng khoá ngày nào đó. Lập danh sách các học viên của một lớp ? Cho biết số lượng học viên của mỗi lớp khai giảng khoá ngày nào đó?

Bài 1.5. QUẢN LÝ COI THI TUYỂN SINH

Một hội đồng coi thi tuyển sinh có nhiều điểm thi, mỗi điểm thi được đặt tại một trường nào đó. Các điểm thi (DIEMTHISO) được đánh số là điểm thi số 1, điểm thi số 2, điểm thi số 3,...Mỗi điểm thi xác định địa chỉ (DIACHIDIEMTHI). Ví dụ: điểm thi số 1, đặt tại trường PTTH Nguyễn Thị Minh Khai, điểm thi số 2 đặt tại trường PTTH Bùi Thị Xuân,... Mỗi thí sinh có một số báo danh (SOBD) duy nhất, mỗi số báo danh xác định các thông tin: họ và tên (HOTEN), ngày sinh (NGAYSINH), phái (PHAI), hộ khẩu thường trú (TINH), đối tượng dự thi (DOITUONG), ngành đăng ký thi, khu vực của thí sinh (KHUVUC), số hiệu phòng thi. Ví dụ: thí sinh Vũ Mạnh Cường, có số báo danh là 02978, sinh ngày 12/12/1984, phái nam, hộ khẩu thường trú tại Chợ Gạo - Tiền Giang, thuộc khu vực 1, đối tượng là 5B, đăng ký dự thi vào ngành có mã ngành là 01, thi tại phòng thi 0178, điểm thi số 1. Mỗi ngành có một mã ngành (MANGANH) duy nhất, mỗi mã ngành xác định tên ngành (TENNGANH). Mỗi điểm thi có nhiều phòng thi - mỗi phòng thi (PHONGTHI) được đánh số khác nhau ở tất cả các điểm thi. Trong một phòng thi, danh sách các thí sinh được sắp xếp theo thứ tự alphabet (do đó trong một phòng thi có thể có thí sinh của nhiều ngành khác nhau). Mỗi phòng thi có thêm cột ghi chú (GHICHU) - ghi thêm các thông tin cần thiết như phòng thi đó nằm tại dãy nhà nào. Ví dụ phòng thi 0060 nằm ở dãy nhà H lầu 2 - điểm thi số 1 - trường PTTH Bùi Thị Xuân. Mỗi môn thi có một mã môn thi duy nhất (MAMT), mỗi mã môn thi biết các thông tin như : tên môn thi (TENMT), ngày thi (NGAYTHI), buổi thi (BUOITHI), thời gian làm bài thi được tính bằng phút (PHUT). Thời gian làm bài thi của các môn tối thiểu là 90 phút và tối đa là 180 phút (tùy theo kỳ tuyển sinh công nhân, trung cấp, cao đẳng hay đại học). Mỗi ngành có một mã ngành, chẳng hạn ngành Công Nghệ Thông Tin có mã ngành là 1, ngành Công Nghệ Hoá Thực Phẩm có mã ngành là 10,...

Mỗi đơn vị có cán bộ tham gia vào kỳ thi có một mã đơn vị duy nhất (MADONVI), mã đơn vị xác định tên đơn vị (TENDONVI). Nếu là cán bộ, công nhân viên của trường thì đơn vị là khoa/phòng quản lý cán bộ đó, nếu là giáo viên từ các trường khác thì ghi rõ tên đơn vị đó. Chẳng hạn cán bộ Nguyễn Thanh Liêm đơn vị Khoa Công Nghệ Thông Tin, cán bộ coi thi Nguyễn Thị Tuyết Mai, đơn vị trường PTTH Ngôi Sao - Quận 1,. Mỗi cán bộ coi thi chỉ làm việc tại một điểm thi nào đó. Mỗi cán bộ có một mã số duy nhất (MACANBO), mỗi MACANBO xác định các thông tin khác như : họ và tên (HOTENCB), đơn vị công tác, chức vụ (CHUCVU) được phân công tại điểm thi, chẳng hạn chức vụ là điểm trưởng, điểm phó, giám sát, thư ký, cán bộ coi thi, phục vụ,... Ví dụ cán bộ Nguyen Van Thanh đơn vị Khoa Công Nghệ Thông Tin, làm nhiệm vụ thi tại điểm thi số 1, chức vụ là giám sát phòng thi.

Chương 2. MÔ HÌNH DỮ LIỆU QUAN HỆ

Mã chương: MHCNTT 10.02

Giới thiệu:

Mô hình dữ liệu quan hệ là dạng mô hình cơ sở dữ liệu cơ bản, được khai thác về mặt lý thuyết khá triệt để, triển khai về mặt ứng dụng hiệu quả nhất so với các mô hình dữ liệu khác. Trong chương này, chúng ta sẽ tìm hiểu sâu hơn một số khái niệm, thuật ngữ, các phép toán trên mô hình dữ liệu quan hệ. Đồng thời tìm hiểu một trong những phương pháp xây dựng (tạo mới) một mô hình dữ liệu quan hệ dựa trên các nguyên tắc chuyển đổi từ mô hình thực thể mối quan hệ đã được học ở chương một.

Mục tiêu:

- Trình bày cụ thể về mô hình dữ liệu quan hệ, các khái niệm, cơ sở lý thuyết của mô hình dữ liệu quan hệ (đại số quan hệ);
- Áp dụng được các quy tắc chuyển đổi từ mô hình ERD sang mô hình dữ liệu quan hệ (ở dạng lược đồ).

Nội dung chính:

1. Một số khái niệm cơ bản.

Mục tiêu: trình bày cụ thể các khái niệm của mô hình dữ liệu quan hệ.

1.1. Thuộc tính

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

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

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

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

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

và hơn), nhưng cũng không nên đặt tên thuộc tính quá ngắn (vì nó sẽ không cho thấy ngữ nghĩa của thuộc tính), đặc biệt không đặt trùng tên hai thuộc tính mang ngữ nghĩa khác nhau thuộc hai đối tượng khác nhau.

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

1.2 Lược Đồ Quan Hệ (relation schema)

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

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

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

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

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

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 này, người ta xác định được tập khoá, siêu khoá của lược đồ quan hệ (sẽ được trình bày trong những mục kế tiếp).

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

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

1.3. Quan Hệ (relation)

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

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

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

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

1.4. Bộ (Tuple)

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

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

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

Mục tiêu: Áp dụng được các quy tắc chuyển đổi mô hình thực thể kết hợp ERD sang mô hình dữ liệu quan hệ (dạng lược đồ).

Quy tắc 1:

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

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

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

Quy tắc 2:

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

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

Phancong(MAGV,MAMH,MALOP)

Quy tắc 3:

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

Chẳng hạn mỗi kết hợp *thuộc* giữa hai loại thực thể Sinhvien và Lop nên lược