Cơ sở dữ liệu

Văn-Giang Nguyễn

Bộ môn Hệ thống thông tin,
Khoa Công nghệ thông tin,
Học viện Kỹ thuật quân sự
Email: giangnv@mta.edu.vn

Slides courtesy of : Đỗ Thị Mai Hường, HTTT, CNTT, HVKTQS

Nội dung môn học

- Chương 1: Tổng quan về cơ sở dữ liệu (5)
- Chương 2: Mô hình liên kết thực thể (5)
- **Chương 3: Mô hình dữ liệu quan hệ (10)**
- Chương 4: Phụ thuộc hàm và các dạng chuẩn CSDL (1
- Chương 5: Hệ quản trị SQL Server (10)
- Chương 6: Ngôn ngữ truy vấn CSDL (15)
- Chương 7: Lập trình T-SQL (15)

Giới thiệu

- Do tiến sĩ E.F. Codd đưa ra
  - "A relation model for large shared data banks", Communications of ACM, 6/1970.
  - A. Turin Award 1981 (~ Nobel Prize in CS)
- Cung cấp một cấu trúc dữ liệu đơn giản và đồng bộ
  - Khái niệm quan hệ
- Có nền tảng lý thuyết vững chắc
  - Lý thuyết tập hơn
- Là cơ sở của các HQT CSDL thương mại
  - Oracle, DB2, SQL Server

Các khái niệm của mô hình quan hệ

- Quan hệ (Relation)
- Thuộc tính (Attribute)
- Lược đồ (Schema)
- Bộ (Tuple)
- Miền giá trị (Domain)

Quan hệ

❖Các thông tin lưu trữ trong CSDL được tổ chức thành <u>bảng (table) 2 chiều</u> gọi là quan hệ

1 cột là 1 thuộc tính của nhân viên

TENNV	HONV	NS	DIACHI	GT	LUONG	PHG
Tùng	Nguyễn	12/08/1955	638 HQV CG	Nam	6000	5
Chuyên	Bùi	07/04/1970	255 XT CG	Nữ	5500	4
Dũng	Hoàng	09/05/1965	51 NTH BĐ	Nam	6000	4

1 dòng là 1 nhân viên

Tên quan hệ là NHANVIEN

Quan hệ

- Quan hệ gồm
  - Tên
  - Tập hợp các cột
    - Cố định
    - Được đặt tên
    - Có kiểu dữ liệu
  - Tập hợp các dòng
    - Thay đổi theo thời gian

- Một dòng ~ Một thực thể
- Quan hệ ~ Tập thực thể

Thuộc tính

- Tên các cột của quan hệ
- Mô tả ý nghĩa cho các giá trị tại cột đó

Thuộc tính

TENNV	HONV	NS	DIACHI	GT	LUONG	PHG
Tùng	Nguyễn	12/08/1955	638 HQV CG	Nam	6000	5
Chuyên	Bùi	07/04/1970	255 XT CG	Nữ	5500	4
Dũng	Hoàng	09/05/1965	51 NTH BĐ	Nam	6000	4

- Tất cả các dữ liệu trong cùng 1 cột có cùng kiểu dữ liệu.

Lược đồ

- Lược đồ quan hệ
  - Tên của quan hệ
  - Tên của tập thuộc tính

Lược đồ quan hệ

NHANVIEN(MANV, TENNV, HONV, NS, DIACHI, GT, LUONG, PHG)

Là tập hợp

Lược đồ

❖Lược đồ CSDL: Gồm nhiều lược đồ quan hệ

Lược đồ CSDL

↓

NHANVIEN	**(MANV, TENNV, HONV, NS, DIACHI, GT, LUONG, PHG)**
PHONGBAN	**(MAPHG, TENPHG, TRPHG, NG_NHANCHUC)**
DIADIEM_PHG	**(MAPHG, DIADIEM)**
THANNHAN	**(MA_NVIEN, TENTN, GT, NS, QUANHE)**
DEAN	**(TENDA, MADA, DDIEM_DA, PHONG)**

Bộ

- Là các dòng của quan hệ (trừ dòng tiêu đề - tên của các thuộc tính
- Thể hiện dữ liệu cụ thể của các thuộc tính trong quan hệ

<Tùng, Nguyễn, 12/08/1955, 638 HQV CG, Nam, 6000, 5>

Dữ liệu cụ thể của thuộc tính

Miền giá trị

- Là tập các giá trị gắn liền với một thuộc tính
  - Kiểu dữ liệu cơ sở
    - Chuối ký tự
    - Số
  - Các kiểu dữ liệu phức tạp
    - Tập hợp (set)
    - Danh sách (list)
    - Mảng (array)
    - Bản ghi (record)

Không được chấp nhận

- Ví dụ
  - TENNV: string
  - Luong: integer

Định nghĩa hình thức

- Lược đồ quan hệ
  - Cho $A_1, A_2, \ldots, A_n$ là các thuộc tính
  - Có các miền giá trị $D_1, D_2, \ldots, D_n$ tương ứng
  - Ký hiệu $R(A_1{:}D_1, A_2{:}D_2, \ldots, A_n{:}D_n)$
    Hoặc $R(A_1, A_2, \ldots, A_n)$ là một lược đồ quan hệ
  - <u>Bậc</u> của lược đồ quan hệ là số lượng thuộc tính trong lược đồ.

- NHANVIEN (MANV: integer, TENNV: string, HONV: string, NGSINH:date. DCHI:string, GT:string, LUONG: integer, DONVI: integer)
  - NHANVIEN là một lược đồ bậc 8 mô tả đối tượng nhân viên
  - MANV là một thuộc tính có miền giá trị là số nguyên

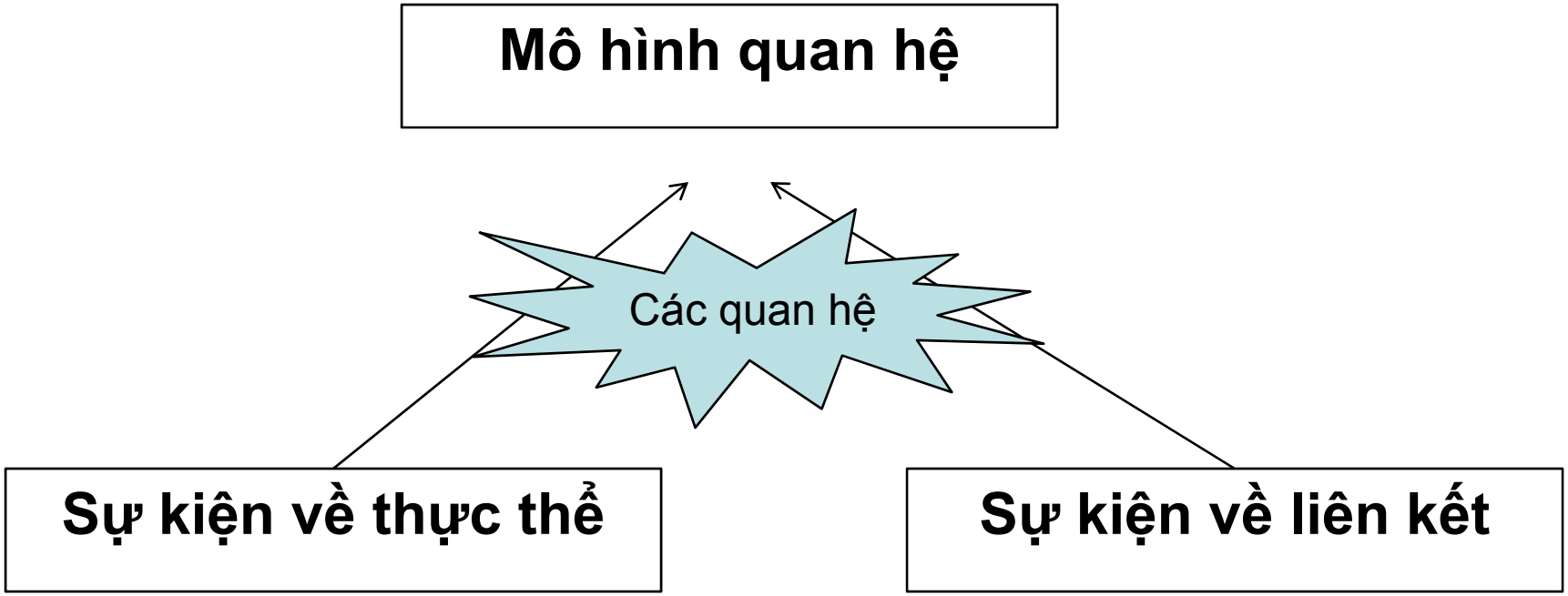
Định nghĩa hình thức

- Quan hệ (hay thể hiện quan hệ)
  - Một quan hệ r của lược đồ quan hệ $R(A_1,A_2,\ldots,A_n)$, ký hiệu r(R) là một tập các bộ $r=\{t_1,t_2,\ldots,t_k\}$
  - Trong đó mỗi t_i là 1 danh sách có thứ tự của n giá trị $t_i = <v_1,v_2,\ldots,v_n>$
    - Mỗi v_j là một phần tử của miền giá trị $DOM(A_j)$ hoặc giá trị rỗng.

	TENNV	HONV	NS	DIACHI	GT	LUONG	PHG
t_1	Tùng	Nguyễn	12/08/1955	638 HQV CG	Nam	6000	5
t_2	Chuyên	Bùi	07/04/1970	255 XT CG	Nữ	5500	4
t_3	Dũng	Hoàng	09/05/1965	51 NTH BĐ	Nam	6000	4
t_4	Minh	Lê	15/12/1960	Null	Nam	3900	5

v_j

Thể hiện mô hình quan hệ



Tóm tắt các ký hiệu

- Lược đồ quan hệ R bậc n: $R(A_1,A_2,\ldots,A_n)$
- Tập thuộc tính của R: R^+
- Quan hệ (Thể hiện quan hệ): R, S, P, Q
- Bộ: t, u, v
- Miền giá trị của thuộc tính A: DOM(A), MGT(A)
- Giá trị tại thuộc tính A của bộ thứ t: t.A, t[A]

Đại số quan hệ

- Giới thiệu
- Đại số quan hệ
- Phép toán tập hợp
- Phép chọn
- Phép chiếu
- Phép tích Cartesian
- Phép nối
- Phép chia
- Các phép toán khác

Giới thiệu

- Xét một số xử lý trên quan hệ NHANVIEN
  - Thêm mới một nhân viên
  - Chuyển nhân viên có tên Tùng sang phòng số 1
  - Cho biết họ tên và ngày sinh của các nhân viên có lương thấp hơn 5000

TENNV	HONV	NS	DIACHI	GT	LUONG	PHG
Tùng	Nguyễn	12/08/1955	638 HQV CG	Nam	6000	5
Chuyên	Bùi	07/04/1970	255 XT CG	Nữ	5500	4
Dũng	Hoàng	09/05/1965	51 NTH BĐ	Nam	6000	4
Minh	Lê	15/12/1960	10 TĐN CG	Nam	3900	5

Giới thiệu

- Có 2 loại xử lý
  - Làm thay đổi dữ liệu (cập nhật): thêm, sửa, xóa
  - Không làm thay đổi dữ liệu (rút trích): truy vấn

- Thực hiện các xử lý
  - Đại số quan hệ (Relational algebra)
    + Biểu diễn câu truy vấn dưới dạng biểu thức
  - Phép tính quan hệ (Relational Calculus)
    + Biểu diễn kết quả
  - SQL (Structured Query Language).

Đại số quan hệ

- Giới thiệu
- Đại số quan hệ
- Phép toán tập hợp
- Phép chọn
- Phép chiếu
- Phép tích Cartesian
- Phép nối
- Phép chia
- Các phép toán khác

Đại số (nói chung)

- Đại số
  - Toán tử (operator)
  - Toán hạng (operand)
- Trong số học
  - Toán tử: +, -, *, /
  - Toán hạng – biến (variable): x,y,z
  - Hằng số (constant)
  - Biểu thức
    - (x+1)/(y+2)

Đại số quan hệ

- Biến là các quan hệ
  - Tập hợp (set)
- Toán tử là các phép toán (operations)
  - Dựa trên lý thuyết tập hợp
    - Hội $\cup$ (union)
    - Giao $\cap$ (intersect)
    - Trừ - (difference)
  - Rút trích 1 phần của quan hệ
    - Chọn σ (selection)
    - Chiếu π (projection)
  - Kết hợp các quan hệ
    - Tích đề-các $\times$ (Cartesian product)
    - Nối $\bowtie$ (join)
  - Đổi tên ρ

Đại số quan hệ

- Hằng số là thể hiện của quan hệ

- Biểu thức
  - Được gọi là câu truy vấn
  - Là chuỗi các phép toán đại số quan hệ.
  - Kết quả trả về là một thể hiện của quan hệ.

Đại số quan hệ

- Giới thiệu
- Đại số quan hệ
- **Phép toán tập hợp**
- Phép chọn
- Phép chiếu
- Phép tích Cartesian
- Phép nối
- Phép chia
- Các phép toán khác

Phép toán tập hợp

- Quan hệ là tập hợp các bộ
  - Phép hợp $R \cup S$
  - Phép giao $R \cap S$
  - Phép trừ $R - S$
- Tính khả hợp (tương thích đồng nhất – Union Compatibility): Hai lược đồ quan hệ $R(A_1, A_2, ..., A_n)$ và $S(B_1, B_2, ..., B_n)$ là khả hợp (tương thích) nếu
  (i) Cùng bậc n
  (ii) Có $DOM(A_i) = DOM(B_i), 1 \le i \le n$
- Kết quả của $\cup, \cap, -$ là một <u>quan</u> hệ có cùng tên thuộc tính với quan hệ đầu tiên (R)

Ví dụ

NHANVIEN

TENNV	NS	GT
Tùng	12/08/1955	Nam
Hằng	19/06/1968	Nữ
Như	20/09/1972	Nữ
Hùng	15/03/1962	Nam

THANNHAN

TENTN	NS_TN	GT_TN
Chính	11/10/1995	Nam
Khang	10/11/1988	Nữ
Phương	10/01/1998	Nữ
Minh	09/02/1992	Nam
Châu	05/06/1987	Nữ

Bậc n=3

DOM (TENNV) = DOM(TENTN)

DOM(NS) = DOM(NS_TN)

DOM(GT) = DOM(GT_TN)

Phép hợp

- Cho 2 quan hệ R và S khả hợp.
- Phép hợp của R và S
  - Ký hiệu $R \cup S$ hoặc $R + S$
  - Là một quan hệ gồm có các bộ thuộc R hoặc thuộc S, hoặc cả hai (các bộ trùng lặp sẽ bị bỏ)

$$R \cup S = \{t \mid t \in R \text{ or } t \in S\}$$

- Ví dụ

R

A	B
α	1
α	2
β	1

S

A	B
α	2
β	3

$R \cup S$

A	B
α	1
α	2
β	1
β	3

Phép giao

- Cho 2 quan hệ R và S khả hợp.
- Phép giao của R và S
  - Ký hiệu $R \cap S$ hoặc $R * S$
  - Là một quan hệ gồm có các bộ thuộc R đồng thời thuộc S

$$R \cap S = \{t \mid t \in R \text{ and } t \in S\}$$

- Ví dụ

R

A	B
α	1
α	2
β	3

S

A	B
α	2
β	3

$R \cap S$

A	B
α	2
β	3

Phép trừ

- Cho 2 quan hệ R và S khả hợp.
- Phép trừ của R và S
  - Ký hiệu $R - S$
  - Là một quan hệ gồm các bộ thuộc R và không thuộc S

$$R - S = \{t \mid t \in R \text{ and } t \notin S\}$$

- Ví dụ

R

A	B
α	1
α	2
β	1

S

A	B
α	2
β	3

$R - S$

A	B
α	1
β	1

Các tính chất

❖Giao hoán

$$R \cup S = S \cup R$$

$$R \cap S = S \cap R$$

❖Kết hợp

$$R \cup (S \cup T) = (R \cup S) \cup T$$

$$R \cap (S \cap T) = (R \cap S) \cap T$$

Đại số quan hệ

- Giới thiệu
- Đại số quan hệ
- Phép toán tập hợp
- **Phép chiếu**
- Phép chọn
- Phép tích Cartesian
- Phép nối
- Phép chia
- Các phép toán khác

Phép chiếu

- Cho LĐQH $R(A_1,A_2,\ldots,A_n)$, cho r là một quan hệ R, X là một tập con của lược đồ R. Chiếu của r lên tập thuộc tính X là một quan hệ trên lược đồ X, ký hiệu r.X gồm các phần tử của r sau khi đã lược bỏ các thuộc tính không thuộc tập X.
- Vậy $r.X = \{t.X : t \in r\}$, t.X là chiếu của phần tử t lên tập thuộc tính X.

Ví dụ về phép chiếu

❖Cho quan hệ r như sau:

A	B	C	D	E	F
a1	b1	c1	d1	e1	f1
a2	b2	c2	d2	e2	f2
a3	b3	c3	d3	e3	f3
a4	b4	c4	d4	e4	f4

❖X={A,B,C}. Khi đó chiếu r lên X ta được

A	B	C
a1	b1	c1
a2	b2	c2
a3	b3	c3
a4	b4	c4

Phép chiếu – định nghĩa khác

- Được dùng để lấy ra một vài cột của quan hệ R
- Ký hiệu $\pi_{A_1, A_2, \ldots, A_k}(R)$
- Kết quả trả về là một <u>quan hệ</u>
  - Có k thuộc tính
  - Có số bộ luôn ít hơn hoặc bằng số bộ của R

R

A	B	C
a1	b1	c1
a2	b2	c2
a3	b3	c3
a4	b4	c4

$\pi_{A,C}(R)$

A	C
a1	c1
a2	c2
a3	c3
a4	c4

Ví dụ 1

- Cho biết họ tên và lương của các nhân viên
  - Quan hệ: NHANVIEN
  - Thuộc tính: HONV, TENNV, LUONG

$$\pi_{\text{HONV,TENNV,LUONG}}(\text{NHANVIEN})$$

Ví dụ 2

❖Cho biết mã nhân viên có tham gia đề án hoặc có thân nhân

$$\pi_{\text{MANV}}(\text{DEAN})$$

$$\pi_{\text{MANV}}(\text{THANNHAN})$$

$$\pi_{\text{MANV}}(\text{DEAN}) \cup \pi_{\text{MANV}}(\text{THANNHAN})$$

Ví dụ 3

- Cho biết mã nhân viên có người thân và có tham gia đề án

Ví dụ 4

- Cho biết mã nhân viên không có người thân nào

Đại số quan hệ

- Giới thiệu
- Đại số quan hệ
- Phép toán tập hợp
- Phép chiếu
- **Phép chọn**
- Phép tích Cartesian
- Phép nối
- Phép chia
- Các phép toán khác

Phép chọn

- Chọn từ bảng quan hệ ra các phần tử thỏa mãn điều kiện nào đó

- Cho quan hệ r trên LDDQH R. P là mệnh đề logic. Phần tử t thuộc r thỏa mãn điều kiện P được ký hiệu là t(P).
  Phép chọn từ quan hệ r theo điều kiện P cho ta một quan hệ mới ký hiều là r(P) trên đúng lược đồ R và chứa các phần tử thỏa mãn điều kiện P

$$r(P) = \{t : t \in r \,\&\, t(P)\}$$

Ví dụ

- Giả sử điều kiện P là chọn ra các phần tử có giá trị trong thuộc tính B<=30

r

A	B	C
a1	10	1
a1	20	1
a2	30	1
a2	40	2

$r(P)$

A	B	C
a1	10	1
a1	20	1
a2	30	1

Phép chọn – Định nghĩa khác

- Được dùng để lấy ra các bộ của quan hệ R
- Các bộ được chọn phải thỏa mãn <u>điều kiện chọn P</u>
- Ký hiệu $\sigma_P(R)$
- P là biểu thức gồm các mệnh đề có dạng
  - <tên thuộc tính><phép so sánh><hằng số>
  - <tên thuộc tính><phép so sánh><tê thuộc tính>
    - <Phép so sánh> gồm $>, <, \geq, \leq, \neq, =$
    - Các mệnh đề được nối lại nhờ phép $\wedge, \vee, \neg$

Phép chọn

- Kết quả trả về là một <u>quan hệ</u>
  - Có cùng danh sách thuộc tính với R
  - Có số bộ luôn ít hơn hặc bằng số bộ của R
- Ví dụ

$$R$$

A	B	C	D
a1	a1	1	7
a1	a2	5	7
a2	a2	12	3
a2	a2	23	10

$$\sigma_{(A=B)\wedge(D>5)}(R)$$

A	B	C	D
a1	a1	1	7
a2	a2	23	10

Tính chất của phép chọn

❖Giao hoán

$$\sigma_{P_1}(\sigma_{P_2}(R)) = \sigma_{P_2}(\sigma_{P_1}(R))$$

❖Kết hợp nhiều phép chọn thành 1 phép chọn

$$\sigma_{P_1}(\sigma_{P_2}(R)) = \sigma_{P_1 \wedge P_2}(R)$$

Ví dụ 5

❖Cho biết các nhân viên ở phòng số 4

- Quan hệ: NHANVIEN
- Thuộc tính: PHG
- Điều kiện: PHG=4

$$\sigma_{PHG=4}(NHANVIEN)$$

Ví dụ 6

- Tìm các nhân viên có lương trên 2.5 tr ở phòng 4 hoặc các nhân viên có lương trên 3tr ở phòng 5
  - Quan hệ: NHANVIEN
  - Thuộc tính: LUONG, PHG
  - Điều kiện:
    - LUONG >2500000 và PHG=4 hoặc
    - LUONG >3000000 và PHG=5

$$\sigma_{(PHG=4\ AND\ LUONG>2.5tr)\ OR\ (PHG=5\ AND\ LUONG>3tr)}(NHANVIEN)$$

Chuỗi các phép toán

❖Kết hợp các phép toán đại số quan hệ

- Lồng các biểu thức lại với nhau

$$\pi_{A_1, A_2, \ldots, A_k}(\sigma_P(R)) = \sigma_P(\pi_{A_1, A_2, \ldots, A_k}(R))$$

- Thực hiện từng phép toán một
  - B1

$$\sigma_P(R)$$

  - B2

$\pi_{A_1, A_2, \ldots, A_k}$ (Quan hệ kết quả ở B1)

Lưu ý: Cần đặt tên cho quan hệ này

Phép gán

- Được sử dụng đề nhận lấy kết quả trả về của một phép toán
  - Thường là kết quả trung gian trong chuỗi các phép toán
  - Ký hiệu: $\leftarrow$
- Ví dụ
  - B1

$$s \leftarrow \sigma_P(R)$$

  - B2

$$KQ \leftarrow \pi_{A_1, A_2, \ldots, A_k}(s)$$

Phép đổi tên

❖Được dùng để đổi tên

- Quan hệ

  Xét quan hệ $R(B,C,D)$

  $\rho_S(R)$: (đọc là rhô) đổi tên quan hệ R thành S

- Thuộc tính

  $\rho_{X,C,D}(R)$ Đổi tên thuộc tính B thành X

  $\rho_{S(X,C,D)}(R)$ Đổi tên quan hệ R thành S và thuộc tính B thành X

Ví dụ 7

- Cho biệt họ và tên nhân viên làm việc ở phòng số 4
  - Quan hệ: NHANVIEN
  - Thuộc tính: HONV, TENNV
  - Điều kiện: PHG=4

- Cách 1

$$\pi_{\text{HONV,TENNV}}(\sigma_{\text{PHG=4}}(\text{NHANVIEN}))$$

- Cách 2

$$\text{NV_P4} \leftarrow \sigma_{\text{PHG=4}}(\text{NHANVIEN})$$

$$\text{KQ} \leftarrow \pi_{\text{HONV,TENNV}}(\text{NV_P4})$$

$$\text{KQ(HO,TEN)} \leftarrow \pi_{\text{HONV,TENNV}}(\text{NV_P4})$$

$$\rho_{\text{KQ(HO,TEN)}}(\pi_{\text{HONV,TENNV}}(\text{NV_P4}))$$

Đại số quan hệ

- Giới thiệu
- Đại số quan hệ
- Phép toán tập hợp
- Phép chiếu
- Phép chọn
- **Phép tích Cartesian**
- Phép nối
- Phép chia
- Các phép toán khác

Phép tích Đề các

- Được dùng để kết hợp các bộ của quan hệ lại với nhau

- Ký hiệu $R \times S$

- Kết quả trả về là một quan hệ Q
  - Mỗi bộ của Q là <u>tổ hợp</u> giữa 1 bộ trong R và 1 bộ trong S
  - Nếu R có u bộ và S có v bộ thì Q sẽ có u×v
  - Nếu R có n thuộc tính và S có m thuộc tính thì Q sẽ có (n+m) thuộc tính $(R^{+} \cap S^{+} = \emptyset)$

Tập thuộc tính của R

Phép tích Đề các

❖Ví dụ

R

A	B
α	1
β	2

S

B	C	D
α	10	+
β	10	+
β	20	-
γ	10	-

R×S

A	R.B	S.B	C	D
α	1	α	10	+
α	1	β	10	+
α	1	β	20	-
α	1	γ	10	-
β	2	α	10	+
β	2	β	10	+
β	2	β	20	-
β	2	γ	10	-

Phép tích Đề các

❖Thông thường sau phép tích Đềcác là phép chọn

$\mathbf{R} \times \mathbf{S}$

A	R.B	S.B	C	D
α	1	α	10	+
α	1	β	10	+
α	1	β	20	-
α	1	γ	10	-
β	2	α	10	+
β	2	β	10	+
β	2	β	20	-
β	2	γ	10	-

$\sigma_{A=S.B}(\mathbf{R} \times \mathbf{S})$

A	R.B	S.B	C	D
α	1	α	10	+
β	2	β	10	+
β	2	β	20	-

Ví dụ

❖Với mỗi phòng ban, cho biết thông tin của người trưởng phòng

- Quan hệ: PHONGBAN, NHANVIEN
- Thuộc tính: TRPHG, MAPHG, TENNV, HONV,…

TENPHG	MAPHG	TRPHG	NG_NCHUC
Nghien cuu	5	111222333	01/01/2012
Dieu hanh	4	123456789	01/06/2011
Quan ly	1	987654321	01/01/2011

MANV	TENNV	HONV	NS	DCHI	GT	LUONG	PHG
111222333	Tùng	Nguyễn	12/08/1955				
123456789	Hoàng	Bùi					
987654321	Như	Lê					
444555666	Hùng	Nguyễn					

Ví dụ 8

- B1: Tích Đề-các PHONGBAN và NHANVIEN

$$PB_NV \leftarrow (NHANVIEN \times PHONGBAN)$$

- B2: Chọn ra những bộ thỏa mãn TRPHG=MANV

$$KQ \leftarrow \sigma_{TRPHG=MANV}(PB_NV)$$

Ví dụ 9

❖Cho biết các phòng ban có cùng địa điểm với phòng số 5

- Quan hệ: DIADIEM_PHG
- Thuộc tính: DIADIEM, MAPHG
- Điều kiện: MAPHG=5

Phòng 5 có tập hợp những địa điểm nào?

MAPHG	DIADIEM
1	TP HCM
4	HA NOI
5	VUNG TAU
5	NHA TRANG
5	TP HCM

Phòng nào có địa điểm năm trong tập hợp đó?

MAPHG	DIADIEM
1	TP HCM
4	HA NOI
5	VUNG TAU
5	NHA TRANG
5	TP HCM

Ví dụ 9

❖B1: Tìm các địa điểm của phòng 5

$$DD_P5(DD) \leftarrow \pi_{DIADIEM}(\sigma_{MAPHG=5}(DIADIEM_PHG))$$

❖B2: Lấy ra các phòng có cùng địa điểm với DD_P5

$$R1 \leftarrow (\sigma_{MAPHG \neq 5}(DIADIEM_PHG)$$

$$R2 \leftarrow (\sigma_{DIADIEM=DD}(R1 \times DD_P5)$$

$$KQ \leftarrow \pi_{MAPHG}(R2)$$

Đại số quan hệ

- Giới thiệu
- Đại số quan hệ
- Phép toán tập hợp
- Phép chiếu
- Phép chọn
- Phép tích Cartesian
- **Phép nối: Theta join, Equi join, Natural join**
- Phép chia
- Các phép toán khác

Nhận xét

- Tập các phép toán $\sigma, \pi, \times, -, \cup$ được gọi là tập đầy đủ các phép toán đại số quan hệ
  - Các phép toán có thể biển diễn qua chúng

- Chúng ta cùng xem xét các phép nối sau đây, mặc dù chúng không cung cấp thêm sức mạnh nào cho đại số quan hệ - nhưng chúng là các câu lệnh truy vấn đã được đơn giản hóa và hay được sử dụng.
- Đọc kỹ các lưu ý về điều kiện sử dụng các phép nối.

Phép nối

- Với θ là phép so sánh $\neq, =, <, >, \leq, \geq$, ta có định nghĩa phép nối.
- Cho r và s là hai quan hệ tương ứng trên các lược đồ rời nhau R và S
- Phép kết nối của các quan hệ r và s, ký hiệu: $r \mid\!><\!\mid_{i\theta j} s$ là một quan hệ trên $R \cup S$ gồm những bộ thuộc tích Đềcác của r và s sao cho thành phần thứ i của quan hệ r có liên hệ θ với thành phần thứ j của quan hệ s
- Vậy kết nối θ: $r \mid\!><\!\mid_{i\theta j} s$ là chọn trong r×s các bộ mà các thành phần thứ *i, j* của các quan hệ r, s tương ứng thỏa mãn iθj, tức là

$$r \mid\!><\!\mid_{i\theta j} s = \{t \in r \times s : t(\theta)\}$$

Lưu ý: $R \cup S$ ở đây ký hiệu là lược đồ quan hệ tạo bởi các thuộc tính của R hợp với các thuộc tính của S.

Phép nối – Một cách định nghĩa khác

- Được dùng để tổ hợp 2 bộ có liên quan từ 2 quan hệ thành 1 bộ
- Ký hiệu $R \mid\!><\!\mid S$

$$R(A_1, A_2, ..., A_n) \mid\!><\!\mid S(B_1, B_2, ..., B_m)$$

- Kết quả của phép nối là một <u>quan hệ</u> Q
  - Có n+m thuộc tính $Q(A_1, A_2, ..., A_n, B_1, B_2, ..., B_m)$
  - Mỗi bộ của Q là tổ hợp của 2 bộ trong R và S, thỏa mãn <u>một số điều kiện nối</u> nào đó
    - Có dạng $A_i \theta B_j$ hoặc $i \theta j$
    - A_i là thuộc tính của R, B_j là thuộc tính của S
    - A_i và B_j có cùng miền giá trị
    - i, j là số thứ tự của thuộc tính trên các quan hệ R và S tương ứng
    - θ là phép so sánh $\neq, =, <, >, \leq, \geq$

Phân loại phép nối

❖Nối theta (theta join) là phép nối có điều kiện

- Ký hiệu $R \bowtie_C S$
- C gọi là điều kiện nối trên thuộc tính

(Nối theta == Tích đề các + phép chọn ở mục trước)

❖Nối bằng (equi join) khi C là điều kiện so sánh bằng

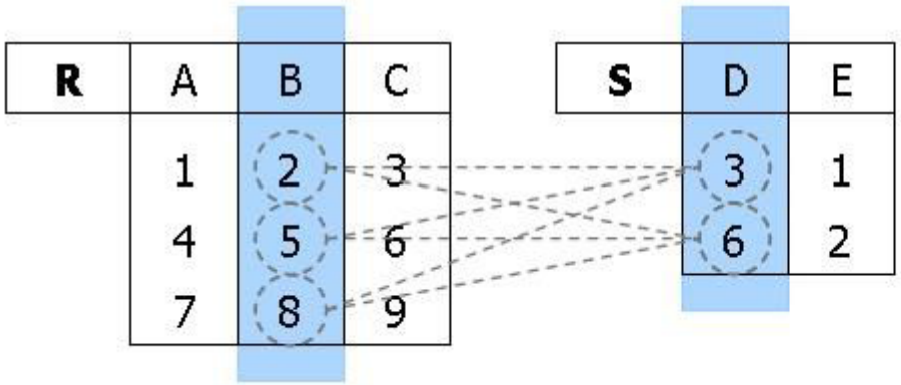
Lưu ý: Với **phép nối theta** và **nối bằng**, yêu cầu R và S không có thuộc tính chung.

❖Nối tự nhiên (natural join)

- Ký hiệu $R \bowtie S$ là quan hệ trên lược đồ $R \cup S$ gồm các phần tử t mà t chiếu lên R là phần tử thuộc r còn chiếu của t lên S là phần tử của s
- Vậy $r \bowtie s = \{t : t.R \in r, t.S \in s\}$

Lưu ý: Với **phép nối tự nhiên**, thường yêu cầu R và S phải có ít nhất một thuộc tính chung. Khi thực hiện nối, ta chỉ lấy các bộ trên R và S có cùng giá trị trên thuộc tính chung này. **Nếu** R và S không có thuộc tính chung, kết quả trả về là **tích đề các**.

Ví dụ phép nối theta



$R \bowtie_{B<D} S$

hoặc

$R \bowtie_{2<1} S$

Lưu ý: Với **phép nối theta** và **nối bằng**, yêu cầu R và S không có thuộc tính chung.

Ví dụ phép nối theta (wikipedia)

❖Giả sử có 2 quan hệ xe(car) và thuyền(boat), giả sử 1 khách hàng muốn mua cả xe và thuyền nhưng không muốn tiêu nhiền tiền vào thuyền hơn xe.

Car

CarModel	CarPrice
CarA	20,000
CarB	30,000
CarC	50,000

Boat

BoatModel	BoatPrice
Boat1	10,000
Boat2	40,000
Boat3	60,000

$$Car \bowtie_{CarPrice > BoatPrice} Boat$$

CarModel	CarPrice	BoatModel	BoatPrice
CarA	20,000	Boat1	10,000
CarB	30,000	Boat1	10,000
CarC	50,000	Boat1	10,000
CarC	50,000	Boat2	40,000

Lưu ý: Với **phép nối theta** và **nối bằng**, yêu cầu R và S không có thuộc tính chung.

Ví dụ phép nối bằng

R	A	B	C
	1	2	3
	4	5	6
	7	8	9

S	D	E
	3	1
	6	2

$R \bowtie_{C=D} S$

Hoặc

$R \bowtie_{3=1} S$

R	A	B	C
	1	2	3
	4	5	6
	7	8	9

S	C	D
	3	1
	6	2

$R \bowtie_{C=S.C} S$

Lưu ý: Với **phép nối theta** và **nối bằng**, yêu cầu R và S không có thuộc tính chung.

Ví dụ phép nối tự nhiên

Employee

Name	EmpId	DeptName
Harry	3415	Finance
Sally	2241	Sales
George	3401	Finance
Harriet	2202	Sales

Dept

DeptName	Manager
Finance	George
Sales	Harriet
Production	Charles

Employee ⋈ Dept

Name	EmpId	DeptName	Manager
Harry	3415	Finance	George
Sally	2241	Sales	Harriet
George	3401	Finance	George
Harriet	2202	Sales	Harriet

Lưu ý: Với **phép nối tự nhiên**, thường yêu cầu R và S phải có ít nhất một thuộc tính chung. Khi thực hiện nối, ta chỉ khớp các bộ trên R và S có cùng giá trị trên thuộc tính chung này. Nếu R và S không có thuộc tính chung, kết quả trả về là tích đề các.

Phép nối nửa

- Cho các quan hệ r và s trên các lược đồ R và S tương ứng
- Nối nửa của các quan hệ r và s, ký hiệu $r \ltimes s$ là một quan hệ trên lược đồ R gồm các bộ của $r \bowtie s$ chiếu lên R. Tức là

$$r \ltimes s = \{t : t \in (r \bowtie s).R\}$$

- Ví dụ:

r

A	B	C
a	b	c
d	b	c
d	b	f
c	a	d

s

B	C	D
b	c	d
b	c	e
a	d	f

$r \ltimes s$

A	B	C
a	b	c
d	b	c
c	a	d

Lưu ý: Do **nối nửa** dựa trên **phép nối tự nhiên**, yêu cầu R và S phải có ít nhất một thuộc tính chung. Khi thực hiện nối, ta chỉ khớp các bộ trên R và S có cùng giá trị trên thuộc tính chung này.

Ví dụ 10

❖Cho biết nhân viên có lương hơn lương của nhân viên "Tùng"

- Quan hệ: NHANVIEN
- Thuộc tính: LUONG

$$R(L_TUNG) \leftarrow \pi_{LUONG}(\sigma_{TENNV='TUNG'}(NHANVIEN))$$

$$KQ \leftarrow NHANVIEN |\!><\!|_{LUONG>L_TUNG} R$$

[Nối theta]

Ví dụ 11

❖Với mỗi nhân viên, hãy cho biết thông tin của phòng ban mà họ đang làm việc

- Quan hệ: NHANVIEN, PHONGBAN

[(1) Sử dụng nối tự nhiên; (2) sử dụng nối bằng
Lưu ý điều kiện sử dụng 2 phép nối này]

Ví dụ 12

- Với mỗi phòng ban, hãy cho biết các địa điểm của phòng ban đó
  - Quan hệ: PHONGBAN, DDIEM_PHG

Ví dụ 13

❖Với mỗi phòng ban, hãy cho biết thông tin của người trưởng phòng

- Quan hệ: PHONGBAN, NHANVIEN

Ví dụ 14

❖Cho biết lương cao nhất trong công ty

- Quan hệ: NHANVIEN
- Thuộc tính: LUONG

Ví dụ 15

- Cho biết phòng ban có cùng địa điểm với phòng 5
  - Quan hệ: DDIEM_PHG

Tập đầy đủ các phép toán ĐSQH

- Tập các phép toán $\sigma, \pi, \times, -, \cup$ được gọi là tập đầy đủ các phép toán đại số quan hệ
  - Các phép toán có thể biển diễn qua chúng
  - Ví dụ:

$$R \cap S = R \cup S - ((R - S) \cup (S - R))$$

$$R \mid\!><\!\mid_C S = \sigma_C (R \times S)$$

Đại số quan hệ

- Giới thiệu
- Đại số quan hệ
- Phép toán tập hợp
- Phép chiếu
- Phép chọn
- Phép tích Cartesian
- Phép nối: Theta join, Equi join, Natural join
- **Phép chia**
- Các phép toán khác

Phép chia

- Cho lược đồ quan hệ $R(A_1,A_2,\ldots,A_n)$, S là lược đồ con của R. Giả sử r,s là các quan hệ trên R, S tương ứng.

- Phép chia của quan hệ r cho quan hệ s, ký hiệu là $r \div s$ là quan hệ trên lược đồ R-S gồm các phần tử r sao cho mọi phần tử $u \in s$ và ghép t với u ta được phần tử thuộc r.

$$r \div s = \{t : \forall u \in s \& < t,u > \in r\}$$

- Một cách vắn tắt: Kết quả trả về là [các bộ với *các thuộc tính chỉ có* trong R] sao cho sự kết hợp của nó với [các bộ trong S] có mặt trong R

Phép chia – Một cách định nghĩa khác

- Được dùng để lấy ra một số bộ trong quan hệ R sao cho thỏa mãn với <u>tất cả</u> các bộ trong quan hệ S

- Ký hiệu $R \div S$
  - R(Z) và S(X)
    - Z là tập thuộc tính của R, X là tập thuộc tính của S
    - $X \subseteq Z$

- Kết quả của phép chia là một quan hệ T(Y)
  - Với Y = Z-X
  - Có t là một của T nếu <u>với mọi bộ</u> $t_S \in S$, tồn tại bộ $t_R \in R$ thỏa 2 điều kiện

$$t_R(Y) = t$$

$$t_R(X) = t_S(X)$$

Ví dụ

Completed

Student	Task
Fred	Database1
Fred	Database2
Fred	Compiler1
Eugene	Database1
Eugene	Compiler1
Sarah	Database1
Sarah	Database2

DBProject

Task
Database1
Database2

Completed
÷
DBProject

Student
Fred
Sarah

Ví dụ

R

A	B	C	D	E
α	a	α	a	1
α	a	γ	a	1
α	a	γ	b	1
β	a	γ	a	1
β	a	γ	b	3
γ	a	γ	a	1
γ	a	γ	b	1
γ	a	β	b	1

S

D	E
a	1
b	1

$R \div S$

Ví dụ 16

❖Cho biết mã nhân viên tham gia tất cả các đề án

- Quan hệ: PHANCONG, DEAN
- Thuộc tính: MANV

$$DA \leftarrow \pi_{MADA}(DEAN)$$

$$NV_DEAN \leftarrow \pi_{MANV,MADA}(PHANCONG)$$

$$MA_NV \leftarrow \pi_{MANV}(NV_DEAN \div DA)$$

Ví dụ 17

❖Cho biết mã nhân viên tham gia tất cả các đề án do phòng số 4 phụ trách

- Quan hệ: NHANVIEN, PHANCONG, DEAN
- Thuộc tính: MANV
- Điều kiện: PHONG=4

$$\text{P4_DA} \leftarrow \pi_{\text{MADA}}(\sigma_{PHG=4}(\text{DEAN}))$$

$$\text{NV_DEAN} \leftarrow \pi_{\text{MANV,MADA}}(\text{PHANCONG})$$

$$\text{MA_NV} \leftarrow \pi_{MANV}(\text{NV_DEAN} \div \text{P4_DA})$$

Phép chia

❖Biểu diễn phép chia thông qua tập đầy đủ của các phép toán ĐSQH

$$\mathrm{T}_1 \leftarrow \pi_\gamma(R)$$

$$\mathrm{T}_2 \leftarrow \mathrm{T}_1 \times S$$

$$\mathrm{T}_3 \leftarrow \pi_\gamma(\mathrm{T}_2 - R)$$

$$\mathrm{T} \leftarrow \mathrm{T}_1 - \mathrm{T}_2$$

Đại số quan hệ

- Giới thiệu
- Đại số quan hệ
- Phép toán tập hợp
- Phép chiếu
- Phép chọn
- Phép tích Cartesian
- Phép nối: Theta join, Equi join, Natural join
- Phép chia
- **Các phép toán khác**
  - Hàm kết hợp (Aggregation function)
  - Phép gom nhóm (Grouping)
  - Phép kết ngoài (Outer join)

Hàm kết hợp

- Nhận vào tập hợp các giá trị và trả về một giá trị đơn
  - AVG
  - MIN
  - MAX
  - SUM
  - COUNT

Ví dụ

A	B
1	2
3	4
1	2
1	2

SUM(B)=10
AVG(A)=1.5
MIN(A)=1
MAX(B)=4
COUNT(A)=4

Phép gom nhóm

❖Được dùng để phân chia quan hệ thành nhiều nhóm dựa trên điều kiện gom nhóm nào đó

❖Ký hiệu

$$G_1, G_2, ..., G_n \Im_{F_1(A_1), F_2(A_2), ..., F_n(A_n)}(E)$$

- E là biểu thức đại số quan hệ
- $G_1, G_2, ..., G_n$ các thuộc tính gom nhóm
- $F_1, F_2, ..., F_n$ các hàm
- $A_1, A_2, ..., A_n$ các thuộc tính tính toán trong hàm F

Ví dụ

R

A	B	C
α	2	7
α	4	7
β	2	3
β	2	10

$$\Im_{SUM(C)}(R)$$

$$_{A}\Im_{SUM(C)}(R)$$

Ví dụ 18

- Tính số lượng nhân viên và lương trung bình của cả công ty

$$\mathfrak{I}_{\text{COUNT(),AVERAGE(LUONG)}}(\text{NHANVIEN})$$

Ví dụ 19

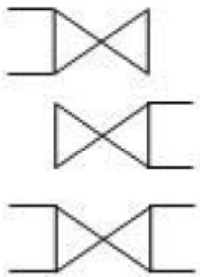
❖Tính số lượng nhân viên và lương trung bình của từng phòng ban

$$_{MAPHG}\mathfrak{I}_{COUNT(),AVERAGE(LUONG)}(NHANVIEN)$$

Phép nối ngoài

❖Mở rộng phép nối ngoài để tránh mất mát thông tin
- Thực hiện phép nối
- Lấy thêm các bộ không thỏa điều kiện nối

❖Có 3 hình thức nối
- Nối ngoài trái
- Nối ngoài phải
- Nối ngoài đầy đủ



(chỉ các bộ bên trái được điền null)

(chỉ các bộ bên phải được điền null)

(cả 2 phía được điền null)

Đôi khi còn được ký hiệu ⋈ LEFT ⋈ RIGHT ⋈ FULL

Phép nối ngoài

❖Ví dụ:

r_1

Employee	Department
Smith	sales
Black	production
White	production

r_2

Department	Head
production	Mori
purchasing	Brown

$r_1 \bowtie_{LEFT} r_2$

Employee	Department	Head
Smith	Sales	NULL
Black	production	Mori
White	production	Mori

$r_1 \bowtie_{RIGHT} r_2$

Employee	Department	Head
Black	production	Mori
White	production	Mori
NULL	purchasing	Brown

$r_1 \bowtie_{FULL} r_2$

Employee	Department	Head
Smith	Sales	NULL
Black	production	Mori
White	production	Mori
NULL	purchasing	Brown

VD Phép nối ngoài vs nối tự nhiên

❖Cho 2 quan hệ *instructor1* và *teaches1*

ID	*name*	*dept_name*
10101	Srinivasan	Comp. Sci.
12121	Wu	Finance
15151	Mozart	Music

ID	*course_id*
10101	CS-101
12121	FIN-201
76766	BIO-101

❖Nối tự nhiên

ID	*name*	*dept_name*	*course_id*
10101	Srinivasan	Comp. Sci.	CS-101
12121	Wu	Finance	FIN-201

❖Nối ngoài trái

ID	*name*	*dept_name*	*course_id*
10101	Srinivasan	Comp. Sci.	CS-101
12121	Wu	Finance	FIN-201
15151	Mozart	Music	*null*

❖Nối ngoài đầy đủ

ID	*name*	*dept_name*	*course_id*
10101	Srinivasan	Comp. Sci.	CS-101
12121	Wu	Finance	FIN-201
15151	Mozart	Music	*null*
76766	null	null	BIO-101

Ví dụ 20

- Cho biết tên nhân viên và tên phòng ban mà họ phụ trách (nếu có) – tức là nếu không có thì ta để trống mục phòng ban họ phụ trách!
  - Quan hệ: NHANVIEN, PHONGBAN
  - Thuộc tính: TENNV, TENPH

$$R1 \leftarrow NHANVIEN \; ⟕_{MANV=TRPHG} PHONGBAN$$

$$KQ \leftarrow \pi_{HONV,TENNV,TENPHG}(R_1)$$

Bài tập 1 – Thực hiện các phép toán

T_1

P	Q	R
10	a	5
15	b	8
25	a	6

T_2

A	B	C
10	b	6
25	c	3
10	b	5

$$T_1 \mid\!\times\!\mid_{T_1.P=T_2.A} T_2$$

$$T_1 \mid\!\times\!\mid_{T_1.Q=T_2.B} T_2$$

$$T_1 \mid\!\times\!\mid_{(T_1.P=T_2.A \text{ AND } T_1.R=T_2.C)} T_2$$

$$T_1 \sqsupset\!\bowtie_{T_1.P=T_2.A} T_2$$

$$T_1 \bowtie\!\sqsubset_{T_1.Q=T_2.B} T_2$$

Bài tập 2

❖Cho CSDL COONGTY gồm các lược đồ

- NHANVIEN(MANV, HONV, TENNV, NS, GT, DCHI, LUONG, MANGS, MADV)
- DONVI(MADV, TENDV, MANQL, NGAY_BD)
- DEAN(MADA, TENDA, DD_DA, MADV)
- THANNHAN(MANV, TEN_TN, NS, GT, QUANHE)
- NV_DEAN(MANV, MADA, SOGIO) [revised on 09/20]
- DONVI_DD(MADV, DD)

Bài tập 2 – Yêu cầu

- Đưa ra tên và địa chỉ của tất cả các nhân viên làm việc cho đơn vị.
- Với mỗi dự án có địa điểm hà Nôi, liệt kê mã số dự án, mã số của đơn vị kiểm soát, tên, địa chỉ và ngày sinh của người quản lý đơn vị.
- Tìm tên của nhân viên làm việc trên tất cả các dự án do đơn vị có mã số 5 kiểm soát.
- Tạo ra một danh sách các mã số dự án đối vwois các dự án có 1 nhân viên hoặc một người quản lý đơn vị kiểm soát dự án có tên là 'Long'
- Đưa ra tên của tất cả các nhân viên có nhiều hơn hoặc bằng 2 người phụ thuộc.
- Đưa ra các nhân viên không có người phụ thuộc.
- Đưa ra tên của những người quản lý dự án có ít nhất một người phụ thuộc.