

"Don't study, don't know - Studying you will know!"

NGUYEN TRUNG HOA

MỤC LỤC

Chương 1	3
MÔ HÌNH QUAN HỆ	3
I NGUYÊN NHÂN RA ĐỜI CỦA MÔ HÌNH QUAN HỆ	3
II CƠ SỞ DỮ LIỆU VÀ HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU.....	3
1 CSDL là gì?.....	4
2 Hệ quản trị CSDL	4
3 Người dùng (User).....	4
4 CSDLQH và Hệ tập tin theo lối cũ	5
III MÔ HÌNH QUAN HỆ	5
1 Mô hình quan hệ là gì ?.....	5
2 Các khái niệm cơ bản của mô hình quan hệ	6
3 Các phép toán tập hợp	10
4 Các phép toán quan hệ.....	11
5 Các tính chất của đại số quan hệ	12
IV MÔ HÌNH THỰC THỂ KẾT HỢP.....	15
1 Giới thiệu mô hình thực thể kết hợp	15
2 Chuyển từ mô hình thực thể kết hợp sang lược đồ CSDL.....	17
V BÀI TẬP	19
1 Phép toán tập hợp và phép toán quan hệ.....	19
2 Mô hình thực thể kết hợp	20
Chương 2	22
NGÔN NGỮ TRUY VẤN SQL.....	22
I CÁCH TẠO QUAN HỆ BẰNG ACCESS	22
II CÂU LỆNH TRUY VẤN.....	22
1 BIỂU THỨC (EXPRESSION)	22
2 CÂU LỆNH SQL.....	25
III BÀI TẬP	28
Chương 3	32
RÀNG BUỘC TOÀN VỆN QUAN HỆ.....	32
I RÀNG BUỘC TOÀN VỆN - CÁC YẾU TỐ CỦA RÀNG BUỘC TOÀN VỆN	32
1 Ràng Buộc Toàn Vẹn.....	32
2 Các Yếu Tố Của Ràng Buộc Toàn Vẹn	32
II PHÂN LOẠI RÀNG BUỘC TOÀN VỆN	33
1 Ràng buộc toàn vẹn liên bộ	34
2 Ràng buộc toàn vẹn về phụ thuộc tồn tại:	34
3 Ràng buộc toàn vẹn về miền giá trị.....	34
4 Ràng buộc toàn vẹn liên thuộc tính	35
5 Ràng buộc toàn vẹn liên thuộc tính liên quan hệ	35
6 Ràng buộc toàn vẹn về thuộc tính tổng hợp	35
III BÀI TẬP	35
Chương 4	37
PHỤ THUỘC HÀM	37
I KHÁI NIỆM PHỤ THUỘC HÀM	37

1	Định nghĩa phụ thuộc hàm	38
2	Phụ thuộc hàm hiển nhiên.....	38
3	Thuật toán Satifies	38
4	Các phụ thuộc hàm có thể có.....	39
II	HỆ LUẬT DẪN ARMSTRONG.....	42
1	Phụ thuộc hàm được suy diễn logic từ F	42
2	Hệ luật dẫn Armstrong	43
3	Hệ luật dẫn Armstrong là đầy đủ.....	47
III	THUẬT TOÁN TÌM F^+	48
1	Thuật toán cơ bản.....	48
2	Thuật toán cải tiến	48
IV	BÀI TẬP	49
Chương 5		50
PHỦ CỦA TẬP PHỤ THUỘC HÀM.....		50
I	ĐỊNH NGHĨA	50
II	PHỦ TỐI THIỂU CỦA MỘT TẬP PHỤ THUỘC HÀM	50
1	Phụ thuộc hàm có vế trái dư thừa	50
2	Tập phụ thuộc hàm có vế phải một thuộc tính.....	51
3	Tập phụ thuộc hàm không dư thừa.....	51
4	Tập phụ thuộc hàm tối thiểu	51
III	KHÓA CỦA LƯỢC ĐỒ QUAN HỆ	52
1	Định Nghĩa	52
2	Thuật toán tìm tất cả khóa	53
IV	BÀI TẬP	55
Chương 6		57
CHUẨN HÓA CƠ SỞ DỮ LIỆU		57
I	DẠNG CHUẨN CỦA LƯỢC ĐỒ QUAN HỆ	57
1	Định nghĩa các dạng chuẩn	57
II	PHÉP TÁCH KẾT NỐI BẢO TOÀN	62
1	Phép tách kết nối bảo toàn thông tin	62
2	Phép tách bảo toàn phụ thuộc hàm	67
III	THIẾT KẾ CSDL BẰNG CÁCH PHÂN RÃ.....	70
1	Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin	70
2	Phân rã thành dạng chuẩn 3 vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm	76
IV	BÀI TẬP	79

----oOo----

Chương 1 .

MÔ HÌNH QUAN HỆ

I NGUYÊN NHÂN RA ĐỜI CỦA MÔ HÌNH QUAN HỆ (RELATIONAL MODEL)

Trong nhiều năm, công nghệ tính toán và thông tin phát triển từ những hệ thống lớn, đắt tiền, độc quyền đến các hệ thống mở mạnh và không đắt tiền. Sự phát triển này mang lại lợi ích to lớn cho người dùng cuối bởi sự phát triển của các gói ứng dụng số như xử lý văn bản, bảng tính điện tử, văn phòng xuất bản, hệ quản lý cơ sở dữ liệu, máy tính trợ giúp công nghệ phần mềm....

Trước khi máy tính hóa cơ sở dữ liệu được giới thiệu, dữ liệu được lưu trữ theo kiểu điện tử thành nhiều tập tin riêng biệt sử dụng **hệ tập tin** (từ đây về sau ta gọi **hệ tập tin theo lối cũ**). Những tập tin này được xử lý bằng các ngôn ngữ thế hệ thứ ba như COBOL, FORTRAN, PASCAL và ngay cả BASIC để tạo ra các giải pháp cho các vấn đề của doanh nghiệp. Mỗi ứng dụng, chẳng hạn như hệ tính lương, hệ kho hay hệ thống kế toán sẽ có một tập các tập tin riêng chứa dữ liệu riêng. Các ứng dụng như vậy tạo ra ba vấn đề sau:

- Có sự liên kết chặt chẽ giữa cấu trúc luận lý và cấu trúc vật lý của các tập tin và chương trình ứng dụng khai thác chúng. Điều này khiến việc tạo nên các ứng dụng này rất khó khăn, tốn nhiều thời gian và do vậy mà tốn kém trong bảo trì hệ thống.
- Có sự dư thừa dữ liệu rất lớn qua việc trùng lặp các tập tin trong các ứng dụng khác nhau. Điều này tạo ra những vấn đề như: dữ liệu thiếu nhất quán, không gian đĩa bị lãng phí, thời gian bảo trì và lưu phòng hồ các tập tin gia tăng, vấn đề về quản trị như không chú trọng bảo mật và tổ chức dữ liệu thiếu thống nhất. Một ví dụ điển hình về sự trùng lặp dữ liệu là:

Hệ quản lý nguồn nhân lực bao gồm ba hệ chính:

1. Hệ lương, hệ này duy trì ngày công và lương cho tất cả nhân viên.
2. Hệ nhân sự, hệ này duy trì lý lịch cá nhân, dữ liệu về tổ chức, công việc đào tạo và vị trí thăng tiến.
3. Hệ hưu, hệ này quản trị các qui tắc liên quan đến nghỉ hưu, loại nghỉ hưu. Chi tiết về hưu của từng nhân viên.

Vấn đề phức tạp là Hệ lương thông thường được quản lý bởi phòng tài chính, trong khi Hệ nhân sự và Hệ hưu được quản lý bởi phòng tổ chức. Rõ ràng, có nhiều dữ liệu về nhân viên là chung cho cả ba hệ. Thường những hệ này thực hiện và giữ gìn riêng biệt và chúng tạo sự trùng dữ liệu nhân viên mà chúng dùng.

- Người sử dụng có ít khả năng khai thác trực tiếp dữ liệu.

II CƠ SỞ DỮ LIỆU VÀ HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU (DATABASE AND DATABASE MANAGEMENT SYSTEM)

Khởi đầu, sự giới thiệu CSDL và HQTCSDL nhằm giải quyết các vấn đề của hệ thông tin dựa trên các tập tin theo lối cũ (C1.I). Điều này tạo ra việc phát triển trên hai mươi lăm năm qua một hệ CSDL quan hệ thương mại xuất hiện cuối những năm thập niên 70 và các năm đầu của thập niên 80. Trước khi xem xét CSDL và hệ QTCSDLQH giải quyết một vài vấn đề của hệ thông tin theo lối cũ như thế nào chúng ta cần làm rõ vài khái niệm.

1 CSDL là gì?

Một cơ sở dữ liệu có thể định nghĩa tạm như sau: một chỗ chứa có tổ chức tập hợp các tập tin dữ liệu có tương quan, các mẫu tin và các cột.

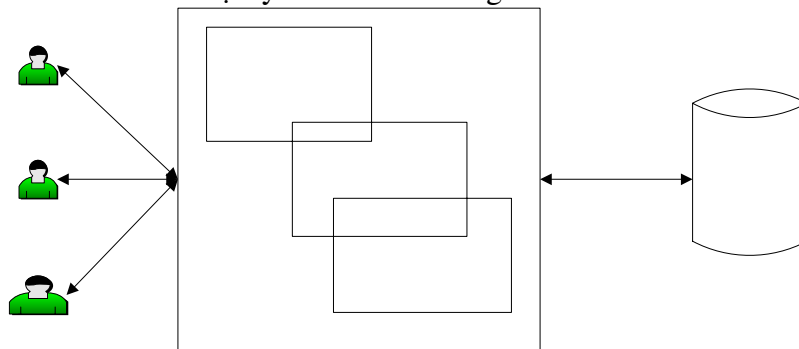
Ngày nay CSDL tồn tại trong mỗi ứng dụng thông dụng, ví dụ:

- Hệ kho và kiểm kê.
- Hệ đặt chỗ máy bay
- Hệ nguồn nhân lực.
- hệ dịch vụ công cộng như cấp nước, điện, khí đốt
- Điều khiển quá trình chế tạo và sản xuất

2 Hệ quản trị CSDL

Một hệ quản trị CSDL (HQTCSDL) là:

- một tập các phần mềm quản lý CSDL và cung cấp các dịch vụ xử lý CSDL cho các những người phát triển ứng dụng và người dùng cuối.
- HQTCSDL cung cấp một giao diện giữa người sử dụng và dữ liệu.
- HQTCSDL biến đổi CSDL vật lý thành CSDL logic.



Dựa vào cách tổ chức dữ liệu, HQTCSDL được chia thành năm loại:

- loại phân cấp như hệ IMS của IBM
- loại mạng như IDMS của Cullinet Software
- Loại tập tin đảo như ADABAS của Software AG
- Loại quan hệ như ORACLE của Oracle, DB2 của IBM, ACCESS của Microsoft Access
- Loại đối tượng là một tiếp cận khá mới trong thiết kế HQTCSDL và việc sử dụng loại này sớm trở nên phổ biến

Hiện tại, loại HQTCSDL chính được sử dụng trong công nghệ là loại HQTCSDL quan hệ (RDBMS). Loại này đã chiếm lĩnh trong công nghệ trên 10-15 năm cuối cùng khi đánh bật loại HQTCSDL phân cấp và gần đây là HQTCSDL mạng.

3 Người dùng (User)

Người dùng khai thác CSDL thông qua HQTCSDL có thể phân thành ba loại: người quản trị CSDL, người phát triển ứng dụng và lập trình, người dùng cuối.

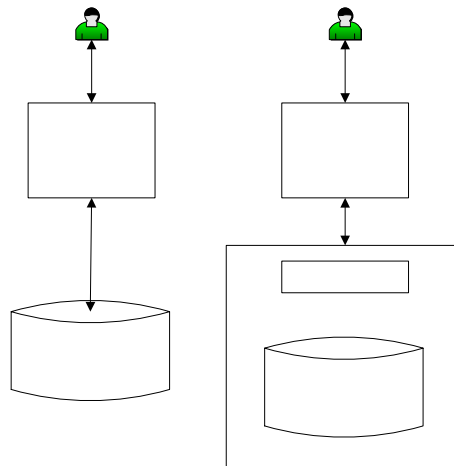
- Người quản trị CSDL, hàng ngày, chịu trách nhiệm quản lý và bảo trì CSDL như:
 - + sự chính xác và toàn vẹn của dữ liệu và ứng dụng trong CSDL, sự bảo mật của CSDL
 - + lưu phòng hồ và phục hồi CSDL
 - + giữ liên lạc với người phát triển ứng dụng, người lập trình và người dùng cuối.
 - + bảo đảm sự hoạt động trôi chảy và hiệu quả của CSDL và HQTCSDL

- Người phát triển và lập trình ứng dụng là những người chuyên nghiệp về máy tính có trách nhiệm thiết kế, tạo dựng và bảo trì hệ thông tin cho người dùng cuối.
- Người dùng cuối là những người không chuyên về máy tính nhưng họ là các chuyên gia trong các lĩnh vực khác có trách nhiệm cụ thể trong tổ chức. Họ khai thác CSDL thông qua hệ được phát triển bởi người phát triển ứng dụng hay các công cụ truy vấn của HQTCSDL.

4 CSDLQH và Hệ tập tin theo lối cũ

Tiếp cận CSDL đã giải quyết 3 vấn đề của hệ tập tin theo lối cũ:

i Vấn đề 1: cấu trúc logic và cấu trúc vật lý



Kiến trúc bên trong HQTCSDL quan hệ tách biệt rõ ràng giữa:

- cấu trúc luận lý của tất cả tập tin và chương trình ứng dụng khai thác tập tin này và
- cấu trúc vật lý của csdl và phần lưu trữ các tập tin.

Tiếp cận này tạo cho người quản trị CSDL có thể thay đổi cấu trúc vật lý hay nơi lưu trữ của tập tin mà không ảnh hưởng đến chương trình ứng dụng.

ii Vấn đề 2: dư thừa dữ liệu

Khi HQTCSDLQH được giới thiệu, nhiều tổ chức mong tích hợp các tập tin đã phân tán khắp trong tổ chức vào một CSDL tập trung. Dữ liệu có thể chia sẻ cho nhiều ứng dụng khác nhau và người sử dụng có thể khai thác đồng thời các tập con dữ liệu liên quan đến họ. Điều này làm hạn chế sự dư thừa dữ liệu.

iii Vấn đề 3: Sự khai thác dữ liệu của người sử dụng

Trong hệ QTCSDLQH người dùng có thể trực tiếp khai thác dữ liệu thông qua việc sử dụng các câu truy vấn hay các công cụ báo cáo được cung cấp bởi hệ QTCSDL.

III MÔ HÌNH QUAN HỆ (RELATIONAL MODEL)

1 Mô hình quan hệ là gì ?

Mô hình Cơ sở dữ liệu Quan hệ (gọi tắt là Mô hình Quan hệ) do E.F Codd đề xuất năm 1971. Mô hình này bao gồm:

- Một hệ thống các ký hiệu để mô tả dữ liệu dưới dạng dòng và cột như **quan hệ, bộ, thuộc tính, khóa chính, khóa ngoại, ...**
- Một tập hợp các phép toán thao tác trên dữ liệu như phép toán tập hợp, phép toán quan hệ.
- ràng buộc toàn vẹn quan hệ.

Các hệ HQTCSDLQH ngày nay được xây dựng dựa vào lý thuyết của mô hình quan hệ.

Mục đích của môn học này giúp cho sinh viên nắm được kiến trúc tổng quát về mô hình quan hệ và áp dụng nó để lập mô hình dữ liệu quan hệ có hiệu quả trong *lưu trữ* và *khai thác*.

MASV	HOTENH	MONHOC	TENKHOA	DIEMTHI
99001	TRAN DAN THU	CO SO DU LIEU	CÔNG NGHỆ THÔNG TIN	3.0
99002	NGUYEN HA DA THAO	CO SO DU LIEU	CÔNG NGHỆ THÔNG TIN	8.0
99001	TRAN DAN THU	TIN HOC VAN PHONG	CÔNG NGHỆ THÔNG TIN	6.0
99005	LE THANH TRUNG	TIN HOC VAN PHONG	ANH VAN	5.0

Chẳng hạn với bài toán quản lý điểm thi của sinh viên; nếu lưu trữ dữ liệu theo dạng bảng với các cột MASV, HOTEN, MONHOC, TENKHOA, DIEMTHI thì các giá trị của các cột HOTEN, MONHOC, TENKHOA sẽ bị trùng lặp. Sự trùng lặp này gây nên một số vấn đề:

- Ta không thể lưu trữ một sinh viên mới khi sinh viên này chưa có điểm thi
- Khi cần sửa đổi họ tên sinh viên thì ta phải sửa tất cả các dòng có liên quan đến sinh viên này. Điều này dễ gây ra tình trạng dữ liệu thiếu nhất quán.
- Khi có nhu cầu xóa điểm thi của một sinh viên kéo theo khả năng xóa luôn họ tên sinh viên đó.

Việc lưu trữ dữ liệu như trên không đúng với mô hình quan hệ. Để lưu trữ đúng với mô hình quan hệ ta phải thay MONHOC bằng MAMH, thay TENKHOA bằng MAKHOA, tách một bảng dữ liệu lớn đó ra thành nhiều bảng con, như mô hình dưới.

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

MAMH	TENMH	SOTIET
CSDL	CO SO DU LIEU	90
THVP	TIN HOC VAN PHONG	90

MAKHOA	TENKHOA
CNTT	CONG NGHE THONG TIN
AV	ANH VAN

2 Các khái niệm cơ bản của mô hình quan hệ

i Thuộc tính (Attribute, Arity)

Chẳng hạn với bài toán quản lý điểm thi của sinh viên; với đối tượng sinh viên ta cần phải chú ý đến các đặc trưng riêng như họ tên, ngày sinh, nữ (giới tính), tỉnh thường trú, học bổng, lớp mà sinh viên theo học,... các đặc trưng này gọi là **thuộc tính**.

Các thuộc tính được phân biệt qua tên gọi và phải thuộc vào một kiểu dữ liệu nhất định (số, chuỗi, ngày tháng, logic, hình ảnh,...). Kiểu dữ liệu ở đây là kiểu đơn. Trong cùng một đối tượng không được có hai thuộc tính cùng tên.

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. Thuộc tính **ngày trong tháng** thì có kiểu dữ liệu là số nguyên, miền giá trị của nó là 1 đến (tối đa là) 31. Hoặc **điểm thi** của sinh viên chỉ là các số nguyên từ 0 đến 10.

Thường người ta dùng các chữ cái hoa A,B,C,... để biểu diễn các thuộc tính, hoặc A_1, A_2, \dots, A_n để biểu diễn một số lượng lớn các thuộc tính.

ii 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 $\{A_1, A_2, \dots, A_n\}$ được viết là $Q(A_1, A_2, \dots, A_n)$. Tập các thuộc tính của Q được ký hiệu là Q^+ . Chẳng hạn lược đồ quan hệ sinh viên (Đặt tên là Sv) với các thuộc tính như trên là:

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

Thường khi thành lập một lược đồ, người thiết kế luôn gán cho nó một ý nghĩa nhất định, ý nghĩa đó gọi là **tân từ** của lược đồ quan hệ đó. Dựa vào tân từ người ta xác định được **tập thuộc tính khóa** của lược đồ quan hệ (khái niệm khóa sẽ được trình bày ở phần sau).

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. Chẳng hạn tân từ của lược đồ quan hệ trên 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), nữ (NU), ngày sinh (NGAYSINH), lớp theo học (MALOP), học bổng (HOCBONG), tỉnh cư trú (TINH).

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

Ví dụ lược đồ cơ sở dữ liệu để 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)

Phần giải thích các thuộc tính:

MASV	Mã sinh viên
HOTEN	Tên sinh viên
NU	Nữ
NGAYSINH	ngày sinh
LOP	lớp
TENLOP	tên lớp
SISO	sĩ số lớp
MAKHOA	mã khoa
HOCBONG	học bổng
TINH	tỉnh
TENKHOA	tên khoa
SOCBGD	số cán bộ giảng dạy
MAMH	mã môn học
TENMH	tên môn học
SOTIET	số tiết
DIEMTHI	điểm thi

iii Quan Hệ (Relation)

Sự thể hiện của lược đồ quan hệ Q ở một thời điểm nào đó được gọi là **quan hệ**, rõ ràng là 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 r, s, q .

iv Bộ (Tuple)

Bộ là tập mỗi giá trị liên quan của tất cả các thuộc tính của một lược đồ quan hệ.

Chẳng hạn quan hệ sau có 2 bộ.

MASV	HOTEN	NU	NGAYSINH	MALOP	TINH	HOCBONG
99001	TRAN DAN THU	TRUE	15-03-1977	CĐTH2B	TIEN GIANG	120000
99002	NGUYEN HA DA THAO	TRUE	25-04-1986	TCTH29C	TPHCM	120000

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: $t \in r$.

Về trực quan thì mỗi quan hệ xem như một bảng, trong đó mỗi cột là thông tin về một thuộc tính, mỗi dòng là thông tin về một bộ. Chẳng hạn sau đây là các thể hiện của các quan hệ định nghĩa trên lược đồ cơ sở dữ liệu quản lý sinh viên ở trên:

rSv

MASV	HOTEN	NU	NGAYSINH	MALOP	TINH	HOCBONG
99001	TRAN DAN THU	TRUE	15-03-1977	CĐTH2B	TIEN GIANG	120000
99002	NGUYEN HA DA THAO	TRUE	25-04-1986	CĐTH2B	TPHCM	120000
99003	PHAM ANH HUY	FALSE	16-08-1977	CĐTH2B	BAC LIEU	
99004	NGUYEN NGOC THUAN	FALSE	24-12-1980	CĐTH2B	CA MAU	
99005	LE THANH TRUNG	FALSE	20-11-1978	CĐAV1	CA MAU	120000
99006	NGUYEN HONG VAN	FALSE	19-09-1979	CĐAV1	SOC TRANG	
99007	VU THI LOAN	TRUE	15-03-1975	CĐAV1	CA MAU	
99008	TRUONG KIM QUANG	FALSE	15-05-1975	CĐTH2B	HA NOI	
99009	TON THAT QUYEN	FALSE	26-06-1976	CĐTH2B	VUNG TAU	60000
99010	HA VAN LONG	FALSE	14-04-1973	CĐAV1	BAC LIEU	
99011	BUI VAN ANH	FALSE	22-12-1972	CĐAV1	AN GIANG	
99012	LE HUU CHI	FALSE	28-08-1977	CĐĐT2	CAN THO	60000
99013	VU THANH CONG	FALSE	29-03-1979	CĐTH2B	KIEN GIANG	60000
99014	TRAN QUANG CUONG	FALSE	30-05-1981	CĐĐT2	DONG THAP	120000
99015	PHAM VAN HAI	FALSE	30-06-1976	CĐĐT2	CA MAU	
99016	HUYNH THANH HOANG	FALSE	29-07-1982	CĐĐT2	TPHCM	80000
99017	TRAN MINH LAM	FALSE	21-08-1977	CĐTH2B	TRA VINH	
99018	PHAN VAN SANG	FALSE	19-05-1979	CĐDL1	DONG THAP	120000
99019	PHAM THI HUYEN	FALSE	16-06-1982	CĐDL1	CAN THO	120000
99020	NGUYEN THI NGAN	TRUE	11-11-1981	CĐTH2B	CA MAU	120000
99021	PHAM TAN QUANG	FALSE	01-01-1980	CĐDL1	CA MAU	
99022	TRAN PHUOC QUYEN	FALSE	12-12-1979	CĐTH2B	BAC LIEU	60000
99023	LE THI THANH VAN	TRUE	11-11-1980	CĐDL1	TPHCM	120000

rKh

MAKHOA	TENKHOA	SOCBGD
CNTT	CONG NGHE THONG TIN	60
AV	ANH VAN	60
HOA	HOA CHAT	20
MÁY TÍNH	MOI TRUONG	10
XD	XAY DUNG	10
DL	DU LICH	5

rMh

MAMH	TENMH	SOTIET
CSDL	CO SO DU LIEU	90
FOX	FOXPRO	120
THVP	TIN HOC VAN PHONG	90
AVTH	ANH VAN TIN HOC	60
KTS	KY THUAT SO	60
CTDL	CAU TRUC DU LIEU	60

TTIN	TOAN - TIN HOC	30
SINH	CONG NGHE SINH HOC	30
VL	VAT LY	20
ĐT	ĐIỆN TỬ	20

TTNT	TRI TUE NHAN TAO	45
MANG	MANG MAY TINH CB	45
VB	VI SUAL BASIC	90
AC	ACCESS	180
LR	LAP RAP MAY TINH	60
INTER	CAC DICH VU INTERNET	45
HDH	HE DIEU HANH	60
KTTLT	KY THUAT LAP TRINH	45
VIFOX	VISUAL FOXPRO	60

rKq

MASV	MAMH	DIEMTHI
99001	CSDL	3.0
99002	CSDL	8.0
99003	CSDL	8.0
99004	CSDL	3.0
99005	CSDL	2.0
99001	THVP	6.0
99002	THVP	3.0
99003	THVP	8.0
99004	THVP	9.0
99005	THVP	5.0

99006	MANG	6.0
99007	MANG	2.0
99008	MANG	7.0
99009	MANG	3.0
99010	TTNT	5.0
99011	FOX	4.0
99012	FOX	5.0
99013	FOX	7.0
99014	VB	7.0
99015	VB	3.0

99016	KTS	7.0
99017	KTS	7.0
99017	FOX	4.0
99017	MANG	8.0
99017	CSDL	8.0
99017	TTNT	6.0
99002	MANG	8.0
99004	MANG	4.0
99018	TTNT	7.0
99019	CSDL	8.0

99020	THVP	7.0
99021	MANG	7.0
99022	MANG	6.0
99023	CSDL	8.0
99023	MANG	6.0

99023	TTNT	3.0
99023	THVP	6.0
99023	FOX	8.0
99023	VB	9.0
99023	KTS	6.0

99021	CSDL	8.0
99021	THVP	9.0
99022	FOX	5.0
99022	TTNT	6.0

rLop

MALOP	TENLOP	SISO	MAKHOA
CĐTH2B	CAO ĐẲNG TIN HỌC KHOÁ 2000B	60	cntt
TCTH29C	TRUNG CẤP TIN HỌC KHOÁ 29 C	121	cntt
CĐAV1	CAO ĐẲNG ANH VĂN 1	120	av
CĐĐT2	CAO ĐẲNG ĐIỆN TỬ 2	80	đt
CĐDL1	CAO ĐẲNG DU LỊCH 1	42	dl

v Khóa (Key, Candidate Key):

Cho lược đồ quan hệ R , $S \subseteq R^+$. S được gọi là một **siêu khóa** (*superkey*) của lược đồ quan hệ R nếu với hai bộ tùy ý trong quan hệ R thì giá trị của các thuộc tính trong S là khác nhau.

Một lược đồ quan hệ có thể có nhiều siêu khóa. Siêu khóa chứa ít thuộc tính nhất được gọi là **khóa chỉ định**, trong trường hợp lược đồ quan hệ có nhiều khóa chỉ định, thì khóa được chọn để cài đặt gọi là **khóa chính** (*Primary key*) (trong các phần sau khóa chính được gọi tắt là khóa)

Các thuộc tính tham gia vào một khóa được gọi là **thuộc tính khóa (prime key)**, ngược lại được gọi là **thuộc tính không khóa (non prime 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.

Ví dụ: Ta hãy xem lược đồ quan hệ sau:

Xe (SODANGBO, QUICACH, INHDANG, MAUSAC, SOSUON, SOMAY, MAXE, QUOCGIA)

Siêu khóa: (SOSUON, QUICACH), ...

Khóa chỉ định: (SODANGBO, QUOCGIA), (SOSUON), (SOMAY), (MAXE)

Khóa chính: MAXE

Thuộc tính khóa: SODANGBO, QUOCGIA, SOSUON, SOMAY, MAXE

Thuộc tính không khóa: QUICACH, HINHDANG, MAUSAC

Khóa của Sv là (MASV), Khoá của Mh là (MAMH), khoá của Kh là (MAKHOA), khóa của Kq là (MASV, MAMH) khóa của Lop là MALOP, trong Lop thuộc tính MAKHOA là khóa ngoại

3 Các phép toán tập hợp (set operation)

i Phép hợp (Union operation)

Cho hai lược đồ quan hệ Q_1 và Q_2 có cùng tập thuộc tính $\{A_1, A_2, \dots, A_n\}$. r_1 và r_2 lần lượt là hai quan hệ trên Q_1 và Q_2 . Phép hợp của hai lược đồ quan hệ Q_1 và Q_2 sẽ tạo thành một lược đồ quan hệ Q_3 . Q_3 được xác định như sau:

$$Q_3^+ = \{A_1, A_2, \dots, A_n\}$$

$$r_3 = r_1 + r_2 = \{t \mid t \in r_1 \text{ hoặc } t \in r_2\}$$

Ví dụ:

r_1			r_2			$r_3 = r_1 + r_2$		
MASV	MAMH	DIEMTHI	MASV	MAMH	DIEMTHI	MASV	MAMH	DIEMTHI
99001	CSDL	5.0	99002	CTDL	2.0	99001	CSDL	5.0
99002	CTDL	2.0	99001	TTNT	5.0	99002	CTDL	2.0
99003	MANG	8.0	99003	CSDL	6.0	99003	MANG	8.0
						99001	TTNT	5.0
						99003	CSDL	6.0

ii Phép Giao (Intersection):

Cho hai lược đồ quan hệ Q_1 và Q_2 có cùng tập thuộc tính $\{A_1, A_2, \dots, A_n\}$. r_1 và r_2 lần lượt là hai quan hệ trên Q_1 và Q_2 . Phép giao của hai lược đồ quan hệ Q_1 và Q_2 sẽ tạo thành một lược đồ quan hệ Q_3 như sau:

$$Q_3^+ = \{A_1, A_2, \dots, A_n\}$$

$$r_3 = r_1 * r_2 = \{t \mid t \in r_1 \text{ và } t \in r_2\}$$

Ví dụ:

r_1			r_2			$r_3 = r_1 * r_2$		
MASV	MAMH	DIEMTHI	MASV	MAMH	DIEMTHI	MASV	MAMH	DIEMTHI
99001	CSDL	5.0	99002	CTDL	2.0	99002	CTDL	2.0
99002	CTDL	2.0	99001	TTNT	5.0			
99003	MANG	8.0	99003	CSDL	6.0			

iii Phép Trừ (Minus, difference)

Cho hai lược đồ quan hệ Q_1 và Q_2 có cùng tập thuộc tính $\{A_1, A_2, \dots, A_n\}$. r_1 và r_2 lần lượt là hai quan hệ trên Q_1 và Q_2 . Phép trừ lược đồ quan hệ Q_1 cho Q_2 sẽ tạo thành một lược đồ quan hệ Q_3 như sau:

$$Q_3^+ = \{A_1, A_2, \dots, A_n\}$$

$$r_3 = r_1 - r_2 = \{t \mid t \in r_1 \text{ và } t \notin r_2\}$$

Ví dụ:

r_1			r_2			$r_3 = r_1 - r_2$		
MASV	MAMH	DIEMTHI	MASV	MAMH	DIEMTHI	MASV	MAMH	DIEMTHI
99001	CSDL	5.0	99002	CTDL	2.0	99001	CSDL	5.0
99002	CTDL	2.0	99001	TTNT	5.0	99003	MANG	8.0
99003	MANG	8.0	99003	CSDL	6.0			

iv Tích Descartes (Cartesian Product, product)

Cho hai lược đồ quan hệ $Q_1(A_1, A_2, \dots, A_n)$, $Q_2(B_1, B_2, \dots, B_m)$. r_1 và r_2 lần lượt là hai quan hệ trên Q_1 và Q_2 . Tích Descartes của hai lược đồ quan hệ Q_1 và Q_2 sẽ tạo thành một lược đồ quan hệ Q_3 như sau:

$$Q_3^+ = Q_1^+ \cup Q_2^+ = \{A_1, \dots, B_1, \dots\}$$

$$r_3 = r_1 \times r_2 = \{(t_1, t_2) \mid t_1 \in r_1 \text{ và } t_2 \in r_2\}$$

Ví dụ:

r_1			$r_3 = r_1 \times r_2$				
MASV	MAMH	DIEMTHI	MASV	MAMH	DIEMTHI	MAMH	TENMH
99001	CSDL	5.0	99001	CSDL	5.0	CSDL	CO SO DU LIEU
99002	CTDL	2.0	99001	CSDL	5.0	FOX	FOXPRO
99003	MANG	8.0	99002	CTDL	2.0	CSDL	CO SO DU LIEU
			99002	CTDL	2.0	FOX	FOXPRO
			99003	MANG	8.0	CSDL	CO SO DU LIEU
			99003	MANG	8.0	FOX	FOXPRO

r_2	
MAMH	TENMH
CSDL	CO SO DU LIEU
FOX	FOXPRO

4 Các phép toán quan hệ**i Phép Chiếu (Projection)**

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

Phép chiếu của Q lên tập thuộc tính X sẽ tạo thành lược đồ quan hệ $Q' = Q[X]$, trong đó Q'^+ chính là X và r' chính là r nhưng chỉ lấy các thuộc tính của X .

$$Q'^+ = X$$

$$r' = r[X] = r.X = \{t' \mid \exists t \in r \text{ và } t.X = t'[X] = t'\}$$

phép chiếu chính là phép rút trích dữ liệu theo cột (chiều dọc)

Ví dụ:

r			r' = r.{MAMH}
MASV	MAMH	DIEMTHI	MAMH
99001	CSDL	5.0	CSDL
99002	CTDL	2.0	CTDL
99003	MANG	8.0	MANG

ii Phép Chọn (Selection)

Cho lược đồ quan hệ $Q(A_1, A_2, \dots, A_n)$, r là một quan hệ trên Q . $X \subset Q^+$ và E là một mệnh đề logic được phát biểu trên tập X . Phần tử $t \in r$ thỏa mãn điều kiện E ký hiệu là $t(E)$.

Phép chọn từ r theo điều kiện E sẽ tạo thành một lược đồ quan hệ Q' như sau:

$$Q'^+ = Q^+$$

$$r' = r(E) = r : E = \{t \mid t \in r \text{ và } t(E)\}$$

phép chọn chính là phép rút trích dữ liệu theo dòng (chiều ngang)

Ví dụ:

r			$r' = r : \text{DIEMTHI} \geq 5$		
MASV	MAMH	DIEMTHI	MASV	MAMH	DIEMTHI
99001	CSDL	5.0	99001	CSDL	5.0
99002	CTDL	2.0			
99003	MANG	8.0	99003	MANG	8.0

iii Phép kết, Phép Kết Tự Nhiên (join, natural join):

Cho hai lược đồ quan hệ $Q_1(A_1, A_2, \dots, A_n)$, $Q_2(B_1, B_2, \dots, B_m)$.

r_1 và r_2 lần lượt là hai quan hệ trên Q_1 và Q_2 .

A_i và B_j lần lượt là các thuộc tính của Q_1 và Q_2 sao cho $\text{MGT}(A_i) = \text{MGT}(B_j)$ (MGT: miền giá trị).

θ là một phép so sánh trên $\text{MGT}(A_i)$.

Phép kết giữa Q_1 và Q_2 sẽ tạo thành một lược đồ quan hệ Q_3 như sau:

$$Q_3^+ = Q_1^+ \cup Q_2^+$$

$$r_3 = r_1 \bowtie_{A_i \theta B_j} r_2 = \{t_{12} \mid \exists t_1 \in r_1, \exists t_2 \in r_2 \text{ sao cho}$$

$$t_{12}.Q_1^+ = t_1$$

$$t_{12}.Q_2^+ = t_2$$

$$t_1.A_i \theta t_2.B_j\}$$

Ta rút ra các bước cụ thể để thực hiện phép kết như sau:

- Tạo tích descartes
- Thực hiện phép chọn theo điều kiện $E = A_i \theta B_j$

Ví dụ:

A_i là thuộc tính B, B_j là thuộc tính F và θ là phép so sánh " \geq ". Ta được kết quả là quan hệ sau:

r_1			r_2			$r_3 = r_1 \bowtie_{B \geq F} r_2$					
A	B	C	E	F	H	A	B	C	E	F	H
6	5	4	1	5	9	6	5	4	1	5	9
7	5	5	4	6	8	6	5	4	7	5	3
4	2	6	7	5	3	7	5	5	1	5	9
						7	5	5	7	5	3

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

Ví dụ: Với $A_i \equiv B_j = \text{MAMH}$

r_1			r_2		$r_3 = r_1 \bowtie_{\text{MAMH}} r_2$			
MASV	MAMH	DIEMTHI	MAMH	TENMH	MASV	MAMH	DIEMTHI	TENMH
99001	CSDL	5.0	CSDL	CO SO DU LIEU	99001	CSDL	5.0	CO SO DU LIEU
99002	CTDL	2.0	CTDL	CAU TRUC DU LIEU	99002	CTDL	2.0	CAU TRUC DU LIEU
99003	MANG	8.0						

iv Phép chia (division):

Cho hai lược đồ quan hệ $Q_1(A_1, A_2, \dots, A_n)$, $Q_2(B_1, B_2, \dots, B_m)$.

r_1 và r_2 lần lượt là hai quan hệ trên Q_1 và Q_2 .

A_i và B_j lần lượt là các thuộc tính của Q_1 và Q_2 sao cho $n > m$.

Phép chia Q_1 và Q_2 sẽ tạo thành một lược đồ quan hệ Q_3 như sau:

$$Q_3^+ = \{A_1, \dots, A_{n-m}\}$$

$$r_3 = r_1 \div r_2 = \{t_3 \mid \forall t_2 \in r_2, \exists t_1 \in r_1 \text{ } t_3 = t_1 \cdot \{A_1, \dots, A_{n-m}\} \\ t_2 = t_1 \cdot \{A_{n-m+1}, \dots, A_n\}\}$$

Ví dụ:

r_1					r_2		$r_3 = r_1 \div r_2$		
A_1	A_2	A_3	A_4	A_5	B_1	B_2	A_1	A_2	A_3
a	b	d	c	g	c	g	a	b	d
a	b	d	e	f	e	f	e	g	c
b	c	e	e	f					
e	g	c	c	g					
e	g	c	e	f					
a	b	e	g	e					

5 Các tính chất của đại số quan hệ

Q là lược đồ quan hệ

q, r, s là quan hệ trên Q ,

E, E_1, E_2 là mệnh đề logic trên Q^+

$X_1 \subseteq X_2 \subseteq Q^+$

Hãy chứng minh các tính chất sau:

$$(r : E_1) : E_2 = (r : E_2) : E_1$$

Chứng minh:

$$\begin{aligned} (r : E_1) : E_2 &= \{t' \mid t' \in (r : E_1) \text{ và } t'(E_2)\} \\ &= \{t' \mid t' \in \{t \mid t \in r \text{ và } t(E_1)\} \text{ và } t'(E_2)\} \\ &= \{t' \in r \mid t'(E_1) \text{ và } t'(E_2)\} \\ &= \{t' \mid t' \in \{t \mid t \in r \text{ và } t(E_2)\} \text{ và } t'(E_1)\} \\ &= \{t' \mid t' \in (r : E_2) \text{ và } t'(E_1)\} \\ &= (r : E_2) : E_1 \end{aligned}$$

$$(r+s) : E = (r : E) + (s : E)$$

Chứng minh:

$$\begin{aligned} (r+s) : E &= \{t \mid t \in (r+s) \text{ và } t(E)\} \\ &= \{t \mid t \in \{t' \mid t' \in r \text{ hoặc } t' \in s\} \text{ và } t(E)\} \\ &= \{t \mid (t \in r \text{ hoặc } t \in s) \text{ và } t(E)\} \\ &= \{t \mid (t \in r \text{ và } t(E)) \text{ hoặc } (t \in s \text{ và } t(E))\} \\ &= \{t \mid t \in \{t' \mid t' \in r \text{ và } t'(E)\} \text{ hoặc } t \in \{t' \mid t' \in s \text{ và } t'(E)\}\} \\ &= \{t \mid t \in (r : E) \text{ hoặc } t \in (s : E)\} \\ &= (r : E) + (s : E) \end{aligned}$$

$$(r*s) : E = (r : E) * (s : E)$$

Chứng minh:

$$(r*s) : E = \{t \mid t \in (r*s) \text{ và } t(E)\}$$

$$\begin{aligned}
&= \{t \mid t \in \{t' \mid t' \in r \text{ và } t' \in s\} \text{ và } t(E)\} \\
&= \{t \mid t \in r \text{ và } t \in s \text{ và } t(E)\} \\
&= \{t \mid (t \in r \text{ và } t(E)) \text{ và } (t \in s \text{ và } t(E))\} \\
&= \{t \mid t \in \{t' \mid t' \in r \text{ và } t'(E)\} \text{ và } t \in \{t' \mid t' \in s \text{ và } t'(E)\}\} \\
&= \{t \mid t \in (r:E) \text{ và } t \in (s:E)\} \\
&= (r:E) * (s:E)
\end{aligned}$$

$$(r-s):E = (r:E) - (s:E)$$

Chứng minh:

$$\begin{aligned}
(r-s):E &= \{t \mid t \in (r-s) \text{ và } t(E)\} \\
&= \{t \mid t \in \{t' \mid t' \in r \text{ và } t' \notin s\} \text{ và } t(E)\} \\
&= \{t \mid t \in r \text{ và } t \notin s \text{ và } t(E)\} \\
&= \{t \mid (t \in r \text{ và } t(E)) \text{ và } (t \notin s \text{ và } t(E))\} \\
&= \{t \mid t \in \{t' \mid t' \in r \text{ và } t'(E)\} \text{ và } t \notin \{t' \mid t' \in s \text{ và } t'(E)\}\} \\
&= \{t \mid t \in (r:E) \text{ và } t \notin (s:E)\} \\
&= (r:E) * (s:E)
\end{aligned}$$

$$\text{Với } X_2 \supseteq X_1 \Rightarrow (r.X_2).X_1 = r.X_1$$

Chứng minh:

$$\begin{aligned}
(r.X_2).X_1 &= \{t.X_1 \mid t \in (r.X_2)\} \\
&= \{t.X_1 \mid t \in \{t'.X_2 \mid t' \in r\}\} \\
&= \{(t'.X_2).X_1 \mid t' \in r\} \\
&= \{t'.X_1 \mid t' \in r\} \text{ vì } X_1 \subseteq X_2 \\
&= r.X_1
\end{aligned}$$

$$E \text{ phát biểu trên } X \Rightarrow (r:E).X = (r.X):E$$

Chứng minh:

$$\begin{aligned}
(r:E).X &= \{t.X \mid t \in (r:E)\} \\
&= \{t.X \mid t \in \{t' \mid t' \in r \text{ và } t'(E)\}\} \\
&= \{t.X \mid t \in r \text{ và } t(E)\} \\
&= \{t' \mid t' \in \{t.X \mid t \in r\} \text{ và } t'(E)\} \\
&= \{t' \mid t' \in (r.X) \text{ và } t'(E)\} \\
&= (r.X):E
\end{aligned}$$

$$q|><|r = r|><|q$$

Chứng minh:

$$\begin{aligned}
(q|><|r) &= \{t_{12} \mid \exists t_1 \in q, \exists t_2 \in r \ t_{12}.Q^+ = t_1, \ t_{12}.R^+ = t_2 \ t_{12}.A_i \theta t_{12}.B_j\} \\
&= r|><|q
\end{aligned}$$

$$A_i \in Q, B_j \in S, C_k \in Q, D_1 \in R \Rightarrow (q|><|r)|><|s = q|><|(r|><|s)$$

Chứng minh:

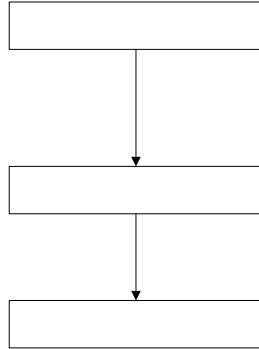
$$\begin{aligned}
(q|><|r)|><|s &= \{t_{12} \mid \exists t_1 \in (q|><|r), \exists t_2 \in s \ t_{12}.Q^+ \cup R^+ = t_1 \\
&\quad t_{12}.S^+ = t_2 \ t_{12}.A_i \theta t_{12}.B_j\} \\
&= \{t_{12} \mid \exists t_1 \in \{u_{12} \mid \exists u_1 \in q, \exists u_2 \in r \ u_{12}.Q^+ = u_1 \ u_{12}.R^+ = u_2 \ u_{12}.C_k \theta u_{12}.D_1\}, \\
&\quad \exists t_2 \in s \ t_{12}.Q^+ \cup R^+ = t_1, t_{12}.S^+ = t_2 \ t_{12}.A_i \theta t_{12}.B_j\} \\
&= \{t_{123} \mid \exists t_1 \in q, \exists t_2 \in r, \exists t_3 \in s \ t_{123}.Q^+ = t_1, t_{123}.R^+ = t_2 \ t_{123}.S^+ = t_3 \\
&\quad t_{123}.A_i \theta t_{123}.B_j \ t_{123}.C_k \theta t_{123}.D_1\} \\
&= \{t_{12} \mid \exists t_1 \in q, \exists t_2 \in \{u_{12} \mid \exists u_1 \in r, \exists u_2 \in s \ u_{12}.R^+ = u_1
\end{aligned}$$

$$\begin{aligned}
& u_{12}.S^+ = u_2 \cdot u_1.C_k \theta_2 u_2.D_1 \}, t_{12}.Q^+ = t_1 \cdot t_{12}.R^+ \cup S^+ = t_2 \cdot t_{12}.A_i \theta_1 t_{12}.B_j \} \\
& = \{ t_{12} \mid \exists t_1 \in q, \exists t_2 \in (r \mid > < \mid s) \}, t_{12}.Q^+ = t_1 \\
& \quad t_{12}.R^+ \cup S^+ = t_2 \cdot t_{12}.A_i \theta_1 t_{12}.B_j \} \\
& = q \mid > < \mid (r \mid > < \mid s)
\end{aligned}$$

IV MÔ HÌNH THỰC THỂ KẾT HỢP

1 Giới thiệu mô hình thực thể kết hợp

Các nhà phân tích thiết kế hệ thống thông tin thường xây dựng *lược đồ cơ sở dữ liệu* (C1.III.2.ii) từ mô hình thực thể kết hợp và mô hình này lại được xây dựng từ phần đặc tả vấn đề của một bài toán thực tế.



Lược đồ cơ sở dữ liệu xây dựng theo hướng này thông thường đạt tối thiểu dạng chuẩn 3 (3NF: third normal form) nghĩa là ở dạng có sự dư thừa dữ liệu ở mức tối thiểu, còn môn CSDL xây dựng lược đồ CSDL đạt dạng chuẩn 3 từ lược đồ cơ sở dữ liệu chưa đạt dạng chuẩn có kèm các *tân từ* (C1.III.2.ii). Ta hãy xem ví dụ sau:

i Ví dụ – Mối quan hệ một-nhiều

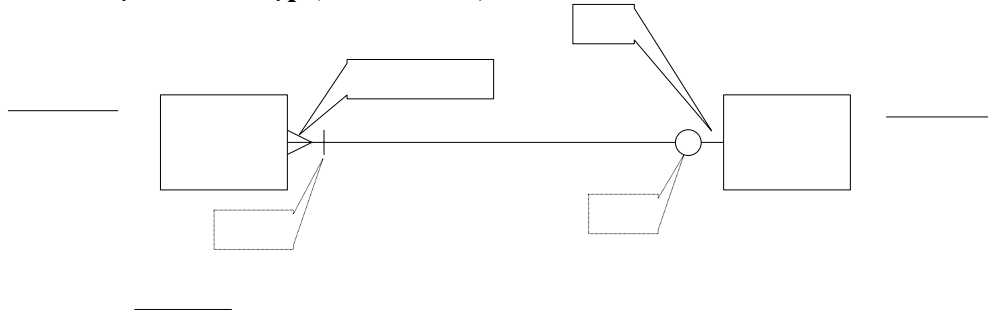
(a) Đặc tả vấn đề

Những người phụ trách đào tạo của *Trường cao đẳng cộng đồng núi Ayers* mong muốn tạo lập một CSDL về các môn đào tạo của trường (như: chứng chỉ leo núi, công nghệ bay) và học viên ghi danh vào những môn học này. Trường cũng có qui định là cùng một lúc, học viên chỉ có thể ghi danh vào một môn học. Họ chỉ quan tâm về dữ liệu của đợt ghi danh hiện tại. Một khi học viên kết thúc môn học thì nhà trường sẽ không còn quan tâm đến họ và những học viên này phải được xóa khỏi CSDL. Thông tin cần lưu trữ về một học viên bao gồm: mã học viên, tên học viên, địa chỉ, ngày sinh, số điện thoại, ngày nhập học

Thông tin về môn học gồm mã môn học, tên môn học, thời lượng

Phân tích:

- phần đặc tả vấn đề chứa đựng các qui tắc quản lý và dữ liệu yêu cầu của vấn đề.
- dữ liệu của vấn đề là: chi tiết về *học viên* có *mã học viên*, *tên học viên*, *địa chỉ*, *ngày sinh*, *số điện thoại* và *ngày nhập học* chi tiết về *môn học* có *mã môn học*, *tên môn học* và *thời lượng*.
- qui tắc quản lý gồm:
 - + Cùng một lúc, một học viên chỉ có thể ghi danh vào một môn học.
 - + Nhiều học viên có thể ghi danh vào một môn học.
 - + Nhà trường chỉ quan tâm đến những học viên của môn học hiện tại.

(b) Mô hình thực thể kết hợp (Mô hình ER)

Các tính chất trong mô hình thực thể kết hợp:

- Hình chữ nhật được gọi là **tập thực thể**. Tên của tập thực thể được ghi bên trong hình chữ nhật và dùng danh từ để đặt tên cho tập thực thể.
- Đường nối giữa hai tập thực thể được gọi là **mối quan hệ (mối kết hợp)**. Mối quan hệ trong vấn đề trên là mối quan hệ một-nhiều (1:M). Nội dung của mối quan hệ được diễn tả theo hai chiều: “ghi danh vào”, “được ghi danh bởi” và chúng diễn tả hai nội dung sau:
 - + Mỗi HỌC VIÊN có thể ghi danh vào một MÔN HỌC
 - + Mỗi MÔN HỌC phải được ghi danh bởi một hay nhiều HỌC VIÊN
- Các dữ liệu ghi bên cạnh tập thực thể được gọi là thuộc tính. Chúng cung cấp thông tin chi tiết về tập thực thể. Có hai loại thuộc tính:
 - Thuộc tính nhận diện là thuộc tính để phân biệt thực thể này với thực thể kia trong tập thực thể.
 - Thuộc tính mô tả là thuộc tính cung cấp thông tin chi tiết hơn về thực thể trong tập thực thể.
- Mối quan hệ của vấn đề trên là mối quan hệ một-nhiều rất phổ biến trong mô hình thực thể kết hợp. Hai loại kết nối còn lại ít phổ biến hơn nhưng cũng không kém phần quan trọng là mối quan hệ một-một và mối quan hệ nhiều-nhiều.

ii Ví dụ – mối quan hệ một-một**(a) Đặc tả vấn đề**

Phòng cảnh sát mong muốn quản lý lý lịch cá nhân những người lái xe và bằng lái của họ. Một người chỉ lấy được một bằng lái và một bằng lái chỉ thuộc về một người. Thông tin về lái xe mà phòng cảnh sát quan tâm là: mã người lái xe, tên, địa chỉ, ngày sinh

Thông tin về bằng lái cần lưu trữ là: mã bằng lái, loại bằng lái, ngày hết hạn

(b) Mô hình thực thể kết hợp

- mỗi NGƯỜI LÁI XE phải sở hữu một BẰNG LÁI
- mỗi BẰNG LÁI phải được sở hữu bởi một NGƯỜI LÁI XE

iii Ví dụ – mối quan hệ nhiều-nhiều**(a) Đặc tả vấn đề**

Người phụ trách đào tạo Trường cao đẳng cộng đồng núi xanh mong muốn thiết lập một csdl về các môn học mà họ cung cấp (như chứng chỉ leo núi, cử nhân công nghệ bay) và các học viên ghi danh vào các môn học này. Nhà trường qui định là một học viên được ghi danh học tối đa ba môn học

Hình 1.4.

trong cùng một lúc. Họ chỉ quan tâm đến dữ liệu của môn học hiện tại. Một khi học viên kết thúc môn học, họ sẽ không còn thuộc diện quản lý của nhà trường và phải được xóa khỏi csdl trừ khi học viên này ghi danh học tiếp môn mới. Thông tin về một học viên gồm: mã học viên, tên học viên, địa chỉ, ngày sinh, số điện thoại, ngày nhập học

Thông tin về môn học gồm: mã môn học, tên môn học, thời lượng

(b) Mô hình ER



+ **Mỗi** HỌC VIÊN có thể ghi danh vào **một hay nhiều** MÔN HỌC

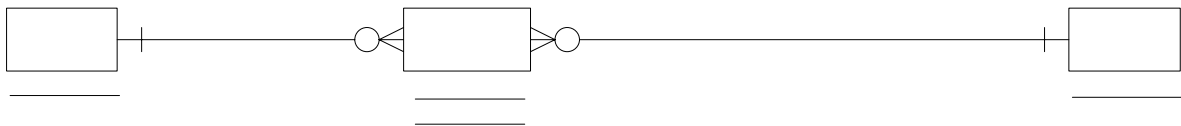
+ **Mỗi** MÔN HỌC phải được ghi danh bởi **một hay nhiều** HỌC VIÊN

Mô hình ER trên có mối quan hệ nhiều-nhiều.

(c) Loại bỏ tính kết nối nhiều-nhiều (nếu được)

Mô hình trên gặp phải khuyết điểm sau:

- **Ngày nhập học** là thuộc tính gắn liền với tập thực thể HỌC VIÊN sẽ không hợp lý vì không diễn tả được trường hợp học viên học cùng lúc nhiều môn học.
- Còn nếu **ngày nhập học** là thuộc tính của MÔN HỌC thì không diễn tả được tình trạng cùng môn học nhưng có các ngày nhập học khác nhau.



Để giải quyết vấn đề này ta phải đưa vào:

- một tập thực thể làm trung gian giữa HỌC VIÊN và MÔN HỌC gọi là tập kết hợp **PHIẾU GHI DANH**.
- Thuộc tính nhận diện của tập kết hợp là sự kết hợp giữa thuộc tính nhận diện của tập thực thể HỌC VIÊN và MÔN HỌC
- thuộc tính mô tả của tập kết hợp **PHIẾU GHI DANH** là ngày nhập học
- tính kết nối của tập kết hợp với tập thực thể là một-nhiều

Nội dung của mối quan hệ giữa các tập thực thể là:

- mỗi HỌC VIÊN có thể có một hay nhiều **PHIẾU GHI DANH**
- mỗi **PHIẾU GHI DANH** phải thuộc về một HỌC VIÊN
- mỗi **PHIẾU GHI DANH** phải ghi nhận đào tạo về một MÔN HỌC
- mỗi MÔN HỌC có thể được ghi nhận đào tạo bởi một hay nhiều **PHIẾU GHI DANH**

Các qui tắc phải tuân thủ khi thêm tập kết hợp làm trung gian để loại bỏ tính kết nối nhiều-nhiều:

- Phải nhận diện được thuộc tính mô tả của tập kết hợp.
- Nếu có thuộc tính mô tả thì tạo tập kết hợp làm trung gian giữa hai tập thực thể.
- Nếu không có thuộc tính mô tả thì vẫn giữ nguyên mô hình như hình 1.4.4

2 Chuyển từ mô hình thực thể kết hợp sang lược đồ CSDL.

i Quy tắc chung

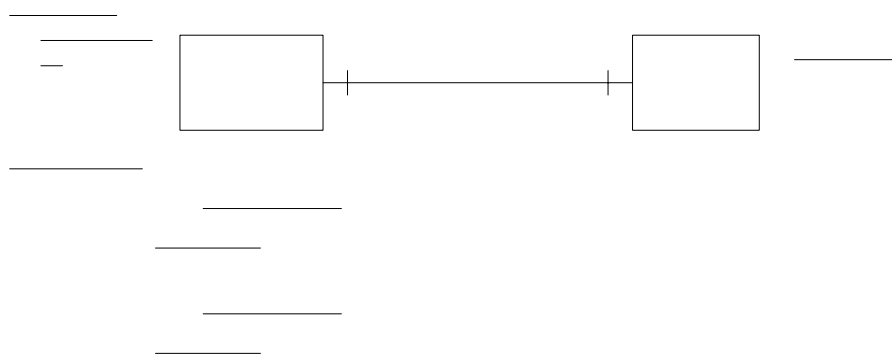
Khi biến đổi mô hình ER thành các mô hình quan hệ ta áp dụng các qui tắc sau:

- Mỗi tập thực thể trong mô hình ER được chuyển thành một lược đồ quan hệ.
- Mỗi thuộc tính trong mô hình ER được chuyển thành thuộc tính trong lược đồ quan hệ tương ứng
- Mỗi thuộc tính nhận diện trong mô hình ER được chuyển thành khóa chính trong lược đồ quan hệ tương ứng.
- Mỗi mối quan hệ trong ER được chuyển thành khóa ngoại theo qui tắc sau

ii Qui tắc chuyển mối quan hệ thành khóa ngoại

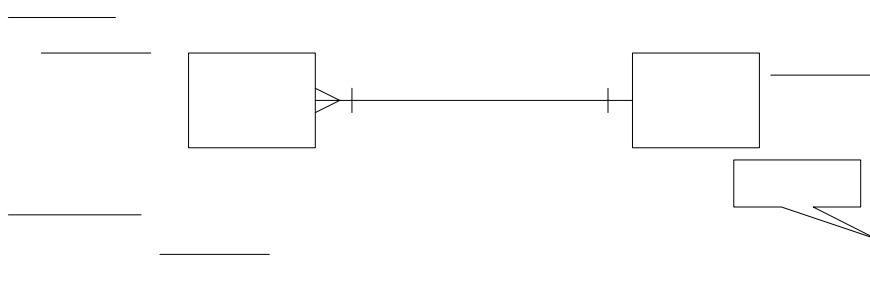
(a) Mối quan hệ một-một

Chuyển khóa chính từ quan hệ 1 sang quan hệ 2 hay ngược lại. Ví dụ vấn đề người lái xe và bằng lái sẽ có mô hình quan hệ là một trong hai mô hình quan hệ sau

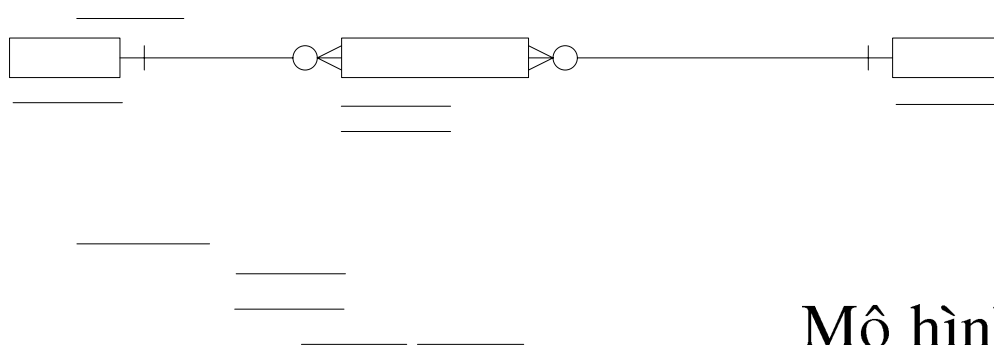


(b) Mối quan hệ một-nhiều

Chuyển khóa chính từ bên một sang bên nhiều.



(c) Mối quan hệ nhiều-nhiều đến tập kết hợp



Trong quan hệ PHIEU GHI DANH có các khóa chính khóa ngoại như sau:
+ mã học viên là khóa ngoại

Mô hình ER

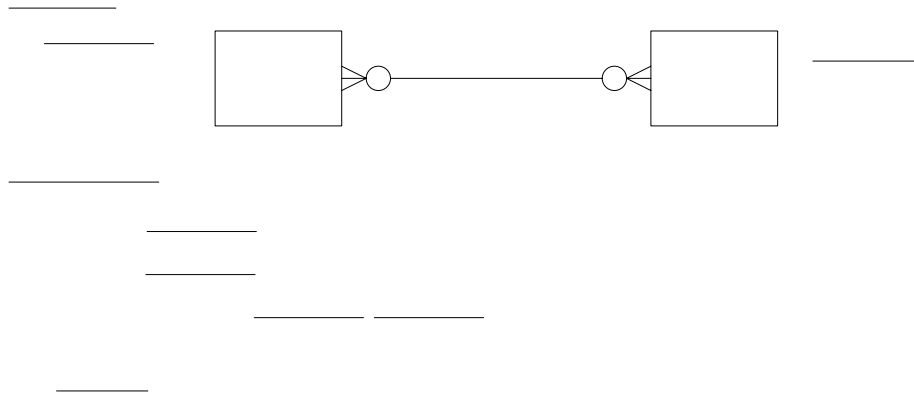
ma người lái
xe
tên
địa chỉ

- + mã môn học là khóa ngoại
- + mã học viên và mã môn học là khóa chính

(d) Mối quan hệ nhiều-nhiều

Tạo một quan hệ mới có khóa chính là sự kết hợp các khóa chính của hai quan hệ có tính kết nối nhiều-nhiều.

Ví dụ giả sử Trường Cao Đẳng Cộng Đồng Núi Xanh không quan tâm đến ngày nhập học của học viên thì mô hình ER sẽ có mối quan hệ nhiều-nhiều như sau:



V BÀI TẬP

1 Phép toán tập hợp và phép toán quan hệ

Cho lược đồ cơ sở dữ liệu dùng để quản lý hồ sơ sinh viên bao gồm các quan hệ Sv(sinh viên), Lop(Lớp), kh(khoa), Mh(môn học), Kq(kết quả) được mô tả bởi các lược đồ quan hệ như sau:

Sv (MASV, HOTEN, NU, NGAYSINH, MALOP, TINH, HOCBONG)

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

Lop (MALOP, TENLOP, SISO, MAKHOA)

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

Kh (MAKHOA, TENKHOA, SOCBGD)

Tên từ: Mỗi khoa có mỗi MAKHOA duy nhất. Mỗi MAKHOA xác định tất cả các thuộc tính còn lại của khoa đó.

Mh (MAMH, TENMH, SOTIET)

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

Kq (MASV, MAMH, DIEMTHI)

Tên từ: Mỗi sinh viên cùng với một môn học xác định duy nhất một điểm thi

YÊU CẦU:

1. Tìm khóa cho mỗi lược đồ quan hệ trên.
2. Hãy thực hiện các câu hỏi sau bằng ngôn ngữ đại số quan hệ
 - a. Lập danh sách sinh viên gồm MASV, HOTEN, HOCBONG
 - b. Lập danh sách sinh viên nữ khoa 'CNTT', danh sách cần MASV, HOTEN, HOCBONG
 - c. Lập bảng điểm cho tất cả sinh viên khoa 'CNTT', bảng điểm gồm MASV, HOTEN, TENMH, DIEMTHI
 - d. Lập phiếu điểm cho sinh viên có MASV="99001"
 - e. Lập danh sách sinh viên gồm MASV, HOTEN, TENLOP, TENKHOA

Mô hình ER

mã học viên

tên học viên

địa chỉ

ngày sinh

số điện thoại

Mô hình quan hệ

HỌC VIÊN(mã

- f. Lập bảng điểm môn học có mã môn học là CSDL cho tất cả sinh viên có mã lớp là “CĐTH2B”
- g. Lập danh sách sinh viên của lớp có mã lớp là “CĐTH2B” và có điểm thi môn học lớn hơn hay bằng 8.

2 Mô hình thực thể kết hợp

Dựa vào các phân tích sơ bộ dưới đây, hãy lập mô hình thực thể kết hợp cho mỗi bài toán quản lý sau:

i QUẢN LÝ LAO ĐỘNG

Để quản lý việc phân công các nhân viên tham gia vào xây dựng các công trình. Công ty xây dựng ABC tổ chức quản lý như sau:

Cùng lúc công ty có thể tham gia xây dựng nhiều công trình, mỗi công trình có một mã số công trình duy nhất (MACT), mỗi mã số công trình xác định các thông tin như: tên gọi công trình (TENCT), địa điểm(ĐIADIEM), ngày công trình được cấp giấy phép xây dựng (NGAYCAPGP), ngày khởi công (NGAYKC), ngày hoàn thành (NGAYHT).

Mỗi nhân viên của công ty ABC có một mã số nhân viên(MANV) duy nhất, một mã số nhân viên xác định các thông tin như: Họ tên (HOTEN), ngày sinh (NGAYSINH), phái (PHAI), địa chỉ (ĐIACHI). Mỗi nhân viên phải chịu sự quản lý hành chánh bởi một phòng ban. Tất nhiên một phòng ban quản lý hành chánh nhiều nhân viên. Công ty có nhiều phòng ban (Phòng kế toán, phòng kinh doanh, phòng kỹ thuật, phòng tổ chức, phòng chuyên môn, Phòng phục vụ,...). Mỗi phòng ban có một mã số phòng ban(MAPB) duy nhất, mã phòng ban xác định tên phòng ban (TENPB).

Công ty phân công các nhân viên tham gia vào các công trình, mỗi công trình có thể được phân cho nhiều nhân viên và mỗi nhân viên cùng lúc cũng có thể tham gia vào nhiều công trình. Với mỗi công trình một nhân viên có một số lượng ngày công (SLNGAYCONG) đã tham gia vào công trình đó.

ii QUẢN LÝ 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 đọc giả được thư viện cấp cho một thẻ thư viện, trong đó có ghi rõ mã đọc giả (MADG), cùng với các thông tin khác như : họ tên (HOTEN), ngày sinh (NGAYSINH), địa chỉ (ĐIACHI), nghề nghiệp(NGHENHIEP).

Cứ mỗi lượt mượn sách, đọc giả phải ghi các quyển sách cần mượn vào một phiếu mượn, mỗi phiếu mượn có một số phiếu mượn (SOPM) duy nhất, mỗi phiếu mượn xác định các thông tin như: ngày mượn (NGAYMUON), đọc giả mượn, các quyển sách mượn và ngày trả (NGAYTRA). Các các quyển sách trong cùng một phiếu mượn không nhất thiết phải trả trong cùng một ngày.

iii QUẢN LÝ BÁN HÀNG

Mỗi khách hàng có một mã khách hàng (MAKH) duy nhất, mỗi MAKH xác định được các thông tin về khách hàng như : họ tên khách hàng (HOTEN), địa chỉ (ĐIACHI), số điện thoại (ĐIENTHOAI). Các mặt hàng được phân loại theo từng nhóm hàng, mỗi nhóm hàng có một mã nhóm (MANHOM) duy nhất, mỗi mã nhóm hàng xác định tên nhóm hàng (TENNHOM), tất nhiên một nhóm hàng có thể có nhiều mặt hàng. Mỗi mặt hàng được đánh một mã số (MAHANG) duy nhất, mỗi mã số này

xác định các thông tin về mặt hàng đó như : tên hàng (TENHANG), đơn giá bán (ĐONGIA), đơn vị tính (ĐVT). Mỗi hóa đơn bán hàng có một số hóa đơn (SOHD) duy nhất, mỗi hóa đơn xác định được khách hàng và ngày lập hóa đơn (NGAYLAPHĐ), ngày bán hàng (NGAYBAN). Với mỗi mặt hàng trong một hóa đơn cho biết số lượng bán (SLBAN) của mặt hàng đó.

iv QUẢN LÝ LỊCH DẠY - HỌC

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 chánh 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 (SOPHONG) 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 (DENTIET), tựa đề bài dạy (BAIDAY), 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.

-----oOo-----

Chương 2 .

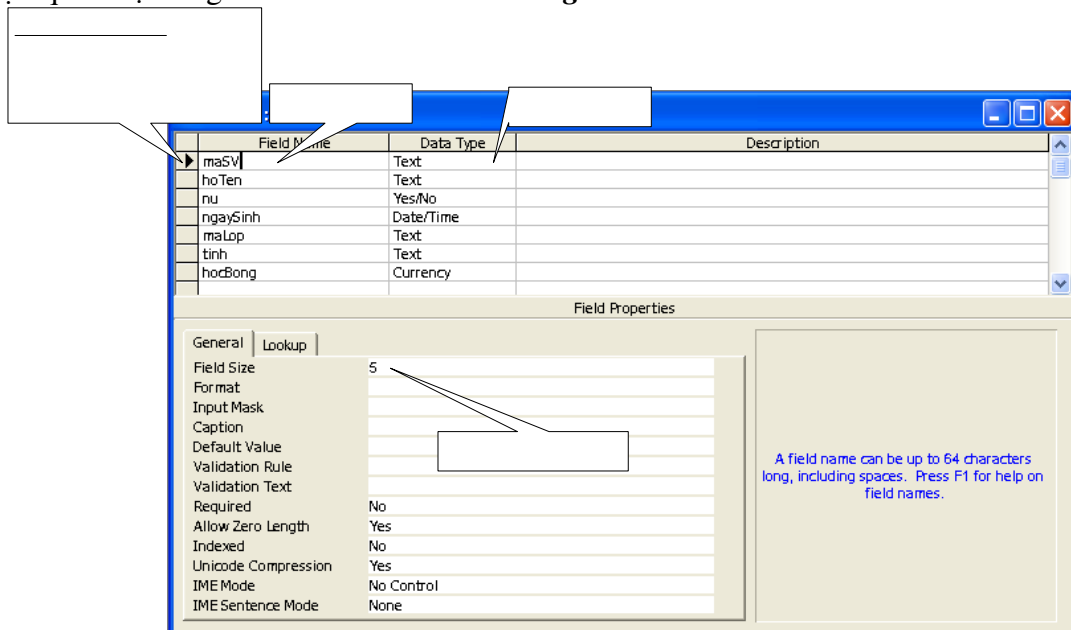
NGÔN NGỮ TRUY VẤN SQL

I CÁCH TẠO QUAN HỆ BẰNG ACCESS

Microsoft Access là một hệ quản trị cơ sở dữ liệu quan hệ. Ta có thể dùng HQTCSDL Access hay gọi tắt là Access để thực hành mô hình quan hệ.

Sau đây ta sẽ tạo lược đồ csdl quản lý sinh viên như ở trang 6 của tài liệu.

- + Khởi động Access: **Start-> Programs-> Microsoft Access.**
- + Tạo lược đồ csdl rỗng có tên là qLSV: **Blank Database->OK->qLSV->Create**
- + Tạo quan hệ bằng cách: **Tables->New->Design View->OK**



- Data type là Text (kiểu chuỗi), Yes/No (Kiểu luận lý), Date/Time (kiểu ngày tháng), Currency (kiểu số)
- Field size là kích thước kiểu dữ liệu
- + Khi tạo xong các thuộc tính và khóa chính ta lưu lại và đặt tên cho quan hệ Sv bằng cách : **File->Close->Yes->Sv->OK.**
- + Tạo các quan hệ còn lại theo bước 3 và 4
- + Nhập dữ liệu cho quan hệ Sv bằng cách: **Tables->Sv->Open**

II CÂU LỆNH TRUY VẤN

1 BIỂU THỨC (EXPRESSION)

Các thành phần tạo nên biểu thức bao gồm:

Literal value

Là các dữ liệu có giá trị đúng như văn bản thể hiện.

Dữ liệu chuỗi có dạng: "New York"

Dữ liệu số có dạng: 1056; 1056.25

Dữ liệu ngày có dạng: #1-Jan-94#; #12/2/2001#

Constant

Là một tên đại diện cho một giá trị không thay đổi như :

Tạo khóa chính:
nhấp chuột vào lề trái
thuộc tính rồi nhấp chuột

Const	Giá trị đại diện
True	-1
False	0
Null	0

Toán tử số học:

Toán tử	Ý nghĩa	Ví dụ	Kết quả
+	Cộng số học Cộng ngày	5+2 #28/08/01# + 4	7 #01/09/01#
-	Trừ số học Ngày trừ số Ngày trừ ngày	5-2 #02/09/01# - 3 #29/9/01#-#24/3/84#	3 #30/08/01# 16.597
*	Phép nhân	5*2	10
/	Phép chia	5/2	2.5
\	Chia nguyên	5\2	2
^	Luỹ thừa	5^2	25
Mod	Lấy số dư của phép chia	5 Mod 2	1

Toán tử luận lý

Toán tử	Ý nghĩa	Ví dụ	Kết Quả
Not	Luật phủ định	Not True Not False	False True
And	Luật và	True And True True And False False And True False And False	True False False False
Or	Luật hay	True Or True True Or False False Or True False Or False	True True True False

Toán tử so sánh

Toán tử	Ý nghĩa	Ví dụ	Kết quả
<	Nhỏ hơn	2 < 5	True
<=	Nhỏ hơn hay bằng	2 <= 5	True
>	Lớn hơn	2 > 5	False
>=	Lớn hơn hay bằng	2 >= 5	False
=	Bằng nhau	2 = 5	False
<>	Khác nhau	2 <> 5	True

Các toán tử khác

Toán tử	Ý nghĩa	Ví dụ	Kết quả
Between ... And ...	Giữa hai giá trị. Dùng trong query	2 Between 1 And 5	True
Like	Giống như	“Hung” Like “Hu*”	True
&	Nối chuỗi	“Nguyễn Văn” & “Hùng”	“Nguyễn Văn Hùng”

Với toán tử like ta có thể dùng các ký tự đại diện sau:

Wildcard characters

Ký tự	Ý nghĩa	Ví dụ
?	Đại diện cho một ký tự bất kỳ	A? -> AN, AM, AC
*	Đại diện cho một chuỗi ký tự bất kỳ.	A* -> AI, ANH,
#	Đại diện cho một ký số	A## ->A13, A24, A35
[]	Đại diện cho các ký tự nằm trong	A[IN] -> AI, AN
-	Đại diện cho các ký tự nằm trong khoảng	A[M-O] -> AM, AN, AO
!	Đại diện cho ký tự không phải ký tự nằm sau !	A[!C] -> AA, AB, AD, AE, ...

Hàm:

Hàm có dạng **tenHam(danhSachDoiSo)**. Hàm luôn luôn đại diện cho một trị gọi là trị trả về.

IIf(điều kiện, trị 1, trị 2)

Kiểm tra điều kiện, nếu điều kiện đúng trả trị 1 ngược lại trả trị 2

Ví dụ: IIf(namNu = 1, “Nam”, ”Nu”)

Date()

Trả về ngày tháng năm của hệ thống.

Now(biểu thức ngày)

Trả về giờ, phút, giây, ngày tháng năm của hệ thống.

Time(biểu thức ngày)

Trả về giờ phút giây của hệ thống.

Day(biểu thức ngày)

Trả về một số từ 1 đến 31 là ngày của Date.

Month(biểu thức ngày)

Trả về một số từ 1 đến 12 là tháng của Date

Year(biểu thức ngày)

Trả về năm của ngày

Len(biểu thức chuỗi)

Trả về chiều dài của chuỗi.

Chr(mã Ascii)

Trả về ký tự có mã ASCII tương ứng.

InStr(Start, s1, s2)

Trả về vị trí chuỗi s2 nằm trong s1

LCase(s), UCase(s)

Đổi chuỗi s thành chuỗi gồm các ký tự thường (hoa)

Left(s, n), Right(s, n)

Trả về chuỗi gồm n ký tự bên trái (phải) của chuỗi s

Mid(s, i, n)

Trả về chuỗi con của chuỗi s, gồm n ký tự kể từ ký tự thứ i

Nz(v1, v2)

Nếu v1 = Null thì Trả về v2, ngược lại trả về v1

Các hàm tính toán trên nhóm:

SUM (thuộc tính)

Tính tổng giá trị của thuộc tính của các bộ trong bảng

MAX(thuộc tính)

tính giá trị lớn nhất của thuộc tính của các bộ trong bảng

MIN(thuộc tính)

tính giá trị nhỏ nhất của thuộc tính của các bộ trong bảng

AVG(thuộc tính>)

tính giá trị trung bình của thuộc tính của các bộ trong bảng

COUNT(thuộc tính)

chỉ đếm những bộ mà giá trị của thuộc tính là khác NULL

Biểu thức

Biểu thức là tổ hợp các toán tử, literal value, hằng, tên hàm, tên thuộc tính.

Biểu thức được lượng giá thành một giá trị.

2 CÂU LỆNH SQL

SQL là ngôn ngữ truy vấn dựa trên đại số quan hệ. Câu lệnh của SQL dùng để rút trích dữ liệu của một một hay nhiều quan hệ. Kết quả của một câu lệnh SQL (truy vấn) là một quan hệ. Để đơn giản trong cách trình bày, ta xem quan hệ mà câu truy vấn sử dụng để tạo ra quan hệ khác gọi là quan hệ nguồn, quan hệ kết quả của truy vấn là quan hệ đích.

i Truy vấn định nghĩa dữ liệu (data definition query)

Tạo lược đồ quan hệ rSV:

```
+ Create Table rSV (MASV Text (10) CONSTRAINT khoaChinh PRIMARY KEY, HOTEN Text (30), NU YesNo, NGAYSINH Date, MALOP Text (10), TINH Text (50), HOCBONG Double )
```

Tạo lược đồ quan hệ rKQ:

```
+ Create Table rKQ (MASV text (10), MAMH Text (10), DIEMTHI Double, CONSTRAINT khoaChinh PRIMARY KEY (MASV, MAMH) )
```

Chú ý: Hai cách tạo trên cho ta thấy hai cách tạo khóa: khóa chỉ gồm một thuộc tính và khóa có nhiều thuộc tính

Thêm cột DTHOAI có kiểu dữ liệu text 20 ký tự vào lược đồ quan hệ rSV:

```
+ Alter Table rSV Add Column DTHOAI Text (20)
```

Sửa kiểu kiểu dữ liệu của một cột:

```
+ Alter Table rSV Alter Column DTHOAI double
```

Xóa cột của một lược đồ đã có:

```
+ Alter Table rSV Drop Column DTHOAI
```

Xóa ràng buộc khóa chính:

```
+ Alter Table rSV Drop Constraint khoaChinh
```

Thêm ràng buộc khóa chính:

```
+ Alter Table rSV Add Constraint khoaChinh Primary Key (MASV)
```

Thêm ràng buộc miền giá trị:

```
+ ALTER TABLE doc_exd WITH NOCHECK ADD CONSTRAINT exd_check CHECK (column_a > 1)
```

ii Truy vấn chọn (select query)

Khi có nhu cầu thể hiện các dòng dữ liệu của một quan hệ hay của nhiều quan hệ dưới dạng một quan hệ có số cột và số dòng theo ý muốn như bảng điểm của sinh viên, danh sách sinh viên thì ta sử dụng truy vấn chọn.

Để truy vấn chọn ta sử dụng câu lệnh SQL sau:

```
Select [Distinct|Top n[%]] field1 [As alias1][,field2 [As alias2][,...]]
From table1 [Inner Join table2 On table1.field1 θ table2.field2] ...
[Where dieuKien]
```

```
[Order By field1[Asc|Desc][,field2[Asc|Desc]][,...]];
```

Distinct: loại bỏ các bộ trùng trong quan hệ đích

Top n[%]: Chọn n hay n% mẫu tin đầu tiên.

table: Tên table hay query chứa dữ liệu.

field: Tên field hay một biểu thức.

Alias: Trường hợp field là một biểu thức thì <Alias> là một tên mới của biểu thức.

Inner Join: mỗi mẫu tin của table1 nối với bất kỳ mẫu tin nào của table2 có dữ liệu của field1 thỏa mãn điều kiện so sánh với dữ liệu của field2 tạo thành mẫu tin của query.

điều kiện: Biểu thức mà dữ liệu mẫu tin phải thỏa mãn

Chú ý:

Khi nêu rõ thuộc tính đó thuộc về quan hệ nào ta viết theo cú pháp *tênQuanHệ.tênThuộcTính*.

Ví dụ: Lập danh sách sinh viên gồm MASV,HOTEN,HOCBONG

```
SELECT MASV,HOTEN,HOCBONG
FROM Sv;
```

Ví dụ: Lập danh sách sinh viên nữ, có học bổng thuộc có mã lớp là “CĐTH2B”.

```
SELECT MASV,HOTEN,HOCBONG
FROM Sv
```

WHERE NU And MALOP='CĐTH2B' And HOCBONG>0;

Ví dụ: Lập danh sách sinh viên gồm MASV,HOTEN,HOCBONG trong đó học bổng được sắp giảm dần

```
SELECT MASV,HOTEN,HOCBONG
FROM Sv
ORDER BY HOCBONG DESC;
```

Ví dụ: Lập danh sách sinh viên gồm MASV,HOTEN,TENLOP

```
SELECT MASV,HOTEN,TENLOP
FROM Sv Inner Join Lop On Sv.MALOP=Kh.LOP;
```

Ví dụ: Lập danh sách sinh viên có MASV,HOTEN,NU,NGAYSINH,DIEMTHI với điểm thi môn ‘CSDL’ >= 8

```
SELECT Kq.MASV,HOTEN,NU,NGAYSINH,DIEMTHI
FROM Kq Inner Sv On Kq.MASV = Sv.MaSV
WHERE MAMH='CSDL' AND DIEMTHI>=8;
```

Ví dụ: Lập danh sách sinh viên có họ là “NGUYEN”

```
SELECT *
FROM Sv
WHERE HOTEN Like "NGUYEN*";
```

iii Truy vấn nhóm dữ liệu (Select query có group by)

Khi có nhu cầu chia các bộ theo từng nhóm rồi tính toán trên từng nhóm này như tính điểm trung bình của tất cả các môn học của từng sinh viên, hay là cần tính số lượng sinh viên của mỗi lớp, hay là cần biết tổng số môn mà một sinh viên đã đăng ký học thì ta sử dụng truy vấn theo nhóm.

Để truy vấn cộng nhóm ta sử dụng câu lệnh SQL sau:

```
Select [Distinct|Top n[%]] field1 [As alias1][,field2 [As alias2]][,...]
```

```

From table1 [Inner Join table2 On table1.field1  $\theta$  table2.field2] ...
[Where dieuKienLocMauTinNgon]
[Group By fieldGroupBy[,fieldGroupBy[,...]]]
[Having dieuKienLocMauTinTongHop]
[Order By field1[Asc|Desc][,field2[Asc|Desc]][,...]]];

```

điềuKiệnLocMẫuTinNgon: điều kiện mà các mẫu tin nguồn phải thỏa mãn (phép chọn)

fieldGroupBy: tên field mà các mẫu tin có dữ liệu giống nhau trên ấy được xếp vào cùng nhóm.

điềuKiệnLocMẫuTinTổngHợp: điều kiện mà các mẫu tin tổng hợp phải thỏa mãn (phép chọn)

Ví dụ: Lập danh sách sinh viên có đăng ký ít nhất là ba môn học

```

SELECT Kq.MASV,HOTEN,NU,NGAYSINH,COUNT(MAMH) As SLMH
FROM Sv Inner Join Kq On Sv.MASV = Kq.MASV
GROUP BY Kq.MASV,HOTEN,NU,NGAYSINH
HAVING COUNT(MAMH)>=3;

```

Ví dụ: Đếm số lượng sinh viên nữ của mỗi khoa

```

SELECT Kh.MAKHOA,TENKHOA,COUNT(Kh.MAKHOA) AS SOLUONG
FROM (Sv Inner Join Lop On Sv.MALOP = Lop.MALOP)
INNER JOIN Kh On Lop.MAKHOA=Kh.MAKHOA
WHERE NU
GROUP BY Kh.MAKHOA,TENKHOA;

```

iv Truy vấn lồng nhau (nested query, sub query)

Là những câu lệnh truy vấn mà trong thành phần WHERE hay HAVING có chứa thêm một câu lệnh Select khác. Câu lệnh select khác này gọi là subquery. Ta lồng câu Select vào phần Where hay Having theo cú pháp sau:

- *bieuthuc toanTuSoSanh* [ANY | ALL | SOME] (cauLenhSQL)

ANY, SOME là bất kỳ, ALL là tất cả

Các mẫu tin của query chính thỏa mãn toán tử so sánh với bất kỳ/ tất cả mẫu tin nào của subquery

- *bieuThuc* [NOT] IN (cauLenhSQL)

Các mẫu tin của query chính có giá trị bằng với một giá trị trong subquery

- [NOT] EXISTS (cauLenhSQL).

Các mẫu tin của query chính thỏa mãn khi subquery có mẫu tin

Ví dụ: Lập danh sách sinh viên có học bổng cao nhất

```

SELECT *
FROM Sv
WHERE HOCBONG>=ALL(SELECT HOCBONG FROM SV);

```

Ví dụ: Lập danh sách sinh viên có điểm thi môn CSDL cao nhất

```

SELECT SV.MASV,HOTEN,NU,NGAYSINH,DIEMTHI
FROM sv Inner Join kq On Sv.MASV = Kq.MASV
WHERE MAMH='CSDL' AND DIEMTHI >= ALL

```

(SELECT DIEMTHI FROM KQ WHERE MAMH='CSDL') ;

Hiểu và vận dụng tốt lệnh truy vấn dữ liệu là một việc làm cực kỳ cần thiết để tạo ra các kết quả cho báo cáo, thống kê số liệu.

v Truy vấn cập nhật dữ liệu (action query, data modification query)

Cú pháp: Update table Set field1 = biểuThức1, field2 = biểuThức2 Where điềuKiện

Cú pháp: Delete From table Where điềuKiện

vi Truy vấn hợp (union query)

Khi có nhu cầu thực hiện truy vấn có kết quả như toán tử hợp, ta sử dụng câu lệnh SQL sau:

```
Select .....Union Select .....
```

Ví dụ: Lệnh sau đây gộp đôi danh sách sinh viên

SELECT MASV,HOTEN,NGAYSINH UNION SELECT MASV,HOTEN,NGAYSINH

III BÀI TẬP

1/ Cho lược đồ CSDL quản lý sinh viên. Hãy thực hiện các câu truy vấn sau

- Lập danh sách những sinh viên nam của tỉnh “LONG AN” học khoa “CNTT”, danh sách cần tất cả các thuộc tính của quan hệ Sv.
- Lập danh sách những sinh viên có điểm thi < 5 (thi lại), danh sách cần MASV,HOTEN,TENMH, DIEMTHI và được sắp tăng dần theo cột MASV.
- Lập danh sách các sinh viên có điểm thi trung bình các môn < 5, danh sách cần MASV,HOTEN, DIEMTRUNGBINH và được sắp tăng dần theo cột MASV.
- Tổng số tiền học bổng của mỗi khoa
- Những sinh viên nào đăng ký học nhiều hơn 3 môn học, danh sách cần MASV,HOTEN,SOLAN_DANGKY
- Lập danh sách sinh viên có điểm trung bình cao nhất, danh sách cần MASV, HOTEN, NGAYSINH, DIEMTRUNGBINH

2/ Cho lược đồ CSDL dùng để quản lý lao động bao gồm các lược đồ quan hệ sau:

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

Tên từ: Mỗi nhân viên có một mã số nhân viên (MANV) duy nhất. Một mã số nhân viên xác định các thông tin như họ tên (HOTEN), ngày sinh (NGAYSINH), phái (PHAI), địa chỉ (DIACHI) và phòng ban (MAPB) nơi quản lý nhân viên.

Phongban(MAPB,TENPB)

Tên từ: Mỗi phòng ban có một mã phòng ban (MAPB) duy nhất, mã phòng ban xác định tên phòng ban (TENPB)

Cong(MACT,MANV,SLNGAYCONG)

Tên từ: Lược đồ quan hệ Cong ghi nhận số lượng ngày công (SLNGAYCONG) của một nhân viên (MANV) tham gia vào công trình (MACT).

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

Tên từ: Mỗi công trình có một mã số công trình (MACT) duy nhất. Mã số công trình xác định các thông tin như tên gọi công trình (TENCT), địa điểm (DIADIEM), ngày công trình được cấp giấy phép xây dựng (NGAYCAPGP), ngày khởi công (NGAYKC), ngày hoàn thành (NGAYHT).

Hãy thực hiện các câu hỏi sau bằng SQL

- Danh sách những nhân viên có tham gia vào công trình có mã công trình (MACT) là X. Yêu cầu các thông tin: MANV, HOTEN, SLNGAYCONG, trong đó MANV được sắp tăng dần.
- Đếm số lượng ngày công của mỗi công trình. Yêu cầu các thông tin: MACT, TENCT, TONGNGAYCONG (TONGNGAYCONG là thuộc tính tự đặt)
- Danh sách những nhân viên có sinh nhật trong tháng 8. yêu cầu các thông tin: MANV, TENNV, NGAYSINH, DIACHI, TENPB, sắp xếp quan hệ kết quả theo thứ tự tuổi giảm dần.
- Đếm số lượng nhân viên của mỗi phòng ban. Yêu cầu các thông tin: MAPB, TENPB, SOLUONG. (SOLUONG là thuộc tính tự đặt.)

3/ Cho các quan hệ sau:

Monhoc (MSMH, TENMH, SOTINCHI, TINHCHAT)

MSMH mã số môn học,

TENMH tên môn học

SOTINCHI số lượng tín chỉ,

TÍNH CHẤT bằng 1 nếu đó là môn học bắt buộc, bằng 0 nếu đó là môn học không bắt buộc

Sinhvien (MSSV, HOTEN, NGAYSINH, LOP)

MSSV mã số sinh viên,

HOTEN họ tên sinh viên

NGAYSINH ngày sinh,

LOP(C,4,0) lớp

Diem (MSSV, MSMH, DIEMTHI)

DIEMTHI điểm thi

Hãy dùng lệnh SQL để thực hiện các câu lệnh sau:

- Hãy cho biết những môn học bắt buộc có SOTINCHI cao nhất.
- Hãy liệt kê danh sách gồm MSSV, HOTEN, LOP, DIEMTHI của những sinh viên thi môn học CSDL, theo thứ tự LOP, DIEMTHI
- Hãy cho biết các sinh viên có điểm thi cao nhất về môn học có mã là CSDL
- Hãy cho biết phiếu điểm của sinh viên có mã số là 9900277
- Hãy liệt kê danh sách gồm MSSV, HOTEN., LOP, ĐIỂM TRUNG BÌNH của những sinh viên có điểm trung bình các môn dưới 5, theo thứ tự LOP, HOTEN.
- Hãy liệt kê danh sách điểm trung bình của sinh viên theo thứ tự, lớp, tên.
- Hãy cho biết điểm của sinh viên theo từng môn.

4/ Dựa vào lược đồ cơ sở dữ liệu

Docgia (MADG, HOTEN, NGAYSINH, DIACHI, NGHENGHIEP)

Phieumuon (SOPM, NGAYMUON, MADG)

Chitietmuon (SOPM, MADAUSACH, NGAYTRA)

Dausach (MADAUSACH, BAN, TAP, MASH)

Sach (MASH, TENSACH, TACGIA, NHAXB, NAMXB)

Hãy thực hiện các câu hỏi sau đây bằng SQL

- Danh sách các đọc giả đã đăng ký mượn sách trong ngày d. Yêu cầu các thông tin: MADG, HOTEN, DIACHI.

- b) Các quyển sách của phiếu mượn có SOPM là x. Yêu cầu các thông tin MASH, TENSACH, TACGIA, NGAYMUON, NGAYTRA.
- c) Tổng số lượt mà mỗi đọc giả đến mượn sách trong năm 2001. Yêu cầu thông tin MAĐG, HOTEN, SOLANMUON (SOLANMUON là thuộc tính tự đặt)
- d) Danh sách các đọc giả cao tuổi nhất đã mượn sách trong ngày d. Yêu cầu các thông tin MAĐG, HOTEN, NGAYSINH, ĐIACHI, NGHENGHIEP.

5/ Dựa vào lược đồ cơ sở dữ liệu

Khach (MAKH, HOTEN, DIACHI, DIENTHOAI)
 Hoadon (SOHD, NGAYLAPHD, NGAYBAN, MAKH)
 DongHoaDon (SOHD, MAHANG, SLBAN)
 Hang (MAHANG, TENHANG, DONGIA, DVT, MANHOM)
 Nhom (MANHOM, TENNHOM)

Hãy thực hiện các câu hỏi sau bằng SQL

- a) Danh sách các khách hàng đã mua hàng trong ngày d. Yêu cầu các thông tin MAKH, HOTEN, ĐIACHI, ĐIENTHOAI.
- b) Danh sách các mặt hàng trong sổ hóa đơn (SOHD) là x. Yêu cầu các thông tin MAHANG, TENHANG, SLBAN, ĐONGIA, THANHTIEN (THANHTIEN= SLBAN*ĐONGIA; THANHTIEN là thuộc tính tự đặt). Yêu cầu sắp xếp tăng dần theo cột TENHANG
- c) Danh sách các mặt hàng thuộc mã nhóm hàng là A có đơn giá cao nhất. Yêu cầu các thông tin : MAHANG, TENHANG, ĐONGIA
- d) Đếm số lượng mặt hàng của mỗi nhóm hàng. Yêu cầu các thông tin : MANHOM, TENNHOM, SOLUONG. (trong đó SOLUONG là thuộc tính tự đặt) (0,75đ)
- e) Danh sách các khách hàng đã mua các mặt hàng có mã nhóm hàng là A trong ngày d. Yêu cầu các thông tin MAKH, HOTEN, ĐIACHI, ĐIENTHOAI, TENHANG.
- f) Thống kê việc mua hàng trong năm 2002 của khách hàng có mã khách hàng là Kh01 (theo từng hóa đơn). Yêu cầu các thông tin MAKH, HOTEN, SOHD, TRIGIAHD trong đó TRIGIAHD là tổng số tiền trong một hóa đơn (TRIGIAHD là thuộc tính tự đặt)

6/ Dựa vào lược đồ cơ sở dữ liệu

Giaovien (MAGV, HOTEN, DTGV, MAKHOA)
 Khoa (MAKHOA, TENKHOA, DTKHOA)
 Lop (MALOP, TENLOP, SISO, MAKHOA)
 Monhoc (MAMH, TENMH)
 Phonghoc (SOPHONG, CHUCNANG)
 Lichbaogiang (MALICH, NGAYDAY, MAGV)
 Dongbaogiang (MALICH, TUTIET, DENTIET, BAIDAY, GHICHU, LYTHUYET, MAMH, MALOP, SOPHONG)

Hãy thực hiện các câu hỏi sau bằng SQL

- a) Xem lịch báo giảng tuần từ ngày 16/09/2002 đến ngày 23/09/2002 của giáo viên có MAGV (mã giáo viên) là TH3A040. Yêu cầu: MAGV, HOTEN, TENLOP, TENMH, SOPHONG, NGAYDAY, TUTIET, DENTIET, BAIDAY, GHICHU

- b) Xem lịch báo giảng ngày 23/09/2002 của các giáo viên có mã khoa là CNTT. Yêu cầu: MAGV ,HOTEN ,TENLOP ,TENMH ,PHONG ,NGAYDAY ,TUTIET ,DENTIET ,BAIDAY ,GHICHU)
- c) Cho biết số lượng giáo viên (SOLUONGGV) của mỗi khoa, kết quả cần sắp xếp tăng dần theo cột tên khoa. yêu cầu: TENKHOA ,SOLUONGGV (SOLUONGGV là thuộc tính tự đặt)

----oOo----

Chương 3 .

RÀNG BUỘC TOÀN VỆN QUAN HỆ (*integrity constraint*)

I RÀNG BUỘC TOÀN VỆN - CÁC YẾU TỐ CỦA RÀNG BUỘC TOÀN VỆN

1 Ràng Buộc Toàn Vện

Trong mỗi CSDL luôn tồn tại nhiều mối liên hệ giữa các thuộc tính, giữa các bộ. Sự liên hệ này có thể xảy ra trong một lược đồ quan hệ hoặc trong các lược đồ quan hệ của một cơ sở dữ liệu. Các mối liên hệ này là những điều kiện bất biến mà tất cả các bộ của những quan hệ có liên quan trong CSDL đều phải thỏa mãn ở mọi thời điểm. Những điều kiện bất biến đó được gọi là ràng buộc toàn vẹn. Trong thực tế ràng buộc toàn vẹn là các quy tắc quản lý được áp đặt trên các đối tượng của thế giới thực.

Nhiệm vụ của người phân tích thiết kế là phải phát hiện càng đầy đủ và chính xác các ràng buộc toàn vẹn càng tốt và mô tả chúng một cách chính xác trong hồ sơ phân tích thiết kế - đó là một việc làm rất quan trọng và rất cần thiết.

Trong một cơ sở dữ liệu, ràng buộc toàn vẹn được xem như là một công cụ để diễn đạt ngữ nghĩa của CSDL. Một CSDL được thiết kế công kênh nhưng nó thể hiện được đầy đủ ngữ nghĩa của thực tế vẫn có giá trị cao hơn rất nhiều so với một cách thiết kế gọn nhẹ nhưng nghèo nàn về ngữ nghĩa vì thiếu các ràng buộc toàn vẹn của cơ sở dữ liệu.

Công việc kiểm tra ràng buộc toàn vẹn thường được tiến hành vào thời điểm cập nhật dữ liệu (thêm, sửa, xóa). Những ràng buộc toàn vẹn phát sinh cần phải được ghi nhận và xử lý một cách tường minh (thường là bởi một hàm chuẩn hoặc một đoạn chương trình).

2 Các Yếu Tố Của Ràng Buộc Toàn Vện

Mỗi ràng buộc toàn vẹn có 3 yếu tố: điều kiện, bối cảnh và tầm ảnh hưởng.

i Điều kiện

Điều kiện của một ràng buộc toàn vẹn R có thể được biểu diễn bằng ngôn ngữ tự nhiên, thuật giải, ngôn ngữ đại số tập hợp, đại số quan hệ,... ngoài ra điều kiện của ràng buộc toàn vẹn cũng có thể được biểu diễn bằng phụ thuộc hàm. Chẳng hạn, với lược đồ quan hệ SV thì có một ràng buộc toàn vẹn như sau:

Với r là một quan hệ của Sv ta có ràng buộc toàn vẹn sau

$$\forall t_1, t_2 \in r \\ t_1.MASV \neq t_2.MASV \\ \text{cuối } \forall$$

ii Bối cảnh

Bối cảnh của một ràng buộc toàn vẹn là những quan hệ mà ràng buộc đó có hiệu lực hay nói một cách khác, đó là những quan hệ cần phải được kiểm tra ràng buộc toàn vẹn. Bối cảnh của một ràng buộc toàn vẹn có thể là một hoặc nhiều quan hệ. Chẳng hạn với ràng buộc toàn vẹn trên thì bối cảnh là một quan hệ Sv

iii Tầm ảnh hưởng

Trong quá trình phân tích thiết kế một CSDL, người phân tích cần lập bảng tầm ảnh hưởng cho một ràng buộc toàn vẹn nhằm xác định thời điểm cần phải tiến hành kiểm tra các ràng buộc toàn vẹn đó. Các thời điểm cần phải kiểm tra RBTV chính là những thời điểm cập nhật dữ liệu (thêm/sửa/xóa)

Một bảng tầm ảnh hưởng của một RBTV có dạng sau:

(Tên RBTV)	Thêm(T)	Sửa(S)	Xóa(X)
r_1	+	-	-
r_2			
...
...
r_n			

Bảng này chứa toàn các ký hiệu + hoặc -

Chẳng hạn + tại ô tương ứng với dòng r_1 , cột thêm thì có nghĩa là khi thêm một bộ vào quan hệ r_1 thì cần phải kiểm tra RBTV

Dấu - Tại ô tương ứng với dòng r_1 , cột sửa thì có nghĩa là khi sửa một bộ trên quan hệ r_1 thì không cần phải kiểm tra RBTV này,...

II PHÂN LOẠI RÀNG BUỘC TOÀN VẸN

Trong quá trình phân tích thiết kế cơ sở dữ liệu, người phân tích phải phát hiện tất cả các ràng buộc toàn vẹn tiềm ẩn trong CSDL đó. Việc phân loại các ràng buộc toàn vẹn là rất có ích, nó nhằm giúp cho người phân tích có được một định hướng, tránh bỏ sót những ràng buộc toàn vẹn. Các ràng buộc toàn vẹn có thể được chia làm hai loại chính như sau:

- + Ràng buộc toàn vẹn trên phạm vi là một quan hệ bao gồm :Ràng buộc toàn vẹn miền giá trị, ràng buộc toàn vẹn liên thuộc tính, ràng buộc toàn vẹn liên bộ.
- + Ràng buộc toàn vẹn trên phạm vi nhiều quan hệ bao gồm :Ràng buộc toàn vẹn phụ thuộc tồn tại, ràng buộc toàn vẹn liên bộ - liên quan hệ, ràng buộc toàn vẹn liên thuộc tính - liên quan hệ.

Để minh họa cho phần lý thuyết của chương này, ta nêu ví dụ sau đây

Ví dụ

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

Q₁: Khách (MAKH, TENKH, DCKH, DT)

Tên từ: Mỗi khách hàng có một mã khách hàng (MAKH) duy nhất, mỗi MAKH xác định một tên khách hàng (TENKH), một địa chỉ (DCKH), một số điện thoại (DT).

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

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

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

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

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

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

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

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

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

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

1 Ràng buộc toàn vẹn liên bộ

Ràng buộc toàn vẹn liên bộ là sự ràng buộc toàn vẹn giữa các bộ trong cùng một quan hệ .

Ràng buộc toàn vẹn liên bộ hay còn gọi là ràng buộc toàn vẹn về khóa. Đây là loại ràng buộc toàn vẹn rất phổ biến, nó có mặt trong mọi lược đồ quan hệ của CSDL và thường được các hệ quản trị CSDL tự động kiểm tra.

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

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

$t_1.MAKH \neq t_2.MAKH$

Cuối \forall

R_1	Thêm	Sửa	Xóa
r	+	+	-

2 Ràng buộc toàn vẹn về phụ thuộc tồn tại:

Ràng buộc toàn vẹn về phụ thuộc tồn tại còn được gọi là ràng buộc toàn vẹn về khóa ngoại. Cũng giống như ràng buộc toàn vẹn về khóa chính, ràng buộc toàn vẹn về phụ thuộc tồn tại rất phổ biến trong CSDL

Ví dụ: Với r, s lần lượt là một quan hệ của Dathang, Khách ta có ràng buộc toàn vẹn sau

$R_2: r[MAKH] \subseteq s[MAKH]$

R_2	Thêm	Sửa	Xóa
r	+	+	-
s	-	+	+

3 Ràng buộc toàn vẹn về miền giá trị

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

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

$R_3: \forall t \in r$

$t.TRIGIAHD > 0$

Cuối \forall

R_3	Thêm	Sửa	Xóa
r	+	+	-

4 Ràng buộc toàn vẹn liên thuộc tính

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

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

$R_4: \forall t \in r$

$t.NGAYLAP \leq t.NGAYXUAT$

Cuối \forall

R_4	Thêm	Sửa	Xóa
r	+	+	-

5 Ràng buộc toàn vẹn liên thuộc tính liên quan hệ

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

Ví dụ: Với r, s lần lượt là quan hệ của Dathang, Hoadon ta có ràng buộc toàn vẹn sau

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

Nếu $t_1.SODH = t_2.SODH$ thì

$t_1.NGAYDH \leq t_2.NGAYXUAT$

Cuối \forall

R_5	Thêm	Sửa	Xóa
r	+	+	+
s	+	+	+

6 Ràng buộc toàn vẹn về thuộc tính tổng hợp

Ràng buộc toàn vẹn về thuộc tính tổng hợp được xác định trong trường hợp mỗi thuộc tính A của một lược đồ quan hệ Q được tính toán giá trị từ các thuộc tính của các lược đồ quan hệ khác.

III BÀI TẬP

1/ Hãy tìm các ràng buộc toàn vẹn có trong CSDL cho các bài tập được liệt kê trong chương 3.

2/ QUẢN LÝ THI TỐT NGHIỆP PTCS

Một phòng giáo dục huyện muốn lập một hệ thống thông tin để quản lý việc làm thi tốt nghiệp phổ thông cơ sở. Công việc làm thi được tổ chức như sau:

Lãnh đạo phòng giáo dục thành lập nhiều hội đồng thi (mỗi hội đồng thi gồm một trường hoặc một số trường gần nhau). Mỗi hội đồng thi có một mã số duy nhất (MAHĐT), một mã số hội đồng thi xác định tên hội đồng thi(TENHĐT), họ tên chủ tịch hội đồng(TENCT), địa chỉ (ĐCHĐT),điện thoại(ĐTHĐT).

Mỗi hội đồng thi được bố trí cho một số phòng thi, mỗi phòng thi có một số hiệu phòng(SOPT) duy nhất, một phòng thi xác định địa chỉ phòng thi (ĐCPT). Số hiệu phòng thi được đánh số khác nhau ở tất cả các hội đồng thi.

Giáo viên của các trường trực thuộc phòng được điều động đến các hội đồng để coi thi, mỗi trường có thể có hoặc không có thí sinh dự thi, mỗi trường có một mã trường duy nhất (MATR), mỗi mã trường xác định một tên trường(TENTR), địa chỉ (ĐCTR), loại hình đào tạo (LHĐT) (Công lập, chuyên, bán công, dân lập, nội trú,...). Giáo viên của một trường có thể làm việc tại nhiều hội đồng thi. Một giáo viên có một mã giáo viên(MAGV), một mã giáo viên xác định tên giáo viên (TENGV), chuyên môn giảng dạy (CHUYENMON), chức danh trong hội đồng thi(CHUCDANH)

Các thí sinh dự thi có một số báo danh duy nhất(SOBD), mỗi số báo danh xác định tên thí sinh(TENTS), ngày sinh (NGSINH), giới tính (PHAI), mỗi thí sinh được xếp thi tại một phòng thi nhất định cho tất cả các môn, mỗi thí sinh có thể có chứng chỉ nghề (CCNGHE) hoặc không (thuộc tính CCNGHE kiểu chuỗi, CCNGHE="x" nếu thí sinh có chứng chỉ nghề và CCNGHE bằng rỗng nếu thí sinh không có chứng chỉ nghề). Thí sinh của cùng một trường chỉ dự thi tại một hội đồng thi. Mỗi môn thi có một mã môn thi duy nhất(MAMT), mỗi mã môn thi xác định tên môn thi(TENMT). Giả sử toàn bộ các thí sinh đều thi chung một số môn do sở giáo dục quy định. Mỗi môn thi được tổ chức trong một buổi của một ngày nào đó.

Ứng với mỗi môn thi một thí sinh có một điểm thi duy nhất(ĐIEMTHI)

Dựa vào phân tích ở trên, giả sử ta có lược đồ CSDL sau:

- Q1: HÐ(MAHĐT,TENHĐT, TENCT, ĐCHĐT,ĐTHĐT)
- Q2: PT(SOPT,ĐCPT,MAHĐT)
- Q3: TS(SOBD, TENTS,NGSINH,PHAI,CCNGHE, MATR,SOPT)
- Q4: MT(MAMT,TENMT,BUOI,NGAY)
- Q5: GV(MAGV,TENGV,CHUYENMON,CHUCDANH,MAHĐT,MATR)
- Q6: TR(MATR,TENTR,ĐCTR,LHĐT)
- Q7: KQ(SOBD,MAMT,ĐIEMTHI)

Yêu cầu:

- a) Hãy xác định khóa cho từng lược đồ quan hệ.
- b) Tìm tất cả các ràng buộc toàn vẹn có trong CSDL trên.
- c) Dựa vào lược đồ CSDL đã thành lập, hãy thực hiện các câu hỏi sau đây bằng ngôn ngữ đại số quan hệ.
 1. Danh sách các thí sinh thi tại phòng thi có số hiệu phòng thi (SOPT) là "100". Yêu cầu các thông tin:SOBD,TENTS,NGSINH,TENTR
 2. Kết quả của môn thi có mã môn thi (MAMT) là "T" của tất cả các thí sinh có mã trường(MATR) là "NTMK", kết quả được sắp theo chiều giảm dần của điểm thi(ĐIEMTHI). Yêu cầu các thông tin:SOBD,TENTS, ĐIEMTHI
 3. Kết quả thi của một học sinh có SOBD là MK01. Yêu cầu : TENMT,ĐIEMTHI
 4. Tổng số thí sinh có chứng chỉ nghề(CCNGHE) của mỗi trường, thông tin cần được sắp theo chiều tăng dần của TENTR. Yêu cầu các thông tin: MATR, TENTR, SOLUONGCC

----oOo----

Chương 4 .

PHỤ THUỘC HÀM

(functional dependency)

Phụ thuộc hàm (functional dependency) là một công cụ dùng để biểu diễn một cách hình thức các ràng buộc toàn vẹn (vấn tắt: ràng buộc). Phương pháp biểu diễn này có rất nhiều ưu điểm, và đây là một công cụ cực kỳ quan trọng, gắn chặt với lý thuyết thiết kế cơ sở dữ liệu.

Phụ thuộc hàm được ứng dụng trong việc giải quyết các bài toán tìm khóa, tìm phủ tối thiểu và chuẩn hóa cơ sở dữ liệu.

I KHÁI NIỆM PHỤ THUỘC HÀM

Cho quan hệ **phanCong** sau:

phanCong (PHICONG, MAYBAY, NGÀYKH, GIOKH)

Cushing	83	9/8	10:15a
Cushing	116	10/8	1:25p
Clark	281	8/8	5:50a
Clark	301	12/8	6:35p
Clark	83	11/8	10:15a
Chin	83	13/8	10:15a
Chin	116	12/8	1:25p
Copely	281	9/8	5:50a
Copely	281	13/8	5:50a
Copely	412	15/8	1:25p

Quan hệ **phanCong** diễn tả phi công nào lái máy bay nào và máy bay khởi hành vào thời gian nào. Không phải sự phối hợp bất kỳ nào giữa phi công, máy bay và ngày giờ khởi hành cũng đều được chấp nhận mà chúng có các điều kiện ràng buộc qui định sau:

- + Mỗi máy bay có một giờ khởi hành duy nhất.
- + Nếu biết phi công, biết ngày giờ khởi hành thì biết được máy bay do phi công ấy lái.
- + Nếu biết máy bay, biết ngày khởi hành thì biết phi công lái chuyến bay ấy.

Các ràng buộc này là các ví dụ về phụ thuộc hàm và được phát biểu lại như sau:

- + MAYBAY xác định GIOKH
- + {PHICONG, NGÀYKH, GIOKH} xác định MABAY
- + {MAYBAY, NGÀYKH} xác định PHICONG

hay

- + GIOKH phụ thuộc hàm vào MAYBAY
- + MABAY phụ thuộc hàm vào {PHICONG, NGÀYKH, GIOKH}
- + PHICONG phụ thuộc hàm vào {MAYBAY, NGÀYKH}

và được ký hiệu như sau:

- + {MAYBAY} → GIOKH
- + {PHICONG, NGÀYKH, GIOKH} → MABAY
- + {MAYBAY, NGÀYKH} → PHICONG

Trong ký hiệu trên ta đã ký hiệu **MAYBAY** thay cho {MAYBAY}.

Một cách tổng quát:

1 Định nghĩa phụ thuộc hàm

$Q(A_1, A_2, \dots, A_n)$ là lược đồ quan hệ.

X, Y là hai tập con của $Q^+ = \{A_1, A_2, \dots, A_n\}$.

r là quan hệ trên Q .

t_1, t_2 là hai bộ bất kỳ của r .

$$X \rightarrow Y \Leftrightarrow (t_1.X = t_2.X \Rightarrow t_1.Y = t_2.Y)$$

(Ta nói X xác định Y hay Y phụ thuộc hàm vào X (X functional determines Y , Y functional dependent on X))

Tính chất:

- + phụ thuộc hàm $X \rightarrow \emptyset$ đúng với mọi quan hệ r
- + phụ thuộc hàm $\emptyset \rightarrow Y$ chỉ đúng trên quan hệ r có cùng giá trị trên Y .

Ví dụ: Quan hệ sau thỏa mãn phụ thuộc hàm $\emptyset \rightarrow \text{GIOKH}$

phanCong	(PHICONG,	MAYBAY,	NGAYKH,	GIOKH)
Cushing	83	9/8	10:15a	
Cushing	116	10/8	10:15a	
Clark	281	8/8	10:15a	
Clark	301	12/8	10:15a	
Clark	83	11/8	10:15a	
Chin	83	13/8	10:15a	
Chin	116	12/8	10:15a	
Copely	281	9/8	10:15a	
Copely	281	13/8	10:15a	
Copely	412	15/8	10:15a	

trên thực tế không có quan hệ r nào thỏa tính chất trên nên từ đây về sau nếu không nói rõ thì với một quan hệ r bất kỳ ta luôn xem phụ thuộc hàm $\emptyset \rightarrow Y$ luôn luôn không thỏa trên r .

2 Phụ thuộc hàm hiển nhiên (Trivial Dependencies)

Hệ quả: Nếu $X \supseteq Y$ thì $X \rightarrow Y$.

Chứng minh:

Giả sử $t_1.X = t_2.X$ do $X \supseteq Y$ nên $t_1.Y = t_2.Y$ theo định nghĩa suy ra $X \rightarrow Y$

Trong trường hợp này $X \rightarrow Y$ được gọi là **phụ thuộc hàm hiển nhiên**.

Ví dụ phụ thuộc hàm $X \rightarrow X$ là phụ thuộc hàm hiển nhiên.

Vậy với r là quan hệ bất kỳ, F là tập phụ thuộc hàm thỏa trên r , ta luôn có $F \supseteq \{\text{các phụ thuộc hàm hiển nhiên}\}$

3 Thuật toán Satifies

Cho quan hệ r và X, Y là hai tập con của Q^+ . Thuật toán SATIFIES sẽ trả về trị true nếu $X \rightarrow Y$ ngược lại là false

SATIFIES

Vào: quan hệ r và hai tập con X, Y

ra: true nếu $X \rightarrow Y$, ngược lại là false

SATIFIES(r, X, Y)

1. Sắp các bộ của quan hệ r theo X để các giá trị giống nhau trên X nhóm lại với nhau

2. Nếu tập các bộ cùng giá trị trên X cho các giá trị trên Y giống nhau thì trả về true ngược lại là False

Ví dụ 1: SATIFIES (phanCong, MAYBAY, GIOKH)

phanCong	(PHICONG, MAYBAY , NGAYKH, GIOKH)
Cushing	83 9/8 10:15a
Clark	83 11/8 10:15a
Chin	83 13/8 10:15a
Cushing	116 10/8 1:25p
Chin	116 12/8 1:25p
Clark	281 8/8 5:50a
Copely	281 9/8 5:50a
Copely	281 13/8 5:50a
Clark	301 12/8 6:35p
Copely	412 15/8 1:25p

cho kết quả là true nghĩa là MAYBAY \rightarrow GIOKH

Ví dụ 2: SATIFIES (phanCong, GIOKH, MAYBAY)

phanCong	(PHICONG, MAYBAY, NGAYKH, GIOKH)
Clark	281 8/8 5:50a
Copely	281 9/8 5:50a
Copely	281 13/8 5:50a
Cushing	83 9/8 10:15a
Clark	83 11/8 10:15a
Chin	83 13/8 10:15a
Cushing	116 10/8 1:25p
Chin	116 12/8 1:25p
Copely	412 15/8 1:25p
Clark	301 12/8 6:35p

cho kết quả là false nghĩa là không có phụ thuộc hàm GIOKH \rightarrow MAYBAY

4 Các phụ thuộc hàm có thể có

i Cách tìm tất cả tập con của Q^+

Lược đồ quan hệ Phancong (PHICONG, MAYBAY, NGAYKH, GIOKH) có tập thuộc tính $\text{Phancong}^+ = \{\text{PHICONG}, \text{MAYBAY}, \text{NGAYKH}, \text{GIOKH}\}$ và tất cả các tập con có thể có của Phancong^+ được cho bởi bảng sau:

	PHICONG	MAYBAY	NGAYKH	GIOKH
\emptyset	{PHICONG}	{MAYBAY}	{NGAYKH}	{GIOKH}
		{PHICONG, MAYBAY}	{PHICONG, NGAYKH}	{PHICONG, GIOKH}
			{MAYBAY, NGAYKH}	{MAYBAY, GIOKH}
			{PHICONG, MAYBAY, NGAYKH}	{PHICONG, MAYBAY, GIOKH}
				{NGAYKH, GIOKH}
				{PHICONG, NGAYKH, GIOKH}
				{MAYBAY, NGAYKH, GIOKH}
				{PHICONG, MAYBAY, NGAYKH, GIOKH}

Số tập con có thể có của $Q^+ = \{A_1, A_2, \dots, A_n\}$ là 2^n

ii Cách tìm tất cả các phụ thuộc hàm có thể có của Q

Ứng với mỗi tập con của Phancong^+ cho $2^n = 2^4 = 16$ phụ thuộc hàm có thể có. Số phụ thuộc hàm có thể có là $2^4 * 2^4 = 16 * 16 = 256$

$\emptyset \rightarrow \emptyset$	PTHHN
$\emptyset \rightarrow \{\text{PHICONG}\}$	F^-
$\emptyset \rightarrow \{\text{MAYBAY}\}$	F^-
$\emptyset \rightarrow \{\text{MAYBAY}, \text{PHICONG}\}$	F^-
$\emptyset \rightarrow \{\text{NGAYKH}\}$	F^-
$\emptyset \rightarrow \{\text{PHICONG}, \text{NGAYKH}\}$	F^-
$\emptyset \rightarrow \{\text{MAYBAY}, \text{NGAYKH}\}$	F^-
$\emptyset \rightarrow \{\text{MAYBAY}, \text{PHICONG}, \text{NGAYKH}\}$	F^-
$\emptyset \rightarrow \{\text{GIOKH}\}$	F^-
$\emptyset \rightarrow \{\text{PHICONG}, \text{GIOKH}\}$	F^-
$\emptyset \rightarrow \{\text{MAYBAY}, \text{GIOKH}\}$	F^-
$\emptyset \rightarrow \{\text{MAYBAY}, \text{PHICONG}, \text{GIOKH}\}$	F^-
$\emptyset \rightarrow \{\text{NGAYKH}, \text{GIOKH}\}$	F^-
$\emptyset \rightarrow \{\text{PHICONG}, \text{NGAYKH}, \text{GIOKH}\}$	F^-
$\emptyset \rightarrow \{\text{MAYBAY}, \text{NGAYKH}, \text{GIOKH}\}$	F^-
$\emptyset \rightarrow \{\text{MAYBAY}, \text{PHICONG}, \text{NGAYKH}, \text{GIOKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \emptyset$	PTHHN
$\{\text{PHICONG}\} \rightarrow \{\text{PHICONG}\}$	PTHHN
$\{\text{PHICONG}\} \rightarrow \{\text{MAYBAY}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{MAYBAY}, \text{PHICONG}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{NGAYKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{PHICONG}, \text{NGAYKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{MAYBAY}, \text{NGAYKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{MAYBAY}, \text{PHICONG}, \text{NGAYKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{GIOKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{PHICONG}, \text{GIOKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{MAYBAY}, \text{GIOKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{MAYBAY}, \text{PHICONG}, \text{GIOKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{NGAYKH}, \text{GIOKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{PHICONG}, \text{NGAYKH}, \text{GIOKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{MAYBAY}, \text{NGAYKH}, \text{GIOKH}\}$	F^-
$\{\text{PHICONG}\} \rightarrow \{\text{MAYBAY}, \text{PHICONG}, \text{NGAYKH}, \text{GIOKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \emptyset$	PTHHN
$\{\text{MAYBAY}\} \rightarrow \{\text{PHICONG}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{MAYBAY}\}$	PTHHN
$\{\text{MAYBAY}\} \rightarrow \{\text{MAYBAY}, \text{PHICONG}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{NGAYKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{PHICONG}, \text{NGAYKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{MAYBAY}, \text{NGAYKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{MAYBAY}, \text{PHICONG}, \text{NGAYKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{GIOKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{PHICONG}, \text{GIOKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{MAYBAY}, \text{GIOKH}\}$	F^+
$\{\text{MAYBAY}\} \rightarrow \{\text{MAYBAY}, \text{PHICONG}, \text{GIOKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{NGAYKH}, \text{GIOKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{PHICONG}, \text{NGAYKH}, \text{GIOKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{MAYBAY}, \text{NGAYKH}, \text{GIOKH}\}$	F^-
$\{\text{MAYBAY}\} \rightarrow \{\text{MAYBAY}, \text{PHICONG}, \text{NGAYKH}, \text{GIOKH}\}$	F^-
$\{\text{PHICONG}, \text{MAYBAY}\} \rightarrow \emptyset$	PTHHN
$\{\text{PHICONG}, \text{MAYBAY}\} \rightarrow \{\text{PHICONG}\}$	PTHHN
$\{\text{PHICONG}, \text{MAYBAY}\} \rightarrow \{\text{MAYBAY}\}$	PTHHN
$\{\text{PHICONG}, \text{MAYBAY}\} \rightarrow \{\text{PHICONG}, \text{MAYBAY}\}$	PTHHN
$\{\text{PHICONG}, \text{MAYBAY}\} \rightarrow \{\text{NGAYKH}\}$	F^-
$\{\text{PHICONG}, \text{MAYBAY}\} \rightarrow \{\text{PHICONG}, \text{NGAYKH}\}$	F^-
$\{\text{PHICONG}, \text{MAYBAY}\} \rightarrow \{\text{MAYBAY}, \text{NGAYKH}\}$	F^-

$\{PHICONG, MAYBAY\} \rightarrow \{MAYBAY, PHICONG, NGAYKH\}$	F^-
$\{PHICONG, MAYBAY\} \rightarrow \{GIOKH\}$	F^+
$\{PHICONG, MAYBAY\} \rightarrow \{PHICONG, GIOKH\}$	F^+
$\{PHICONG, MAYBAY\} \rightarrow \{MAYBAY, GIOKH\}$	F^+
$\{PHICONG, MAYBAY\} \rightarrow \{MAYBAY, PHICONG, GIOKH\}$	F^+
$\{PHICONG, MAYBAY\} \rightarrow \{NGAYKH, GIOKH\}$	F^-
$\{PHICONG, MAYBAY\} \rightarrow \{PHICONG, NGAYKH, GIOKH\}$	F^-
$\{PHICONG, MAYBAY\} \rightarrow \{MAYBAY, NGAYKH, GIOKH\}$	F^-
$\{PHICONG, MAYBAY\} \rightarrow \{MAYBAY, PHICONG, NGAYKH, GIOKH\}$	F^-
$\{NGAYKH\} \rightarrow \emptyset$	F^-
$\{NGAYKH\} \rightarrow \{PHICONG\}$	F^-
$\{NGAYKH\} \rightarrow \{MAYBAY\}$	F^-
$\{NGAYKH\} \rightarrow \{PHICONG, MAYBAY\}$	F^-
$\{NGAYKH\} \rightarrow \{NGAYKH\}$	PTHHN
$\{NGAYKH\} \rightarrow \{PHICONG, NGAYKH\}$	F^-
$\{NGAYKH\} \rightarrow \{MAYBAY, NGAYKH\}$	F^-
$\{NGAYKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH\}$	F^-
$\{NGAYKH\} \rightarrow \{GIOKH\}$	F^-
$\{NGAYKH\} \rightarrow \{PHICONG, GIOKH\}$	F^-
$\{NGAYKH\} \rightarrow \{MAYBAY, GIOKH\}$	F^-
$\{NGAYKH\} \rightarrow \{MAYBAY, PHICONG, GIOKH\}$	F^-
$\{NGAYKH\} \rightarrow \{NGAYKH, GIOKH\}$	F^-
$\{NGAYKH\} \rightarrow \{PHICONG, NGAYKH, GIOKH\}$	F^-
$\{NGAYKH\} \rightarrow \{MAYBAY, NGAYKH, GIOKH\}$	F^-
$\{NGAYKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH, GIOKH\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \emptyset$	PTHHN
$\{PHICONG, NGAYKH\} \rightarrow \{PHICONG\}$	PTHHN
$\{PHICONG, NGAYKH\} \rightarrow \{MAYBAY\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{PHICONG, MAYBAY\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{NGAYKH\}$	PTHHN
$\{PHICONG, NGAYKH\} \rightarrow \{PHICONG, NGAYKH\}$	PTHHN
$\{PHICONG, NGAYKH\} \rightarrow \{MAYBAY, NGAYKH\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{GIOKH\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{PHICONG, GIOKH\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{MAYBAY, GIOKH\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{MAYBAY, PHICONG, GIOKH\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{NGAYKH, GIOKH\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{PHICONG, NGAYKH, GIOKH\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{MAYBAY, NGAYKH, GIOKH\}$	F^-
$\{PHICONG, NGAYKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH, GIOKH\}$	F^-
$\{MAYBAY, NGAYKH\} \rightarrow \emptyset$	PTHHN
$\{MAYBAY, NGAYKH\} \rightarrow \{PHICONG\}$	F
$\{MAYBAY, NGAYKH\} \rightarrow \{MAYBAY\}$	PTHHN
$\{MAYBAY, NGAYKH\} \rightarrow \{PHICONG, MAYBAY\}$	F^+
$\{MAYBAY, NGAYKH\} \rightarrow \{NGAYKH\}$	PTHHN
$\{MAYBAY, NGAYKH\} \rightarrow \{PHICONG, NGAYKH\}$	F^+
$\{MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, NGAYKH\}$	PTHHN
$\{MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH\}$	F^+
$\{MAYBAY, NGAYKH\} \rightarrow \{GIOKH\}$	F^+
$\{MAYBAY, NGAYKH\} \rightarrow \{PHICONG, GIOKH\}$	F^+
$\{MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, GIOKH\}$	F^+
$\{MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, PHICONG, GIOKH\}$	F^+
$\{MAYBAY, NGAYKH\} \rightarrow \{NGAYKH, GIOKH\}$	F^+
$\{MAYBAY, NGAYKH\} \rightarrow \{PHICONG, NGAYKH, GIOKH\}$	F^+
$\{MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, NGAYKH, GIOKH\}$	F^+
$\{MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH, GIOKH\}$	F^+
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \emptyset$	PTHHN
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{PHICONG\}$	PTHHN
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{MAYBAY\}$	PTHHN

$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{PHICONG, MAYBAY\}$	PTHHN
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{NGAYKH\}$	PTHHN
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{PHICONG, NGAYKH\}$	PTHHN
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, NGAYKH\}$	PTHHN
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{PHICONG, MAYBAY, NGAYKH\}$	PTHHN
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{GIOKH\}$	F ⁺
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{PHICONG, GIOKH\}$	F ⁺
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, GIOKH\}$	F ⁺
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, PHICONG, GIOKH\}$	F ⁺
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{NGAYKH, GIOKH\}$	F ⁺
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{PHICONG, NGAYKH, GIOKH\}$	F ⁺
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, NGAYKH, GIOKH\}$	F ⁺
$\{PHICONG, MAYBAY, NGAYKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH, GIOKH\}$	F ⁺
.....	
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \emptyset$	PTHHN
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{PHICONG\}$	PTHHN
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{MAYBAY\}$	F
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{PHICONG, MAYBAY\}$	F ⁺
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{NGAYKH\}$	PTHHN
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{PHICONG, NGAYKH\}$	PTHHN
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, NGAYKH\}$	F ⁺
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH\}$	F ⁺
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{GIOKH\}$	PTHHN
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{PHICONG, GIOKH\}$	PTHHN
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, GIOKH\}$	F ⁺
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, PHICONG, GIOKH\}$	F ⁺
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{NGAYKH, GIOKH\}$	PTHHN
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{PHICONG, NGAYKH, GIOKH\}$	PTHHN
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, NGAYKH, GIOKH\}$	F ⁺
$\{PHICONG, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH, GIOKH\}$	F ⁺
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \emptyset$	PTHHN
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{PHICONG\}$	F ⁺
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{MAYBAY\}$	PTHHN
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{PHICONG, MAYBAY\}$	F ⁺
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{NGAYKH\}$	PTHHN
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{PHICONG, NGAYKH\}$	F ⁺
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, NGAYKH\}$	PTHHN
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH\}$	F ⁺
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{GIOKH\}$	PTHHN
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{PHICONG, GIOKH\}$	F ⁺
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, GIOKH\}$	PTHHN
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, PHICONG, GIOKH\}$	F ⁺
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{NGAYKH, GIOKH\}$	PTHHN
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{PHICONG, NGAYKH, GIOKH\}$	F ⁺
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, NGAYKH, GIOKH\}$	PTHHN
$\{MAYBAY, NGAYKH, GIOKH\} \rightarrow \{MAYBAY, PHICONG, NGAYKH, GIOKH\}$	F ⁺
.....

Số phụ thuộc hàm có thể có của $Q(A_1, A_2, \dots, A_n)$ là $2^n \times 2^n = 2^{2n}$

II HỆ LUẬT DẪN ARMSTRONG (Armstrong inference rule)

Người ta thường dùng F để chỉ tập các phụ thuộc hàm của lược đồ quan hệ Q. Ta có thể đánh số các phụ thuộc hàm của F là f_1, f_2, \dots, f_m . Quy ước rằng chỉ cần mô tả các phụ thuộc hàm không hiển nhiên trong tập F (các phụ thuộc hàm hiển nhiên được ngầm hiểu là đã có trong F).

1 Phụ thuộc hàm được suy diễn logic từ F

Nói rằng phụ thuộc hàm $X \rightarrow Y$ được suy diễn logic từ F nếu một quan hệ r thỏa mãn tất cả các phụ thuộc hàm của F thì cũng thỏa phụ thuộc hàm $X \rightarrow Y$. Ký hiệu $F \models X \rightarrow Y$.

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

Các tính chất của tập F^+

1. **Tính phản xạ:** Với mọi tập phụ thuộc hàm F^+ ta luôn luôn có $F \subseteq F^+$
2. **Tính đơn điệu:** Nếu $F \subseteq G$ thì $F^+ \subseteq G^+$
3. **Tính lũy đẳng:** Với mọi tập phụ thuộc hàm F ta luôn luôn có $(F^+)^+ = F^+$.

Gọi G là tập tất cả các phụ thuộc hàm có thể có của r , phần phụ của F ký hiệu $F^- = G - F^+$

Chứng minh

1. $X \rightarrow Y \in F \Rightarrow r$ thỏa $X \rightarrow Y \Rightarrow X \rightarrow Y \in F^+$
2. Nếu $X \rightarrow Y$ là phụ thuộc hàm thuộc F^+ ta phải chứng minh $X \rightarrow Y$ thuộc G^+
 Giả sử r thỏa tất cả các phụ thuộc hàm của G (1)
 $\Rightarrow r$ thỏa tất cả phụ thuộc hàm của F vì $F \subseteq G$
 $\Rightarrow r$ thỏa phụ thuộc hàm $X \rightarrow Y$ (2) vì $X \rightarrow Y \in F^+$
 (1) và (2) $\Rightarrow X \rightarrow Y \in G^+ \Rightarrow F^+ \subseteq G^+$
3. $F \subseteq F^+$ (tính phản xạ) $\Rightarrow F^+ \subseteq (F^+)^+$ (1)
 Nếu $X \rightarrow Y \in (F^+)^+$ (2) $\Rightarrow X \rightarrow Y \in F^+$ thật vậy: (3)
 Giả sử r thỏa tất cả các phụ thuộc hàm của F (4)
 $\Rightarrow r$ thỏa tất cả các phụ thuộc hàm của F^+ (theo định nghĩa)
 $\Rightarrow r$ thỏa tất cả các phụ thuộc hàm của $(F^+)^+$ (theo định nghĩa)
 $\Rightarrow r$ thỏa $X \rightarrow Y$ (vì (2)) $\Rightarrow X \rightarrow Y \in F^+$
 (1) và (3) $\Rightarrow (F^+)^+ = F^+$

2 Hệ luật dẫn Armstrong

Hệ luật dẫn là một phát biểu cho biết nếu một quan hệ r thỏa mãn một vài phụ thuộc hàm thì nó phải thỏa mãn phụ thuộc hàm khác.

Với X, Y, Z, W là tập con của Q^+ . r là quan hệ bất kỳ của Q . Ta có 6 luật dẫn sau:

1. **Luật phản xạ (reflexive rule):** $X \rightarrow X$
2. **Luật thêm vào (augmentation rule):** Cho $X \rightarrow Y \Rightarrow XZ \rightarrow Y$
3. **Luật hợp (union rule):** Cho $X \rightarrow Y, X \rightarrow Z \Rightarrow X \rightarrow YZ$
4. **Luật phân rã (decomposition rule):** Cho $X \rightarrow YZ \Rightarrow X \rightarrow Y$
5. **Luật bắc cầu (transitive rule):** Cho $X \rightarrow Y, Y \rightarrow Z \Rightarrow X \rightarrow Z$
6. **Luật bắc cầu giả (pseudo transitive rule):** Cho $X \rightarrow Y, YZ \rightarrow W \Rightarrow XZ \rightarrow W$

Hệ tiên đề Armstrong (Armstrong's Axioms) gồm 3 luật: (1), (2) và (5)

Chứng minh

Với t_1, t_2 là hai bộ bất kỳ của quan hệ r .

Luật phản xạ: Ta có $(t_1.X = t_2.X \Rightarrow t_1.X = t_2.X)$ theo định nghĩa suy ra $X \rightarrow X$

Luật thêm vào: giả sử có $t_1.XZ = t_2.XZ$ (1)

$$\begin{aligned} &\Rightarrow t_1.X = t_2.X \\ &\Rightarrow t_1.Y = t_2.Y \text{ (do } X \rightarrow Y \text{)} \quad (2) \\ &\Rightarrow XZ \rightarrow Y \text{ (do (1) } \Rightarrow (2)) \end{aligned}$$

Luật hợp: giả sử có

$$\begin{aligned} &t_1.X = t_2.X \quad (1) \\ &\Rightarrow t_1.X = t_2.X \text{ và } t_1.Z = t_2.Z \\ &\Rightarrow t_1.XZ = t_2.XZ \quad (2) \\ &\Rightarrow X \rightarrow YZ \text{ (do (1) } \Rightarrow (2)) \end{aligned}$$

Luật phân rã: giả sử có:

$$\begin{aligned} &t_1.X = t_2.X \quad (1) \\ &\Rightarrow t_1.YZ = t_2.YZ \text{ (do } X \rightarrow YZ \text{)} \end{aligned}$$

$$\Rightarrow t_1.Y = t_2.Y \quad (2)$$

$$\Rightarrow X \rightarrow Y \text{ (do (1) } \Rightarrow (2) \text{)}$$

Luật bắc cầu: giả sử có $t_1.X = t_2.X \quad (1)$

$$\Rightarrow t_1.Y = t_2.Y$$

$$\Rightarrow t_1.Z = t_2.Z \quad (2)$$

$$\Rightarrow X \rightarrow Z \text{ (do (1) } \Rightarrow (2) \text{)}$$

Luật bắc cầu giả: giả sử có: $t_1.XZ = t_2.XZ \quad (1)$

$$\Rightarrow t_1.X = t_2.X \text{ và } t_1.Z = t_2.Z \quad (2)$$

$$\Rightarrow t_1.Y = t_2.Y \text{ (do } X \rightarrow Y \text{)} \quad (3)$$

$$\Rightarrow t_1.YZ = t_2.YZ \text{ (Kết hợp (2) và (3))}$$

$$\Rightarrow t_1.W = t_2.W \text{ (do } YZ \rightarrow W \text{)} \quad (4)$$

$$\Rightarrow XZ \rightarrow W$$

Trong 6 luật trên, chỉ cần 3 luật 1, 2 và 6 là đủ, nghĩa là các luật còn lại có thể suy dẫn từ ba luật này.

i Hệ luật dẫn Armstrong là đúng

Nói rằng $X \rightarrow Y$ là phụ thuộc hàm được suy diễn nhờ vào luật dẫn Armstrong nếu tồn tại các tập phụ thuộc hàm $F_0 \subset F_1 \subset \dots \subset F_n$ sao cho $X \rightarrow Y \in F_n$ với F_0, F_1, \dots, F_n lần lượt được hình thành thỏa phương pháp sau:

Bước 1: $F_0 = F$

Bước 2: chọn một số phụ thuộc hàm trong F_i áp dụng hệ luật dẫn Armstrong để thu được một số phụ thuộc hàm mới. Đặt $F_{i+1} = F_i \cup \{\text{các phụ thuộc hàm mới}\}$

Ví dụ: Cho $F = \{AB \rightarrow C, C \rightarrow B, BC \rightarrow A\}$ thì có $F_0 \subset F_1 \subset F_2$ sao cho $C \rightarrow A \in F_2$

$F_0 = \{AB \rightarrow C, C \rightarrow B, BC \rightarrow A\}$ áp dụng luật hợp cho $C \rightarrow B$ và $C \rightarrow C$

$F_1 = \{AB \rightarrow C, C \rightarrow B, BC \rightarrow A, C \rightarrow BC\}$ áp dụng luật bắc cầu.

$F_2 = \{AB \rightarrow C, C \rightarrow B, BC \rightarrow A, C \rightarrow BC, C \rightarrow A\}$

Hệ quả: Hệ luật dẫn Armstrong là đúng nghĩa là nếu F là tập các phụ thuộc hàm đúng trên quan hệ r và $X \rightarrow Y$ là một phụ thuộc hàm được suy diễn từ F nhờ hệ luật dẫn Armstrong thì $X \rightarrow Y$ đúng trên quan hệ r . Vậy $X \rightarrow Y$ là phụ thuộc hàm được suy diễn logic từ F

Phần tiếp theo chúng ta sẽ chứng minh hệ luật dẫn Armstrong là đầy đủ, nghĩa là mọi phụ thuộc hàm $X \rightarrow Y$ được suy diễn logic từ F sẽ được suy diễn từ F nhờ hệ luật dẫn Armstrong.

ii Bao đóng của tập thuộc tính X (closures of attribute sets)

(a) Định nghĩa

Q là lược đồ quan hệ, r là quan hệ tương ứng, F là tập các phụ thuộc hàm trong Q . X, A_i là các tập con của Q^+ .

Bao đóng của tập thuộc tính X đối với F ký hiệu là X^+ (hoặc X_F^+) được định nghĩa như sau:

$X^+ = \cup A_i$ với $X \rightarrow A_i$ là phụ thuộc hàm được suy diễn từ F nhờ hệ tiên đề Armstrong

Tính chất:

+ bao đóng của Q là Q^+

(b) Các tính chất của bao đóng

Nếu X, Y là các tập con của tập thuộc tính Q^+ thì ta có các tính chất sau đây:

1. Tính phản xạ: $X \subseteq X^+$

2. **Tính đơn điệu:** Nếu $X \subseteq Y$ thì $X^+ \subseteq Y^+$
 3. **Tính lũy đẳng:** $X^{++} = X^+$
 4. $(XY)^+ \supseteq X^+Y^+$
 5. $(X^+Y)^+ = (XY^+)^+ = (X^+Y^+)^+$
 6. $X \rightarrow Y \Leftrightarrow Y^+ \subseteq X^+$
 7. $X \rightarrow X^+$ và $X^+ \rightarrow X$
 8. $X^+ = Y^+ \Leftrightarrow X \rightarrow Y$ và $Y \rightarrow X$

Chứng minh:

1. $X \rightarrow X \Rightarrow X^+ \supseteq X$
 2. $A \in X^+ \Rightarrow X \rightarrow A \Rightarrow Y \rightarrow A \Rightarrow A \in Y^+$
 3. $A \in X^{++} \Rightarrow X^+ \rightarrow A$ và $X \rightarrow X^+$ (áp dụng 8) $\Rightarrow X \rightarrow A \Rightarrow A \in X^+ \Rightarrow X^{++} \subseteq X^+$. Áp dụng 1 $\Rightarrow X^{++} \supseteq X^+$

 7. $X \rightarrow A_1$ và $X \rightarrow A_2 \Rightarrow X \rightarrow A_1 \cup A_2 \dots X \rightarrow \cup A_i = X^+$
 $X^+ \supseteq X \Rightarrow X^+ \rightarrow X$ (Phụ thuộc hàm hiển nhiên)

(c) Thuật toán tìm bao đóng

Tính liên tiếp tập các tập thuộc tính X_0, X_1, X_2, \dots theo phương pháp sau:

Bước 1: $X_0 = X$

Bước 2: lần lượt xét các phụ thuộc hàm của F

Nếu $Y \rightarrow Z$ có $Y \subseteq X_i$ thì $X_{i+1} = X_i \cup Z$

Loại phụ thuộc hàm $Y \rightarrow Z$ khỏi F

Bước 3: Nếu ở bước 2 không tính được X_{i+1} thì X_i chính là bao đóng của X

Ngược lại lặp lại bước 2

Ví dụ 1:

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

$$F = \left\{ \begin{array}{lll} f_1: & B & \rightarrow A \\ f_2: & DA & \rightarrow CE \\ f_3: & D & \rightarrow H \\ f_4: & GH & \rightarrow C \\ f_5: & AC & \rightarrow D \end{array} \right\}$$

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

Giải:

Bước 1: $X_0 = AC$

Bước 2: Do f_1, f_2, f_3, f_4 không thỏa. f_5 thỏa vì $X^+ \supseteq AC$

$$X_1 = AC \cup D = ACD$$

Lặp lại bước 2: f_1 không thỏa, f_2 thỏa vì $X_1 \supseteq AD$:

$$X_2 = ACD \cup CE = ACDE$$

f_3 thỏa vì $X_2 \supseteq D$

$$X_3 = ACDE \cup H = ACDEH$$

f_4 không thỏa, f_5 không xét vì đã thỏa

Lặp lại bước 2: f_2, f_3 không xét vì đã thỏa, f_1, f_4 không thỏa, f_5 không xét vì đã thỏa. Trong bước này X_3 không thay đổi $\Rightarrow X^+ = X_3 = \{ACDEH\}$ là bao đóng của X

Ví dụ 2:

$Q(A, B, C, D, E, G)$

$F = \{f_1: A \rightarrow C;$
 $f_2: A \rightarrow EG;$
 $f_3: B \rightarrow D;$
 $f_4: G \rightarrow E\}$

$X = \{A, B\}; Y = \{C, G, D\}$

Kết quả:

$X^+ = \{ABCDEG\}$

$Y^+ = \{CGDE\}$

(d) Định lý

Thuật toán tìm bao đóng cho kết quả $X_i = X^+$

Chứng minh

1. Ta chứng minh $X_i \subseteq X^+$ bằng phương pháp qui nạp.

Bước cơ sở chứng minh $X \rightarrow X_0$

Theo tính phản xạ của hệ luật dẫn thì $X \rightarrow X$ theo thuật toán thì $X_0 = X \Rightarrow X \rightarrow X_0$

Vậy $X_0 \subseteq X^+$

Bước qui nạp giả sử có $X \rightarrow X_{i-1}$ (1) ta phải chứng minh $X \rightarrow X_i$

Theo thuật toán tìm bao đóng thì có $f_j = X_j \rightarrow Y_j$ để $X_{i-1} \supseteq X_j$ và $X_i = X_{i-1} \cup Y_j$

$\Rightarrow X_{i-1} \rightarrow Y_j$ (2). (1) và (2) $\Rightarrow X \rightarrow Y_j$ (3)

(1) và (3) $\Rightarrow X \rightarrow X_{i-1}Y_j = X_i \Rightarrow X \rightarrow X_i$

Vậy $X_i \subseteq X^+$

2. Ta chứng minh $A \subseteq X^+ \Rightarrow A \subseteq X_i$ để kết luận $X_i \supseteq X^+$

$A \subseteq X^+$ nên có một phụ thuộc hàm $X \rightarrow A$. Theo thuật toán tìm bao đóng thì $X \subseteq X_i \Rightarrow A \subseteq X_i$

(e) Hệ quả

1. Q là lược đồ quan hệ. F là tập phụ thuộc hàm, A là thuộc tính chỉ xuất hiện ở vế phải của các phụ thuộc hàm trong F thì $X^+ = (X - A)^+ \cup A$

2. Q là lược đồ quan hệ. F là tập phụ thuộc hàm, X là tập con của Q^+ và $Y = \{\text{các thuộc tính xuất hiện ở vế phải của các phụ thuộc hàm trong } F\}$ thì $X^+ \subseteq X \cup Y$.

Chứng minh

1. Theo thuật toán tìm bao đóng thì bao đóng X^+ hay $(X-A)^+$ được hình thành qua một số bước.

Ta chứng minh biểu thức $X^+ = (X - A)^+ \cup A$ theo qui nạp.

Bước cơ sở: $X_0 = X, (X-A)_0 = X - A \Rightarrow X_0 = (X - A)_0 \cup A$ đúng

Bước qui nạp: giả sử ta có $X_{i-1} = (X - A)_{i-1} \cup A$. Bao đóng X_i được hình thành do có $f_j = X_j \rightarrow Y_j$ để:

$X_{i-1} \supseteq X_j$ và $X_i = X_{i-1} \cup Y_j = (X - A)_{i-1} \cup A \cup Y_j$ (1).

Sự hình thành X_i luôn kéo theo sự hình thành $(X-A)_i$ vì:

$X_{i-1} = (X - A)_{i-1} \cup A \supseteq X_j$ do X_j không chứa A nên:

$(X - A)_{i-1} \supseteq X_j$ vậy $(X - A)_i = (X - A)_{i-1} \cup Y_j$ (2)

(1) và (2) cho:

$X_i = (X - A)_i \cup A$ là điều phải chứng minh

2. **Bước cơ sở:** $X_0 = X \Rightarrow X_0 \subseteq X \cup Y$

Bước qui nạp: giả sử có $X_{i-1} \subseteq X \cup Y$ ta chứng minh $X_i \subseteq X \cup Y$.

Bao đóng X_i được hình thành do có $f_j = X_j \rightarrow Y_j$ để:

$X_{i-1} \supseteq X_j$ và $X_i = X_{i-1} \cup Y_j \subseteq X \cup Y \cup Y_j$ do Y_j là vế phải của phụ thuộc hàm nên $Y \cup Y_j = Y$ vậy $X_i \subseteq X \cup Y$

3 Hệ luật dẫn Armstrong là đầy đủ

i Định lý

Hệ luật dẫn Armstrong là đầy đủ nghĩa là mọi phụ thuộc hàm $X \rightarrow Y$ được suy diễn logic từ F sẽ được suy diễn từ F nhờ hệ luật dẫn Armstrong.

Chứng minh:

Để chứng minh $X \rightarrow Y$ được suy diễn từ F nhờ hệ luật dẫn Armstrong ta chứng minh bằng phương pháp phản chứng nghĩa là nếu $X \rightarrow Y$ không suy diễn được từ hệ luật dẫn Armstrong thì có quan hệ r thỏa các phụ thuộc hàm F nhưng không thỏa phụ thuộc hàm $X \rightarrow Y$ (điều này nghịch lý với giả thuyết là mọi quan hệ r thỏa các phụ thuộc hàm trong F thì r cũng thỏa phụ thuộc hàm $X \rightarrow Y$).

Thật vậy giả sử $Q(A_1, A_2, \dots, A_n)$ là lược đồ quan hệ, a_i, b_i là các giá trị khác nhau trên miền giá trị A_i . r là quan hệ trên Q có hai bộ t và t' được xác định như sau:

$$t = (a_1, a_2, \dots, a_n)$$

$$t'.A_i = \begin{cases} a_i & \text{Nếu } A_i \in X^+ \\ b_i & \text{Ngược lại} \end{cases}$$

Vậy quan hệ r có $t.X = t'.X$ nhưng $t.Y \neq t'.Y$ ($t.Y$ gồm các giá trị a_i còn $t'.Y$ phải có ít nhất một b_i nếu không $Y \subseteq X^+ \Rightarrow X \rightarrow Y$ được suy diễn từ hệ luật dẫn Armstrong). Như vậy r không thỏa phụ thuộc hàm $X \rightarrow Y$.

Bây giờ ta chứng minh quan hệ r thỏa mọi phụ thuộc hàm trong F . Gọi $W \rightarrow Z$ là phụ thuộc hàm trong F .

Nếu $W \not\subseteq X^+ \Rightarrow t.W \neq t'.W \Rightarrow$ mệnh đề $(t.W = t'.W \Rightarrow t.Z = t'.Z)$ đúng

Nếu $W \subseteq X^+ \Rightarrow t.Z = t'.Z =$ bộ các a_i

\Rightarrow mệnh đề $(t.W = t'.W \Rightarrow t.Z = t'.Z)$ đúng

ii Hệ quả:

Bao đóng của tập thuộc tính X đối với F là:

$X^+ = \cup A_i$ với $X \rightarrow A_i$ là phụ thuộc hàm được suy diễn logic từ F

Tính chất

$$X \rightarrow Y \in F^+ \Leftrightarrow Y \subseteq X^+$$

Chứng minh

$X \rightarrow Y \Rightarrow$ có k sao cho $Y = A_k \subseteq \cup A_i = X^+$

$Y \subseteq X^+ \Rightarrow X^+ = Y \cup (X^+ - Y) \Rightarrow X \rightarrow Y \cup (X^+ - Y) \Rightarrow X \rightarrow Y$

Bài toán thành viên

Nói rằng $X \rightarrow Y$ là thành viên của F nếu $X \rightarrow Y \in F^+$

Một vấn đề quan trọng khi nghiên cứu lý thuyết CSDL là khi cho trước tập các phụ thuộc hàm F và một phụ thuộc hàm $X \rightarrow Y$, làm thế nào để biết $X \rightarrow Y$ có thuộc F^+ hay không bài toán này được gọi là bài toán thành viên. Để trả lời câu hỏi này ta có thể tính F^+ rồi xác định xem $X \rightarrow Y$ có thuộc F^+ hay không. Việc tính F^+ là một công việc đòi hỏi thời gian và công sức. Tuy nhiên, thay vì tính F^+ chúng ta có thể dùng thuật toán sau để xác định $X \rightarrow Y$ có là thành viên của F hay không. Thuật toán này sử dụng tính chất vừa chứng minh trên.

Thuật toán xác định $F = X \rightarrow Y$ có là thành viên của F hay không

Bước 1: tính X^+

Bước 2: so sánh X^+ với Y nếu $X^+ \supseteq Y$ thì ta khẳng định $X \rightarrow Y$ là thành viên của F

Bạn đọc hãy nắm thật kỹ thuật toán này – nó mở đầu cho một loạt ứng dụng về sau.

III THUẬT TOÁN TÌM F^+ **1 Thuật toán cơ bản**

Để tính bao đóng F^+ của tập các phụ thuộc hàm F ta thực hiện các bước sau:

Bước 1: Tìm tất cả tập con của Q^+

Bước 2: Tìm tất cả các phụ thuộc hàm có thể có của Q .

Bước 3: Tìm bao đóng của tất cả tập con của Q .

Bước 4: Dựa vào bao đóng của tất cả các tập con đã tìm để xác định phụ thuộc hàm nào thuộc F^+

Ví dụ 3:

Cho lược đồ quan hệ $Q(A, B, C)$ $F = \{AB \rightarrow C, C \rightarrow B\}$ là tập phụ thuộc hàm trên Q . F^+ được tính lần lượt theo các bước trên là như sau:

Bước 1: Các tập con của Q^+

\emptyset	A	B	C
\emptyset	$\{A\}$	$\{B\}$	$\{C\}$
		$\{A, B\}$	$\{A, C\}$
			$\{B, C\}$
			$\{A, B, C\}$

Bước 2: các phụ thuộc hàm có thể có của F (không kê các phụ thuộc hàm hiển nhiên)

$A \rightarrow B$	$A \rightarrow BC$	$B \rightarrow C$	$AB \rightarrow C \in F$	$C \rightarrow A$	$C \rightarrow BC \in F^+$	$AC \rightarrow BC \in F^+$	$BC \rightarrow AC$
$A \rightarrow AB$	$A \rightarrow ABC$	$B \rightarrow AC$	$AB \rightarrow AC \in F^+$	$C \rightarrow B \in F$	$C \rightarrow ABC$	$AC \rightarrow ABC \in F^+$	$BC \rightarrow ABC$
$A \rightarrow C$	$B \rightarrow A$	$B \rightarrow BC$	$AB \rightarrow BC \in F^+$	$C \rightarrow AB$	$AC \rightarrow B \in F^+$	$BC \rightarrow A$	
$A \rightarrow AC$	$B \rightarrow AB$	$B \rightarrow ABC$	$AB \rightarrow ABC \in F^+$	$C \rightarrow AC$	$AC \rightarrow AB \in F^+$	$BC \rightarrow AB$	

Bước 3: bao đóng của các tập con của Q đối với F

$\emptyset^+ = \emptyset$	$A^+ = A$	$C^+ = BC$
$ABC^+ = ABC$	$B^+ = B$	$AC^+ = ABC$
	$AB^+ = ABC$	$BC^+ = BC$

Bước 4: $F = \{AB \rightarrow C, C \rightarrow B\}$ suy ra:

$F^+ = \{AB \rightarrow C, AB \rightarrow AC, AB \rightarrow BC, AB \rightarrow ABC, C \rightarrow B, C \rightarrow BC, AC \rightarrow B, AC \rightarrow AB, AC \rightarrow BC, AC \rightarrow ABC\}$

2 Thuật toán cải tiến

Dựa vào thuật toán cơ bản trên, ta nhận thấy có thể tính F^+ theo các bước sau:

Bước 1: Tìm tất cả tập con của Q^+

Bước 2: Tìm bao đóng của tất cả tập con của Q^+ .

Bước 4: Dựa vào bao đóng của các tập con đã tìm để suy ra các phụ thuộc hàm thuộc F^+ .

Ví dụ bao đóng $A^+ = A$ chỉ gồm các phụ thuộc hàm hiển nhiên

bao đóng $\{AB\}^+ = ABC$ cho các phụ thuộc hàm không hiển nhiên sau

$AB \rightarrow C, AB \rightarrow AC, AB \rightarrow BC, AB \rightarrow ABC$

(Tìm tất cả các tập con của $\{ABC\}$ rồi bỏ các tập con của $\{AB\}$)

Các tập con của $\{ABC\}$ là: $\emptyset, \{A\}, \{B\}, \{AB\}, \{C\}, \{AC\}, \{BC\}, \{ABC\}$

Bỏ các tập con của $\{AB\}$ là: $\emptyset, \{A\}, \{B\}, \{AB\}, \{C\}, \{AC\}, \{BC\}, \{ABC\}$

Các tập còn lại chính là vế phải của phụ thuộc hàm có vế trái là AB

IV BÀI TẬP

1/ Cho quan hệ sau:

r (A	B	C	D	E)
	a ₁	b ₁	c ₁	d ₁	e ₁
	a ₁	b ₂	c ₂	d ₂	d ₁
	a ₂	b ₁	c ₃	d ₃	e ₁
	a ₂	b ₁	c ₄	d ₃	e ₁
	a ₃	b ₂	c ₅	d ₁	e ₁

Phụ thuộc hàm nào sau đây thỏa r:

$A \rightarrow D, AB \rightarrow D, C \rightarrow BDE, E \rightarrow A, A \rightarrow E$

2/ Cho $Q^+ = \{ABCD\}$

a) Tìm tất cả các tập con của Q

b) Tìm tất cả các phụ thuộc hàm có thể có của Q (không liệt kê phụ thuộc hàm hiển nhiên)

3/ Tìm bao đóng F^+ của quan hệ phảnCong (PHICONG, MAYBAY, NGAYKH, GIOKH)

4/ Cho $F = \{AB \rightarrow C, B \rightarrow D, CD \rightarrow E, CE \rightarrow GH, G \rightarrow A\}$

a) Hãy chứng tỏ phụ thuộc hàm $AB \rightarrow E, AB \rightarrow G$ được suy diễn từ F nhờ luật dẫn Armstrong

b) Tìm bao đóng của AB (với bài toán không nói gì về lược đồ quan hệ Q ta ngầm hiểu $Q^+ = \{ABCDEGH\}$)

5/ Cho $F = \{A \rightarrow D, AB \rightarrow DE, CE \rightarrow G, E \rightarrow H\}$. Hãy tìm bao đóng của AB.

6/ Cho $F = \{AB \rightarrow E, AG \rightarrow I, BE \rightarrow I, E \rightarrow G, GI \rightarrow H\}$.

a) Hãy chứng tỏ phụ thuộc hàm $AB \rightarrow GH$ được suy diễn từ F nhờ luật dẫn Armstrong

b) Tìm bao đóng của $\{AB\}$

7/ Cho $F = \{A \rightarrow D, AB \rightarrow E, BI \rightarrow E, CD \rightarrow I, E \rightarrow C\}$ tìm bao đóng của $\{AE\}^+ = \{ACDEI\}$

----oOo----

Chương 5 .**PHỦ CỦA TẬP PHỤ THUỘC HÀM****I ĐỊNH NGHĨA**

Nói rằng hai tập phụ thuộc hàm F và G là tương đương (Equivalent) nếu $F^+ = G^+$ ký hiệu $F \equiv G$.
Ta nói F phủ G nếu $F^+ \supseteq G^+$

Thuật toán xác định F và G có tương đương không

Bước 1: Với mỗi phụ thuộc hàm $X \rightarrow Y$ của F ta xác định xem $X \rightarrow Y$ có là thành viên của G không

Bước 2: Với mỗi phụ thuộc hàm $X \rightarrow Y$ của G ta xác định xem $X \rightarrow Y$ có là thành viên của F không

Nếu cả hai bước trên đều đúng thì $F \equiv G$

Ví dụ 1: Cho lược đồ quan hệ $Q(ABCDE)$ hai tập phụ thuộc hàm:

$$F = \{A \rightarrow BC, A \rightarrow D, CD \rightarrow E\} \text{ và } G = \{A \rightarrow BCE, A \rightarrow ABD, CD \rightarrow E\}$$

a) F có tương đương với G không?

b) F có tương đương với $G' = \{A \rightarrow BCDE\}$ không?

Giải:

a) Ta có $A_G^+ = ABCDE \Rightarrow$ trong G^+ có $A \rightarrow BC$ và $A \rightarrow D \Rightarrow F \subseteq G^+ \Rightarrow F^+ \subseteq G^+ \quad (1)$.

$A_F^+ = ABCDE \Rightarrow$ trong F^+ có $A \rightarrow BCE$ và $A \rightarrow ABD \Rightarrow F^+ \supseteq G \Rightarrow F^+ \supseteq G^+ \quad (2)$

(1) và (2) $\Rightarrow F^+ = G^+ \Rightarrow F \equiv G$.

b) Do $(CD)_{G'}^+ = CD \Rightarrow G'^+$ không chứa phụ thuộc hàm $CD \rightarrow E \Rightarrow F$ không tương đương với G'

II PHỦ TỐI THIỂU CỦA MỘT TẬP PHỤ THUỘC HÀM (minimal cover)**1 Phụ thuộc hàm có vế trái dư thừa**

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

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

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

Ví dụ 2: $Q(A, B, C)$ $F = \{AB \rightarrow C; B \rightarrow C\}$

$$F \equiv F - \{AB \rightarrow C\} \cup \{(AB - A) \rightarrow C\} = \{B \rightarrow C\}$$

$AB \rightarrow C$ là phụ thuộc hàm không đầy đủ

$B \rightarrow C$ là phụ thuộc hàm đầy đủ

Chú ý: phụ thuộc hàm có vế trái chứa một thuộc tính là phụ thuộc hàm đầy đủ.

Ví dụ 3: cho tập phụ thuộc hàm $F = \{A \rightarrow BC, B \rightarrow C, AB \rightarrow D\}$ thì phụ thuộc hàm $AB \rightarrow D$ có vế trái dư thừa B vì:

$$F \equiv F - \{AB \rightarrow D\} \cup \{A \rightarrow D\}$$

$$\equiv \{A \rightarrow BC, B \rightarrow C, A \rightarrow D\}$$

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

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

Bước 1: lần lượt thực hiện bước 2 cho các phụ thuộc hàm $X \rightarrow Y$ của F

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

Nếu $X' \rightarrow Y \in F^+$ thì thay $X \rightarrow Y$ trong F bằng $X' \rightarrow Y$ thực hiện lại bước 2

Ví dụ: Ở ví dụ 3 phụ thuộc hàm $AB \rightarrow D$ có $A^+ = ABCD \Rightarrow A \rightarrow D \in F^+$. Trong F ta thay $AB \rightarrow D$ bằng $A \rightarrow D \Rightarrow F = \{A \rightarrow BC, B \rightarrow C, A \rightarrow D\}$

2 Tập phụ thuộc hàm có vế phải một thuộc tính (the right sides of dependencies has a single attribute)

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

Ví dụ 4: cho $F = \{A \rightarrow BC, B \rightarrow C, AB \rightarrow D\}$ ta suy ra
 $F = \{A \rightarrow B, A \rightarrow C, B \rightarrow C, AB \rightarrow D\} = G$

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

3 Tập phụ thuộc hàm không dư thừa

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

Ví dụ: cho $F = \{A \rightarrow BC, B \rightarrow D, AB \rightarrow D\}$ thì F dư thừa vì
 $F = F' = \{A \rightarrow BC, B \rightarrow D\}$

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

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

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

Bước 3: thực hiện bước 2 cho các phụ thuộc hàm tiếp theo của F

4 Tập phụ thuộc hàm tối thiểu (minimal cover)

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

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

Thuật toán tìm phủ tối thiểu của một tập phụ thuộc hàm

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

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

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

Chú ý: Theo thuật toán trên, từ một tập phụ thuộc hàm F luôn tìm được ít nhất một phủ tối thiểu F_{tt} để $F = F_{tt}$ và nếu thứ tự loại các phụ thuộc hàm trong tập F là khác nhau thì có thể sẽ thu được những phủ tối thiểu khác nhau.

Ví dụ 5: Cho lược đồ quan hệ $Q(A, B, C, D)$ và tập phụ thuộc F như sau:

$$F = \{AB \rightarrow CD, B \rightarrow C, C \rightarrow D\}$$

Hãy tính phủ tối thiểu của F .

Giải:

Bước 1: $AB \rightarrow CD$ là phụ thuộc hàm có vế trái dư thừa?

$B \rightarrow CD \in F^+$? trả lời: $B^+ = BCD \Rightarrow B \rightarrow CD \in F^+$

Vậy $AB \rightarrow CD$ là phụ thuộc hàm có vết trái dư thừa $A \Rightarrow$ kết quả của bước 1 là:

$$F \equiv \{B \rightarrow CD; B \rightarrow C; C \rightarrow D\}$$

Bước 2: kết quả của bước 2 là:

$$F \equiv \{B \rightarrow D; B \rightarrow C; C \rightarrow D\} = F_{1tt}$$

Bước 3: trong F_{1tt} , $B \rightarrow C$ là phụ thuộc hàm dư thừa?

$$B \rightarrow C \in G^+? \text{ với } G = F_{1tt} - \{B \rightarrow C\} = \{B \rightarrow D; C \rightarrow D\}$$

$$B_G^+ = BD \Rightarrow B \rightarrow C \notin G^+ \Rightarrow \text{trong } F_{1tt} \text{ } B \rightarrow C \text{ không dư thừa.}$$

trong F_{1tt} , $B \rightarrow D$ là phụ thuộc hàm dư thừa?

$$B \rightarrow D \in G^+? \text{ với } G = F_{1tt} - \{B \rightarrow D\} = \{B \rightarrow C; C \rightarrow D\}$$

$$B_G^+ = BCD \Rightarrow B \rightarrow D \in G^+ \Rightarrow \text{trong } F_{1tt}, B \rightarrow D \text{ dư thừa.}$$

kết quả của bước 3 cho phủ tối thiểu:

$$F \equiv \{B \rightarrow C; C \rightarrow D\} = F_{tt}$$

Ví dụ 6: Cho lược đồ quan hệ $Q(MSCD, MSSV, CD, HG)$ và tập phụ thuộc F như sau:

$$F = \{ \begin{array}{ll} MSCD & \rightarrow CD; \\ CD & \rightarrow MSCD; \\ CD, MSSV & \rightarrow HG; \\ MSCD, HG & \rightarrow MSSV; \\ CD, HG & \rightarrow MSSV; \\ MSCD, MSSV & \rightarrow HG \end{array} \}$$

Hãy tìm phủ tối thiểu của F

kết quả:

$$F_{tt} = \{ \begin{array}{ll} MSCD & \rightarrow CD; \\ CD & \rightarrow MSCD; \\ CD, HG & \rightarrow MSSV; \\ MSCD, MSSV & \rightarrow HG \end{array} \}$$

III KHÓA CỦA LƯỢC ĐỒ QUAN HỆ (Key)

1 Định Nghĩa

$Q(A_1, A_2, \dots, A_n)$ là lược đồ quan hệ.

Q^+ là tập thuộc tính của Q .

F là tập phụ thuộc hàm trên Q .

K là tập con của Q^+ .

Nói rằng K là một khóa của Q nếu:

1. $K^+ = Q^+$ và
2. Không tồn tại $K' \subset K$ sao cho $K'^+ = Q^+$

Tập thuộc tính S được gọi là **siêu khóa** nếu $S \supseteq K$

Thuộc tính A được gọi là **thuộc tính khóa** nếu $A \in K$ với K là khóa bất kỳ của Q . Ngược lại A được gọi là **thuộc tính không khóa**.

Một lược đồ quan hệ có thể có nhiều khóa và tập thuộc tính không khóa cũng có thể bằng rỗng.

(Khi thiết kế một hệ thống thông tin, thì việc lập lược đồ cơ sở dữ liệu đạt đến một tiêu chuẩn nào đó là một việc làm quan trọng. Việc xác định chuẩn cho một lược đồ quan hệ có liên quan mật thiết với thuật toán tìm khóa).

Thuật toán tìm một khóa của một lược đồ quan hệ Q

Bước 1: gán $K = Q^+$

Bước 2: A là một thuộc tính của K, đặt $K' = K - A$. Nếu $K'^+ = Q^+$ thì gán $K = K'$ thực hiện lại bước 2

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

Ví dụ 7:

$Q(A, B, C, D, E, G, H, I) F = \{AC \rightarrow B; BI \rightarrow ACD; ABC \rightarrow D; H \rightarrow I; ACE \rightarrow BCG; CG \rightarrow AE\}$

Tìm K

Lần lượt loại các thuộc tính trong K theo thứ tự sau:

A, B, D, E, I

Ta được một khóa là của lược đồ quan hệ là $\{C, G, H\}$

(Lưu ý là thuật toán này chỉ nên sử dụng trong trường hợp chỉ cần tìm một khóa).

2 Thuật toán tìm tất cả khóa

i Thuật toán cơ bản

Bước 1: Xác định tất cả các tập con khác rỗng của Q^+ . Kết quả tìm được giả sử là các tập thuộc tính $X_1, X_2, \dots, X_{2^n-1}$

Bước 2: Tìm bao đóng của các X_i

Bước 3: Siêu khóa là các X_i có bao đóng đúng bằng Q^+ . Giả sử ta đã có các siêu khóa là $S = \{S_1, S_2, \dots, S_m\}$

Bước 4: Xây dựng tập chứa tất cả các khóa của Q từ tập S bằng cách xét mọi S_i, S_j con của S ($i \neq j$), nếu $S_i \subset S_j$ thì ta loại S_j ($i, j=1..n$), kết quả còn lại của S chính là tập tất cả các khóa cần tìm.

Ví dụ 8: Tìm tất cả các khóa của lược đồ quan hệ và tập phụ thuộc hàm như sau:

$Q(C, S, Z); F = \{f_1: CS \rightarrow Z; f_2: Z \rightarrow C\}$

X_i	X_i^+	Siêu khóa	khóa
C	C		
S	S		
CS	CSZ	CS	CS
Z	ZC		
CZ	CZ		
SZ	SZC	SZ	SZ
CSZ	CSZ	CSZ	

Vậy lược đồ quan hệ Q có hai khóa là: $\{C, S\}$ và $\{S, Z\}$

Thuật toán trên thì dễ hiểu, dễ cài đặt, tuy nhiên nếu với n khá lớn thì phép duyệt để tìm ra tập tất cả các tập con của tập Q^+ là không hiệu quả. Do vậy cần thu hẹp không gian duyệt. Chúng ta sẽ nghiên cứu thuật toán cải tiến theo hướng giảm số thuộc tính của tập cần duyệt tất cả các tập con.

ii Thuật toán cải tiến

Trước khi đi vào thuật toán cải tiến, ta cần quan tâm một số khái niệm sau:

- + Tập thuộc tính nguồn (TN) chứa tất cả các thuộc tính có xuất hiện ở vế trái và không xuất hiện ở vế phải của các phụ thuộc hàm và các thuộc tính không xuất hiện ở cả vế trái lẫn vế phải của các phụ thuộc hàm.
- + Tập thuộc tính đích (TD) chứa tất cả các thuộc tính có xuất hiện ở vế phải và không xuất hiện ở vế trái của các phụ thuộc hàm.
- + Tập thuộc tính trung gian (TG) chứa tất cả các thuộc tính xuất hiện ở cả vế trái lẫn vế phải của các phụ thuộc hàm.

Hệ quả: Nếu K là khóa của Q thì $TN \subseteq K$ và $TD \cap K = \emptyset$

Chứng minh $TN \subseteq K$

Theo hệ quả 2 của thuật toán tìm bao đóng ta có $K^+ \subseteq K \cup TD \cup TG$

Ta chứng minh $A \in TN \Rightarrow A \in K$. Thật vậy:

Nếu $A \notin K \Rightarrow K^+ \subseteq K \cup TD \cup TG \subseteq Q^+ - A \Rightarrow K$ không là khóa \Rightarrow mâu thuẫn

Chứng minh $TD \cap K = \emptyset$

Giả sử có thuộc tính $A \in TD \cap K$ ta sẽ dẫn đến điều mâu thuẫn. Thật vậy:

Theo hệ quả 1 của thuật toán tìm bao đóng thì $K^+ = (K-A)^+ \cup A$

$A \in TD \Rightarrow$ có X là vế trái của một phụ thuộc hàm trong F sao cho $X \rightarrow A$ (1) và $A \notin X \Rightarrow X$

$\subseteq K^+ = (K-A)^+ \cup A$ vì $A \notin X \Rightarrow X \subseteq (K-A)^+ \Rightarrow (K-A) \rightarrow X$ (2)

(1) và (2) cho $(K-A) \rightarrow A \Rightarrow A \in (K-A)^+ \Rightarrow (K-A)^+ \cup A = (K-A)^+ \Rightarrow K^+ = (K-A)^+$ mâu thuẫn với điều K là khóa.

Dựa vào hệ quả trên ta có thuật toán tìm tất cả khóa sau:

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

Dữ liệu ra: Tất cả các khóa của quan hệ

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

Bước 1: tạo tập thuộc tính nguồn TN , tập thuộc tính trung gian TG

Bước 2: if $TG = \emptyset$ then lược đồ quan hệ chỉ có một khóa K

$$K = TN$$

kết thúc

Ngược lại

Qua bước 3

Bước 3: tìm tất cả các tập con X_i của tập trung gian TG

Bước 4: tìm các siêu khóa S_i bằng cách $\forall X_i$

if $(TN \cup X_i)^+ = Q^+$ then

$$S_i = TN \cup X_i$$

Bước 5: tìm khóa bằng cách loại bỏ các siêu khóa không tối thiểu

$$\forall S_i, S_j \in S$$

if $S_i \subset S_j$ then Loại S_j ra khỏi Tập siêu khóa S

S còn lại chính là tập khóa cần tìm.

Ví dụ 9: Giải lại bài tập ví dụ 8

Áp dụng thuật toán cải tiến ta có lời giải như sau:

$$TN = \{S\}; \quad TG = \{C, Z\}$$

Gọi X_i là các tập con của tập TG :

X_i	$(TN \cup X_i)$	$(TN \cup X_i)^+$	Siêu khóa	khóa
\emptyset	S	S		

C	SC	Q^+	SC	SC
Z	SZ	Q^+	SZ	SZ
CZ	SCZ	Q^+	SCZ	

Kết quả quan hệ trên có hai khóa là : $\{S, C\}$ và $\{S, Z\}$

IV BÀI TẬP

1/ Chứng minh các tính chất sau:

a) Tính cộng đầy đủ $X \rightarrow Y$ và $Z \rightarrow W \Rightarrow XZ \rightarrow YW$

b) Tính tích lũy $X \rightarrow Y$ và $Y \rightarrow ZW \Rightarrow X \rightarrow YZW$

2/ Cho $G = \{AB \rightarrow C, A \rightarrow B, B \rightarrow C, A \rightarrow C\}$. $F = \{AB \rightarrow C, A \rightarrow B, B \rightarrow C\}$ có tương đương với G không?

3/ Cho lược đồ CSDL

Kehoach (NGAY, GIO, PHONG, MONHOC, GIAOVIEN)

$F = \{NGAY, GIO, PHONG \rightarrow MONHOC$

$MONHOC, NGAY \rightarrow GIAOVIEN$

$NGAY, GIO, PHONG \rightarrow GIAOVIEN$

$MONHOC \rightarrow GIAOVIEN\}$

a) Tính $\{NGAY, GIO, PHONG\}^+; \{MONHOC\}^+$

b) Tìm phủ tối thiểu của F

c) Tìm tất cả các khóa của Kehoach

4/ Cho lược đồ CSDL

$Q (TENTAU, LOAITAU, MACHUYEN, LUONGHANG, BENCANG, NGAY)$

$F = \{TENTAU \rightarrow LOAITAU$

$MACHUYEN \rightarrow TENTAU, LUONGHANG$

$TENTAU, NGAY \rightarrow BENCANG, MACHUYEN\}$

a) Hãy tìm tập phủ tối thiểu của F

b) Tìm tất cả các khóa của Q

5/ $Q(A, B, C, D, E, G)$

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

$X = \{B, D\}, X^+ = ?$

$Y = \{C, G\}, Y^+ = ?$

6/ cho lược đồ quan hệ Q và tập phụ thuộc hàm F

a) $F = \{AB \rightarrow E; AG \rightarrow I; BE \rightarrow I; E \rightarrow G; GI \rightarrow H\}$ chứng minh rằng $AB \rightarrow GH$.

b) $F = \{AB \rightarrow C; B \rightarrow D; CD \rightarrow E; CE \rightarrow GH; G \rightarrow A\}$ chứng minh rằng $AB \rightarrow E; AB \rightarrow G$

7/ Cho quan hệ r

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

Trong các phụ thuộc hàm sau đây, PTH nào không thỏa

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

8/ Hãy tìm tất cả các khóa cho lược đồ quan hệ sau:

$Q(BROKER, OFFICE, STOCK, QUANTITY, INVESTOR, DIVIDENT)$

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

9/ Xét lược đồ quan hệ và tập phụ thuộc dữ liệu:

$Q(C, T, H, R, S, G)$

$f = \{$
 $f_1: C \rightarrow T; \quad f_2: HR \rightarrow C; \quad f_3: HT \rightarrow R;$
 $f_4: CS \rightarrow G; \quad f_5: HS \rightarrow R\}$

Tìm phủ tối thiểu của F

10/ $Q(A, B, C, D, E, H)$

$F = \{A \rightarrow E; C \rightarrow D; E \rightarrow DH\}$

Chứng minh $K = \{A, B, C\}$ là khóa duy nhất của Q

11/ $Q(A, B, C, D)$

$F = \{AB \rightarrow C; D \rightarrow B; C \rightarrow ABD\}$

Hãy tìm tất cả các khóa của Q

12/ $Q(A, B, C, D, E, G)$

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

Hãy tìm tất cả các khóa của Q.

13/ Xác định phủ tối thiểu của tập phụ thuộc hàm sau:

a) $Q(A, B, C, D, E, G),$

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

b) $Q(A, B, C)$

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

14/ Xác định phủ tối thiểu của các tập phụ thuộc hàm sau:

a) $Q_1(ABCDEFGH)$

$F_1 = \{A \rightarrow H, AB \rightarrow C, BC \rightarrow D; G \rightarrow B\}$

b) $Q_2(ABCSXYZ)$

$F_2 = \{S \rightarrow A; AX \rightarrow B; S \rightarrow B; BY \rightarrow C; CZ \rightarrow X\}$

c) $Q_3(ABCDEFGH I J)$

$F_3 = \{BG \rightarrow D; G \rightarrow J; AI \rightarrow C; CE \rightarrow H; BD \rightarrow G; JH \rightarrow A; D \rightarrow I\}$

d) $Q_4(ABCDEFGH I J)$

$F_4 = \{BH \rightarrow I; GC \rightarrow A; I \rightarrow J; AE \rightarrow G; D \rightarrow B; I \rightarrow H\}$

-----oOo-----

Chương 6 .**CHUẨN HÓA CƠ SỞ DỮ LIỆU*****I DẠNG CHUẨN CỦA LƯỢC ĐỒ QUAN HỆ (normal forms for relation schemes)***

Trong thực tế, một ứng dụng cụ thể có thể được thiết kế thành nhiều lược đồ cơ sở dữ liệu khác nhau, và tất nhiên chất lượng thiết kế của các lược đồ CSDL này cũng khác nhau. Chất lượng thiết kế của một lược đồ CSDL có thể được đánh giá dựa trên nhiều tiêu chuẩn trong đó sự trùng lặp thông tin và chi phí kiểm tra các ràng buộc toàn vẹn là hai tiêu chuẩn quan trọng.

Sau đây là một số tiêu chuẩn để đánh giá độ tốt/xấu của một lược đồ quan hệ. Trước tiên ta tìm hiểu một số khái niệm liên quan:

1 Định nghĩa các dạng chuẩn**i Dạng Chuẩn Một (First Normal Form)**

Một lược đồ quan hệ Q ở dạng chuẩn 1 nếu toàn bộ các thuộc tính của mọi bộ đều mang giá trị đơn.

Ví dụ 1: Xét quan hệ

MASV	HOVATEN	KHOA	TENMONHOC	DIEMTHI
99023	NGUYENTHITHU	CONG NGHE THONG TIN	KY THUAT LAP TRINH TOAN ROI RAC CO SO DU LIEU	6 8 4
99030	LE VAN THANH	DIEN TU	VI XULY	4

Quan hệ này không đạt chuẩn 1 vì các thuộc tính TENMONHOC, DIEMTHI của bộ thứ nhất không mang giá trị đơn (chẳng hạn sinh viên NGUYEN THI THU có thuộc tính TENMONHOC là KY THUAT LAP TRINH, TOAN ROI RAC, CO SO DU LIEU).

Ta hoàn toàn có thể đưa quan hệ trên về dạng chuẩn 1 như sau:

MASV	HOVATEN	KHOA	TENMONHOC	DIEMTHI
99023	NGUYENTHITHU	CONG NGHE THONG TIN	KY THUAT LAP TRINH	6
99023	NGUYENTHITHU	CONG NGHE THONG TIN	TOAN ROI RAC	8
99023	NGUYENTHITHU	CONG NGHE THONG TIN	CO SO DU LIEU	4
99030	LE VAN THANH	DIEN TU	VI XULY	4

Chú ý rằng khi xét các dạng chuẩn, nếu ta không nói gì thêm, ta hiểu dạng chuẩn đang xét ít nhất là đạt dạng chuẩn 1.

ii Dạng Chuẩn 2 (Second Normal Form)

Một lược đồ quan hệ Q ở dạng chuẩn 2 nếu Q đạt chuẩn 1 và mọi thuộc tính không khóa của Q đều phụ thuộc đầy đủ vào khóa.

Thuật toán kiểm tra dạng chuẩn 2

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

Ra: khẳng định Q đạt chuẩn 2 hay không đạt chuẩn 2.

Bước 1: Tìm tất cả khóa của Q

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

Bước 3: Nếu có bao đóng S^+ chứa thuộc tính không khóa thì Q không đạt chuẩn 2

Ngược lại thì Q đạt chuẩn 2

Ví dụ 2: Cho lược đồ quan hệ $Q(A, B, C, D)$ và tập phụ thuộc hàm

$F = \{AB \rightarrow C; B \rightarrow D; BC \rightarrow A\}$. Hỏi Q có đạt chuẩn 2 không?

Giải:

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

X_i	$(TN \cup X_i)$	$(TN \cup X_i)^+$	Siêu khóa	khóa
ϕ	B	BD		
A	AB	ABCD	AB	AB
C	BC	ABCD	BC	BC
AC	ABC	ABCD	ABC	

Khóa là $K_1 = AB$ và $K_2 = BC$. Ta thấy $B \subset K_1$, $B \rightarrow D$, D là thuộc tính không khóa \Rightarrow thuộc tính không khóa không phụ thuộc đầy đủ vào khóa $\Rightarrow Q$ không đạt chuẩn 2.

Ví dụ 3: Quan hệ sau đạt chuẩn 2.

$Q(G, M, V, N, H, P)$ $F = \{G \rightarrow M; G \rightarrow N; G \rightarrow H; G \rightarrow P; M \rightarrow V; NHP \rightarrow M\}$

Giải:

$$TN = \{G\} \quad TG = \{M, N, H, P\}$$

X_i	$(TN \cup X_i)$	$(TN \cup X_i)^+$	Siêu khóa	khóa
ϕ	G	Q^+	G	G
M	GM	Q^+	GM	
N	GN	Q^+	GN	
MN	GMN	Q^+	GMN	
H	GH	Q^+	GH	
MH	GMH	Q^+	GMH	
NH	GNH	Q^+	GNH	
MNH	GMNH	Q^+	GMNH	
P	GP	Q^+	GP	
MP	GMP	Q^+	GMP	
NP	GNP	Q^+	GNP	
MNP	GMNP	Q^+	GMNP	
HP	GHP	Q^+	GHP	
MHP	GMHP	Q^+	GMHP	
NHP	GNHP	Q^+	GNHP	
MNHP	GMNHP	Q^+	GMNHP	

Lược đồ quan hệ Q chỉ có một khóa và khóa chỉ có một thuộc tính nên mọi thuộc tính đều phụ thuộc đầy đủ vào khóa $\Rightarrow Q$ đạt chuẩn 2

Hệ quả:

- Nếu Q đạt chuẩn 1 và tập thuộc tính không khóa của Q bằng rỗng thì Q đạt chuẩn 2
- Nếu tất cả khóa của quan hệ chỉ gồm một thuộc tính thì quan hệ đó ít nhất đạt chuẩn 2.

Ví dụ 4: $Q(A, B, C, D, E, H)$ $F = \{A \rightarrow E; C \rightarrow D; E \rightarrow DH\}$

Giải:

$$TN = \{ACB\} \quad TG = \{E\}$$

X_i	$(TN \cup X_i)$	$(TN \cup X_i)^+$	Siêu khóa	khóa
ϕ	ACB	ABCDEH	ACB	ACB
E	ACBE	ABCDEH	ACBE	

\Rightarrow khóa của Q là $K = \{ABC\}$. $C \subset K$, $C \rightarrow D$, D là thuộc tính không khóa $\Rightarrow D$ phụ thuộc không đầy đủ vào khóa nên Q không đạt chuẩn 2.

iii Dạng Chuẩn 3 (Third Normal Form)

Thuộc tính phụ thuộc bậc cao

Q là lược đồ quan hệ, X, Y là hai tập con của Q^+ , A là một thuộc tính.

Nói rằng A phụ thuộc bậc cao vào X nếu cả ba điều sau thỏa:

- + $X \rightarrow Y, Y \rightarrow A$
- + $Y \not\rightarrow X$
- + $A \notin XY$

Định nghĩa 1:

Lược đồ quan hệ Q ở dạng chuẩn 3 nếu mọi phụ thuộc hàm $X \rightarrow A \in F^+$ với $A \notin X$ đều có:

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

Định nghĩa 2:

Lược đồ quan hệ Q ở dạng chuẩn 3 nếu mọi thuộc tính không khóa của Q đều không phụ thuộc bậc cao vào một khóa bất kỳ của Q

Hai định nghĩa trên là tương đương, tuy nhiên việc cài đặt thuật toán kiểm tra dạng chuẩn 3 theo định nghĩa 1 thì hiệu quả hơn nhiều vì không phải kiểm tra tính phụ thuộc bậc cao.

Ta chứng minh hai định nghĩa tương đương bằng cách:

Từ định nghĩa 1 \Rightarrow không có phụ thuộc bậc cao vào một khóa bất kỳ của Q . Thật vậy:

Giả sử có phụ thuộc bậc cao vào khóa nghĩa là có $K \rightarrow Y, Y \rightarrow A, Y \not\rightarrow K$ và $A \notin KY$. $Y \rightarrow A$ là một phụ thuộc hàm nên theo định nghĩa 1 có hai trường hợp xảy ra cho Y :

- + Y là siêu khóa $\Rightarrow Y \rightarrow K$ điều này mâu thuẫn với $Y \not\rightarrow K$.
- + Y không là siêu khóa $\Rightarrow A$ là thuộc tính khóa \Rightarrow điều này trái với giả thiết $A \notin KY$

Từ định nghĩa 2 \Rightarrow nếu $X \rightarrow A \in F^+$ với $A \notin X$ thì X là siêu khóa hoặc A là thuộc tính khóa

Nếu $X \rightarrow A \in F^+$ với $A \notin X$ có X không là siêu khóa và A không là thuộc tính khóa thì dẫn đến một số điều sau:

A không là thuộc tính khóa $\Rightarrow A \notin K$

X không là siêu khóa $\Rightarrow X \not\rightarrow K$

Tóm lại ta có $K \rightarrow X, X \rightarrow A, X \not\rightarrow K$ và $A \notin KX \Rightarrow$

A phụ thuộc bậc cao vào K điều này mâu thuẫn với định nghĩa 2.

Hệ quả 1: Nếu Q đạt chuẩn 3 thì Q đạt chuẩn 2

Hệ quả 2: Nếu Q không có thuộc tính không khóa thì Q đạt chuẩn 3.

Chứng minh:

Hệ quả 1: Giả sử Q đạt dạng chuẩn 3 và có thuộc tính không khóa A không phụ thuộc hàm đầy đủ vào khóa $K \Rightarrow K' \subset K$ sao cho $K' \rightarrow A$ như vậy ta có $K \rightarrow K', K' \rightarrow A, K' \not\rightarrow K, A \notin KK' \Rightarrow Q$ có phụ thuộc bậc cao.

Hệ quả 2: mọi phụ thuộc hàm trong Q đều có vế phải là thuộc tính khóa $\Rightarrow Q$ đạt dạng chuẩn 3

Định lý:

Q là lược đồ quan hệ

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

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

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

Chứng minh:

Q đạt dạng chuẩn 3 theo định nghĩa ta suy ra mọi phụ thuộc hàm $X \rightarrow A \in F$ với $A \notin X$ có X là siêu khóa hoặc A là thuộc tính khóa.

Ngược lại ta phải chứng minh nếu mọi phụ thuộc hàm $X \rightarrow A \in F$ với $A \notin X$ có X là siêu khóa hoặc A là thuộc tính khóa thì mọi phụ thuộc hàm $X \rightarrow A \in F^+$ với $A \notin X$ cũng có X là siêu khóa hoặc A là thuộc tính khóa

Giả sử có phụ thuộc hàm $X \rightarrow A \in F^+$ với $A \notin X$ sao cho X không là siêu khóa và A không là thuộc tính khóa sẽ dẫn đến $A \in X^+ \subseteq X \cup \{\text{các thuộc tính khóa}\}$ điều này mâu thuẫn với $A \notin K$. Trước khi chứng minh $A \in X^+ \subseteq X \cup \{\text{các thuộc tính khóa}\}$ ta có nhận xét sau:

X không là siêu khóa $\Rightarrow X^+$ cũng không là siêu khóa. Theo thuật toán tìm bao đóng, X^+ được hình thành từ các $X_i \Rightarrow$ ở mỗi bước X_i cũng không là siêu khóa.

Bước cơ sở: $X_0 = X \Rightarrow X_0 \subseteq X \cup \{\text{các thuộc tính khóa}\}$

Bước qui nạp: giả sử có $X_{i-1} \subseteq X \cup \{\text{các thuộc tính khóa}\}$. Bao đóng X_i được hình thành do có $f_j = X_j \rightarrow Y_j$ để $X_{i-1} \supseteq X_j$ và $X_i = X_{i-1} \cup Y_j \Rightarrow f_j = X_j \rightarrow Y_j$ là phụ thuộc hàm có X_j không là siêu khóa $\Rightarrow f_j = X_j \rightarrow Y_j$ là phụ thuộc hàm có Y_j là thuộc tính khóa $\Rightarrow X_i = X_{i-1} \cup Y_j \subseteq X \cup \{\text{các thuộc tính khóa}\}$

Qua chứng minh trên $\Rightarrow A \in X^+ \subseteq X \cup \{\text{các thuộc tính khóa}\} \Rightarrow A \in X \cup \{\text{các thuộc tính khóa}\} \Rightarrow A \in \{\text{các thuộc tính khóa}\}$ điều này nghịch lý với điều $A \notin K$.

Thuật toán kiểm tra dạng chuẩn 3

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

Ra: khẳng định Q đạt chuẩn 3 hay không đạt chuẩn 3.

Bước 1: Tìm tất cả khóa của Q

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

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

Ví dụ 5: Cho lược đồ quan hệ $Q(A, B, C, D)$ $F = \{AB \rightarrow C; D \rightarrow B; C \rightarrow ABD\}$. Hỏi Q có đạt chuẩn 3 không?

Giải:

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

X_i	$(TN \cup X_i)$	$(TN \cup X_i)^+$	Siêu khóa	khóa
\emptyset	\emptyset	\emptyset		
A	A	A		
B	B	B		
AB	AB	ABCD	AB	AB
C	C	ABCD	C	C
AC	AC	ABCD	AC	
BC	BC	ABCD	BC	
ABC	ABC	ABCD	ABC	
D	D	BD		
AD	AD	ABCD	AD	AD
BD	BD	BD		
ABD	ABD	ABCD	ABD	

CD	CD	ABCD	CD	
ACD	ACD	ABCD	ACD	
BCD	BCD	ABCD	BCD	
ABCD	ABCD	ABCD	ABCD	

$K_1 = \{AB\}; K_2 = \{AD\}; K_3 = \{C\}$ là các khóa \Rightarrow mọi phụ thuộc hàm $X \rightarrow A \in F$ đều có A là thuộc tính khóa. Vậy Q đạt chuẩn 3

Ví dụ 6: Quan hệ sau đạt chuẩn 3. $Q(N, G, P, M) \quad F = \{NGP \rightarrow M, M \rightarrow P\}$

iv Dạng Chuẩn BC (Boyce-Codd Normal Form)

Một quan hệ Q ở dạng chuẩn BC nếu mọi phụ thuộc hàm $X \rightarrow A \in F^+$ với $A \notin X$ đều có X là siêu khóa.

Hệ quả 1: Nếu Q đạt chuẩn BC thì Q đạt chuẩn 3 (hiển nhiên do định nghĩa)

Hệ quả 2: Mỗi lược đồ có hai thuộc tính đều đạt chuẩn BC (xét phụ thuộc hàm có thể có của Q)

Định lý:

Q là lược đồ quan hệ

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

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

Chứng minh:

Q đạt dạng chuẩn BC theo định nghĩa ta suy ra mọi phụ thuộc hàm $X \rightarrow A \in F$ với $A \notin X$ có X là siêu khóa.

Ngược lại ta phải chứng minh nếu mọi phụ thuộc hàm $X \rightarrow A \in F$ với $A \notin X$ có X là siêu khóa thì mọi phụ thuộc hàm $Z \rightarrow B \in F^+$ với $B \notin Z$ cũng có Z là siêu khóa. Thật vậy, do $Z \rightarrow B$ không là phụ thuộc hàm hiển nhiên nên theo thuật toán tìm bao đóng phải có $X \rightarrow A \in F$ sao cho $Z \supseteq X$ (X là siêu khóa) \Rightarrow Z là siêu khóa.

Thuật toán kiểm tra dạng chuẩn BC

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

Ra: khẳng định Q đạt chuẩn BC hay không đạt chuẩn BC.

Bước 1: Tìm tất cả khóa của Q

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

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

Ví dụ 7: $Q(A, B, C, D, E, I) \quad F = \{ACD \rightarrow EBI; CE \rightarrow AD\}$. Hỏi Q có đạt chuẩn BC không?

Giải: $TN = \{C\} \quad TG = \{ADE\}$

X_i	$(TN \cup X_i)$	$(TN \cup X_i)^+$	Siêu khóa	khóa
ϕ	C	C		
A	AC	AC		
D	CD	CD		
AD	ACD	ABCDEI	ACD	ACD
E	CE	ABCDEI	CE	CE
AE	ACE	ABCDEI	ACE	
DE	CDE	ABCDEI	CDE	
ADE	ACDE	ABCDEI	ACDE	

$$F \equiv F_{1tt} = \{ACD \rightarrow E, ACD \rightarrow B, ACD \rightarrow I, CE \rightarrow A, CE \rightarrow D\}$$

Mọi phụ thuộc hàm của F_{1tt} đều có vế trái là siêu khóa $\Rightarrow Q$ đạt dạng chuẩn BC

Ví dụ 8: $Q(SV, MH, THAY)F = \{SV, MH \rightarrow THAY; THAY \rightarrow MH\}$

Quan hệ trên đạt chuẩn 3 nhưng không đạt chuẩn BC..

Ví dụ 9:

Chẳng hạn cho $Q(A, B, C, D)$ và $F = \{AB \rightarrow C; D \rightarrow B; C \rightarrow ABD\}$

thì Q là 3NF nhưng không là BCNF

Nếu $F = \{B \rightarrow D, A \rightarrow C, C \rightarrow ABD\}$ là 2 NF nhưng không là 3 NF

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

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

Ra: khẳng định Q đạt chuẩn gì?

Bước 1: Tìm tất cả khóa của Q

Bước 2: Kiểm tra chuẩn BC nếu đúng thì Q đạt chuẩn BC, kết thúc thuật toán ngược lại qua bước 3

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

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

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

II PHÉP TÁCH KẾT NỐI BẢO TOÀN

1 Phép tách kết nối bảo toàn thông tin (lossless-join decomposition)

Cho lược đồ quan hệ $Q(TENNCC, DIACHI, SANPHAM, DONGIA)$ có quan hệ tương ứng là r

Đặt r_1 là quan hệ có được bằng cách chiếu r lên $Q_1(TENNCC, SANPHAM, DONGIA)$,

Đặt r_2 là quan hệ có được bằng cách chiếu r lên $Q_2(TENNCC, DIACHI)$

Đặt r' là quan hệ có được bằng cách kết tự nhiên giữa r_1 và r_2 qua $TENNCC$.

chẳng hạn:

r			
TENNCC	DIACHI	SANPHAM	DONGIA
Hung	12 Nguyễn Kiệm	Gạch ống	200
Hung	12 Nguyễn Kiệm	Gạch thẻ	250
Hung	40 Nguyễn Oanh	Gạch ống	200

$r_2 = r.Q_2^+$		$r_1 = r.Q_1^+$		
TENNCC	DIACHI	TENNCC	SANPHAM	DONGIA
Hung	12 Nguyễn Kiệm	Hung	Gạch ống	200
Hung	40 Nguyễn Oanh	Hung	Gạch thẻ	250

$r' = r_1 \bowtie_{TENNCC} r_2$			
TENNCC	DIACHI	SANPHAM	DONGIA
Hung	12 Nguyễn Kiệm	Gạch ống	200

Hung	12 Nguyễn Kiệm	Gạch thẻ	250
Hung	40 Nguyễn Oanh	Gạch ống	200
Hung	40 Nguyễn Oanh	Gạch thẻ	250

Kết quả là $r \neq r'$ hay $r \neq r.Q_1 \mid \rangle \langle \mid r.Q_2$.

Với kết quả trên, ta nói phép tách $\rho(Q_1, Q_2)$ tách Q thành Q_1, Q_2 là tách-kết nối (phân rã) mất mát thông tin.

Nếu $r = r.Q_1 \mid \rangle \langle \mid r.Q_2$ ta nói phép tách $\rho(Q_1, Q_2)$ là tách-kết nối không mất mát thông tin (tách kết nối bảo toàn thông tin hay phân rã bảo toàn thông tin).

Vậy với điều kiện nào thì phép tách trở thành tách-kết nối không mất mát thông tin?

i Định nghĩa phép tách Q thành 2 lược đồ con

Q là lược đồ quan hệ, Q_1, Q_2 hai lược đồ con có:

$$Q_1^+ \cap Q_2^+ = X$$

$$Q_1^+ \cup Q_2^+ = Q^+$$

Nói rằng lược đồ quan hệ Q được tách thành hai lược đồ con Q_1, Q_2 theo phép tách $\rho(Q_1, Q_2)$ là phép tách kết nối không mất (hay phép tách bảo toàn thông tin) nếu với r là quan hệ bất kỳ của Q ta có:

$$r = r.Q_1 \mid \rangle \langle \mid r.Q_2$$

Tức là r được tạo nên từ phép kết nối tự nhiên của các hình chiếu của nó trên các Q_1, Q_2

ii Tính chất

Nếu Q là một lược đồ quan hệ, Q_1, Q_2 là hai lược đồ quan hệ con có

$$Q_1^+ \cap Q_2^+ = X$$

$$Q_1^+ \cup Q_2^+ = Q^+$$

$$X \rightarrow Q_2^+$$

Thì $r = r.Q_1 \mid \rangle \langle \mid r.Q_2$

Chứng minh:

$$\underline{t \in r \Rightarrow t \in r.Q_1 \mid \rangle \langle \mid r.Q_2}$$

$$t \in r \Rightarrow \exists t_1 \in r_1 \ t_1 = t.Q_1 \ \exists t_2 \in r_2 \ t_2 = t.Q_2 \ t_1.X = t_2.X = t.X$$

$$\Rightarrow t \in r.Q_1 \mid \rangle \langle \mid r.Q_2 \text{ (Theo định nghĩa)}$$

$$\underline{t \in r.Q_1 \mid \rangle \langle \mid r.Q_2 \Rightarrow t \in r}$$

$$t \in r.Q_1 \mid \rangle \langle \mid r.Q_2 \Rightarrow \exists t_1 \in r_1 \ t_1 = t.Q_1 \text{ (1)}$$

$$\text{mà } t_1 \in r_1 = r.Q_1 \text{ nên theo định nghĩa phép chiếu ta lại có } \exists t' \in r \ t_1 = t'.Q_1 \text{ (2)}$$

$$(1) \text{ và } (2) \Rightarrow t'.Q_1 = t.Q_1 \Rightarrow t'.X = t.X \Rightarrow t'.Q_2 = t.Q_2 \text{ (do } X \rightarrow Q_2)$$

$$\Rightarrow t' = t \Rightarrow t \in r$$

Ví dụ 10: cho $Q(SAIP)$, $Q_1 = (SA)$, $Q_2 = (SIP)$ $F = \{S \rightarrow A, SI \rightarrow P\}$. Hỏi việc tách Q thành Q_1 và Q_2 có gây ra mất mát thông tin không?

Áp dụng tính chất trên, ta có

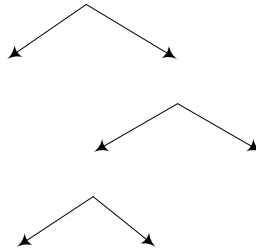
$$Q_1^+ \cap Q_2^+ = S$$

$$Q_1^+ \cup Q_2^+ = SAIP = Q^+$$

$$S \rightarrow SA = Q_1^+$$

Theo tính chất trên, với mọi quan hệ r của Q ta luôn có $r = r.Q_1 \bowtie^S r.Q_2$. Suy ra phép tách trên là phép tách kết nối bảo toàn thông tin.

iii Phép tách Q thành n lược đồ con



Q là một lược đồ quan hệ, F là tập phụ thuộc hàm. Q được tách thành các lược đồ con $Q_1, Q_2, Q_3, \dots, Q_n$ theo từng bước mà ở mỗi bước một lược đồ được tách thành hai lược đồ con và thỏa mãn điều kiện của tính chất bảo toàn thông tin thì với r là quan hệ bất kỳ của Q ta luôn có:

$$r = r.Q_1 \bowtie^S r.Q_2 \bowtie^S r.Q_3 \dots \bowtie^S r.Q_n$$

Chứng minh:

Ta chứng minh bằng phương pháp qui nạp.

Ở bước $i = 1$ thì $r = r.Q_1 \bowtie^S r.Q_{1m}$ đúng theo định lý bảo toàn thông tin

Giả sử biểu thức trên đúng ở bước $i = k$ nghĩa là ta có:

$$r = r.Q_1 \bowtie^S r.Q_2 \bowtie^S r.Q_3 \dots \bowtie^S r.Q_k \bowtie^S r.Q_{km} \quad (1)$$

ta phải chứng minh $r = r.Q_1 \bowtie^S r.Q_2 \bowtie^S r.Q_3 \dots \bowtie^S r.Q_k \bowtie^S r.Q_{k+1} \bowtie^S r.Q_{k+1m}$

Với Q_{km} được tách thành hai lược đồ con Q_{k+1} và Q_{k+1m} theo đúng điều kiện của tính chất bảo toàn thông tin nghĩa nếu s là quan hệ của Q_{km} thì $s = s.Q_{k+1} \bowtie^S s.Q_{k+1m} \Rightarrow$

$$r.Q_{km} = (r.Q_{km}).Q_{k+1} \bowtie^S (r.Q_{km}).Q_{k+1m} = r.Q_{k+1} \bowtie^S r.Q_{k+1m} \Rightarrow$$

$$r = r.Q_1 \bowtie^S r.Q_2 \bowtie^S r.Q_3 \dots \bowtie^S r.Q_k \bowtie^S r.Q_{k+1} \bowtie^S r.Q_{k+1m}$$

iv Thuật toán kiểm tra phép tách kết nối bảo toàn thông tin

(a) Thuật toán

Dữ liệu vào: lược đồ quan hệ $Q(A_1, A_2, \dots, A_n)$, tập phụ thuộc hàm F , phép tách $\rho = (Q_1, Q_2, \dots, Q_k)$.

Dữ liệu ra: kết luận phép tách ρ có phải là phép tách bảo toàn thông tin?

1. Thiết lập bảng với $k+1$ dòng, $n+1$ cột. Cột j ứng với thuộc tính A_j ($j=1 \dots n$), hàng i ứng với lược đồ quan hệ Q_i ($i=1 \dots k$). Tại vị trí hàng i , cột j ta điền ký hiệu A_j nếu $A_j \in Q_i$, nếu không ta đặt ký hiệu b_t vào vị trí đó. (với t đầu tiên bằng 1) và sau đó tăng t lên một đơn vị.
2. Xét lần lượt các phụ thuộc hàm trong F , áp dụng cho bảng vừa mới thành lập ở trên. Giả sử xét $(X \rightarrow Y) \in F$, chúng ta tìm những hàng giống nhau ở tất cả các thuộc tính của X , nếu thấy những hàng như vậy ta sẽ làm cho các ký hiệu của hai hàng này bằng nhau ở tất cả các thuộc tính của Y . Khi làm cho 2 ký hiệu này bằng nhau, nếu một trong hai ký hiệu là a_j thì cho ký hiệu kia trở thành a_j , nếu hai ký hiệu là b_k hoặc b_t thì có thể cho chúng trở thành b_t hoặc b_l (với $t = \min(k, l)$). Bước này được tiếp tục cho các phụ thuộc hàm còn lại của F cho đến khi không còn áp dụng được nữa.
3. Xét bảng kết quả, nếu thấy trong bảng này có một hàng chứa toàn a_j ($i=1 \dots n$) thì kết luận đó là phép kết nối bảo toàn thông tin, ngược lại là phép kết nối mất mát thông tin.

Chú ý: một điều quan trọng cần phải nhớ là khi cho hai ký hiệu bằng nhau thì phải cho bằng nhau ở tất cả các xuất hiện của chúng trong bảng chứ không phải chỉ cho bằng nhau ở những ký hiệu trong phạm vi các phụ thuộc $X \rightarrow Y \in F$.

Ví dụ 11: Với $Q(ABCDE)$

$Q_1 = (AD), Q_2 = (AB), Q_3 = (BE), Q_4 = (CDE), Q_5 = (AE)$

$F = \{A \rightarrow C, B \rightarrow C, A \rightarrow D, DE \rightarrow C, CE \rightarrow A\}$

Kiểm tra tính bảo toàn thông tin của phép phân rã Q thành Q_1, Q_2, Q_3, Q_4, Q_5 .

Bước 1:

	a_1	a_2	a_3	a_4	a_5
	A	B	C	D	E
$Q_1(AD)$	a_1			a_4	
$Q_2(AB)$	a_1	a_2			
$Q_3(BE)$		a_2			a_5
$Q_4(CDE)$			a_3	a_4	a_5
$Q_5(AE)$	a_1				a_5

Bước 2: Điền b_1, b_2, b_3, \dots

	A	B	C	D	E
$Q_1(AD)$	a_1	b_1	b_2	a_4	b_3
$Q_2(AB)$	a_1	a_2	b_4	b_5	b_6
$Q_3(BE)$	b_7	a_2	b_8	b_9	a_5
$Q_4(CDE)$	b_{10}	b_{11}	a_3	a_4	a_5
$Q_5(AE)$	a_1	b_{12}	b_{13}	b_{14}	a_5

Sửa bảng giá trị để nó thỏa $A \rightarrow C$

Sửa b_4, b_{13} thành b_2

	A	B	C	D	E
$Q_1(AD)$	a_1	b_1	b_2	a_4	b_3
$Q_2(AB)$	a_1	a_2	b_2	b_5	b_6
$Q_3(BE)$	b_7	a_2	b_8	b_9	a_5
$Q_4(CDE)$	b_{10}	b_{11}	a_3	a_4	a_5
$Q_5(AE)$	a_1	b_{12}	b_2	b_{14}	a_5

Sửa bảng giá trị để nó thỏa $B \rightarrow C$

Sửa b_8 thành b_2

	A	B	C	D	E
$Q_1(AD)$	a_1	b_1	b_2	a_4	b_3
$Q_2(AB)$	a_1	a_2	b_2	b_5	b_6
$Q_3(BE)$	b_7	a_2	b_2	b_9	a_5
$Q_4(CDE)$	b_{10}	b_{11}	a_3	a_4	a_5
$Q_5(AE)$	a_1	b_{12}	b_2	b_{14}	a_5

Sửa bảng giá trị để nó thỏa $A \rightarrow D$

Sửa b_5, b_{14} thành a_4

	A	B	C	D	E
$Q_1(AD)$	a_1	b_1	b_2	a_4	b_3
$Q_2(AB)$	a_1	a_2	b_2	a_4	b_6
$Q_3(BE)$	b_7	a_2	b_2	b_9	a_5
$Q_4(CDE)$	b_{10}	b_{11}	a_3	a_4	a_5
$Q_5(AE)$	a_1	b_{12}	b_2	a_4	a_5

Sửa bảng giá trị để nó thỏa $DE \rightarrow C$

sửa b_2 thành $a_3 \Rightarrow$ sửa tất cả b_2

thành a_3

	A	B	C	D	E
$Q_1(AD)$	a_1	b_1	a_3	a_4	b_3
$Q_2(AB)$	a_1	a_2	a_3	a_4	b_6
$Q_3(BE)$	b_7	a_2	a_3	b_9	a_5
$Q_4(CDE)$	b_{10}	b_{11}	a_3	a_4	a_5
$Q_5(AE)$	a_1	b_{12}	a_3	a_4	a_5

Sửa bảng giá trị để nó thỏa $CE \rightarrow A$

Sửa b_7, b_{10} thành a_1 .

	A	B	C	D	E
$Q_1(AD)$	a_1	b_1	a_3	a_4	b_3
$Q_2(AB)$	a_1	a_2	a_3	a_4	b_6
$Q_3(BE)$	a_1	a_2	a_3	b_9	a_5
$Q_4(CDE)$	a_1	b_{11}	a_3	a_4	a_5
$Q_5(AE)$	a_1	b_{12}	a_3	a_4	a_5

Lần lượt xét lại các phụ thuộc hàm trong F , nếu bảng giá trị chưa thỏa phụ thuộc hàm nào thì tiếp tục làm cho nó thỏa.

Sửa bảng giá trị để nó thỏa $A \rightarrow D$

	A	B	C	D	E
$Q_1(AD)$	a_1	b_1	a_3	a_4	b_3
$Q_2(AB)$	a_1	a_2	a_3	a_4	b_6
$Q_3(BE)$	a_1	a_2	a_3	a_4	a_5
$Q_4(CDE)$	a_1	b_{11}	a_3	a_4	a_5
$Q_5(AE)$	a_1	b_{12}	a_3	a_4	a_5

Dòng thứ Q_3 (BE) của bảng chứa toàn giá trị a_j ($j=1..n$) nên phép phân rã trên là bảo toàn thông tin.

(b) Định lý

Bảng kết quả của thuật toán trên cho phép ta kết luận được tính bảo toàn hay không bảo toàn thông tin của phép tách.

Chứng minh:

Ta chứng minh nếu bảng kết quả thuật toán không có hàng chỉ chứa toàn giá trị a thì phép tách không bảo toàn thông tin. Thật vậy:

Ta xây dựng một quan hệ r có các giá trị như bảng kết quả của thuật toán, các hàng là các bộ. Quan hệ r thỏa tập phụ thuộc F vì thuật toán đã sửa các giá trị của r để nó khỏi vi phạm các phụ thuộc hàm trong $F \Rightarrow r$ là một quan hệ của lược đồ Q . Ta tách quan hệ r thành các quan hệ r_i với $r_i = r.Q_i$ và dùng phép kết tự nhiên để kết chúng lại. Nếu:

- + $\exists k Q_k^+ \cap Q_i^+ = \emptyset \forall i \Rightarrow r_1 |><| r_2 \dots |><| r_k$ không tồn tại \Rightarrow phép tách không bảo toàn thông tin.
- + $\forall i, \exists k Q_i^+ \cap Q_k^+ = X_{ik} \neq \emptyset$ mà mỗi r_i đều có một bộ t_i chứa toàn $a \Rightarrow$ các t_i nối được với nhau vì có cùng giá trị trên $X_{ik} \Rightarrow$ có một bộ $t \in r_1 |><| r_2 \dots |><| r_k$ có toàn giá trị a , bộ này lại không có trong $r \Rightarrow r \neq r_1 |><| r_2 \dots |><| r_k \Rightarrow$ phép tách không bảo toàn thông tin.

Ta chứng minh nếu bảng kết quả thuật toán có hàng chỉ chứa toàn giá trị a thì phép tách bảo toàn thông tin. Ta chứng minh điều này qua 2 bước:

+ **Bước 1:** chứng minh nếu $t \in r \Rightarrow t \in r_1 |><| r_2 \dots |><| r_k$. Suy ra $r \subseteq r_1 |><| r_2 \dots |><| r_k$. Giả sử $t = (a_1, \dots, a_n) \in r$. Ta tách quan hệ r thành các $r_i = r.Q_i$ với $t_i = t.Q_i$. Có hai trường hợp:

- o $\forall i, \exists k Q_i^+ \cap Q_k^+ = X_{ik} \neq \emptyset \Rightarrow$ các t_i nối được với nhau vì có cùng giá trị trên $X_{ik} \Rightarrow$ bộ $t \in r_1 |><| r_2 \dots |><| r_k \Rightarrow r \subseteq r_1 |><| r_2 \dots |><| r_k$.
- o $\exists k Q_k^+ \cap Q_i^+ = \emptyset \forall i$. Suy ra bảng kiểm tra bảo toàn thông tin ở giai đoạn chưa thỏa các phụ thuộc hàm, có dạng:

	A_1	A_2	...	A_k	A_{k+1}	...
Q_1				b_{k1}	b_{k2}	$b..$
Q_2				$b..$
...				$b..$
$Q_k(A_k, A_{k+1}, \dots)$	$b..$	$b..$	$b..$	a_k	a_{k+1}	...

Với mọi $X \subseteq Q^+ \quad t_k.X \neq t_i.X$ với $i \neq k$ nên khi làm bằng các giá trị theo các phụ thuộc hàm $X \rightarrow Y$ thì các giá trị b ở dòng Q_k không thay đổi còn các giá trị b ở các cột A_k, A_{k+1}, \dots không đổi thành a được. Suy ra bảng kết quả của thuật toán không bao giờ chứa dòng có toàn giá trị a . Vậy trường hợp $\exists k Q_k^+ \cap Q_i^+ = \emptyset \forall i$ không xảy ra khi bảng kiểm tra bảo toàn thông tin có một dòng toàn a .

+ **Bước 2:** chứng minh nếu $t \in r_1 |><| r_2 \dots |><| r_k \Rightarrow t \in r$. Suy ra $r_1 |><| r_2 \dots |><| r_k \subseteq r$. Giả sử $t = (a_1, \dots, a_n) \in r_1 |><| r_2 \dots |><| r_k$ theo định nghĩa suy ra $\forall i \exists t_i \in r_i$ sao cho $t.Q_i^+ = t_i$. Nhưng $r_i = r.Q_i^+ \Rightarrow \exists t_i' \in r$ sao cho $t_i'.Q_i^+ = t_i = t.Q_i^+ \forall i$. Trường hợp xấu nhất là các t_i' là các dòng khác nhau. Trong trường hợp này, ta có thể xem t_i' là dòng Q_i của bảng kiểm tra bảo toàn thông tin với các giá trị b xem như chưa biết. Nhưng các dòng Q_i phải thỏa các phụ thuộc hàm trong F , phép làm bằng các giá trị theo các phụ thuộc hàm đã dần dần xác định

được tất cả các giá trị b của một dòng t_i , nào đó, là dòng có toàn giá trị a. Vậy có một i' để $t_{i'} = t \Rightarrow t \in r \Rightarrow r \supseteq r_1 | > < | r_2 \dots | > < | r_n$ (2)

(1) và (2) $\Rightarrow r = r_1 | > < | r_2 \dots | > < | r_n$. Nói cách khác phép tách bảo toàn thông tin.

2 Phép tách bảo toàn phụ thuộc hàm (decompositions that preserve dependencies)

i Tập phụ thuộc hàm F_i của Q_i

Phần trên chỉ đề cập vấn đề tách một lược đồ quan hệ $Q(A_1, A_2, \dots, A_n)$ thành các lược đồ con Q_1, Q_2, \dots, Q_k còn không đề cập đến tập phụ thuộc hàm của các lược đồ con này. Nếu $Q(A_1, A_2, \dots, A_n)$ là lược đồ quan hệ, F phụ thuộc hàm, $\rho = (Q_1, Q_2, \dots, Q_k)$ là phép phân rã bảo toàn thông tin, r_i là quan hệ của Q_i thì tính chất sau thỏa:

+ r_i chỉ thỏa các phụ thuộc hàm $X \rightarrow Y \in F^+$ với $XY \subseteq Q_i^+$

Nói cách khác, tập phụ thuộc hàm của Q_i chính là F_i có $F_i^+ = \{X \rightarrow Y \in F^+ \mid XY \subseteq Q_i^+\}$. Ta có thể hiểu F được phân rã thành các F_1, \dots, F_k

Chứng minh tính chất trên:

Do r_i được tách từ r mà r thỏa $F^+ \Rightarrow r_i$ thỏa các phụ thuộc hàm $X \rightarrow Y \in F^+$ với $XY \subseteq Q_i^+$. Theo định nghĩa phụ thuộc hàm, đương nhiên r_i không thỏa các phụ thuộc hàm $X \rightarrow Y \in F^+$ với $XY \not\subseteq Q_i^+$. Ngoài ra r_i không thỏa bất kỳ một phụ thuộc hàm nào $X \rightarrow Y \notin F^+$. Thật vậy nếu có $X \rightarrow Y$ như vậy thì $r = r_1 | > < | r_2 \dots | > < | r_n$ cũng phải thỏa $X \rightarrow Y \notin F^+$. Điều này mâu thuẫn với định nghĩa của tập F^+ .

ii Định nghĩa:

Cho phân rã $\rho = (Q_1, Q_2, \dots, Q_k)$ của một lược đồ quan hệ, và một tập phụ thuộc hàm F . Hình chiếu của F trên một tập các thuộc tính Q_i^+ ký hiệu $\Pi_{Q_i}(F)$ là tập các phụ thuộc hàm $X \rightarrow Y \in F^+$ sao cho $XY \subseteq Q_i$.

$$\Pi_{Q_i}(F) = F_i^+ = \{X \rightarrow Y \mid X \rightarrow Y \in F^+ \text{ và } XY \subseteq Q_i\}$$

Ta nói phân rã ρ bảo toàn tập phụ thuộc hàm F nếu

$$F \equiv \cup \Pi_{Q_i}(F) \Leftrightarrow F^+ = (\cup \Pi_{Q_i}(F))^+ \text{ với } i=1..k$$

Hệ quả: $F^+ \supseteq (\cup \Pi_{Q_i}(F))^+$ với $i=1..k$

Nhận xét: từ hệ quả trên ta suy ra: để xác định phép phân rã $\rho = (Q_1, Q_2, \dots, Q_k)$ có bảo toàn phụ thuộc hàm hay không, với mỗi phụ thuộc hàm $X \rightarrow Y \in F$ ta xác định xem nó có là thành viên của tập phụ thuộc hàm $G = \cup \Pi_{Q_i}(F)$ hay không. Ta không cần xác định chiều ngược lại.

Ví dụ 12: Cho lược đồ quan hệ $Q(A, B, C)$ và $F = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$. Phép phân rã $\rho = (Q_1, Q_2)$ tách Q thành hai lược đồ quan hệ $Q_1(A, B)$ và $Q_2(B, C)$. Hãy tính hình chiếu của F trên Q_1^+ và Q_2^+ . Phép phân rã có bảo toàn phụ thuộc hàm F không?

Giải: về nguyên tắc ta có thể giải bài toán theo các bước dưới đây

Bước 1: Kê tất cả tập con của Q^+

	A	B	C
\emptyset	A	B	C
		AB	AC
			BC
			ABC

Bước 2: Tính bao đóng của các tập con của Q^+

$$\emptyset^+ = \emptyset \quad A^+ = ABC \quad B^+ = ABC \quad C^+ = ABC$$

$$\begin{array}{lcl}
 AB^+ & = & ABC \\
 AC^+ & = & ABC \\
 BC^+ & = & ABC \\
 ABC^+ & = & ABC
 \end{array}$$

Bước 3: Tính F^+

$A \rightarrow B$	$B \rightarrow A$	$C \rightarrow A$	$AB \rightarrow ABC$	$AC \rightarrow B$	$BC \rightarrow A$
$A \rightarrow AB$	$B \rightarrow AB$	$C \rightarrow B$	$AB \rightarrow C$	$AC \rightarrow AB$	$BC \rightarrow AB$
$A \rightarrow C$	$B \rightarrow C$	$C \rightarrow AB$	$AB \rightarrow BC$	$AC \rightarrow BC$	$BC \rightarrow AC$
$A \rightarrow AC$	$B \rightarrow AC$	$C \rightarrow AC$	$AB \rightarrow ABC$	$AC \rightarrow ABC$	$BC \rightarrow ABC$
$A \rightarrow BC$	$B \rightarrow BC$	$C \rightarrow BC$			
$A \rightarrow ABC$	$B \rightarrow ABC$	$C \rightarrow ABC$			

Bước 4: Tính $\Pi_{Q_1}(F)$, $\Pi_{Q_2}(F)$

$$\Pi_{Q_1}(F) = F_1^+ = \{A \rightarrow B, A \rightarrow AB, B \rightarrow A, B \rightarrow AB\} \equiv \{A \rightarrow B, B \rightarrow A\} \text{ (chỉ lấy pth có vế phải 1 tt)}$$

$$\Pi_{Q_2}(F) = F_2^+ = \{B \rightarrow C, B \rightarrow BC, C \rightarrow B, C \rightarrow BC\} \equiv \{B \rightarrow C, C \rightarrow B\} \text{ (chỉ lấy pth có vế phải 1 tt)}$$

Bước 5:

$$G = \Pi_{Q_1}(F) \cup \Pi_{Q_2}(F) = \{A \rightarrow B, A \rightarrow AB, B \rightarrow A, B \rightarrow AB, B \rightarrow C, B \rightarrow BC, C \rightarrow B, C \rightarrow BC\}$$

$F = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$ có $A \rightarrow B$, $B \rightarrow C$ đều là thành viên của G , còn $C \rightarrow A$ có là thành viên của G hay không ta tính C_G^+ . $C_G^+ = ABC \Rightarrow C \rightarrow A$ cũng là thành viên của G . Vậy phép phân rã trên bảo toàn phụ thuộc hàm.

Bài toán trên có thể được giải theo các bước đơn giản sau cho từng lược đồ quan hệ con:

Tính cho Q_1

Bước 1: Kê tất cả tập con của Q_1^+

$$\begin{array}{l}
 A \quad B \\
 \emptyset \quad A \quad B \\
 AB
 \end{array}$$

Bước 2: Tính bao đóng của các tập con của Q_1^+

$$\begin{array}{lcl}
 \emptyset^+ = \emptyset & A^+ = & ABC \\
 B^+ & = & ABC \\
 AB^+ & = & ABC
 \end{array}$$

Bước 3: Tính $F_1^+ = \Pi_{Q_1}(F)$

$$\begin{array}{lcl}
 A \rightarrow B & B \rightarrow A \\
 A \rightarrow AB & B \rightarrow AB
 \end{array}$$

Tính cho Q_2

Bước 4: Kê tất cả tập con của Q_2^+

$$\begin{array}{l}
 B \quad C \\
 \emptyset \quad B \quad C \\
 BC
 \end{array}$$

Bước 5: Tính bao đóng của các tập con của Q_2^+

$$\begin{array}{lcl}
 \emptyset^+ = \emptyset & B^+ = & ABC \\
 C^+ & = & ABC \\
 BC^+ & = & ABC
 \end{array}$$

Bước 6: Tính $F_2^+ = \Pi_{Q_2}(F)$

$$\begin{array}{lcl}
 B \rightarrow C & C \rightarrow B \\
 B \rightarrow BC & C \rightarrow BC
 \end{array}$$

Bước 7:

$$G = \Pi_{Q_1}(F) \cup \Pi_{Q_2}(F) = \{A \rightarrow B, A \rightarrow AB, B \rightarrow A, B \rightarrow AB, B \rightarrow C, B \rightarrow BC, C \rightarrow B, C \rightarrow BC\}$$

$F = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$ có $A \rightarrow B$, $B \rightarrow C$ đều là thành viên của G còn $C \rightarrow A$ có là thành viên của G hay không ta tính C_G^+ . $C_G^+ = ABC \Rightarrow C \rightarrow A$ cũng là thành viên của G . Vậy phép phân rã trên bảo toàn phụ thuộc hàm.

iii Ý nghĩa của phân rã có bảo toàn phụ thuộc hàm

Ví dụ 13: Cho lược đồ quan hệ $Q(C, S, Z)$ và $F = \{CS \rightarrow Z, Z \rightarrow C\}$. Phép tách $\rho = (Q_1, Q_2)$ tách Q thành hai lược đồ $Q_1(S, Z)$ và $Q_2(C, Z)$. Hỏi phép tách có bảo toàn phụ thuộc hàm không?

Giải:

Q_1 có các tập thuộc tính con:

	S	Z
\emptyset	S	Z
	SZ	

Bao đóng của các tập thuộc tính con Q_1^+

$$\emptyset^+ = \emptyset \quad S^+ = S \quad Z^+ = ZC \\ SZ^+ = CSZ$$

F_1^+ chỉ gồm các phụ thuộc hàm hiển nhiên vì tất cả các phụ thuộc hàm sau đều không thỏa:

$$\begin{aligned} Z \rightarrow C \quad SZ \rightarrow C \\ Z \rightarrow ZC \quad SZ \rightarrow CS \\ SZ \rightarrow CZ \\ SZ \rightarrow CSZ \end{aligned}$$

Q_2 có các tập thuộc tính con:

	C	Z
\emptyset	C	Z
	CZ	

Bao đóng của các tập thuộc tính con Q_2^+

$$\emptyset^+ = \emptyset \quad C^+ = C \quad Z^+ = ZC \\ CZ^+ = CZ$$

F_2^+ gồm các phụ thuộc:

$$Z \rightarrow C \quad Z \rightarrow ZC$$

$\Pi_{Q_1}(F) \cup \Pi_{Q_2}(F) = \{Z \rightarrow C, Z \rightarrow ZC\} \equiv \{Z \rightarrow C\}$ không tương đương với $F = \{CS \rightarrow Z, Z \rightarrow C\}$

Vậy phép phân rã trên không bảo toàn phụ thuộc hàm, điều này có nghĩa khi ta đưa dữ liệu vào Q_1 và Q_2 sao cho không vi phạm phụ thuộc hàm hình chiếu của nó, nhưng khi kết nối chúng lại thì dữ liệu kết quả của lược đồ quan hệ Q lại vi phạm phụ thuộc hàm $CS \rightarrow Z$

$\Pi_{Q_1}(F) = \{PTHHN\}$	$\Pi_{Q_2}(F) = \{Z \rightarrow C, Z \rightarrow ZC\}$	$F = \{CS \rightarrow Z, Z \rightarrow C\}$
$Q_1 \quad (S \quad Z)$	$Q_2 \quad (C \quad Z)$	$Q \quad (C \quad S \quad Z)$
$s_1 \quad z_1$	$c_1 \quad z_1$	$c_1 \quad s_1 \quad z_1$
$s_1 \quad z_2$	$c_1 \quad z_2$	$c_1 \quad s_1 \quad z_2$

iv Thuật toán kiểm tra bảo toàn phụ thuộc hàm

Thuật toán tìm bao đóng của tập thuộc tính X đối với $G = \cup \Pi_{Q_i}(F)$

Vào: $\rho = (Q_1, Q_2, \dots, Q_k), F, X$

Ra: X_G^+

Bước 1: Với mỗi phụ thuộc hàm $X \rightarrow Y \in F$ ta thực hiện từ bước 2 đến bước 4

Bước 2: đặt $Z' = X$

Bước 3: thế $Z' = Z' \cup ((Z' \cap Q_i^+)^+ \cap Q_i^+)$

Bước 4: nếu ở Q_i , Z' thay đổi thì thực hiện lại bước 3 cho $Q_{đầu tiên}$

Ngược lại kết thúc thuật toán và trả về Z' (là bao đóng X_G^+)

Thuật toán kiểm tra bảo toàn phụ thuộc hàm

Vào: $\rho = (Q_1, Q_2, \dots, Q_k), F$

Ra: kết luận phép tách ρ bảo toàn hay không bảo toàn phụ thuộc hàm

Bước 1: Với mỗi phụ thuộc hàm $X \rightarrow Y \in F$ ta thực hiện từ bước 2 đến bước 3:

Bước 2: Tìm bao đóng X_G^+ với $G = \bigcup \Pi_{Q_i}(F)$

Bước 3: Nếu $Y \subseteq X_G^+$ thì $X \rightarrow Y \in \bigcup \Pi_{Q_i}(F)^+$

Bước 4: Nếu tất cả phụ thuộc $X \rightarrow Y \in F$ đều thuộc $\bigcup \Pi_{Q_i}(F)^+$ thì ta kết luận phân rã ρ bảo toàn phụ thuộc hàm ngược lại ρ không bảo toàn phụ thuộc hàm

Ví dụ 14: thực hiện lại ví dụ 13, nghĩa là kiểm tra phép tách có bảo toàn phụ thuộc hàm không?

Vào: $Q(C, S, Z), F = \{CS \rightarrow Z, Z \rightarrow C\}, Q_1(S, Z)$ và $Q_2(C, Z)$

Đương nhiên $Z \rightarrow C \in G = \Pi_{Q_1}(F) \cup \Pi_{Q_2}(F) \Rightarrow Z \rightarrow C \in (\Pi_{Q_1}(F) \cup \Pi_{Q_2}(F))^+$

1. $Z' = CS$

2. gán $Z' = Z' \cup ((Z' \cap Q_1^+)^+ \cap Q_1^+)$: $Z' = CS \cup (S \cap SZ) = CS$

Bước 1 và 2 có Z' không thay đổi, ta sang lược đồ Q_2 và tính tiếp Z'

3. gán $Z' = Z' \cup ((Z' \cap Q_2^+)^+ \cap Q_2^+)$: $Z' = CS \cup (C \cap CZ) = CS$

Z' không thay đổi và hết lược đồ quan hệ \Rightarrow ngừng không tính tiếp Z'

4. Vậy $CS_G^+ = CS \Rightarrow CS \rightarrow Z \notin (\Pi_{Q_1}(F) \cup \Pi_{Q_2}(F))^+$ phép phân rã không bảo toàn phụ thuộc hàm.

Ví dụ 15: thực hiện lại ví dụ 12 với nội dung kết luận phép tách ρ có bảo toàn phụ thuộc hàm không (không tính F^+)

Vào: $Q(A, B, C), F = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}, Q_1(A, B)$ và $Q_2(B, C)$

Hiển nhiên $G = \Pi_{Q_1}(F) \cup \Pi_{Q_2}(F) \supseteq \{A \rightarrow B, B \rightarrow C\}$

Ta xác định $C \rightarrow A$ có thuộc $(\Pi_{Q_1}(F) \cup \Pi_{Q_2}(F))^+$

1. $Z' = C$

2. gán $Z' = Z' \cup ((Z' \cap Q_1^+)^+ \cap Q_1^+)$: $Z' = C \cup (\emptyset \cap AB) = C$

Bước 1 và 2 có Z' không thay đổi, ta sang lược đồ Q_2 và tính tiếp Z'

3. gán $Z' = Z' \cup ((Z' \cap Q_2^+)^+ \cap Q_2^+)$: $Z' = C \cup (ABC \cap BC) = BC$

Z' thay đổi \Rightarrow tính tiếp Z' bắt đầu từ lược đồ Q_1

4. gán $Z' = Z' \cup ((Z' \cap Q_1^+)^+ \cap Q_1^+)$: $Z' = BC \cup (ABC \cap AB) = ABC$

do $Z' = Q^+ \Rightarrow Z'$ sẽ không bao giờ thay đổi.

5. vậy $C_G^+ = ABC \Rightarrow C \rightarrow A \in (\Pi_{Q_1}(F) \cup \Pi_{Q_2}(F))^+$ phép phân rã bảo toàn phụ thuộc hàm.

III THIẾT KẾ CSDL BẰNG CÁCH PHÂN RÃ

1 Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

i Cách thông thường

Thuật toán phân rã Q, F thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

Bước 1: Tìm tất cả khóa của Q

Bước 2: Tìm phụ thuộc hàm $X \rightarrow Y \in F$ có X không là siêu khóa và Y không chứa thuộc tính khóa.

Nếu tìm thấy thì tách Q thành Q_1 và Q_2 theo quy tắc sau:

$Q_1 = Q[XY]$; $F_1 \equiv \Pi_{Q_1}(F)$ tìm bao đóng của tất cả tập con của XY để suy ra $\Pi_{Q_1}(F) \Rightarrow F_1$
 $Q_2 = Q[Q^+ - Y]$ $F_2 \equiv \Pi_{Q_2}(F)$ tìm bao đóng của tất cả tập con của $Q^+ - Y$ để suy ra $\Pi_{Q_2}(F) \Rightarrow F_2$
 thực hiện thuật toán phân rã (Q_1, F_1)
 thực hiện thuật toán phân rã (Q_2, F_2)

Ngược lại nếu không tìm thấy thì có hai trường hợp:

Trường hợp 1: mọi phụ thuộc hàm trong F_i đều có vế trái là siêu khóa thì Q_i đạt chuẩn BC

Trường hợp 2: nếu có phụ thuộc hàm có vế trái không là siêu khóa và vế phải là thuộc tính khóa thì Q_i đạt chuẩn 3.

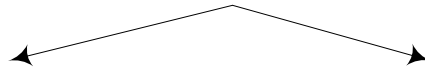
Ví dụ 16: cho $Q(S, D, I, M)$ $F = \{SI \rightarrow D; SD \rightarrow M\}$ hãy phân rã Q thành các lược đồ con đạt chuẩn BC bảo toàn thông tin

Giải:

Bước 1: tìm tất cả khóa của Q

X_i	$TN \cup X_i$	$(TN \cup X_i)^+$	Siêu khóa	Khóa
\emptyset	SI	SDIM	SI	SI
D	SID	SDIM	SID	

Bước 2: phụ thuộc hàm $SD \rightarrow M \in F$ có SD không là siêu khóa.



Chú ý: để tính được F_1, F_2, K_1, K_2 như hình trên, ta phải tính bao đóng của tất cả tập con của $\{SDM\}$ và $\{SDI\} \Rightarrow F_1, F_2$ rồi tìm tất cả khóa của Q_1 và Q_2 .

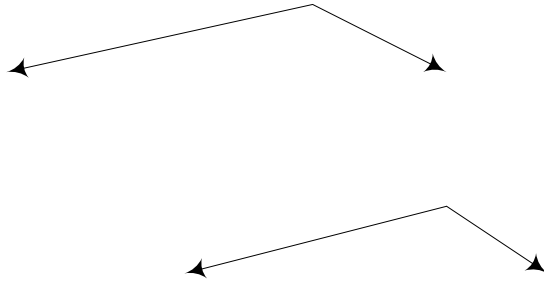
$$\begin{array}{llll}
 S^+ = S & D^+ = D & M^+ = M & S^+ = S \quad D^+ = D \quad I^+ = I \\
 SD^+ = SDM & SM^+ = SM & & SD^+ = SDM \quad SI^+ = SDIM \\
 & DM^+ = DM & & DI^+ = DI \\
 & SDM^+ = SDM & & SDI^+ = SDIM
 \end{array}$$

$$F_1^+ = \Pi_{Q_1}(F) = \{SD \rightarrow M, SD \rightarrow SM, SD \rightarrow DM, SD \rightarrow SDM\} \equiv \{SD \rightarrow M\} = F_1$$

$$F_2^+ = \Pi_{Q_2}(F) = \{SI \rightarrow D, SI \rightarrow SD, SI \rightarrow DI, SI \rightarrow SDI\} \equiv \{SI \rightarrow D\} = F_2$$

Q_1 và Q_2 đều đạt dạng chuẩn BC vì trong Q_i chỉ có phụ thuộc hàm có vế trái là khóa. F_1 được tạo thành bằng cách lấy các phụ thuộc hàm của $\Pi_{Q_1}(F)$ có vế phải một thuộc tính. Tương tự cho F_2

Ví dụ 17: cho $Q(CTHRSG)$, $F = \{C \rightarrow T; HR \rightarrow C; HT \rightarrow R; CS \rightarrow G; HS \rightarrow R\}$ hãy phân rã Q thành các lược đồ con đạt chuẩn BC bảo toàn thông tin. (giải như ví dụ trên)



Tính chất: Theo thuật toán trên, khi phân rã Q thành $Q_1(XY)$ với $X \rightarrow Y$ và Q_2 thì tập khóa S_Q của Q luôn luôn bằng với tập khóa S_{Q_2} của Q_2 .

Chứng minh

Thật vậy, K là một khóa của $Q \Rightarrow K$ là một siêu khóa của Q_2 . Giả sử có $K' \subset K$ và K' là khóa của $Q_2 \Rightarrow K' \rightarrow (Q^+ - Y)$ mà $X \rightarrow Y \Rightarrow K' \rightarrow Q^+$. Điều này mâu thuẫn với K là khóa của $Q \Rightarrow K$ là khóa của Q_2 . Ngược lại cũng đúng.

Dựa vào tính chất trên, ta cải tiến thuật toán phân rã nhằm giảm bớt khối lượng tính các phụ thuộc hàm của tập F^+

Thuật toán phân rã Q, F thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

Bước 1: Tìm tập tất cả khóa S_K của Q

Bước 2: Tìm phụ thuộc hàm $X \rightarrow Y \in F$ có X không là siêu khóa và Y không chứa thuộc tính khóa. Nếu tìm thấy thì tách Q thành Q_1 và Q_2 theo quy tắc sau:

$Q_1 = Q[XY]$; Tính F_1 bằng cách tính bao đóng tất cả tập con của XY

$Q_2 = Q[Q^+ - Y]$ S_K cũng là tập khóa của Q_2

thực hiện bước 1 cho Q_1

thực hiện bước 2 cho Q_2

Ngược lại nếu không tìm thấy thì có hai trường hợp:

Trường hợp 1: mọi phụ thuộc hàm trong F đều có vế trái là siêu khóa thì Q_1 đạt chuẩn BC

Trường hợp 2: nếu có phụ thuộc hàm có vế trái không là siêu khóa và vế phải là thuộc tính khóa thì Q_1 đạt chuẩn 3.

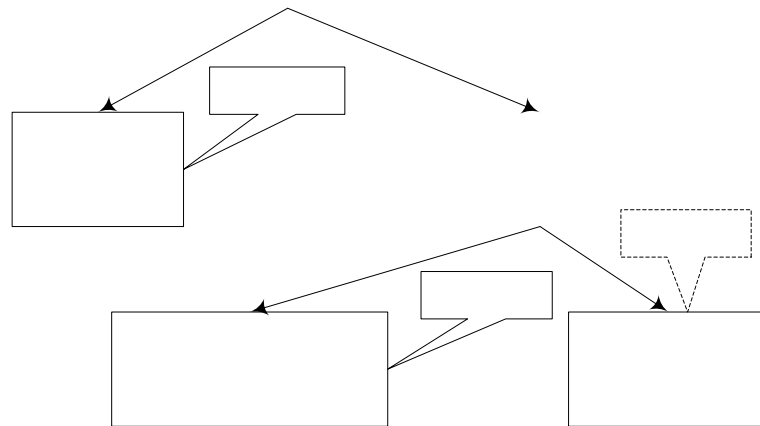
Chú ý: Thuật toán này chỉ tiện trong trường hợp khối lượng tính toán trong việc tìm tất cả khóa của lược đồ quan hệ Q không lớn. Nói cách khác tập trung gian TG có ít thuộc tính. Ngược lại ta phải dùng thuật toán của phần tiếp theo.

Ví dụ 18: phân rã lược đồ ở ví dụ trên thành các lược đồ con ở dạng chuẩn BC bảo toàn thông tin.

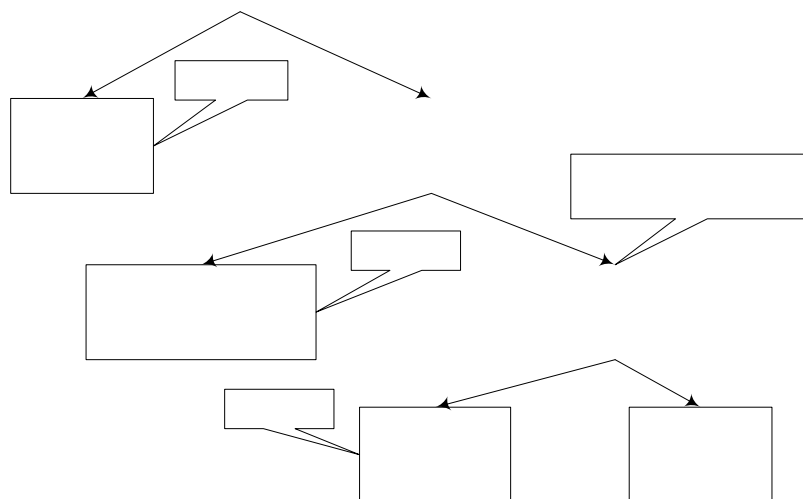
$$F_1 = \{$$

$$Q_1 (C$$

$$K_1 =$$



Trong F có 4 phụ thuộc hàm $C \rightarrow T, HR \rightarrow C, HT \rightarrow R, CS \rightarrow G$ làm Q không đạt dạng chuẩn 3 hay BC và phép phân rã trên đã chọn ngẫu nhiên phụ thuộc hàm $C \rightarrow T$ để phân rã thành Q_1 và tập phụ thuộc tính của Q_{12} chính là tập thuộc tính của Q bỏ thuộc tính T . Tập phụ thuộc hàm F_{12} sẽ chứa các phụ thuộc hàm của F bỏ đi các phụ thuộc hàm có vế trái hay vế phải chứa thuộc tính T . Như vậy tùy theo cách chọn phụ thuộc hàm để phân rã thành Q_1 mà số lượng phụ thuộc hàm mang xuống Q_{12} khác nhau và chất lượng phân rã cũng khác nhau. Kết quả của phép phân rã trên chính là Q_1, Q_2, Q_3 của hình trên. Phép phân rã bảo toàn thông tin, và các lược đồ con đạt chuẩn BC nhưng phép phân rã không bảo toàn phụ thuộc hàm vì $G = F_1 \cup F_2 \cup F_3 = \{C \rightarrow T; HR \rightarrow C; CH \rightarrow R; HS \rightarrow RG\}$ không tương đương với F ($HT \rightarrow R \notin G^+$ và $CS \rightarrow G \notin G^+$). Ta hãy xem phép phân rã sau sẽ cho kết quả tốt hơn.



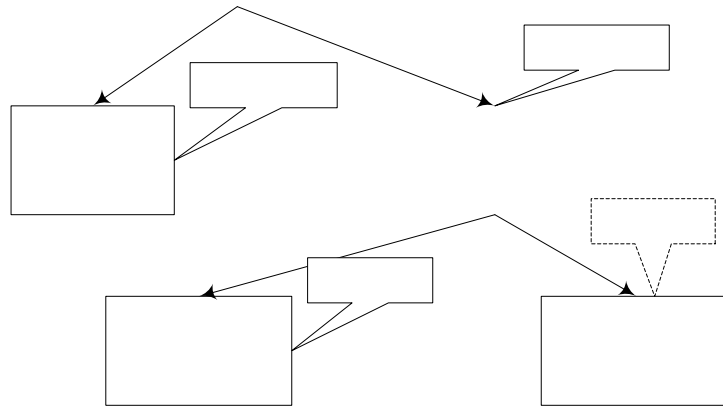
$$F_1 = \{C \rightarrow T, HR \rightarrow C, HT \rightarrow R, CS \rightarrow G\}$$

$$Q_1(C, T)$$

$$K_1 = \{C \rightarrow T, HR \rightarrow C, HT \rightarrow R, CS \rightarrow G\}$$

Phép phân rã cũng cho kết quả phép phân rã bảo toàn thông tin, các lược đồ con Q_1, Q_2, Q_3, Q_4 đạt chuẩn BC và phép phân rã không bảo toàn phụ thuộc hàm vì $G = F_1 \cup F_2 \cup F_3 \cup F_4 = \{CS \rightarrow G; HR \rightarrow C; CH \rightarrow R; C \rightarrow T; HS \rightarrow C\}$ không tương đương với F ($HT \rightarrow R \notin G^+$). Phép phân rã này tốt hơn vì chỉ có một phụ thuộc hàm $HT \rightarrow R$ không thuộc G^+ trong khi phép phân rã trên có tới 2 phụ thuộc hàm $HT \rightarrow R$ và $CS \rightarrow G$ không thuộc G^+ . Sở dĩ phép phân rã thứ 2 tốt hơn vì ở bước chọn phụ thuộc hàm để phân rã thành Q_1 phép phân rã đã chọn phụ thuộc hàm sao cho khi chiếu F xuống Q_{12} số phụ thuộc hàm mang xuống càng nhiều càng tốt.

Ví dụ 19: cho $Q(A, B, C, D, E, G)$, $F = \{AE \rightarrow C; CG \rightarrow A; BD \rightarrow G; GA \rightarrow E\}$ hãy phân rã Q thành các lược đồ con đạt chuẩn BC bảo toàn thông tin.



Nếu Q được phân rã thành:

$(Q_1(BDG), Q_2(A, B, C, D, E))$ lược đồ cơ sở dữ liệu đạt chuẩn 3

$(Q_1(BDG), Q_2(A, C, E), Q_3(A, B, D, E))$ lược đồ cơ sở dữ liệu đạt chuẩn BC

ii **Bổ đề:**

Nếu Q không ở dạng chuẩn BC thì có thuộc tính A, B thuộc Q^+ sao cho $(Q^+ - AB) \rightarrow A$

Chứng minh:

Q không ở dạng chuẩn BC \Rightarrow có $X \rightarrow A$ sao cho X không là siêu khóa \Rightarrow có thuộc tính $B \notin XA$ (vì 1 nếu không có $B \notin XA$ thì X phải là siêu khóa) $\Rightarrow (Q^+ - AB) \supseteq X \Rightarrow (Q^+ - AB) \rightarrow A$

Nhận xét:

- + Một lược đồ Q ở dạng chuẩn BC vẫn có thể có AB sao cho $(Q^+ - AB) \rightarrow A$
- + Một lược đồ Q không có AB sao cho $(Q^+ - AB) \rightarrow A$ thì Q ở dạng chuẩn BC

iii **Thuật toán**

Thuật toán phân rã sau không cần tìm tất cả khóa của lược đồ quan hệ Q

Thuật Toán phân rã Q, F thành dạng chuẩn BC bảo toàn thông tin

Bước 1: $Z' = Q^+$

Bước 2: phân rã Z' theo thuật toán chi tiết để được 2 lược đồ $Z' - A$ và XA trong đó XA ở dạng chuẩn BC và $X \rightarrow A$

Nếu thuật toán chi tiết cho kết quả thì qua bước 3

Ngược lại kết thúc thuật toán

Bước 3: nhận XA là một lược đồ con của các lược đồ kết quả Q_1, \dots, Q_k

Bước 4: thực hiện phân rã $Z' - A, F$

Thuật toán chi tiết

Bước 1: nếu Z' không chứa AB sao cho $(Z' - AB) \rightarrow A$ thì báo không phân rã được.

Ngược lại qua bước 2

Bước 2: đặt $Y' = Z'$

Bước 3: nếu Y' chứa AB sao cho $(Y' - AB) \rightarrow A$ thì gán $Y' = Y' - B$ thực hiện lại bước 2

Bước 4: bước 3 cho kết quả $Y' = XA$ với XA ở dạng chuẩn BC và $X \rightarrow A$. Trả về XA

Nhân xét

Ở mỗi bước 2 của thuật toán phân rã Q, F ta thu được 2 lược đồ $Q_1^+ = Z' - A, Q_1^+ = XA$ với $Q_1^+ \cap Q_1^+ = (Z' - A) \cap XA = X$ và $X \rightarrow Q_1^+$ và Q_1 là lược đồ ở dạng chuẩn BC. Thuật toán lại tiếp tục phân rã Q_1 theo đúng cách đã làm \Rightarrow thuật toán phân rã bảo toàn thông tin và các lược đồ con Q_i đạt dạng chuẩn BC.

Thuật toán chi tiết tìm Q_1 đạt chuẩn BC sao cho Q_1^+ chứa nhiều thuộc tính nhất. Để tìm được Q_1 như vậy **thuật toán chi tiết** tìm hai thuộc tính $AB \in Q^+$ sao cho $(Q^+ - AB) \rightarrow A$. Nếu tìm thấy chứng tỏ Q chưa đạt chuẩn BC và thuật toán giảm B trong Q với hy vọng thu được lược đồ con Q_1 đạt chuẩn BC và thỏa phụ thuộc hàm $(Q^+ - AB) \rightarrow A$. Thuật toán chi tiết tiếp tục tìm và giảm cho tới khi thu được lược đồ con không có hai thuộc tính AB sao cho $(Q^+ - AB) \rightarrow A \Rightarrow Q_1$ là lược đồ con đạt chuẩn BC cần tìm.

Ví dụ 19: Cho quan hệ $Q(B, O, S, Q, I, D)$ và tập phụ thuộc hàm F

$$F = \{ S \rightarrow D, \\ I \rightarrow B \\ IS \rightarrow Q \\ B \rightarrow O \}$$

Hãy phân rã Q thành các lược đồ con đạt dạng chuẩn BC và bảo toàn thông tin.

Giải

***Đặt $Z' = Q^+ = BOSQID$

Thực hiện thuật toán chi tiết

$Y' = BOSQID$

Chọn 2 thuộc tính. Tìm bao đóng của tập hợp thuộc tính còn lại. Nếu bao đóng chứa 1 trong 2 thuộc tính chọn chẳng hạn A , nghĩa là ta đã tìm được 2 thuộc tính AB sao cho $(Y' - AB) \rightarrow A$

Chọn $BO: (SQID)^+ \supset B$

Giảm O trong Y' ta được $Y' = BSQID$

Chọn $BS: (QID)^+ \supset B$

Giảm S trong Y' ta được $Y' = BQID$

Chọn $BQ: (ID)^+ \supset B$

Giảm Q trong Y' ta được $Y' = BID$

Chọn $BD: I^+ \supset B$

Giảm D trong Y' ta được $Y' = BI \Rightarrow Q_1 = (BI)$ và $F_1 = \{I \rightarrow B\}$

Để tính F_1 ta phải tính bao đóng của tất cả tập con của $\{BI\} \Rightarrow F_1$

***Giảm B trong Z' ta được $Z' = OSQID$

Đặt $Y' = OSQID$

Chọn $OD: (SQI)^+ \supset D;$

Giảm O trong Y' ta được $Y' = SQID$

chọn $QD: (SI)^+ \supset D$

giảm Q trong Y' ta được $Y' = SID$

chọn $ID: S^+ \supset D;$

giảm I trong Y' ta được $Y' = SD \Rightarrow Q_2 = (SD)$ và $F_2 = \{S \rightarrow D\}$

Để tính F_2 ta phải tính bao đóng của tất cả tập con của $\{SD\} \Rightarrow F_2$

*** Giảm D trong Z' ta được $Z' = OSQI$

Đặt $Y' = OSQI$

chọn OQ : $(SI)^+ \supset Q$

giảm O trong Y' ta được $Y' = SQI \Rightarrow Q_3 = (SQI)$ và $F_3 = \{SI \rightarrow Q\}$

Ở bước trên không chọn AB để bao đóng tập hợp thuộc tính còn lại chứa A hay B

Để tính F_3 ta phải tính bao đóng của tất cả tập con của $\{SQI\} \Rightarrow F_3$

*** Giảm Q trong Z' ta được $Z' = OSI$

Đặt $Y' = OSI$

Chọn OS : $I^+ = IBO \supset O$

giảm S trong Y' ta được $Y' = OI \Rightarrow Q_4 = (OI)$ và $F_4 = \{I \rightarrow O\}$

*** Giảm O trong Z' ta được $Z' = SI \Rightarrow Q_5 = (SI)$ và $F_5 = \{PTHNN\}$

Ta có thể hiểu $Q_3(SQI)$ là tổ hợp của 2 lược đồ con $Q_5(SI)$ và $Q_3(SQI)$

Vậy kết quả phân rã là:

1: $Q_1(BI)$ $F_1 = \{I \rightarrow B\}$

2: $Q_2(SD)$ $F_2 = \{S \rightarrow D\}$

3: $Q_3(SQI)$ $F_3 = \{SI \rightarrow Q\}$

4: $Q_4(OI)$ $F_4 = \{I \rightarrow O\}$

iv Chú ý

- + Nên tránh phân rã nếu lược đồ đã ở dạng chuẩn mong muốn.
- + Nên xem xét tổ hợp các lược đồ quan hệ con thành lược đồ lớn hơn nếu lược đồ lớn hơn vẫn đạt dạng chuẩn mong muốn.
- + Một kết quả phân rã bảo toàn phụ thuộc hàm sẽ có giá trị hơn kết quả phân rã không bảo toàn phụ thuộc hàm. Giữa hai kết quả phân rã đều không bảo toàn phụ thuộc hàm thì kết quả phân rã thỏa nhiều phụ thuộc hàm trong F sẽ có giá trị hơn.
- + Không có thuật toán phân rã lược đồ Q thành các lược đồ con ở dạng chuẩn BC vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm.
- + Vẫn có lược đồ Q được phân rã thành các lược đồ con ở dạng chuẩn BC vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm.

Ví dụ 20: cho lược đồ $Q(CSZ)$ có $F = \{CS \rightarrow Z, Z \rightarrow C\}$. Q không thể phân rã thành các lược đồ con ở dạng chuẩn BC vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm. Thật vậy:

$TN = \{S\}$ $TG = \{CZ\}$

Tất cả khóa của Q là:

X_i	$TN \cup X_i$	$(TN \cup X_i)^+$	siêu khóa	khóa
\emptyset	S	S		
Z	SZ	SZC	SZ	SZ
C	SC	SZC	SC	SC
ZC	SZC	SZC	SZC	

Vậy Q đạt dạng chuẩn 3 nhưng không ở dạng chuẩn BC vì có $Z \rightarrow C$ có vế trái không là siêu khóa. Nhưng nếu ta phân rã Q thành các lược đồ con có ít hơn 3 thuộc tính thì phụ thuộc $CS \rightarrow Z$ không suy ra được từ các phụ thuộc hình chiếu.

2 Phân rã thành dạng chuẩn 3 vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm

Thuật Toán phân rã Q, F thành dạng chuẩn 3, bảo toàn thông tin, bảo toàn phụ thuộc hàm

Dữ liệu vào: lược đồ quan hệ Q và tập phụ thuộc hàm F .

Dữ liệu ra: một phân rã sao cho mỗi lược đồ quan hệ con đều đạt chuẩn 3 vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm.

Tìm phủ tối thiểu F_u của F

Nếu có một phụ thuộc hàm nào của F_u mà liên quan đến tất cả các thuộc tính của Q thì kết quả phân rã chính là Q (Q không thể phân rã)

Nếu có những thuộc tính của Q không nằm trong một phụ thuộc nào của F_u - dù ở vế phải hay vế trái của F thì chúng tạo thành một lược đồ cần tìm.

Cứ mỗi phụ thuộc hàm $X \rightarrow A \in F_u$ thì XA là một lược đồ cần tìm

Nếu có một lược đồ con chứa khóa K của Q thì kết thúc thuật toán

Ngược lại tạo một lược đồ con K

Ví dụ 21: cho lược đồ $Q(CTHRSG)$, $F = \{C \rightarrow T, HR \rightarrow C, TH \rightarrow R, CS \rightarrow G, HS \rightarrow R\}$. Hãy phân rã Q thành các lược đồ con đạt dạng chuẩn 3 vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm.

Giải:

- + $F = F_{tt} = \{C \rightarrow T, HR \rightarrow C, TH \rightarrow R, CS \rightarrow G, HS \rightarrow R\}$ là phủ tối thiểu.
- + Áp dụng thuật toán trên Q được phân rã thành các lược đồ con $Q_1(CT)$, $Q_2(HRC)$, $Q_3(THR)$, $Q_4(CSG)$, $Q_5(HSR)$
- + Khóa của Q

X_i	$TN \cup X_i$	$(TN \cup X_i)^+$	siêu khóa	khóa
\emptyset	HS	CTHRSG	HS	HS
C	HSC	CTHRSG	HSC	
T	HST	CTHRSG	HST	
CT	HSCT	CTHRSG	HSCT	
R	HSR	CTHRSG	HSR	
CR	HSCR	CTHRSG	HSCR	
TR	HSTR	CTHRSG	HSTR	
CTR	HSCTR	CTHRSG	HSCTR	

- + Q_5 chứa khóa của Q nên Q_1, Q_2, Q_3, Q_4, Q_5 là kết quả của phân rã.

Định lý: Thuật toán trên tạo ra một phân rã ở dạng chuẩn 3 vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm

Chứng minh:

1. Nếu F_{tt} có phụ thuộc hàm f_i liên quan đến tất cả thuộc tính thì Q đạt chuẩn 3. Thật vậy:

$f_i \in F_{tt} \Rightarrow f_i$ là phụ thuộc hàm có vế phải 1 thuộc tính $\Rightarrow f_i$ có dạng $K \rightarrow A \Rightarrow K$ là siêu khóa. Nếu khóa của Q là $K' \subset K$ thì ta có $K' \rightarrow A \Rightarrow K \rightarrow A$ là phụ thuộc hàm có vế trái dư thừa điều này mâu thuẫn với $f_i \in F_{tt}$. Vậy K là khóa của $Q \Rightarrow$ nếu Q có thuộc tính không khóa thì A là thuộc tính không khóa duy nhất của Q và mọi phụ thuộc hàm có vế phải là A phải có vế trái là $K \Rightarrow$ lược đồ quan hệ Q không có phụ thuộc hàm có vế trái không là siêu khóa và vế phải không là thuộc tính khóa $\Rightarrow Q$ đạt chuẩn 3.

2. Nếu lược đồ $Q'(W)$ gồm các thuộc tính không xuất hiện trong F_{tt} thì Q đạt chuẩn 3. Thật vậy:

V là tập con bất kỳ của W ta có $V^+ = V \Rightarrow F'$ của Q' chỉ gồm các phụ thuộc hàm hiển nhiên \Rightarrow trong F' không có phụ thuộc hàm có vế trái không là siêu khóa và vế phải là thuộc tính không khóa $\Rightarrow Q'$ đạt chuẩn 3.

3. Ta chứng minh mỗi lược đồ con ở dạng chuẩn 3. Thật vậy:

Theo thuật toán thì mỗi lược đồ con Q_i có dạng $Y \rightarrow B \Rightarrow Y$ là siêu khóa. Giả sử trong Q_i có phụ thuộc hàm $X \rightarrow A$ có vế trái không là siêu khóa và vế phải không là thuộc tính khóa. Ta phân làm hai trường hợp:

Trường hợp 1: $A=B \Rightarrow X \rightarrow B \Rightarrow X \subset Y \Rightarrow Y \rightarrow B$ là phụ thuộc có vế trái dư thừa, điều này trái với $Y \rightarrow B$ là phụ thuộc hàm trong phủ tối thiểu.

Trường hợp 2: $A \neq B \Rightarrow A \in Y$ (1). Gọi K là khóa của $Q_i \Rightarrow K \subseteq Y$ (2). A là thuộc tính không khóa nên $A \notin K$ (3). (1)(2)(3) $\Rightarrow K \subset Y$ (4). K là khóa nên $K \rightarrow B \Rightarrow Y \rightarrow B$ là phụ thuộc hàm có vế trái dư thừa. Điều này trái với điều phụ thuộc hàm $Y \rightarrow B$ là phụ thuộc hàm của phủ tối thiểu F_{tt}

4. Ta chứng minh phép phân rã bảo toàn phụ thuộc hàm. Thật vậy:

Hiển nhiên $F_{tt} \subseteq G = \cup \Pi_{Q_i}(F_{tt}) \Rightarrow F_{tt}^+ \subseteq G^+$ (1)

Hơn nữa $F_{tt}^+ \supseteq G = \cup \Pi_{Q_i}(F_{tt}) \Rightarrow F_{tt}^{++} \supseteq G^+ \Rightarrow F_{tt}^+ \supseteq G^+$ (2)

(1) và (2) $\Rightarrow F_{tt}^+ = G^+$

5. Ta chứng minh phép phân rã bảo toàn thông tin. Thật vậy:

Lập bảng kiểm tra bảo toàn thông tin. Ta lần lượt đồng nhất các giá trị của bảng trên theo các phụ thuộc hàm được phát hiện ở mỗi bước của thuật toán tìm bao đóng của tập thuộc tính Q_i^+ với Q_i^+ chứa khóa K của lược đồ Q . Phụ thuộc hàm đầu tiên được phát hiện là $Y \rightarrow A_j \in F_{tt}$ sao cho $Q_i^+ \supseteq Y$ và $A_j \notin Q_i^+$. Ở dòng của lược đồ $Q_1(YA_j)$ có giá trị a_j ở cột A_j nên khi làm bằng giá trị kết quả là ở cột A_j của dòng có lược đồ Q_i có thêm giá trị a_j . Tiếp tục cho các phụ thuộc hàm phát hiện tiếp theo ta sẽ có thêm các giá trị a ở các cột khác của dòng Q_i . Do Q_i chứa khóa nên các giá trị a mới thêm vào của dòng Q_i sẽ xuất hiện ở tất cả các thuộc tính của lược đồ Q . Suy ra hàng của lược đồ Q_i sẽ chứa toàn a là điều phải chứng minh. Để làm sáng tỏ ý tưởng của phần chứng minh này ta xét trường hợp cụ thể của ví dụ 21 :

Bước 1: ta lập bảng kiểm tra bảo toàn thông tin:

	C	T	H	R	S	G
$Q_1(CT)$	a_1	a_2				
$Q_2(HRC)$	a_1		a_3	a_4		
$Q_3(THR)$		a_2	a_3	a_4		
$Q_4(CSG)$	a_1				a_5	a_6
$Q_5(HSR)$			a_3	a_4	a_5	

Bước 2: Ta chứng minh dòng Q_5 của bảng trên sẽ chứa toàn giá trị a . Thật vậy: ta lần lượt đồng nhất các giá trị của bảng trên theo các phụ thuộc hàm được phát hiện theo thuật toán tìm bao đóng của $X = \{HSR\} \supseteq K$; $F = \{C \rightarrow T, HR \rightarrow C, TH \rightarrow R, CS \rightarrow G, HS \rightarrow R\}$

$X_0 = HSR$

$X_1 = HSRC$

$X_2 = HSRCT$

$X_3 = HSRCTG$

do $HR \rightarrow C$. Đồng nhất các giá trị theo phụ thuộc hàm này. Trên dòng Q_2 ở cột C chứa giá trị a nên trên dòng Q_5 sẽ có thêm giá trị a ở cột C
do $C \rightarrow T$. Đồng nhất các giá trị theo phụ thuộc hàm này.
do $CS \rightarrow G$ đồng nhất các giá trị theo phụ thuộc hàm này.

	C	T	H	R	S	G
$Q_1(CT)$	a_1	a_2				

Q_2 (HRC)	a_1	a_2	a_3	a_4		
Q_3 (THR)	a_1	a_2	a_3	a_4		
Q_4 (CSG)	a_1	a_2			a_5	a_6
Q_5 (HSR)	a_1	a_2	a_3	a_4	a_5	a_6

Do $X^+ = Q^+$ nên dòng Q_5 chứa toàn giá trị a

Ví dụ 22: Cho $Q(ABCDEFGHI)$, $F = \{AB \rightarrow D; EH \rightarrow G; G \rightarrow C; D \rightarrow C\}$ hãy phân rã Q thành các lược đồ con ở dạng chuẩn 3 vừa bảo toàn thông tin vừa bảo toàn phụ thuộc.

Giải:

Tìm phủ tối thiểu F_{tt} của F

$$F_{tt} = F = \{AB \rightarrow D; EH \rightarrow G; G \rightarrow C; D \rightarrow C\}$$

Áp dụng thuật toán, Q được phân rã thành lược đồ CSDL sau:

$$Q_1(ABD), Q_2(EHG), Q_3(GC), Q_4(DC)$$

Tìm khóa của Q

$$TN = \{ABEH\} \quad TG = \{GD\}$$

X_i	$TN \cup X_i$	$(TN \cup X_i)^+$	Siêu khóa	Khóa
\emptyset	ABEH	ABCDEFGHI	ABEH	ABEH
G	ABEHG	ABCDEFGHI	ABEHG	
D	ABEHD	ABCDEFGHI	ABEHD	
GD	ABEHGD	ABCDEFGHI	ABEHGD	

Q_1, Q_2, Q_3, Q_4 không chứa khóa \Rightarrow để bảo toàn thông tin ta cần có $Q_5(A, B, E, H)$. Vậy kết quả của phân rã là Q_1, Q_2, Q_3, Q_4, Q_5

IV BÀI TẬP

1/ Cho biết dạng chuẩn của các lược đồ quan hệ sau:

- $Q(ABCDEG); F = \{A \rightarrow BC, C \rightarrow DE, E \rightarrow G\}$
- $Q(ABCDEFGH); F = \{C \rightarrow AB, D \rightarrow E, B \rightarrow G\}$
- $Q(ABCDEFGH); F = \{A \rightarrow BC, D \rightarrow E, H \rightarrow G\}$
- $Q(ABCDEG); F = \{AB \rightarrow C, C \rightarrow B, ABD \rightarrow E, G \rightarrow A\}$
- $Q(ABCDEFGHI); F = \{AC \rightarrow B, BI \rightarrow ACD, ABC \rightarrow D, H \rightarrow I, ACE \rightarrow BCG, CG \rightarrow AE\}$

2/ Kiểm tra sự bảo toàn thông tin ?

$$Q(ABCDE) \quad R_1(AD); R_2(AB); R_3(BE); R_4(CDE); R_5(AE)$$

$$F = \{A \rightarrow C; B \rightarrow C; C \rightarrow D; DE \rightarrow C; CE \rightarrow A\}$$

3/ Cho lược đồ quan hệ $Q(A, B, C, D)$ và tập phụ thuộc hàm $F = \{A \rightarrow B; B \rightarrow C; A \rightarrow D; D \rightarrow C\}$

Và một lược đồ CSDL như sau: $C = \{Q_1(AB); Q_2(AC); Q_3(BD)\}$

- C có bảo toàn thông tin đối với F
- C có bảo toàn phụ thuộc hàm ?

4/ Kiểm tra dạng chuẩn $Q(C, S, Z) \quad F = \{CS \rightarrow Z; Z \rightarrow C\}$

5/ Phân rã $Q(G, H, A, B, C, D) \quad F = \{GH \rightarrow AD; AG \rightarrow B; CD \rightarrow GH; C \rightarrow A; BH \rightarrow C\}$

6/ Cho lược đồ CSDL

$$Kehoach(NGAY, GIO, PHONG, MONHOC, GIAOVIEN)$$

$F = \{ \text{NGAY, GIO, PHONG} \rightarrow \text{MONHOC}$
 $\text{MONHOC, NGAY} \rightarrow \text{GIAOVIEN}$
 $\text{NGAY, GIO, PHONG} \rightarrow \text{GIAOVIEN}$
 $\text{MONHOC} \rightarrow \text{GIAOVIEN} \}$

- Xác định dạng chuẩn cao nhất của Kehoach
- Nếu Kehoach chưa đạt dạng chuẩn 3, hãy phân rã Kehoach thành lược đồ CSDL dạng chuẩn 3 vừa bảo toàn phụ thuộc hàm vừa bảo toàn thông tin.
- Nếu Kehoach chưa đạt dạng chuẩn BC, hãy phân rã KeHoach thành lược đồ CSDL dạng BC

7/ Cho lược đồ quan hệ $Q(A, B, C, D)$ và tập phụ thuộc hàm F

$F = \{A \rightarrow B; B \rightarrow C; D \rightarrow B\}$ $C = \{Q_1(A, C, D); Q_2(B, D)\}$

- Xác định các F_i (những phụ thuộc hàm F được bao trong Q_i)
- Lược đồ CSDL C có đạt dạng chuẩn BC? Nếu không có thể phân rã tiếp các Q_i của C để biến C thành dạng chuẩn BC?

8/ Giả sử ta có lược đồ quan hệ $Q(C, D, E, G, H, K)$ và tập phụ thuộc hàm F như sau;

$F = \{CK \rightarrow H; C \rightarrow D; E \rightarrow C; E \rightarrow G; CK \rightarrow E\}$

- Từ tập F , hãy chứng minh $EK \rightarrow DH$
- Tìm tất cả các khóa của Q .
- Xác định dạng chuẩn của Q .
- Hãy tìm cách phân rã Q thành một lược đồ CSDL đạt dạng chuẩn BC (hoặc dạng chuẩn 3). tìm tập phụ thuộc hàm và khóa cho mỗi lược đồ quan hệ con.

9/ Cho lược đồ quan hệ $Q(S, I, D, M)$

$F = \{f_1: SI \rightarrow DM; f_2: SD \rightarrow M; f_3: D \rightarrow M\}$

- Tính bao đóng D^+, SD^+, SI^+
- Tìm tất cả các khóa của Q
- Tìm phủ tối thiểu của F
- Xác định dạng chuẩn cao nhất của Q
- Nếu Q chưa đạt dạng chuẩn 3, hãy phân rã Q thành lược đồ CSDL dạng chuẩn 3 vừa bảo toàn phụ thuộc hàm vừa bảo toàn thông tin.
- Nếu Q chưa đạt dạng chuẩn BCNF, hãy phân rã Q thành lược đồ CSDL dạng BCNF
- Kiểm tra phép tách Q thành các lược đồ con (SID, SIM) có bảo toàn thông tin?
- Kiểm tra phép tách Q thành các lược đồ con (SID, SIM) có bảo toàn phụ thuộc hàm?

10/ Cho lược đồ quan hệ

$R(W, A, Z, Y, Q, P)$
 $R_1(A, Z);$
 $R_2(W, Y, Q, P)$
 $R_3(Y, Q, P, A)$
 $F = \{W \rightarrow AYQP, A \rightarrow Z, YQP \rightarrow A\}$

Hãy kiểm tra tính kết nối không mất thông tin.

11/ Cho lược đồ quan hệ Q(Môn, GiảngViên, Giờ giảng, Phòng, SinhViên, Hạng) với

 $F = \{M \rightarrow GV; G, P \rightarrow M; G, GV \rightarrow P; M, SV \rightarrow H; G, SV \rightarrow P\}$
 $C = \{Q_1(M, G, P); Q_2(M, GV); Q_3(M, SV, H)\}$

Kiểm tra xem lược đồ cơ sở dữ liệu sau đây có bảo toàn thông tin đối với F ?

12/ Kiểm Tra Dạng Chuẩn

a) $Q(A, B, C, D)$ $F = \{CA \rightarrow D; A \rightarrow B\}$

b) $Q(S, D, I, M)$ $F = \{SI \rightarrow D; SD \rightarrow M\}$

c) $Q(N, G, P, M, GV)$ $F = \{N, G, P \rightarrow M; M \rightarrow GV\}$

d) $Q(S, N, D, T, X)$ $F = \{S \rightarrow N; S \rightarrow D; S \rightarrow T; S \rightarrow X\}$

13/ Phân rã lược đồ thành dạng BCK

a) $Q(S, D, I, M)$ $F = \{S, I \rightarrow D; S, D \rightarrow M\}$

b) $Q(A, B, C, D)$ $F = \{A \rightarrow B; B \rightarrow C; D \rightarrow B\}$

c) $Q(C, S, Z)$ $F = \{C, S \rightarrow Z; Z \rightarrow C\}$

14/ Phân rã lược đồ thành dạng 3NF vừa bảo toàn phụ thuộc hàm vừa bảo toàn thông tin

a) $Q(A, B, C)$, $F = \{A \rightarrow B; A \rightarrow C; B \rightarrow A; C \rightarrow A; B \rightarrow C\}$

b) $Q(MSCD, MSSV, CD, HG)$

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

c) $Q(A, B, C, D)$ $F = \{AB \rightarrow C; C \rightarrow B\}$

-----oOo-----

ĐỀ THI MẪU MÔN CƠ SỞ DỮ LIỆU

(thời gian 60 phút)

Đề 1**BÀI 1: (6 điểm)**

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 (TUTIET) nào đến tiết (DENTIET) nào, 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. Giả sử ta có lược đồ cơ sở dữ liệu để quản lý bài toán trên như sau:

Giaovien(MAGV, HOTEN, DTGV, MAKHOA)

Monhoc(MAMH, TENMH)

Phonghoc(PHONG, CHUCNANG)

Khoa(MAKHOA, TENKHOA, DTKHOA)

Lop(MALOP, TENLOP, SISO, MAKHOA)

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

1. Hãy xác định khóa cho mỗi lược đồ quan hệ trên. (2,0 đ)

2. Phát biểu các ràng buộc toàn vẹn miền giá trị, ràng buộc toàn vẹn liên thuộc tính (1.0 đ)

3. Dựa vào lược đồ CSDL trên, hãy thực hiện các câu hỏi sau bằng SQL (3,0 đ)

a. Xem lịch báo giảng tuần từ ngày 16/09/2002 đến ngày 23/09/2002 của giáo viên có MAGV (mã giáo viên) là TH3A040. Yêu cầu: MAGV, HOTEN, TENLOP, TENMH, PHONG, NGAYDAY, TUTIET, DENTIET, BAIDAY, GHICHU)

b. Xem lịch báo giảng ngày 23/09/2002 của các giáo viên có mã khoa là CNTT. Yêu cầu: MAGV, HOTEN, TENLOP, TENMH, PHONG, NGAYDAY, TUTIET, DENTIET, BAIDAY, GHICHU

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

BÀI 2: (4 điểm)

Cho lược đồ quan hệ $Q(A, B, C, D, E, G, H, K)$ và tập phụ thuộc hàm F như sau;

$F = \{C \rightarrow AD; E \rightarrow BH; B \rightarrow K; CE \rightarrow G\}$

1. Kiểm tra xem các phụ thuộc hàm $E \rightarrow K; E \rightarrow G$ có thuộc tập F^+ ? (1,0đ)
2. Tìm tất cả các khóa của Q . (1,0đ)
3. Xác định dạng chuẩn của Q . (1,0đ)
4. Nếu Q chưa đạt chuẩn BC. Hãy phân rã Q thành lược đồ CSDL đạt chuẩn BC (1,0đ)

ĐÁP ÁN**BÀI 1: Câu 1:**

Giaovien (MAGV, HOTEN, DTGV, MAKHOA)

Monhoc (MAMH, TENMH)

Phonghoc (PHONG, CHUCNANG)

Khoa (MAKHOA, TENKHOA, DTKHOA)

Lop (MALOP, TENLOP, SISO, MAKHOA)

Lichday (MAGV, MAMH, PHONG, MALOP, NGÀYDAY, TUTIET, DENTIET, BAIDAY, LYTHUYET, GHICHU)

Câu 2:

$\forall t \in r_{\text{Giaovien}}$

$t.HOTEN \neq \text{NULL}$

RBTV miễn giá trị

cuối \forall

$\forall t \in r_{\text{Monhoc}}$

$t.TENMH \neq \text{NULL}$

RBTV miễn giá trị

cuối \forall

$\forall t \in r_{\text{Khoa}}$

$t.TENKHOA \neq \text{NULL}$

RBTV miễn giá trị

cuối \forall

$\forall t \in r_{\text{LOP}}$

$t.TENLOP \neq \text{NULL}$ và $t.SISO > 0$

RBTV miễn giá trị

cuối \forall

$\forall t \in r_{\text{Lichday}}$

$t.TUTIET < t.DENTIET$ và

RBTV liên thuộc tính

$t.NGAYDAY \neq \text{NULL}$ và

RBTV miễn giá trị

$t.BAIDAY \neq \text{NULL}$ và

RBTV miễn giá trị

$(t.LYTHUYET = 1 \text{ Or } t.LYTHUYET = 2)$ và

RBTV miễn giá trị

$(t.TUTIET \geq 1 \text{ và } t.TUTIET \leq 16)$ và

RBTV miễn giá trị

$(t.DENTIET \geq 1 \text{ và } t.DENTIET \leq 16)$

RBTV miễn giá trị

cuối \forall

Câu 3:

- SELECT** giaovien.magv,hoten,tenlop,tenmh,phong,ngayday,tutiet,đentiet,baiday,ghichu
FROM ((lichday **INNER JOIN** giaovien **ON** lichday.magv = giaovien.magv)
 INNER JOIN lop **ON** Lichday.malop = lop.malop)
 INNER JOIN monhoc **ON** lichday.mamh = monhoc.mamh
WHERE ngayday >=#16/09/2002# **AND** ngayday <=#23/09/2002# **AND** magv= "TH3A040"
- SELECT** giaovien.magv,hoten,tenlop,tenmh,phong,ngayday,tutiet,đentiet,baiday,ghichu
FROM ((lichday **INNER JOIN** giaovien **ON** lichday.magv = giaovien.magv)
 INNER JOIN lop **ON** Lichday.malop = lop.malop)
 INNER JOIN monhoc on lichday.mamh = monhoc.mamh

WHERE ngayday = #23/09/2002# **AND** makh= "CNTT"

- c. **SELECT** tenkhoa, **COUNT**(giaovien.makh) **AS** soluonggv
FROM giaovien **INNER JOIN** khoa **ON** giaovien.makh=khoa.makh
GROUP BY giaovien.makh,tenkhoa

BÀI 2.

1. $E^+ = E, B, H, K \supseteq K$ nên $E \rightarrow K \in F^+$

$E^+ = E, B, H, K \not\supseteq G$ nên $E \rightarrow G \notin F^+$

2. $TN = \{CE\}; TG = \{B\}$

X_i	$TN \cup X_i$	$(TN \cup X_i)^+$	Siêu khóa	Khóa
\emptyset	CE	Q^+	CE	CE
B	CEB	Q^+	CEB	

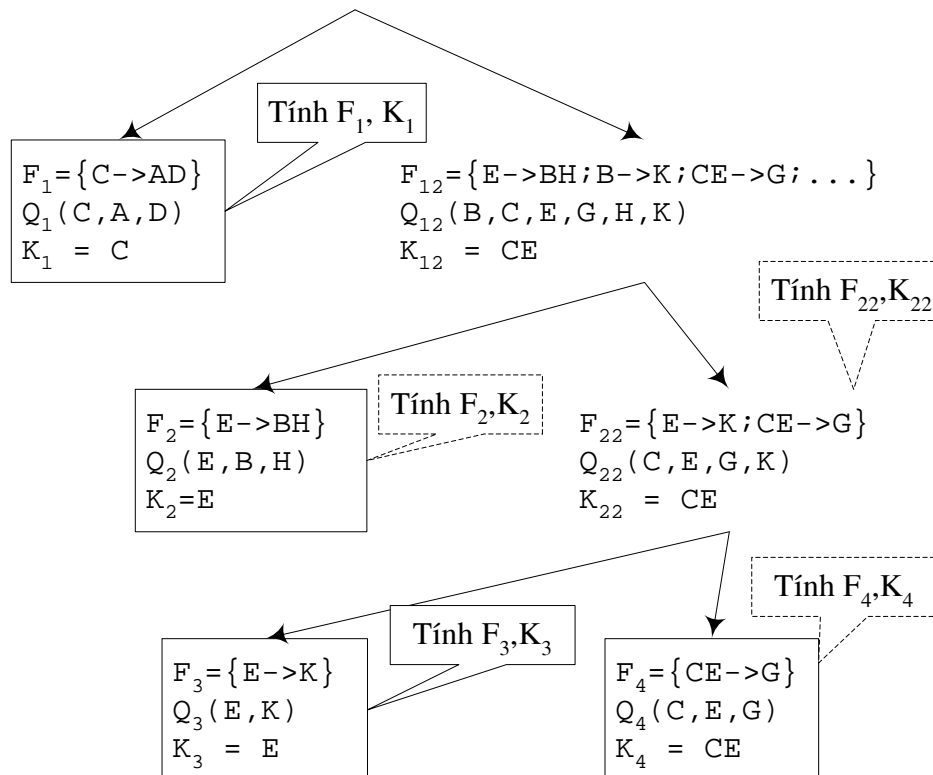
vậy Q có khóa duy nhất là $K = \{C, E\}$

3. $C \subset K$ và $C^+ = CAD \supseteq$ thuộc tính không khóa $A \Rightarrow Q$ không đạt chuẩn 2 $\Rightarrow Q$ đạt chuẩn 1

$F = \{C \rightarrow AD; E \rightarrow BH; B \rightarrow K; CE \rightarrow G\}$

$Q(A, B, C, D, E, G, H, K)$

$K = CE$



4. $Q(A, B, C, D, E, G, H, K); F = \{C \rightarrow AD; E \rightarrow BH; B \rightarrow K; CE \rightarrow G\}$

Vậy lược đồ quan hệ Q được tách thành lược đồ cơ sở dữ liệu đạt chuẩn BC sau:

$Q_1(C, A, D) \quad F_1 = \{C \rightarrow AD\}$ đạt chuẩn BC

$Q_2(E, B, H) \quad F_2 = \{E \rightarrow BH\}$ đạt chuẩn BC

$Q_3(E, K) \quad F_3 = \{E \rightarrow K\}$ đạt chuẩn BC

$Q_4(C, E, G) \quad F_4 = \{CE \rightarrow G\}$ đạt chuẩn BC

----oOo----

Đề 2

Cho một lược đồ cơ sở dữ liệu C dùng để quản lý hoạt động sửa chữa, bảo trì xe của một gara xe hơi. Lược đồ cơ sở dữ liệu C gồm các lược đồ quan hệ như sau:

Q₁: Tho (MATHO, TENTHO, NHOM, NHOM_TRUONG)

Tân từ: Mỗi người thợ đều có mã số là MATHO để nhận diện. Mỗi thợ chỉ có một tên (TENTHO) và chỉ thuộc một nhóm (NHOM). Nhóm trưởng (NHOM_TRUONG) của mỗi nhóm là một trong số những người thợ của nhóm đó.
MGT (MATHO) = MGT (NHOM_TRUONG)

Q₂: Cong_viec (MACV, NOIDUNGCV)

Tân từ: Dịch vụ sửa chữa xe được chia nhỏ thành nhiều công việc để dễ dàng tính toán chi phí với khách hàng. Mỗi công việc đều có mã riêng (MACV) và nội dung của công việc được mô tả qua NOIDUNGCV.

Q₃: Hop_dong (SOHD, NGAYHD, MAKH, TENKH, DCHI, SOXE, TRIGIAHD, NG_GIAO_DK, NG_NGTHU)

Tân từ: Mỗi hợp đồng sửa xe ký kết với khách hàng đều có mã số (SOHD) để phân biệt. NGAYHD là ngày ký hợp đồng. Mỗi khách hàng có một mã số (MAKH), một tên (TENKH) và một địa chỉ (DCHI) để theo dõi công nợ. SOXE là số đăng bộ của xe đem đến sửa chữa, số này do phòng CSGT đường bộ cấp (nếu xe đổi chủ thì xem như một xe khác). Khách hàng ký hợp đồng chính là chủ xe sửa chữa. Một khách hàng có thể ký nhiều hợp đồng sửa chữa nhiều xe khác nhau hoặc hợp đồng sửa chữa nhiều lần của cùng một xe nhưng trong cùng một ngày. Những công việc sửa chữa cho một đầu xe chỉ ký hợp đồng một lần. TRIGIAHD là tổng trị giá của hợp đồng. NG_GIAO_DK là ngày dự kiến phải giao trả xe cho khách. NG_NGTHU là ngày nghiệm thu thật sự sau khi đã sửa chữa xong để thanh lý hợp đồng.

Q₄: Chitiet_HD (SOHD, MACV, TRIGIA_CV, MATHO, KHOANTHO)

Tân từ: Mỗi hợp đồng sửa xe có thể gồm nhiều công việc. MACV là mã số của từng công việc. TRIGIA_CV là chi phí về vật tư, phụ tùng, thiết bị, công thợ ... đã tính toán với khách. Mỗi công việc của hợp đồng sẽ giao cho một người thợ phụ trách (MATHO) và một người thợ có thể tham gia vào nhiều công việc của một hay nhiều hợp đồng khác nhau. KHOANTHO là số tiền giao khoán lại cho người thợ sửa chữa.

Q₅: Phieu_thu (SOPH, NGAYPH, SOHD, MAKH, HOTEN, SOTIENTHU)

Tân từ: Khách hàng (MAKH) có thể thanh toán tiền của một hợp đồng (SOHD) làm nhiều lần trước hoặc sau khi nghiệm thu (trong cùng ngày hoặc khác ngày). Mỗi lần thanh toán đều có số phiếu để phân biệt (SOPH), NGAYPH là ngày phát hành phiếu và SOTIENTHU là số tiền thanh toán. HOTEN là họ tên của người mang tiền đến thanh toán (có thể khác với tên của khách hàng đứng ra ký hợp đồng)

Câu hỏi:

1/ Xác định tập hợp F gồm tất cả các phụ thuộc hàm suy ra từ tân từ của các lược đồ quan hệ (không cần liệt kê các phụ thuộc hàm hiển nhiên). Xác định khóa cho từng lược đồ quan hệ.

2/ Mô tả tất cả các ràng buộc toàn vẹn của lược đồ cơ sở dữ liệu C. Lập bảng tầm ảnh hưởng tổng hợp.

3/ Dùng ngôn ngữ SQL để thực hiện những yêu cầu sau:

- Cho biết danh sách những người thợ hiện không tham gia vào một hợp đồng sửa chữa nào.
- Cho biết danh sách những hợp đồng hiện đã thanh lý (đã giao tra xe cho khách) nhưng chưa được thanh toán đầy đủ.
- Giả sử hôm nay là ngày 21/12/95 cho biết danh sách những hợp đồng cần phải hoàn tất trước ngày 31/12/95.
- Cho biết người thợ nào thực hiện nhiều công việc nhất.
- Cho biết người thợ nào thực hiện tổng giá trị công việc (tổng số tiền) cao nhất.

4/ Lược đồ cơ sở dữ liệu C ở dạng chuẩn mấy (cao nhất). Hãy dùng thuật toán phân rã để nâng cấp lược đồ cơ sở dữ liệu trên.

Lưu ý: Các thuộc tính đều được xem như thuộc tính đơn.

Đáp án:

Câu 1:

$F_1 = \{ \text{MATHO} \rightarrow \text{TENTHO}, \text{NHOM}, \text{NHOM_TRUONG} \}$

$Q_1: \text{Tho}(\text{MATHO}, \text{TENTHO}, \text{NHOM}, \text{NHOM_TRUONG})$

$F_2 = \{ \text{MACV} \rightarrow \text{NOIDUNGCV} \}$

$Q_2: \text{Congviiec}(\text{MACV}, \text{NOIDUNGCV})$

$F_3 = \{ \text{SOHD} \rightarrow \text{NGAYHD}, \text{MAKH}, \text{SOXE}, \text{TRIGIAHD}, \text{NG_GIAO_DK}, \text{NG_NGTHU};$

$\text{MAKH} \rightarrow \text{TENKH}, \text{DCHI} \}$

$Q_3: \text{Hopdong}(\text{SOHD}, \text{NGAYHD}, \text{MAKH}, \text{TENKH}, \text{DCHI}, \text{SOXE}, \text{TRIGIAHD}, \text{NG_GIAO_DK}, \text{NG_NGTHU})$

$F_4 = \{ \text{SOHD}, \text{MACV} \rightarrow \text{TRIGIA_CV}, \text{MATHO}, \text{KHOANTHO} \}$

$Q_4: \text{ChiTiet_hd}(\text{SOHD}, \text{MACV}, \text{TRIGIA_CV}, \text{MATHO}, \text{KHOANTHO})$

$F_5 = \{ \text{SOPH} \rightarrow \text{NGAYPH}, \text{SOHD}, \text{HOTEN}, \text{SOTIENTHU}; \text{SOHD} \rightarrow \text{MAKH} \}$

$Q_5: \text{Phieu_thu}(\text{SOPH}, \text{NGAYPH}, \text{SOHD}, \text{MAKH}, \text{HOTEN}, \text{SOTIENTHU})$

Câu 2: mô tả tất cả các ràng buộc toàn vẹn:

$R_{11} \quad \forall t_1, t_2 \in r\text{Tho}$

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

cuối \forall

RBTV khóa chính

$R_{12} \quad r\text{Tho}[\text{NHOM_TRUONG}] \subseteq r\text{Tho}[\text{MATHO}]$

RBTV khóa ngoại

$R_{13} \quad \forall t \in r\text{Tho}$

$t.\text{TENTHO} \neq \text{NULL}$

$t.\text{NHOM} \neq \text{NULL}$

cuối \forall

RBTV miền giá trị

RBTV miền giá trị

$R_{21} \quad \forall t_1, t_2 \in r\text{Congviiec}$

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

cuối \forall

RBTV khóa chính

- R₂₂ $\forall t \in rCongviec$
 $t.NOIDUNGCV \neq NULL$
 cuối \forall RBTV miền giá trị.
- R₃₁ $\forall t_1, t_2 \in rHopdong$
 $t_1.SOHD \neq t_2.SOHD$
 cuối \forall RBTV khóa chính.
- R₃₂ $\forall t \in rHopdong$
 $t.NGAYHD \neq NULL$ RBTV miền giá trị.
 $t.MAKH \neq NULL$ RBTV miền giá trị.
 $t.TENKH \neq NULL$ RBTV miền giá trị.
 $t.SOXE \neq NULL$ RBTV miền giá trị.
 $t.TRIGIAHD > 0$ RBTV miền giá trị.
 $t.NGAYHD \leq t.NG_NGTHU$ RBTV liên thuộc tính.
 $t.NG_NGTHU \leq t.NG_GIAO_DK$ RBTV liên thuộc tính.
 cuối \forall
- R₄₁ $\forall t_1, t_2 \in rChiTiet_hd$
 $t_1.\{SOHD, MACV\} \neq t_2.\{SOHD, MACV\}$
 cuối \forall RBTV khóa chính.
- R₄₂ $rChitiet_HD[MATHO] \subseteq rTho[MATHO]$ RBTV khóa ngoại
 R₄₃ $rChitiet_HD[SOHD] \subseteq rHopdong[SOHD]$ RBTV khóa ngoại
 R₄₄ $rChitiet_HD[MACV] \subseteq rCongviec[MACV]$ RBTV khóa ngoại
 R₄₅ $\forall t \in rChiTiet_hd$
 $t.TRIGIA_CV > t.KHOANTHO$ RBTV liên thuộc tính.
 $t.KHOANTHO > 0$ RBTV miền giá trị.
 cuối \forall
- R₅₁ $\forall t_1, t_2 \in rPhieu_thu$
 $t_1.SOPH \neq t_2.SOPH$
 cuối \forall RBTV khóa chính
- R₅₂ $rPhieu_thu[SOHD] \subseteq rHopdong[SOHD]$ RBTV khóa ngoại
 R₅₃ $\forall t \in rPhieu_thu$
 $t.NGAYPH \neq NULL$ RBTV miền giá trị
 $t.MAKH \neq NULL$ RBTV miền giá trị
 $t.HOTEN \neq NULL$ RBTV miền giá trị
 $t.SOTIENTHU > 0$ RBTV miền giá trị
 cuối \forall
- R₅₄ $\forall t \in rPhieu_thu \exists t' \in rHopdong$
 $t.SOHD = t'.SOHD$ và
 $t.NGAYPH \leq t'.NGAYHD$
 cuối \forall RBTV liên thuộc tính liên quan hệ.

Bảng tầm ảnh hưởng tổng hợp:

	rTho			rCongviec			rHopdong			rChiTiet_hd			rPhieu_thu		
	T	S	X	T	S	X	T	S	X	T	S	X	T	S	X
R ₁₁	+	+	-												
R ₁₂	+	+	+												
R ₁₃	+	+	-												
R ₂₁				+	+	-									

R ₂₂				+	+	-									
R ₃₁							+	+	-						
R ₃₂							+	+	-						
R ₄₁										+	+	-			
R ₄₂	-	+	+							+	+	-			
R ₄₃							-	+	+	+	+	-			
R ₄₄				-	+	+				+	+	-			
R ₄₅										+	+	-			
R ₅₁													+	+	-
R ₅₂							-	+	+				+	+	-
R ₅₃													+	+	-
R ₅₄							-	+	+				+	+	-

Câu 3:

a) **SELECT** matho,tentho

FROM tho

WHERE matho **NOT IN**

(**SELECT** matho

FROM hop_dong **INNER JOIN** chitiet_HD **ON** hop_dong.sohd = chitiet_HD.sohd

WHERE ng_ngthu > date() **OR ISNULL**(ng_ngthu))

b) **SELECT**

sohd,ngayhd,makh,tenkh,dchi,soxe,trigiahd,ng_giao_dk,ng_ngthu

FROM hop_dong

WHERE ng_giao_dk <= **DATE**() **AND** (TRIGIAHD >

(**SELECT SUM**(SOTIENTHU) **FROM** phieu_thu

WHERE phieu_thu.sohd = hop_dong.sohd) **OR** sohd Not In (Select sohd From phieu_thu))

c) **SELECT** sohd,ngayhd,makh,tenkh,dchi,soxe,trigiahd,ng_giao_dk,ng_ngthu

FROM hop_dong

WHERE ng_giao_dk > #12/21/95# **AND** ng_giao_dk <= #12/31/95#

d) **SELECT** chitiet_hd.matho,tentho,COUNT(macv) **AS** soluongcv

FROM chiTiet_hd **INNER JOIN** tho **ON** chiTiet_hd.matho = tho.matho

GROUP BY chiTiet_hd.matho,tentho

HAVING COUNT(macv) >= **ALL** (**SELECT COUNT**(macv) **FROM** chiTiet_hd **GROUP BY** matho)

e) **SELECT** chiTiet_hd.matho,tentho,SUM(trigia_cv) **AS** congtrigia_cv

FROM chiTiet_hd **INNER JOIN** tho **ON** chiTiet_hd.matho = tho.matho

GROUP BY chiTiet_hd.matho,tentho

HAVING SUM(trigia_cv) >= **ALL** (**SELECT SUM**(trigia_cv) **FROM** chiTiet_hd **GROUP BY** matho)

Câu 4:

$F_1 = \{ \text{MATHO} \rightarrow \text{TENTHO}, \text{NHOM}, \text{NHOM_TRUONG} \}$

$Q_1 : \text{Tho}(\text{MATHO}, \text{TENTHO}, \text{NHOM}, \text{NHOM_TRUONG})$

$K_1 = \text{MATHO}$

$\Rightarrow Q_1$ ở dạng chuẩn BC

$F_2 = \{ \text{MACV} \rightarrow \text{NOIDUNGCV} \}$

Q_2 : Congviec (MACV, NOIDUNGCV)

$K_2 = \text{MACV}$

$\Rightarrow Q_2$ ở dạng chuẩn BC

$F_3 = \{ \text{SOHD} \rightarrow \text{NGAYHD}, \text{MAKH}, \text{SOXE}, \text{TRIGIAHD}, \text{NG_GIAO_DK}, \text{NG_NGTHU};$
 $\text{MAKH} \rightarrow \text{TENKH}, \text{DCHI} \}$

Q_3 : Hopdong (SOHD, MACV, TRIGIAHD, NG_GIAO_DK, NG_NGTHU)

$K_3 = \text{SOHD}$

$\Rightarrow Q_3$ ở dạng chuẩn 2

$F_4 = \{ \text{SOHD}, \text{MACV} \rightarrow \text{TRIGIA_CV}, \text{MATHO}, \text{KHOANTHO} \}$

Q_4 : ChiTiet_hd (SOHD, MACV, TRIGIA_CV, MATHO, KHOANTHO)

$K_4 = \{ \text{SOHD}, \text{MACV} \}$

$\Rightarrow Q_4$ ở dạng chuẩn BC

$F_5 = \{ \text{SOPH} \rightarrow \text{NGAYPH}, \text{SOHD}, \text{HOTEN}, \text{SOTIENTHU}; \text{SOHD} \rightarrow \text{MAKH} \}$

Q_5 : Phieu_thu (SOPH, NGAYPH, SOHD, MAKH, HOTEN, SOTIENTHU)

$K_5 = \text{SOPH}$

$\Rightarrow Q_5$ ở dạng chuẩn 2

Vậy lược đồ cơ sở dữ liệu C đạt dạng chuẩn 2.

Để nâng cấp lược đồ cơ sở dữ liệu trên ta phải phân rã Q_3 và Q_5 thành:

$F_{31} = \{ \text{SOHD} \rightarrow \text{NGAYHD}, \text{MAKH}, \text{SOXE}, \text{TRIGIAHD}, \text{NG_GIAO_DK}, \text{NG_NGTHU} \}$

Q_{31} : Hopdong (SOHD, NGAYHD, MAKH, SOXE, TRIGIAHD, NG_GIAO_DK, NG_NGTHU)

$K_{31} = \text{SOHD}$

$\Rightarrow Q_{31}$ đạt chuẩn BC

$F_{32} = \{ \text{MAKH} \rightarrow \text{TENKH}, \text{DCHI} \}$

Q_{32} : Khachhang (MAKH, TENKH, DCHI)

$K_{32} = \text{MAKH}$

$\Rightarrow Q_{32}$ đạt chuẩn BC

$F_{51} = \{ \text{SOPH} \rightarrow \text{NGAYPH}, \text{SOHD}, \text{HOTEN}, \text{SOTIENTHU} \}$

Q_{51} : Phieu_thu (SOPH, NGAYPH, SOHD, HOTEN, SOTIENTHU)

$K_{51} = \text{SOPH}$

$\Rightarrow Q_{51}$ đạt chuẩn BC

$F_{52} = \{ \text{SOHD} \rightarrow \text{MAKH} \}$

Q_{52} : Hopdong (SOHD, MAKH)

$K_{52} = \text{SOHD} \Rightarrow Q_{52}$ đạt chuẩn BC

Lược đồ Q_{52} là lược đồ con của Q_{31} nên ta loại Q_{52} khỏi lược đồ cơ sở dữ liệu C. Vậy lược đồ cơ sở dữ liệu C được phân rã thành các lược đồ $Q_1, Q_2, Q_{31}, Q_{32}, Q_4, Q_{51}$

----oOo----

Đề 3

Cho một lược sơ đồ cơ sở dữ liệu C dùng để quản lý hoạt động kinh doanh kiều hối của một đơn vị. Lược đồ cơ sở dữ liệu C gồm các lược đồ quan hệ như sau :

Q₁: Nguyen_te (LOAINT, NGAY, TIGIA, TILE_HH)

Tân từ : Mỗi nguyên tệ được nhận diện duy nhất qua LOAINT. Các loại nguyên tệ có thể là: USD, FF, DM, YEN, ... Thuộc tính TIGIA chỉ tỉ giá so với tiền đồng VN của mỗi nguyên tệ trong ngày (do Ngân hàng Ngoại thương quyết định vào đầu mỗi ngày và cố định trong ngày). Thuộc tính TILE_HH là tỉ lệ % huê hồng mà công ty được hưởng trên giá trị chi trả của mỗi nguyên tệ (tỉ lệ này cố định trong năm).

Lưu ý : Quan hệ này chỉ chứa các nguyên tệ mà công ty có chi trả kiều hối. Mỗi khi cần chi trả một loại nguyên tệ mới, công ty bắt đầu lưu tỉ giá nguyên tệ mới kể từ ngày chi trả trở đi.

Q₂: Don_vi (MADV, NUOC)

Tân từ : Công ty làm đại diện cho khoảng 50 đơn vị của nước ngoài để chi trả kiều hối cho khách hàng tại VN. Mỗi đơn vị có mã riêng để phân biệt (MADV) và đặt trụ sở chính tại 1 nước (NUOC) .

Q₃: Danh_sach (MADV, SODS, NGAYDS)

Tân từ : Mỗi danh sách liên quan đến 1 đơn vị, có một số thứ tự (SODS) để phân biệt với các danh sách khác của cùng đơn vị. Thuộc tính NGAYDS là ngày công ty nhận được danh sách, cũng là ngày mà đơn vị tại nước ngoài gửi danh sách cho công ty. Trong một ngày, một đơn vị tại nước ngoài chỉ gửi tối đa một danh sách.

Q₄: Ctiet_ds (MADV, SODS, HOTENKH, DIACHI, LOAINT, TIENNT)

Tân từ : Mỗi danh sách chi trả của một đơn vị có thể gồm nhiều khách hàng. Giả sử rằng thuộc tính HOTENKH có thể thêm một số thông tin phụ đủ để phân biệt với các khách hàng khác trong cùng danh sách. Mỗi khách hàng chỉ có một địa chỉ (DIACHI) và nhiều khách hàng có thể có chung một địa chỉ. Trong một danh sách, mỗi khách hàng chỉ nhận một loại nguyên tệ với số tiền nguyên tệ là TIENNT

Q₅: Giay_bao (SOGB, MADV, SODS, HOTENKH, NGAYGB, LAN)

Tân từ : Sau khi nhận danh sách của một đơn vị, công ty gửi giấy báo cho các khách hàng tại VN. Mỗi giấy báo có số thứ tự là SOGB (đánh số tăng dần) để phân biệt với các giấy báo khác (của cùng đơn vị hoặc khác đơn vị). Nếu sau 3 lần gửi giấy báo (mỗi lần cách nhau 1 tuần) mà khách hàng không đến nhận tiền, công ty sẽ gửi trả cho đơn vị tại nước ngoài).

Q₆: Chi_tra (SOPCHI, MADV, SODS, HOTENKH, NGAYCTRA, SOTIENVN)

Tân từ : Mỗi khách hàng trong danh sách của 1 đơn vị, sau khi nhận giấy báo, sẽ đến nhận tiền đồng VN 1 lần tại công ty và mỗi phiếu chi tiền cho khách sẽ có số là SOPCHI để phân biệt với bất kỳ phiếu chi khác. Thuộc tính SOTIENVN chỉ số tiền ĐVN mà khách hàng nhận được tương đương với TIENNT ghi trong danh sách. Tỉ giá qui đổi được tính vào ngày chi trả (NGAYCTRA). Số tiền huê hồng mà công ty được hưởng sẽ được tính toán dựa trên số tiền thực chi (SOTIENVN) và tỉ lệ huê hồng của nguyên tệ.

Câu hỏi :

1. Xác định tập F gồm tất cả các phụ thuộc hàm suy ra từ tập của các lược đồ quan hệ. Xác định khóa cho từng lược đồ quan hệ.
2. Mô tả tất cả các ràng buộc toàn vẹn của lược đồ cơ sở dữ liệu. Lập bảng tầm ảnh hưởng tổng hợp.
3. Dùng ngôn ngữ SQL để thực hiện những yêu cầu sau:
 - a) Cho biết tỉ giá của các nguyên tệ trong ngày 21/12/95
 - b) Cho biết những danh sách chi trả kiểu hối của các đơn vị có trụ sở chính đặt tại nước Pháp.
 - c) Cho biết những khách hàng không đến nhận tiền.
 - d) Cho tổng số tiền huê hồng mà công ty được trong khoảng thời gian từ ngày d_1 đến ngày d_2 .
 - e) Cho biết đơn vị nước ngoài có tổng số tiền chi trả (tính theo tiền đồng VN) cao nhất.
4. Lược đồ cơ sở dữ liệu C ở dạng chuẩn mấy (cao nhất) ? Hãy dùng thuật toán phân rã để nâng cấp cơ sở dữ liệu trên.

Lưu ý : Các thuộc tính có miền giá trị là ngày dương lịch xem như thuộc tính đơn.

----oOo----

Đề 4

Cho một lược đồ cơ sở dữ liệu C dùng để quản lý việc thuê mượn phòng tại một khách sạn. Lược đồ cơ sở dữ liệu C gồm các lược đồ quan hệ như sau :

Q₁: Phong (MAPH, SO_NGUOI, DACDIEM, GIA_PHONG)

Tân từ: Các phòng của khách sạn được phân biệt với nhau qua MAPH. SO_NGUOI là khả năng chứa tối đa của phòng. DACDIEM mô tả số đặc điểm của phòng. GIA_PHONG là giá cả thuê phòng trong 1 ngày.

Q₂: Tien_nghi (LOAI_TN, TEN_TN)

Tân từ: Ngoài các vật dụng tối thiểu, khách sạn có thể trang bị thêm một số tiện nghi khác cho các phòng như : điện thoại, tivi, tủ lạnh, ... LOAI_TN là mã số để phân biệt từng loại tiện nghi. TEN_TN là tên gọi của loại tiện nghi.

Q₃: Tai_san (LOAI_TN, STT, MAPH, NGAY_TB)

Tân từ : Mỗi loại tiện nghi, khách sạn có thể mua một số lượng lớn và STT dùng để phân biệt các vật dụng trong cùng loại tiện nghi. Một vật dụng có thể được sắp xếp trang bị cho nhiều phòng khác nhau nhưng trong một ngày vật dụng chỉ trang bị cho một phòng. MAPH là phòng được trang bị và NGAY_TB là ngày bắt đầu trang bị.

Lưu ý : Mỗi khi một vật dụng được thay đổi phòng thì cập nhật lại MAPH và NGAY_TB của vật dụng đó.

Q₄: Thue_phong (MAPH, HOTEN, NGAYBD, NGAYKT, NGAYTRA, LOAIDV, NGAYDV, TIENDV)

Tân từ : HOTEN là họ tên của khách thuê phòng MAPH. Giả sử rằng họ tên các khách thuê phòng trong cùng một phòng trong một ngày luôn luôn khác nhau. NGAYBD và NGAYKT là ngày bắt đầu và ngày kết thúc (dự kiến) thuê phòng. NGAYTRA là ngày trả thật sự. Giả sử rằng không có trường hợp khách trả phòng và thuê lại chính phòng đó trong cùng một ngày. Số tiền thuê phòng được chia đều cho số khách thuê trong cùng phòng.

Khách thuê phòng có thể sử dụng thêm các dịch vụ (gọi điện thoại đường dài, thuê xe, thủ tục hành chính, ...) LOAI_DV là mã số của loại dịch vụ sử dụng. NGAYDV ngày dịch vụ thực hiện. TIENDV là số tiền khách thuê phải trả cho dịch vụ.

Nếu trong cùng một ngày khách thuê phòng sử dụng 1 dịch vụ nhiều lần thì tiền dịch vụ được cộng dồn lại thành một lần và tạo thành một bộ (ví dụ trong ngày gọi điện thoại 3 cuộc với số tiền phải trả lần lượt là : 5000ĐVN, 4500ĐVN, 2000ĐVN thì sẽ được tính chung một lần là 11500ĐVN). Các dịch vụ được tính riêng đối với từng khách. Nếu là dịch vụ chung cho một số khách thì sẽ tính tiền cho một đơn vị khách đại diện nào đó.

Câu hỏi :

1. Xác định tập F gồm tất cả các phụ thuộc hàm suy ra từ tân từ của các lược đồ quan hệ. Xác định các khóa cho từng lược đồ quan hệ.
2. Mô tả tất cả các ràng buộc toàn vẹn của lược đồ cơ sở dữ liệu C. Lập bảng tầm ảnh hưởng tổng hợp của các ràng buộc toàn vẹn.
3. Dùng ngôn ngữ SQL để thực hiện những yêu cầu sau :
 - a) Cho biết các thông tin của các phòng có khả năng chứa trên 3 người.
 - b) Cho biết các thông tin của các phòng có trang bị máy lạnh (LOAITN='ML')

- c) Cho biết các thông tin của các phòng hiện nay (02/01/96) có trang bị máy lạnh.
 - d) Giả sử hôm nay là ngày 02/01/96. Tính tổng số tiền phải trả (tiền thuê phòng và tiền dịch vụ) của từng khách đã thuê phòng X từ ngày 21/12/95 và trả phòng vào hôm nay .
 - e) Cho biết doanh số thu được của từng phòng (không tính tiền dịch vụ)
1. Lược đồ cơ sở dữ liệu C ở dạng chuẩn mấy (cao nhất) ?. Hãy dùng thuật toán phân rã để nâng cấp lược đồ cơ sở dữ liệu C.

Lưu ý : Các thuộc tính có miền giá trị là ngày dương lịch xem như thuộc tính đơn.

----oOo----

Đề 5

Cho một lược đồ cơ sở dữ liệu C dùng để quản lý việc cho mượn sách tại một thư viện (xem tại chỗ hoặc mang về nhà). Lược đồ cơ sở dữ liệu C gồm các lược đồ quan hệ như sau :

Q₁ : The_loai (MATL, TENTL)

Tân từ : Sách của thư viện được phân chia theo thể loại để bạn đọc dễ dàng tra cứu. MATL là mã số của từng thể loại và dùng để phân biệt giữa các thể loại. TENTL là tên gọi của thể loại.

Q₂ : Sach (MASH, TENSCH, NGUYEN_TAC, TAC_GIA, MATL)

Tân từ : MASH dùng để phân biệt các quyển sách. TENSCH là tên (tựa) bằng tiếng Việt của sách và NGUYEN_TAC là tên nguyên tác (tiếng Việt hoặc tiếng nước ngoài). TAC_GIA là tên tác giả (hay nhóm các tác giả) của sách. Nếu sách có nhiều tập hay nhiều bản thì cũng xem như các đầu sách khác nhau và có mã số khác nhau. MATL là mã thể loại của sách.

Q₃ : phieu_muon (MADG, TENDG, DCHI, NGAYCAP, MASH, NGAYMUON, NGAYTRA, TAI_CHO)

Tân từ : Mỗi độc giả chỉ có một phiếu mượn sách với mã số là MADG để phân biệt với các độc giả khác. Các thuộc tính TENDG, DCHI là tên và địa chỉ của độc giả. NGAYCAP là ngày cấp thẻ cho độc giả. MASH là mã số của sách mượn. Giả sử không có trường hợp mượn rồi trả lại cùng 1 quyển sách trong cùng 1 ngày. Nếu sách mượn đọc tại chỗ thì thuộc tính TAI_CHO có giá trị True và NGAYMUON=NGAYTRA. Nếu sách mượn về nhà thì thuộc tính TAI_CHO có giá trị False và NGAYTRA sẽ có giá trị trống cho đến khi sách được mang trả lại cho thư viện. Mỗi độc giả chỉ được giữ tại nhà tối đa 3 quyển sách và mỗi quyển sách chỉ được giữ tại nhà tối đa 30 ngày (không cần lưu ý đến biện pháp xử lý nếu khách vi phạm nội qui)

Q₄ : Le_phi (MADG, NAM, NGAY_NOP, SOTIEN)

Tân từ : Độc giả phải đóng lệ phí hằng năm (NAM) để gia hạn thẻ mới được mượn sách. NGAY_NOP, SOTIEN là ngày và số tiền nộp lệ phí cho NAM.

Câu hỏi :

1. Xác định tập F gồm tất cả các phụ thuộc hàm suy ra từ tân từ của các lược đồ quan hệ. Xác định các khóa cho từng lược đồ quan hệ.
2. Mô tả tất cả các ràng buộc toàn vẹn của lược đồ cơ sở dữ liệu C. Lập bảng tầm ảnh hưởng tổng hợp của các ràng buộc toàn vẹn.
3. Dùng ngôn ngữ SQL để thực hiện những yêu cầu sau :
 - a) Cho biết danh sách độc giả và những quyển sách mượn quá 20 ngày (kể từ ngày 02/01/96).
 - b) Cho biết những quyển sách có tên thể loại là “Tin học” và có sự tham gia biên soạn của tác giả “X”.
 - c) Cho biết tổng số lần mượn của từng quyển sách.
 - d) Cho biết tổng số lần mượn của từng thể loại sách.
 - e) Cho biết thể loại sách nào được mượn nhiều nhất.
4. Lược đồ cơ sở dữ liệu C ở dạng chuẩn mấy (cao nhất) ? Hãy dùng thuật toán phân rã để nâng cấp lược đồ cơ sở dữ liệu C.

Lưu ý : Các thuộc tính có miền giá trị là ngày dương lịch xem như thuộc tính đơn.

TÀI LIỆU THAM KHẢO

- [1].Nhập môn cơ sở dữ liệu
Đồng Thị Bích Thủy - Nguyễn An Tế
- [2].Cơ sở dữ liệu – lý thuyết và thực hành
Nguyễn Bá Tường
- [3].Cơ sở dữ liệu
Đỗ Trung Tuấn
- [4].Mô hình dữ liệu và ngôn ngữ truy vấn
JEFFREY D. ULLMAN
- [5].Lý thuyết thiết kế cơ sở dữ liệu
JEFFREY D. ULLMAN
- [6].The Theory Of Relational Databases
DAVID MAIER

----oOo----