


Mô hình dữ liệu quan hệ và đại số quan hệ

- ❑ CSDL quan hệ
- ❑ Ngôn ngữ truy vấn quan hệ - đại số quan hệ

Lịch sử phát triển của Mô hình dữ liệu quan hệ

- do E.F Codd đưa ra.
- cung cấp một cấu trúc dữ liệu đơn giản và đồng bộ - quan hệ - và nền tảng lý thuyết vững chắc.
- là cơ sở của hầu hết DBMS thương mại: Oracle, DB2, Sysbase ...
- đặc biệt được dùng trong thiết kế quan niệm: hoặc trực tiếp (tạo bảng sử dụng SQL DDL) hoặc phát sinh từ một lược đồ ER cho trước.



Cấu trúc cơ sở của mô hình quan hệ

Định nghĩa không hình thức về quan hệ

- Tất cả thông tin được lưu trữ trong CSDL Quan hệ được tổ chức thành các quan hệ
- Quan hệ -relation: có thể xem là một bảng các giá trị.
- Một quan hệ có:
 - ▣ Name
 - ▣ Columns
 - ▣ Rows
- Mỗi dòng trong bảng thể hiện dữ kiện cụ thể tương ứng với một thực thể hay mỗi kết hợp trong thể giới thực.
- Tên bảng và cột giúp diễn đạt ngữ nghĩa của các giá trị tại mỗi dòng.
- Tất cả các giá trị trong cùng một cột đều có cùng kiểu dữ liệu.
- Ex: KHACHHANG(MA_KH, TEN_KH, DC_KH)

Định nghĩa hình thức về Quan hệ

- Một quan hệ r trên một tập hợp D_1, D_2, \dots, D_n là một **tập con** của tích Cartesian ($D_1 \times D_2 \times \dots \times D_n$)
- \Rightarrow Một quan hệ là một tập hợp các bộ có n thuộc tính (d_1, d_2, \dots, d_n) , trong đó $d_i \in D_i$
- **Ex1:**
 - ▣ Cho $S_1 = \{0, 1\}, S_2 = \{a, b, c\}$, R là tập con của $S_1 \times S_2$.
 - ▣ Ta có: $r(R) = \{(0, a), (0, b), (1, c)\}$ là một quan hệ trên $S_1 \times S_2$.

Lược đồ quan hệ, lược đồ CSDL và Thuộc tính

- A_1, A_2, \dots, A_n là các thuộc tính(attribute) có các miền giá trị(domain) D_1, D_2, \dots, D_n thì $R(A_1:D_1, A_2:D_2, \dots, A_n:D_n)$ là một lược đồ quan hệ (relation schema).
- Ex:
 - ▣ Sinhvien(MASV:number,TENSV:string,KHOA:string) \Rightarrow Một lược đồ quan hệ được hình thành từ tên quan hệ R và một cấu trúc của quan hệ.

Lược đồ CSDL quan hệ ... (tt)

- Lược đồ CSDL quan hệ - **S** - (Relation database schema): là một tập hợp các lược đồ quan hệ của cùng một CSDL + tập hợp các ràng buộc toàn vẹn trên CSDL đó.
- $S = \{R_1, R_2, \dots, R_m\}$
- Thể hiện quan hệ (Relation Instance) hay Trạng thái quan hệ (Relation State) $r(R)$ của lược đồ quan hệ $R(A_1:D_1, A_2:D_2, \dots, A_n:D_n)$ có thể được xem là một bảng có n cột và có nhiều dòng. \Rightarrow Thay vì nói thể hiện của quan hệ, ta thường dùng từ quan hệ.

Thể hiện CSDL quan hệ ... (tt)

- Thể hiện CSDL (Relation Database Instance) - DB – của một lược đồ CSDL quan hệ là một tập hợp các quan hệ thể hiện một lược đồ CSDL quan hệ.
- $DB = \{r_1, r_2, \dots, r_m\}$
- sao cho mỗi thể hiện quan hệ r_i là một thể hiện của R_i và mỗi thể hiện quan hệ r_i thỏa mãn ràng buộc toàn vẹn trên CSDL đó.

Lược đồ quan hệ ... (tt)

Tên quan hệ

↓

Thuộc tính


↙ ↓ ↘

KHACHHANG	TEN_KH	DC_KH	TP_KH
Bộ	Phan Thanh Binh	45 Le Loi	DT
	Le Tan Tai	11 Tran Huy Lieu	NT
	Phan Van Tai Em	34 Xuan Loc	LA

- Mỗi một dòng $t \in r(R)$ được gọi là một bộ - tuple.
- Trong một quan hệ, thứ tự của các dòng không quan trọng; ngoài ra, không có dòng nào trùng nhau trong một quan hệ.

Lược đồ quan hệ ... (tt)

- Mỗi bộ t có n thuộc tính là một danh sách có thứ tự t gồm n giá trị $t = \langle v_1, v_2, \dots, v_n \rangle$
- Mỗi giá trị v_i ; $1 < i < n$, là một phần tử của $dom(A_i)$ hay là một giá trị đặc biệt *null*.
- Giá trị thứ i trong bộ t , tương ứng với thuộc tính A_i , được tham chiếu là $t[A_i]$
- Bậc (degree) của một quan hệ là số lượng thuộc tính n của quan hệ.
- Cardinality: là số lượng các bộ, ký hiệu *card*
- Quan hệ cơ sở (base relational): là quan hệ thực sự được lưu xuống một tập tin lưu trữ (khác với các quan hệ tạm thời hay khung nhìn).



Ràng buộc toàn vẹn trong Mô hình Quan hệ

Ràng buộc toàn vẹn – Integrity Constraint (IC)

- Phải luôn đúng với mọi thể hiện của một lược đồ quan hệ. Ex., các ràng buộc về miền giá trị.
 - ▣ ICs được mô tả khi định nghĩa lược đồ.
 - ▣ ICs được kiểm tra bởi DBMS khi các quan hệ (thể hiện) được cập nhật.
- Sự khác nhau giữa các bộ trong một quan hệ được biểu diễn bằng các thuộc tính.

Ràng buộc khóa chính

- **Super key (siêu khóa): SK** của r là một thuộc tính hay tập hợp các thuộc tính để xác định duy nhất một bộ trong một quan hệ (không có hai bộ nào trong bất kỳ thể hiện quan hệ $r(R)$ có cùng giá trị SK).
- i.e., Với mọi bộ phân biệt t_1 và t_2 trong $r(R)$, $t_1[SK] \neq t_2[SK]$.
- Mọi quan hệ đều có ít nhất một SK.

Candidate Key

- **Candidate key K** của $r(R)$ thỏa mãn 2 điều kiện:
 - ▣ K là một SK, và
 - ▣ K là SK bé nhất ($\forall K' \subset K, K' \neq K, K'$ không phải là SK).
- Quan hệ r có thể có nhiều hơn một khóa K
- **Ex:**
 - ▣ $XE\{TINH_TP, BSXE, SOSUON, SOMAY, NOISX, KIEUDANG, NAM\}$
 - ▣ Có hai SK là $K1 = \{TINH_TP, BIENSOXE\}$, $K2 = \{SOSUON, SOMAY\}$, nhưng $K1$ không là Khóa.

Primary Key – Khóa chính

- Nếu có nhiều Khóa (candidate key) trong một quan hệ, người thiết kế hoặc người quản trị CSDL sẽ chọn ra một khóa là khóa chính để nhận biết các bộ trong một quan hệ.
- Các thuộc tính khóa chính được gạch dưới.
- Các thuộc tính khóa chính PK của quan hệ r không thể mang giá trị $null$ ở bất kỳ bộ nào vì giá trị của khóa chính được dùng để xác định các bộ.
 - ▣ $t[PK] \neq null \quad \forall t \in r(R)$
- Các thuộc tính khác của r cũng có thể có ràng buộc quy định không được có giá trị $null$, cho dù chúng không tham gia vào khóa chính.
- **Ex:** Sinhvien(MaSv, TenSv, Khoa)

Alternate Key

- Là Candidate Key nhưng không được chọn là primary key.
- Được khai báo ràng buộc duy nhất (uniqueness constraint)

Ràng buộc khóa ngoại và toàn vẹn tham chiếu

KHOA	MAKHOA	TENKHOA
	CNTT	Cong nghe Thong tin
	VT	Vien thong
	DT	Dien tu

SVIEN	TENSV	MASV	MAKH
	Nhat	11	CNTT
	Minh	16	DT

- Tập hợp các thuộc tính trong một quan hệ (quan hệ con) được dùng để tham chiếu tới một bộ của quan hệ khác (quan hệ cha). Khóa ngoại phải tham chiếu tới khóa chính của quan hệ được tham chiếu.

Foreign Key – Khóa ngoại

- Cho $FK \subseteq r_1$, $FK \neq \emptyset$, FK được gọi là khóa ngoại của r_1 (tham chiếu đến r_2) khi:
- Các thuộc tính trong FK phải có cùng miền giá trị với các thuộc tính khóa chính PK trong r_2 .
- Giá trị tại FK của một bộ $t_1 \in r_1$ bằng giá trị tại PK của một bộ $t_2 \in r_2$ hoặc bằng giá trị rỗng *null*. Trường hợp đầu, ta nói t_1 tham chiếu tới bộ t_2 .
($t_1[FK] = t_2[PK]$)
- Nếu tất cả các ràng buộc khóa ngoại được bảo đảm cho một quan hệ, ta nói đã đạt được toàn vẹn tham chiếu - referential integrity.

Các đặc điểm của Mô hình Quan hệ

- Không có bộ nào trùng nhau – phải được kiểm tra khi:
 - ▣ Thêm 1 bộ mới.
 - ▣ Sửa giá trị.
 - ▣ Một quan hệ mới được tạo ra là bản thu gọn của quan hệ cũ.
- Ngầm định là khóa chính luôn tồn tại.
- Thứ tự của các bộ trong r không quan trọng.
 - ▣ Một bảng chỉ là một đại diện của một quan hệ
 - ▣ Tuy nhiên, lưu trữ vật lý của một quan hệ phải có thứ tự.
- Giá trị trong một bộ:
 - ▣ Tất cả các giá trị được xem là atomic (không thể phân chia được nữa).
- Giá trị không biết cũng phải được thể hiện:
 - ▣ Chúng được thay thế bằng giá trị đặc biệt null.
 - ▣ Null có nhiều ý nghĩa: “không biết”, “không quan tâm”, “không thể áp dụng một bộ đó”

Thao tác trên CSDL Quan hệ

- Có hai loại thao tác:
 - ▣ Cập nhật dữ liệu (thay đổi dữ liệu)
 - ▣ Truy vấn dữ liệu (rút trích dữ liệu)
- Các cách thực hiện:
 - ▣ Ngôn ngữ hình thức: Đại số quan hệ - cho phép mô tả các câu truy vấn dưới dạng chuỗi các phép toán đại số
 - ▣ Ngôn ngữ thực sự: SQL – ngôn ngữ thao tác CSDL quan hệ chuẩn.

Một ví dụ minh họa

- Lược đồ CSDL quan hệ: Quản lý Đề án của Công ty
 - NHANVIEN(HONV,TENLOT,TENNV,MANV,NGSINH,DCHI,PHAI,LUONG,MA_NGS,PHG)
 - PHONGBAN(TENPHG,MAPHG,TRPHG,NG_NHAMCHUC)
 - DIADIEM_PHG(MAPHG,DIADIEM)
 - THANNHAN(MA_NVVIEN,TENTN,PHAI,NGSINH,QUANHE)
 - DEAN(TENDA,MADA,DDIEM_DA,PHONG)
 - PHANCONG(MA_NVVIEN,SODA,THOIGIAN)

Ví dụ ... (tt)

NHANVIEN									
HONV	TENLOT	TENNV	MANV	NGSINH	DCHI	PHAI	LUONG	MA_NGS	PHG
Dinh	Van	Thong	12345	14/2/77	Hue	Nam	500	54321	1
Tran	null	Hung	23451	24/5/68	HCM	Nam	700	18794	5
Le	Minh	Nhat	67543	12/7/75	HN	Nam	450	null	2
Huynh	Chi	Thanh	78694	6/12/70	BD	Nam	500	79864	null
Mai	Ngoc	Nga	67543	13/2/81	LD	Nu	550	12345	1
Nguyen	Tu	Anh	87964	4/2/78	HCM	Nu	null	18794	5
Le	Xuan	Tung	75864	12/13/83	HP	Nam	250	54321	2

Ví dụ ... (tt)

PHONGBAN			
TENPHG	MAPHG	TRPHG	NG_NHAM CHUC
Nghien cuu	5	12345	5/12/99
Quan ly	1	67895	3/1/04
Thuc hien	2	67543	1/5/02

DIADIEM_PHG	
MAPHG	DIADIEM
1	TP HCM
2	HN
4	DN
4	CT

Ví dụ ... (tt)

PHANCONG		
MA_NV	SODA	THOIGIAN
12345	1	32.5
23451	4	24
67543	1	21
78694	2	2
87964	4	3
67543	2	4
12345	1	12
75864	9	4
78789	2	5

THANNHAN				
MA_NVIEN	TENTN	PHAI	NGSINH	QUANHE
12345	Trinh	Nam	2/3/70	Anh trai
75864	Minh	Nam	4/6/85	Em trai
57853	Luc	Nu	Null	Chi
23451	Tuan	Nam	4/2/56	Cha
75864	Mai	Nu	5/8/89	Em gai

DEAN			
TENDA	MADA	DDIEM_DA	PHONG
112CP	1	HCM	1
Cap quang	2	VT	4
Mouse Exam	4	HN	2
PM Tinh cuoc	8	CT	5
PM Ban do	10	HCM	2
Nghien cuu Mobile IP	16	DN	5



Đại số Quan hệ

Ngôn ngữ truy vấn

- Query Language (QL): ngôn ngữ cho phép user thực hiện và rút trích dữ liệu từ CSDL.
- QL \neq PL (ngôn ngữ lập trình)
 - ▣ QL không nhằm mục đích dùng cho các ứng dụng phức tạp.
 - ▣ QL hỗ trợ truy xuất dễ dàng tới các tập tin dữ liệu lớn.
- Ngôn ngữ truy vấn (toán học) là cơ sở của ngôn ngữ thực sự (vd., SQL).
- Hiểu được đại số quan hệ là chìa khóa để hiểu cách xử lý và tối ưu hóa câu truy vấn.

Đại số Quan hệ

- Ngôn ngữ thủ tục
- Bao gồm tập hợp các phép toán được áp dụng trên các thể hiện quan hệ, kết quả của một câu truy vấn là một thể hiện quan hệ.
- Chuỗi các phép toán đại số quan hệ hình thành nên biểu thức đại số quan hệ mà kết quả của nó cũng trả về một thể hiện quan hệ.

Các phép toán quan hệ

- Phép chọn
- Phép chiếu
- Phép gán
- Phép đổi tên
- Các phép toán được hình thành từ lý thuyết tập hợp toán học:
 - ▣ Phép hội
 - ▣ Phép giao
 - ▣ Phép trừ
 - ▣ Phép tích Cartesian
- Phép kết
- Phép chiếu tổng quát
- Các phép toán quan hệ khác: Phép chia, OUTER JOIN, AGGREGATE FUNCTION và GROUPING.

Phép chọn (Selection Operator)

- Dùng để trích chọn các bộ từ quan hệ r . Các bộ được trích chọn phải thỏa mãn điều kiện P .
- Ký hiệu: $\sigma_P(r) = \{t \mid t \in r \wedge P(t)\}$
- r : quan hệ được chọn
- P : biểu thức mệnh đề điều kiện:
 - $\langle \text{attr} \rangle \text{ op } \langle \text{attr} \rangle$ hay $\langle \text{const} \rangle$
 - op : $=, \neq, <, >, \dots$
 - các điều kiện có thể liên kết với nhau thông qua \wedge, \vee, \neg
- Kết quả trả về là một quan hệ, có cùng danh sách thuộc tính với quan hệ r .
- Phép chọn có tính giao hoán:
 - $\sigma_{P_1}(\sigma_{P_2}(r)) = \sigma_{P_2}(\sigma_{P_1}(r)) = \sigma_{P_1 \wedge P_2}(r)$

Example

A	B	C	D
α	α	1	7
α	β	5	7
β	β	12	3
β	β	23	10

Quan hệ r

A	B	C	D
α	α	1	7
β	β	23	10

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

- Cho biết:
 - Những nhân viên thuộc phòng số 4.
 - Những nhân viên có lương > 3000
 - Những nhân viên thuộc phòng số 4 và lương > 3000, hoặc những nhân viên thuộc phòng số 5.

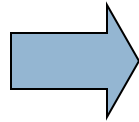
Phép chiếu (Project Operation)

- Dùng để trích chọn một vài cột của quan hệ r được chỉ ra trong danh sách thuộc tính.
- Ký hiệu: $\pi_{A_1, A_2, \dots, A_k}(r)$
- r : quan hệ.
- A_1, A_2, \dots, A_k : các thuộc tính được chiếu.
- Kết quả trả về là một quan hệ có k thuộc tính theo thứ tự được liệt kê.
- Số lượng các bộ trong một quan hệ kết quả của phép chiếu luôn luôn ít hơn hoặc bằng số lượng các bộ trong r . Các dòng trùng nhau sẽ được loại bỏ khỏi quan hệ kết quả.
- $\pi_{A_1, A_2, \dots, A_k}(\pi_{A_1, A_2, \dots, A_l}(r)) = \pi_{A_1, A_2, \dots, A_k}(r)$, với $k \leq l$
- Phép chiếu không có tính giao hoán

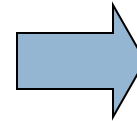
Example

A	B	C
α	10	1
α	20	1
β	30	1
β	40	2

Quan hệ r



A	C
α	1
α	1
β	1
β	2



A	C
α	1
β	1
β	2

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

- Cho biết:
 - Họ tên và lương của tất cả nhân viên.
 - Mã phòng và lương của tất cả nhân viên

Phép gán (Assignment Operation)

- Cung cấp một cách thức để diễn tả câu truy vấn phức tạp.
- Ý tưởng: viết câu truy vấn như là một chương trình tuần tự gồm một chuỗi các phép gán, theo sau là một biểu thức có giá trị được trình bày như là kết quả của câu truy vấn.
- Ký hiệu: \leftarrow
- Việc gán luôn luôn thực hiện cho một biến quan hệ tạm.
- Biến nằm bên trái của biểu thức gán có thể được sử dụng trong các biểu thức theo sau.

Phép đổi tên (Rename Operation)

- Cho phép đặt tên, và do đó dễ tham chiếu tới kết quả của biểu thức đại số quan hệ.
- Cho phép tham chiếu tới một quan hệ bằng nhiều tên (vd., nếu cùng một quan hệ được dùng hai lần trong một biểu thức đại số quan hệ).
- **Đổi tên quan hệ và tên thuộc tính:**
- Cho một biểu thức đại số quan hệ E có n thuộc tính, biểu thức

$$\rho_{\chi(A_1, A_2, \dots, A_n)}(E)$$

trả về kết quả biểu thức E dưới tên χ , và với các tên thuộc tính đã được thay đổi thành A_1, A_2, \dots, A_n .

Đổi tên quan hệ và tên thuộc tính

□ Đổi tên quan hệ:

- Cho một biểu thức đại số quan hệ E , biểu thức $\rho_\chi(E)$ trả về kết quả biểu thức E dưới tên χ .

□ Đổi tên thuộc tính:

- Cho một biểu thức đại số quan hệ E có n thuộc tính, biểu thức $\rho_{(A_1, A_2, \dots, A_n)}(E)$ trả về kết quả biểu thức E với các tên thuộc tính đã được thay đổi thành A_1, A_2, \dots, A_n .

Chuỗi các phép toán

- Kết hợp các phép toán đại số quan hệ với nhau bằng một trong hai cách:
 - ▣ Tạo một biểu thức đại số quan hệ bằng cách lồng các phép toán với nhau.
 - ▣ Áp dụng từng phép toán một, ở mỗi lần áp dụng phép toán cho ra quan hệ kết quả.

Example

- Tìm họ tên và lương nhân viên làm việc ở phòng số 4.

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

- hoặc:

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

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

- hoặc có thể đổi tên bằng cách liệt kê các tên thuộc tính mới trong dấu ngoặc:

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

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

Các phép toán tập hợp

- Các phép toán thực hiện trên 2 quan hệ xuất phát từ lý thuyết tập hợp của toán học
 - ▣ Phép hội: $r \cup s$
 - ▣ Phép giao: $r \cap s$
 - ▣ Phép trừ: $r - s$
 - ▣ Phép tích Cartesian: $r \times s$
- Đối với các phép toán \cup , \cap , hay $-$, các quan hệ r và s phải khả hợp.

Quan hệ Khả hợp

- Các quan hệ $R(A_1, A_2, \dots, A_n)$ và $S(B_1, B_2, \dots, B_n)$ gọi là khả hợp khi:
- Số lượng thuộc tính của r , s phải bằng nhau (cùng bậc n)
- Miền giá trị của thuộc tính phải tương thích ($dom(A_i) = dom(B_i)$), với $1 \leq i \leq n$.
- Quan hệ kết quả của \cup , \cap , hay $-$ có cùng tên thuộc tính với quan hệ đầu tiên r (quy ước).

Phép hội (Union Operation)

- Ký hiệu: $r \cup s$
- Định nghĩa:
$$r \cup s = \{t \mid t \in r \vee t \in s\}$$
- r và s là hai quan hệ khả hợp.

A	B
α	1
α	2
β	1

Quan hệ r

A	B
α	2
β	3

Quan hệ s

- Ex1: \Rightarrow
- Ex2: Mã nhân viên có tham gia đề án hoặc có thân nhân hoặc cả hai.

A	B
α	1
α	2
β	1
β	3

$r \cup s$

Phép Giao (Set-Intersection Operation)

- Ký hiệu: $r \cap s$
- Định nghĩa:
$$r \cap s = \{t \mid t \in r \wedge t \in s\}$$
- r và s là hai quan hệ khả hợp.
- Ex1: \Rightarrow
- Ex2: Mã nhân viên có người thân và được phân công vào đề án.

A	B
α	1
α	2
β	1

Quan hệ r

A	B
α	2
β	3

Quan hệ s

A	B
α	2

$r \cap s$

Phép trừ (Set Difference Operation)

- Ký hiệu: $r - s$.
- Định nghĩa:
- $r - s = \{t \mid t \in r \wedge t \notin s\}$
- r và s là hai quan hệ khả hợp.
- Ex1: \Rightarrow
- Ex2: Mã nhân viên không có thân nhân nào cả.

A	B
α	1
α	2
β	1

Quan hệ r

A	B
α	2
β	3

Quan hệ s

A	B
α	1
β	1

$r - s$

Phép tích (Cartesian Product Operation)

- Ký hiệu: $r \times s$
- Định nghĩa:
$$r \times s = \{t.q \mid t \in r \wedge q \in s\}$$
- Các thuộc tính của $r(R)$ và $s(S)$ khác nhau (i.e. $R \cap S = \emptyset$).
- Nếu các thuộc tính của $r(R)$ và $s(S)$ không khác nhau thì phải đổi tên chúng.
- Mỗi bộ t của tập kết quả là tổ hợp giữa 1 bộ t_1 trong r và 1 bộ t_2 trong s sao cho:

$$t[A_1, A_2, \dots, A_{nr}] = t_1 \text{ và } t[B_1, B_2, \dots, B_{ms}] = t_2$$

- $KQ(A_1, A_2, \dots, A_{nr}, B_1, B_2, \dots, B_{ms}) \leftarrow r(A_1, A_2, \dots, A_{nr}) \times s(B_1, B_2, \dots, B_{ms})$
- Nếu r có n_r bộ và s có m_s bộ thì $r \times s$ có $n_r \times m_s$ bộ.

Example 1

A	B
α	1
β	2

Quan hệ r

C	D	E
α	10	+
β	10	+
β	20	–
γ	10	–

Quan hệ s

A	B	C	D	E
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	–

$r \times s$

- **Nhân xét:** Phép tích Cartersian là một phép tính vô nghĩa nếu đứng riêng. Tuy nhiên, nó có thể kết hợp các bộ liên quan từ hai quan hệ nếu tiếp theo sau phép toán chọn σ phù hợp (xem ví dụ trang sau).

Example 2

A	B	C	D	E
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	−

$\mathbf{r} \times \mathbf{s}$

A	B	C	D	E
α	1	α	10	+
β	2	β	10	+
β	2	β	20	−

$\sigma_{A=C}(\mathbf{r} \times \mathbf{s})$

Example 3

- Kết hợp từng bộ PHONGBAN với bộ NHANVIEN của người quản lý (trưởng phòng). Với mỗi phòng ban, cho biết thông tin về người trưởng phòng.

$PB_NV \leftarrow PHONGBAN \times NHANVIEN$

$PB_TRPHG_INFO \leftarrow \sigma_{MANV=TRPHG} PB_NV$

- Hoặc:

$\sigma_{MANV=TRPHG}(PHONGBAN \times NHANVIEN)$

Example 4

- Cho biết lương cao nhất trong công ty.

- Quan hệ chứa những lương không phải là lớn nhất

$\text{Temp} = \sigma_{\text{NHANVIEN.LUONG} < \text{NV.LUONG}}(\text{NHANVIEN} \times \rho_{\text{NV}}(\text{NHANVIEN}))$

- Thực hiện trừ giữa quan hệ và quan hệ tạm

$\pi_{\text{LUONG}}(\text{NHANVIEN}) - \pi_{\text{NHANVIEN.LUONG}}(\text{temp})$

Example 5

- Cho biết những phòng ban có cùng địa điểm với phòng số 5.

- Tìm địa điểm của phòng số 5.

$\rho_{PHG_5(DD)}(\pi_{DIADIEM}(\sigma_{MAPHG=5}(DIADIEM_PHG)))$

- Đổi tên thành PHG_5 và đổi tên thuộc tính DIADIEM thành DD.

$\pi_{DIADIEM_PHG.MAPHG}(\sigma_{DIADIEM_PHG.DIADIEM=PHG_5.DD}(DIADIEM_PHG \times PHG_5)).$

Phép kết (Join Operation)

- Natural Join
- Condition Join (Theta Join)
- Equi-Join

Natural Join

- Ký hiệu: $r \bowtie s$
- r, s là các quan hệ tương ứng trên các lược đồ R và S . Kết quả trả về là quan hệ trên lược đồ $R \cup S$. các kết quả có được bằng cách xét đến từng cặp bộ $t_r \in r$ và $t_s \in s$.
- Nếu t_r và t_s có cùng giá trị tại các thuộc tính trong $R \cap S$ (cùng tên thuộc tính), thì một bộ t được thêm vào quan hệ kết quả sao cho:
 - ▣ t có cùng giá trị như t_r trên R .
 - ▣ t có cùng giá trị như t_s trên S .

Example

- Cho các quan hệ R(A,B,C,D) và S(B,D,E)
- Phép kết có thể áp dụng được vì $R \cap S \neq \emptyset$.
- Lược đồ kết quả là (A,B,C,D,E)
- Và kết quả của $r \bowtie s$ được định nghĩa là:

$$\pi_{R.A, R.B, R.C, R.D, S.E}((\sigma_{R.B=S.B \wedge R.D=S.D}(r \times s)))$$

A	B	C	D
α	1	α	b
β	2	γ	a
γ	4	β	b
α	1	γ	a
δ	2	β	b

r

B	D	E
a	1	α
3	a	β
1	a	γ
2	b	δ

s

A	B	C	D	E
α	1	γ	a	γ
δ	2	β	b	δ

r \bowtie s

Condition Join

- Ký hiệu: $r \bowtie_c s$
- C: điều kiện tồn tại trên các thuộc tính trong $R \cup S$, lược đồ kết quả cũng giống như trong phép tích Cartesian. Nếu $R \cap S \neq \emptyset$ và điều kiện c tham chiếu tới các thuộc tính này, thì một số thuộc tính phải được đổi tên.
- Đôi khi còn gọi là Theta join $r \bowtie_\theta s$
- Nguồn gốc: $r \bowtie_c s = \sigma_c(r \times s)$
- Chú ý: không giống như điều kiện chọn P, c là một điều kiện trên các thuộc tính của cả r và s.

Example

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

r

D	E
3	1
6	2

s

A	B	C	D	E
1	2	3	3	1
1	2	3	6	2
4	5	6	6	2

r $\triangleright \triangleleft$ **B < D** **s**

EQUI-JOIN

- Nếu θ là phép so sánh bằng “=” thì phép kết được gọi là phép kết bằng – Equi-Join.
- Ex:
- Tên phòng ban và tên trưởng phòng của mỗi phòng ban.

A	B	C
4	5	6
7	8	9

r

C	D
6	8
10	12

s

A	B	C	SC	D
4	5	6	6	8

r \bowtie **C=SC**(ρ (SC,D)(s)) **s**

Một số lưu ý

- Giữa 2 quan hệ có thể có nhiều hơn một tập thuộc tính kết mang ý nghĩa khác nhau. Vd:

Thuộc tính kết	Mối kết hợp
NHANVIEN.MANV=PHONGBAN.TRPHG	NHANVIEN quản lý PHONGBAN
NHANVIEN.PHG=PHONGBAN.MAPHG	NHANVIEN làm việc ở PHONGBAN

- Ex: Tìm tên nhân viên và tên phòng ban mà nhân viên đó làm việc. (?)
- Một quan hệ có thể có một tập thuộc tính kết để kết với chính quan hệ đó. Trong trường hợp này, phải sử dụng phép đổi tên. Vd:

Thuộc tính kết	Mối kết hợp
NHANVIEN(1).MA_NGS=NHANVIEN(2).MANV	NHANVIEN(2) giám sát NHANVIEN(1)

- Ex: Tìm tên nhân viên và tên người giám sát nhân viên đó (?)

Tập đầy đủ các phép toán

ĐSQH

- Các phép toán cơ sở: phép chọn, chiếu, hội, giao, trừ, tích.
- Tập hợp $\{\sigma, \pi, \cup, \cap, -, \times\}$ được gọi là tập đầy đủ các phép toán ĐSQH.
- Nếu E_1 và E_2 là các biểu thức ĐSQH, thì các biểu thức sau đây cũng là các biểu thức ĐSQH:
 - $E_1 \cup E_2$
 - $E_1 - E_2$
 - $E_1 \times E_2$
 - $\sigma_P(E_1)$, với P là một vị từ (điều kiện) trên các thuộc tính của E_1
 - $\pi_A(E_1)$, với A là danh sách các thuộc tính của E_1
 - $\rho_\chi(E_1)$, với χ là tên mới cho quan hệ kết quả [và các thuộc tính] được xác định bởi E_1 .

Phép chia (Division Operator)

- Áp dụng cho câu truy vấn có từ “tất cả”.
- Mục tiêu: Đưa ra các bộ trong một quan hệ, r , sao cho khớp với tất cả các bộ trong quan hệ khác, s .
- Ký hiệu: $r \div s$
- Điều kiện: các thuộc tính trong S phải là một tập thuộc tính con của R , i.e., $S \subseteq R$. Cho r, s là các quan hệ tương ứng trên lược đồ R, S
$$R = (A_1, A_2, \dots, A_m, B_1, \dots, B_n)$$
$$S = (B_1, \dots, B_n)$$
- Kết quả của phép chia $r \div s$ là một quan hệ trên lược đồ $R-S = (A_1, A_2, \dots, A_m)$.
- Kết quả của phép chia bao gồm tập hợp các bộ lấy từ r được định nghĩa trên các thuộc tính $R-S$ sao cho khớp với việc kết hợp mọi bộ trong s .
$$r \div s = \{t \mid t \in \pi_{R-S}(r) \wedge \forall u \in s : t.u \in r\}$$
- Nghĩa là, $r \div s$, với các thuộc tính A_1, A_2, \dots, A_m , là tập hợp chứa tất cả các bộ t sao cho với mọi bộ u trong s , thì có một bộ $t.u$ trong r .

Examples

A	B
α	1
α	2
α	3
β	1
γ	1
δ	1
δ	3
δ	4
ε	6
ε	1
β	2

r

B
1
2

s

A
α
β

r ÷ s

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

r

D	E
a	1
b	1

s

A	B	C
α	a	γ
γ	a	γ

r ÷ s

- Ex3: Mã nhân viên tham gia tất cả đề án.
- Ex4: Mã nhân viên tham gia tất cả đề án do phòng số 4 phụ trách.

Phép chiếu tổng quát

- Mở rộng phép chiếu bằng cách cho phép sử dụng các hàm số học trong danh sách chiếu.
- $\pi_{F_1, F_2, \dots, F_k}(E)$
- E là biểu thức đại số quan hệ.
- F_1, F_2, \dots, F_k là các biểu thức số học có liên quan đến hằng và thuộc tính trong lược đồ E .
- Ex: Cho quan hệ THETINDUNG(MSTHE, TRIGIATHE, SOTIENSĐ). Tìm số tiền còn lại trong thẻ:
- $\pi_{\text{MSTHE, TRIGIATHE} - \text{SOTIENSĐ}}(\text{THETINDUNG})$

Các phép toán khác

- Hàm kết hợp – Aggregate Functions nhận vào tập hợp giá trị và trả về một giá trị đơn.
 - ▣ Avg (average value): giá trị trung bình
 - ▣ Min (minimum value): giá trị nhỏ nhất.
 - ▣ Max (maximum value): giá trị lớn nhất.
 - ▣ Sum (sum of values): tính tổng các giá trị
 - ▣ Count (number of values): đếm số mẫu tin
- Phép toán gom nhóm – Grouping \mathfrak{S} trong ĐSQH:
$$G_1, G_2, \dots, G_n \mathfrak{S}_{F_1(A_1), F_2(A_2), \dots, F_n(A_n)}(E)$$
- E là biểu thức đại số quan hệ
- G_i là tên thuộc tính gom nhóm (có thể không có)
- F_i là hàm gom nhóm
- A_i là tên thuộc tính tính toán trong hàm gom nhóm F_i .

Examples

Ex1

A	B	C
α	α	7
α	β	7
β	β	3
β	β	10

r

SUM_C
27

$\mathfrak{I}_{\text{SUM}(C)}(r)$

Ex2

- Tính số lượng nhân viên và lương trung bình của nhân viên.
- Quan hệ NHANVIEN sau khi gom nhóm:

...	MANV	...	LUONG	MA_NQL	PHG
...	123456789	...	30000	333445555	5
...	333445555	...	40000	888665555	5
...	666884444	...	38000	333445555	5
...	453453453	...	25000	333445555	5
...	999887777	...	25000	987654321	4
...	987654321	...	43000	888665555	4
...	987987987	...	25000	987654321	4
...	888665555	...	55000	null	1

- Số lượng nhân viên và lương trung bình của cả công ty:

COUNT_MANV	AVG_LUONG
8	35125

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

- Số lượng nhân viên và lương trung bình của mỗi phòng ban, không sử dụng phép đổi tên:

PHG	COUNT_MANV	AVG_LUONG
5	4	33250
4	3	31000
1	1	55000

$\text{PHG} \mathfrak{I}_{\text{COUNT(MANV),AVG(LUONG)}}(\text{NHANVIEN})$

- Số lượng nhân viên và lương trung bình của mỗi phòng ban, có sử dụng phép đổi tên:

MAPHONG	SONV	LUONGTB
5	4	33250
4	3	31000
1	1	55000

$\text{P}_{\text{MAPHONG,SONV,LUONGTB}}(\text{PHG} \mathfrak{I}_{\text{COUNT(MANV),AVG(LUONG)}}(\text{NHANVIEN}))$

Outer Join

- Mở rộng phép kết để tránh sự mất mát thông tin
- Thực hiện phép kết và sau đó thêm vào kết quả của phép kết các bộ của một quan hệ mà không phù hợp với các bộ trong quan hệ kia.
- 3 hình thức:
 - ▣ Left outer join: $_ \triangleright \triangleleft$
 - ▣ Right outer join: $\triangleright \triangleleft _$
 - ▣ Full outer join: $_ \triangleright \triangleleft _$

Example:

HONV	TENLOT	TENNV	TENPHG
Nguyen	Thanh	Tung	Nghien cuu
Tran	Hong	Quang	Dieu hanh
Pham	Van	Vinh	Quan ly

$\pi_{\text{HONV, TENLOT, TENNV, TENPHG}}(\text{NHANVIEN} \bowtie_{\text{MANV=TRPHG}} \text{PHONGBAN})$

HONV	TENLOT	TENNV	TENPHG
Dinh	Ba	Tien	Null
Nguyen	Thanh	Tung	Nghien cuu
Bui	Ngoc	Hang	Null
Le	Quynh	Nhu	Null
Nguyen	Manh	Hung	Null
Tran	Thanh	Tam	Null
Tran	Hong	Quang	Dieu hanh
Pham	Van	Vinh	Quan ly

$\pi_{\text{HONV, TENLOT, TENNV, TENPHG}}(\text{NHANVIEN} \text{ _ } \bowtie_{\text{MANV=TRPHG}} \text{PHONGBAN})$

Các thao tác cập nhật trên quan hệ

- Nội dung của CSDL có thể được cập nhật bằng cách dùng các thao tác: Thêm, Xóa, Sửa.
- Tất cả các thao tác này được diễn đạt thông qua phép toán gán:
- $r_{\text{new}} \leftarrow \{\text{Thêm, Xóa, Sửa}\}(r_{\text{old}})$

Thêm (Insertion)

- Hoặc nêu ra một bộ cần chèn, hoặc viết một câu truy vấn mà kết quả là một tập hợp các bộ cần chèn.

- Trong ĐSQH, thao tác chèn được diễn đạt như sau:

$$r \leftarrow r \cup E$$

- ▣ r là quan hệ và E là biểu thức ĐSQH.

- **Ex:**

- $\text{PHANCONG} \leftarrow \text{PHANCONG} \cup \{('123456789', 20, 10)\}$

Xoá (Deletion)

- Yêu cầu xóa được diễn đạt như câu truy vấn, chỉ khác ở chỗ, thay vì hiển thị các bộ kết quả với người dùng, thì bộ được chọn bị xóa khỏi CSDL.
- Chỉ có thể xóa toàn bộ bộ, không thể chỉ xóa một vài giá trị trên các thuộc tính nào đó.
- Thao tác xóa được diễn đạt trong ngôn ngữ ĐSQH như sau:

$$r \leftarrow r - E$$

- r là quan hệ và E là câu truy vấn ĐSQH.

- **Ex1:**

- $\text{PHANCONG} \leftarrow \text{PHANCONG} - (\sigma_{\text{MA_NV}='123456789'}(\text{PHANCONG}))$

- **Ex2:**

- $r_1 \leftarrow \sigma_{\text{DDIEM_DA}='HANOI'}(\text{PHANCONG} \bowtie_{\text{SODA=MADA}} \text{DEAN})$
- $r_2 \leftarrow \pi_{\text{MA_NV}, \text{SODA}, \text{THOIGIAN}}(r_1)$
- $\text{PHANCONG} \leftarrow \text{PHANCONG} - r_2.$

Sửa (Updating)

- Cơ chế làm thay đổi một giá trị trong một bộ mà không làm thay đổi tất cả giá trị trong bộ đó.
- Để cập nhật, sử dụng phép chiếu tổng quát hóa như sau:

$$r \leftarrow \pi_{F_1, F_2, \dots, F_n}(r)$$

- Có thể được biểu diễn bằng một chuỗi các thao tác xóa và thêm. Phép toán xóa sẽ xóa các bộ có giá trị cũ và phép toán thêm sẽ chèn thêm các bộ có giá trị mới.

- **Ex1:**

- $\text{PHANCONG} \leftarrow \pi_{\text{MA_NV}, \text{SODA}, \text{THOIGIAN} * 1.5}(\text{PHANCONG})$

- **Ex2:**

- $\text{PHANCONG} \leftarrow \pi_{\text{MA_NV}, \text{SODA}, \text{THOIGIAN} * 1.5}(\sigma_{\text{THOIGIAN} > 30}(\text{PHANCONG})) \cup \pi_{\text{MA_NV}, \text{SODA}, \text{THOIGIAN} * 2}(\sigma_{\text{THOIGIAN} \leq 30}(\text{PHANCONG}))$