

Đại số quan hệ

Chương 4

Nội dung trình bày

- Giới thiệu
- Phép toán một ngôi
- Phép toán hai ngôi.
- Phép toán khác.

Tổng quan(1)

▪Đại số quan hệ

- Là tập hợp các phép toán cơ sở của mô hình dữ liệu quan hệ.
- Biểu thức đại số quan hệ là một chuỗi các phép toán.
- Kết quả của một biểu thức là một thể hiện quan hệ.

▪Ý nghĩa

- Cơ sở hình thức cho các phép toán của mô hình quan hệ.
- Cơ sở để cài đặt và tối ưu hóa các truy vấn trong các HQT CSDL quan hệ.
- Được áp dụng trong SQL.

Tổng quan (2)

- Toán hạng
 - Các thể hiện quan hệ.
 - Các tập hợp.
- Toán tử là các phép toán
 - Phép toán tập hợp
 - Hội, giao, hiệu, tích Cartesian.
 - Phép toán quan hệ
 - Chọn, chiếu, kết, chia, đổi tên.
 - Một số phép toán khác.

Phép toán 1 ngôi

- Là các phép toán chỉ tác động lên một quan hệ.
- Gồm
 - Phép chọn (Select): σ
 - Phép chiếu (Project): π
 - Phép đổi tên (Rename): ρ

Phép chọn - Select (σ) (1)

- Để rút trích các bộ dữ liệu thỏa điều kiện chọn từ một quan hệ.

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

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

A	B	C	D
α	α	1	7
β	β	23	10

Phép chọn - Select (σ) (2)

- Ký hiệu: $\sigma_{\langle DK \rangle}(r)$.
 - $\langle DK \rangle$ là biểu thức logic.
- Cú pháp
 - $\sigma_{\langle DK \rangle}(r) = \{t | t \in r \text{ và } DK(t) \text{ đúng}\}$
- Biểu thức điều kiện
 - Là một biểu thức logic
 - Chứa các mệnh đề có dạng
 - $\langle \text{thuộc tính} \rangle \langle \text{toán tử so sánh} \rangle \langle \text{hằng số} \rangle$.
 - $\langle \text{thuộc tính} \rangle \langle \text{toán tử so sánh} \rangle \langle \text{thuộc tính} \rangle$.
 - Toán tử so sánh: $=, <, \leq, >, \geq, \neq$.
 - Các mệnh đề được nối bởi toán tử logic: \wedge, \vee, \neg .

Phép chọn - Select (σ) (3)

▪ Đặc trưng

- Phép chọn có tính giao hoán.
 - $\sigma_{\langle DK1 \rangle}(\sigma_{\langle DK2 \rangle}(r)) = \sigma_{\langle DK2 \rangle}(\sigma_{\langle DK1 \rangle}(r))$.
- Kết quả là một quan hệ
 - Có cùng bậc với r
 - Có số bộ ít hơn hoặc bằng số bộ của r .

Phép chiếu (1)

- Để rút trích các cột ứng với các thuộc tính nào đó của một quan hệ.

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

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

A	D
α	7
β	3
β	10

Phép chiếu (2)

- Ký hiệu

- $\pi_{\langle DSTT \rangle}(r)$.

- Cú pháp

- $\pi_{\langle DSTT \rangle}(r) = \{t[DSTT] \mid t \in r\}$
- $\langle DSTT \rangle$ là danh sách các thuộc tính của R.

- Đặc trưng

- Phép chiếu không có tính giao hoán.
 - $\pi_{\langle DSTT_1 \rangle}(\pi_{\langle DSTT_2 \rangle}(r)) \neq \pi_{\langle DSTT_2 \rangle}(\pi_{\langle DSTT_1 \rangle}(r))$.
- Phép chiếu loại bỏ các bộ trùng nhau.
- Kết quả là một quan hệ
 - Có bậc bằng số thuộc tính của danh sách thuộc tính.
 - Có bậc nhỏ hơn hoặc bằng bậc của r.
 - Có số bộ ít hơn hoặc bằng số bộ của r.

- Mở rộng phép chiếu

- Cho phép sử dụng các phép toán số học trong danh sách thuộc tính.
 - $\pi_{A,2^*C}(r)$.

Chuỗi các phép toán và phép gán

■ Chuỗi các phép toán

- Muốn sử dụng kết quả của phép toán này làm toán hạng của phép toán khác.
- Muốn viết các phép toán lồng nhau.
 - $\pi_{A,C}(\sigma_{A=B \wedge D>5}(r))$

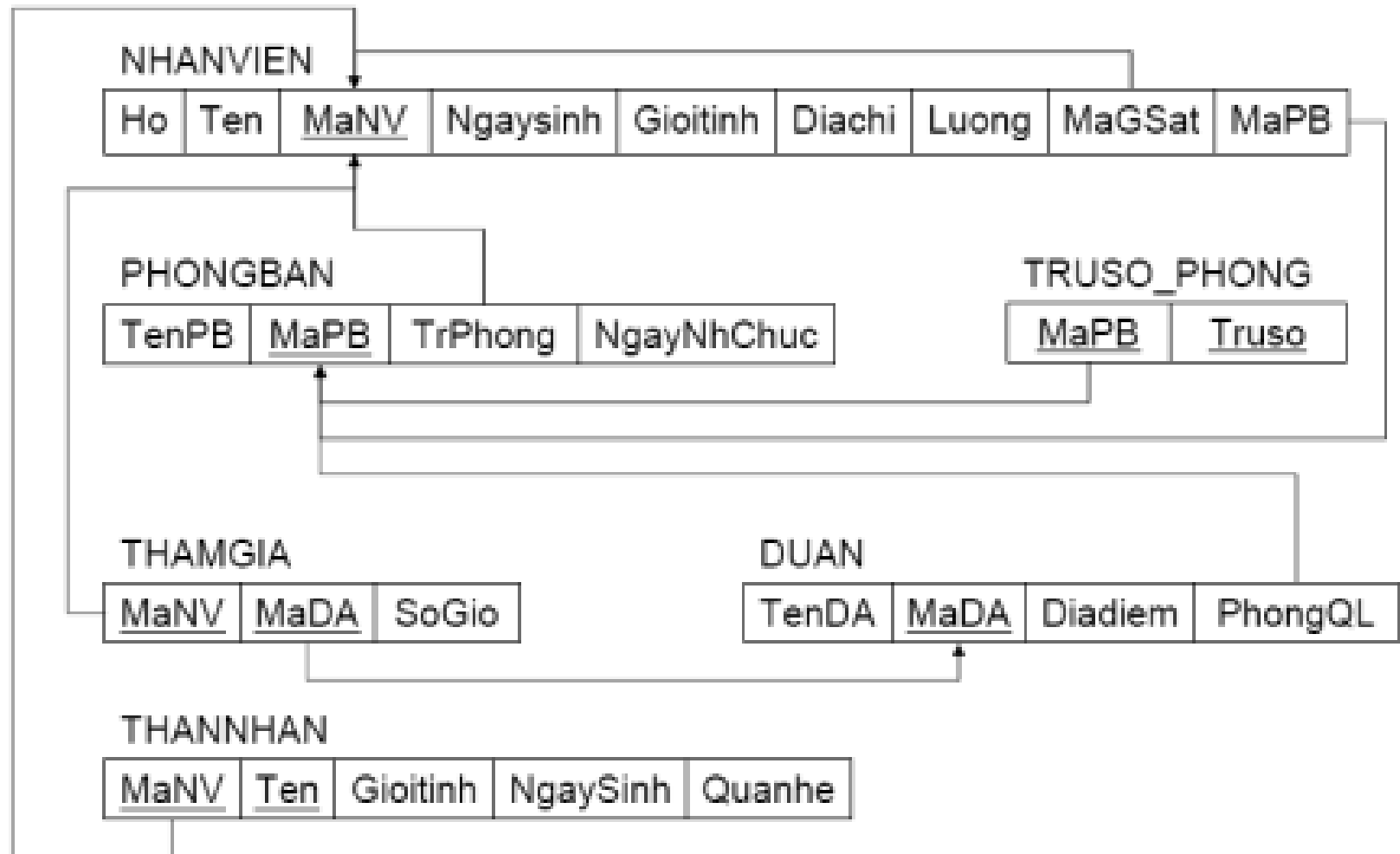
■ Phép gán

- Dùng để lưu lại kết quả của một phép toán.
- Để đơn giản hóa một chuỗi phép toán phức tạp.
- Cú pháp
 - $R' \leftarrow E$
 - E là biểu thức đại số quan hệ.
- Ví dụ
 - $r' \leftarrow \sigma_{A=B \wedge D>5}(r)$
 $\pi_{A,C}(r')$

Phép đổi tên

- Để đổi tên quan hệ và các thuộc tính.
- Cú pháp: cho quan hệ $R(A_1, \dots, A_n)$
 - Đổi tên quan hệ R thành S
 - $\rho_S(R)$.
 - Đổi tên quan hệ R thành S và các thuộc tính A_i thành B_i
 - $\rho_{S(B_1, B_2, \dots, B_n)}(R)$.
 - Đổi tên các thuộc tính A_i thành B_i
 - $\rho_{(B_1, B_2, \dots, B_n)}(R)$.
 - Đổi tên quan hệ R thành S và thuộc tính A_1 thành B_1
 - $\rho_{S(B_1, A_2, A_3, \dots, A_n)}(R)$.
 - Đổi tên thuộc tính A_1 thành B_1
 - $\rho_{(B_1, A_2, A_3, \dots, A_n)}(R)$.

Một số ví dụ



Một số ví dụ

- Tìm các nhân viên làm việc trong phòng số 4.
- Tìm các nhân viên làm việc trong phòng số 4 và có mức lương từ 25.000 đến 40.000.
- Cho biết họ, tên, giới tính và mức lương của các nhân viên.
- Cho biết họ, tên, giới tính và mức lương của các nhân viên của phòng số 5.

Phép toán 2 ngôi

- Là các phép toán tác động lên hai quan hệ.
- Gồm 2 loại
 - Phép toán tập hợp
 - Phép hội (Union).
 - Phép giao (Intersection).
 - Phép hiệu (Difference).
 - Phép tích Đề các (Cartesian).
 - Phép toán phi tập hợp
 - Phép kết (Join).
 - Phép chia (Division).

Phép toán tập hợp (1)

- Chỉ được sử dụng khi hai quan hệ được tác động là *khả hợp*.
- Hai quan hệ $R(A_1, \dots, A_n)$ và $S(B_1, \dots, B_n)$ gọi là khả hợp nếu
 - $\text{Bậc } R = \text{Bậc } S$.
 - Miền giá trị $A_i \equiv \text{Miền giá trị } B_i$, với $i = 1, \dots, n$.

Phép hội

- Hội của R và S
 - $R \cup S$
 - Là quan hệ gồm các bộ thuộc R hoặc thuộc S.
 - Các bộ trùng nhau thì chỉ giữ lại 1 bộ.
- $R \cup S = \{t \mid t \in R \vee t \in S\}$

R	A	C
	α	1
	α	5
	β	12
	β	23

S	A	C
	α	1
	γ	12
	β	23

$R \cup S$

A	C
α	1
α	5
β	12
β	23
γ	12

Phép giao

- Giao của R và S
 - $R \cap S$
 - Là quan hệ gồm các bộ thuộc R đồng thời thuộc S.
- $R \cap S = \{t \mid t \in R \wedge t \in S\}$

R	A	C
	α	1
	α	5
	β	12
	β	23

S	A	C
	α	1
	γ	12
	β	23

$R \cap S$



A	C
α	1
β	23

Phép hiệu

- Hiệu của R và S
 - $R - S$
 - Là quan hệ gồm các bộ thuộc R nhưng không thuộc S.
- $R - S = \{t \mid t \in R \wedge t \notin S\}$

R	A	C
	α	1
	α	5
	β	12
	β	23

S	A	C
	α	1
	γ	12
	β	23

$R - S$ →

A	C
α	5
β	12

Phép toán tập hợp (2)

▪ Đặc trưng

- Phép hội và giao có tính giao hoán
 - $R \cup S = S \cup R$ và $R \cap S = S \cap R$.
- Phép hội và giao có tính kết hợp
 - $R \cup (S \cap T) = (R \cup S) \cap T$ và $R \cap (S \cup T) = (R \cap S) \cup T$.

Phép tích Đề các (Cartesian)

- Tích Cartesian của R và S (không nhất thiết khả hợp).
 - $R \times S$
 - Là quan hệ Q mà mỗi bộ là một tổ hợp của một thuộc R và một bộ thuộc S.
 - Bậc Q = Bậc R + Bậc S.
 - Số bộ Q = Số bộ R \times Số bộ S.

$$R \times S = \{(a_1, \dots, a_m, b_1, \dots, b_n) \mid (a_1, \dots, a_m) \in R \wedge (b_1, \dots, b_n) \in S\}$$

R	A	B	C		S	D	E	
	α	α	1			1	7	
	α	β	5			5	7	
	β	β	12					

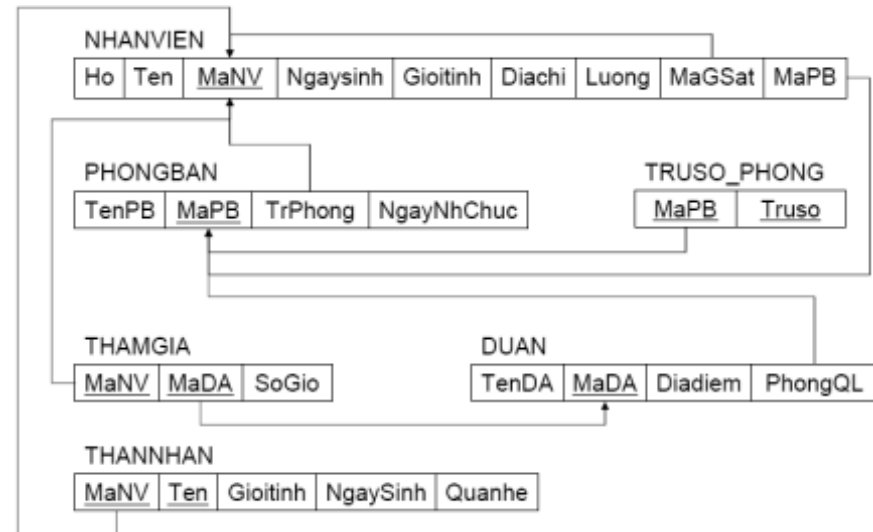
$R \times S$

A	B	C	D	E
α	α	1	1	7
α	α	1	5	7
α	β	5	1	7
α	β	5	5	7
β	β	12	1	7
β	β	12	5	7

Một số ví dụ

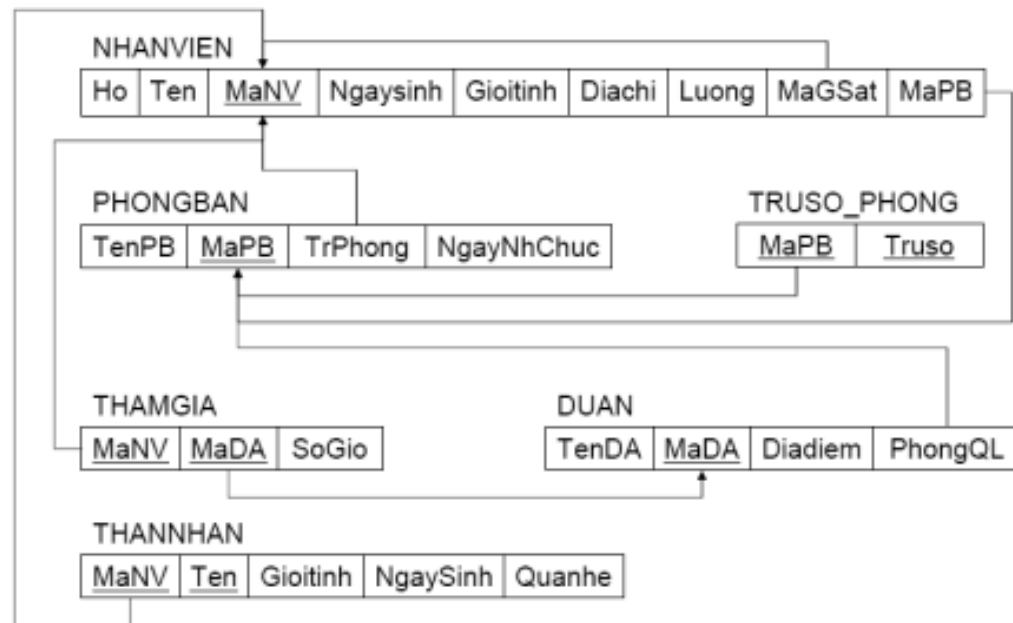
- Tìm mã số các nhân viên của phòng số 5 hoặc giám sát trực tiếp các nhân viên phòng số 5.

- Cho biết họ, tên của các nhân viên nữ và tên các thân nhân của họ.



Một số ví dụ

- Cho biết tên, địa chỉ của các nhân viên của phòng có tên Nghiên cứu.



Phép toán phi tập hợp

- Phép kết nối
- Phép chia

Phép kết nối - Join

- Để kết hợp các bộ có liên quan từ hai quan hệ.
- Có 3 loại
 - Kết theta (Theta Join)
 - $R \bowtie_{\langle \text{ĐK} \rangle} S$
 - $\langle \text{ĐK} \rangle$ là biểu thức logic.
 - Kết bằng (Equi Join)
 - Kết tự nhiên (Natural Join)
 - $R \bowtie S$ hoặc $R * S$.

Phép kết theta ($R \bowtie_{\langle \mathcal{DK} \rangle} S$)

■ Biểu thức điều kiện

- Chứa các mệnh đề có dạng
 - $A_i \langle \text{toán tử so sánh} \rangle B_j$.
 - A_i là thuộc tính của R.
 - B_j là thuộc tính của S.
 - Miền giá $A_i \equiv$ Miền giá trị B_j .
- Toán tử so sánh: $=, <, \leq, >, \geq, \neq$.
- Các mệnh đề được nối bởi toán tử logic: \wedge, \vee, \neg .

R	A	B	C
	α	α	1
	α	β	5
	β	α	5
	β	β	12

S	E	F
	α	1
	α	4
	β	12

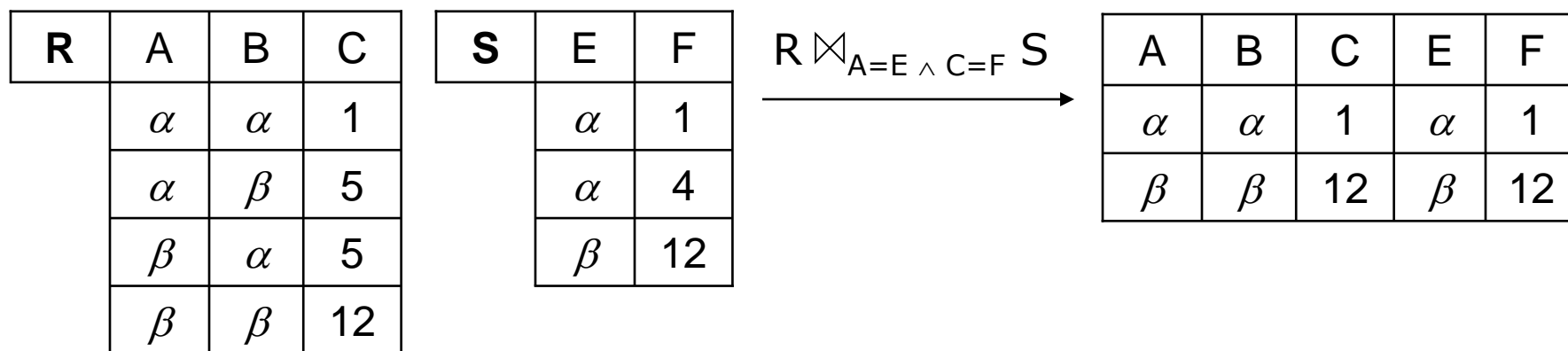
$R \bowtie_{A=E \wedge C < F} S$

A	B	C	E	F
α	α	1	α	4
β	α	5	β	12

A	B	C	E	F
α	α	1	α	1
α	α	1	α	4
α	α	1	β	12
α	β	5	α	1
α	β	5	α	4
α	β	5	β	12
β	α	5	α	1
β	α	5	α	4
β	α	5	β	12
β	β	12	α	1
β	β	12	α	4
β	β	12	β	12

Phép kết bằng

- Tất cả các toán tử so sánh trong biểu thức điều kiện đều là =.



- Trong mỗi bộ luôn có một hoặc nhiều cặp thuộc tính có giá trị giống nhau.

Phép kết tự nhiên

- Là phép kết bảng và các cặp thuộc tính trong các mệnh đề phải cùng tên và cùng miền giá trị.

R	A	B	C
	α	α	1
	α	β	5
	β	α	5
	β	β	12

S	A	C
	α	1
	α	4
	β	12

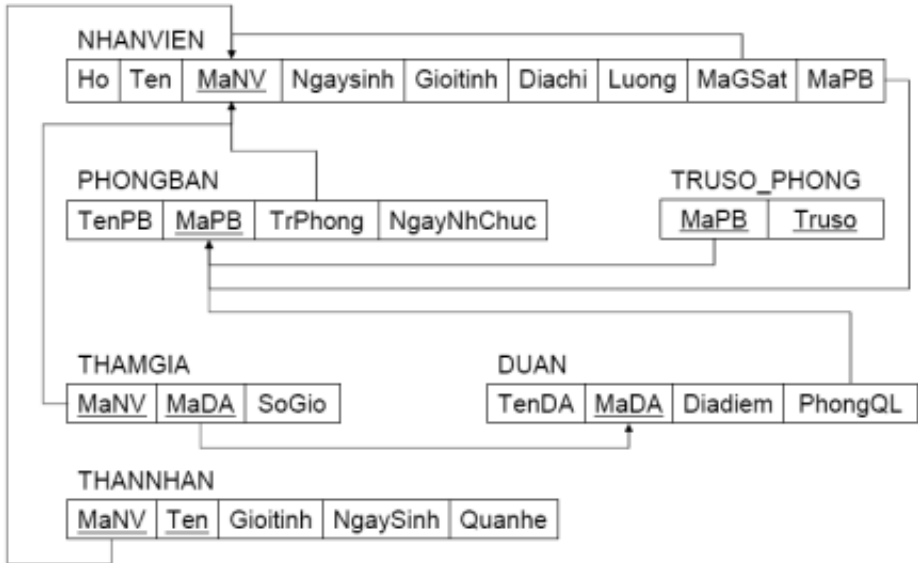
 $\xrightarrow{R \bowtie S}$

	A	B	C
	α	α	1
	β	β	12

- Nếu các cặp thuộc tính không cùng tên thì phải thực hiện phép toán đổi tên trước khi kết.
 - $R(A, B, C)$ và $S(E, F)$, muốn kết tự nhiên trên 2 cặp thuộc tính (A, E) và (C, F) .
 - $R \bowtie (\rho_{(A,C)}(S))$.

■ Cho biết họ, tên của các nhân viên nữ và tên các thân nhân của họ.

■ Cho biết tên các nhân viên tham gia dự án do phòng số 5 quản lý



■ Cho biết tên, địa chỉ của các nhân viên của phòng Nghiên cứu.

Phép chia - Division (1)

- Để rút trích các bộ của một quan hệ liên quan với tất cả các bộ của quan hệ còn lại.
- Cho 2 quan hệ $R(Z)$ và $S(X)$
 - Z tập hợp các thuộc tính của quan hệ R .
 - X tập hợp các thuộc tính của quan hệ S .
 - $X \subseteq Z$.
 - R chia S là quan hệ $T(Y)$ với $Y = Z - X$.
 - $T(Y) = \{t \mid t \in \pi_Y(R) \wedge \forall u \in S \Rightarrow (t, u) \in R\}$.
- Cú pháp
 - $R \div S$

Phép chia - Division (2)

A	B	C
α	α	1
α	β	12
β	α	23
β	β	3
β	β	23
β	β	12

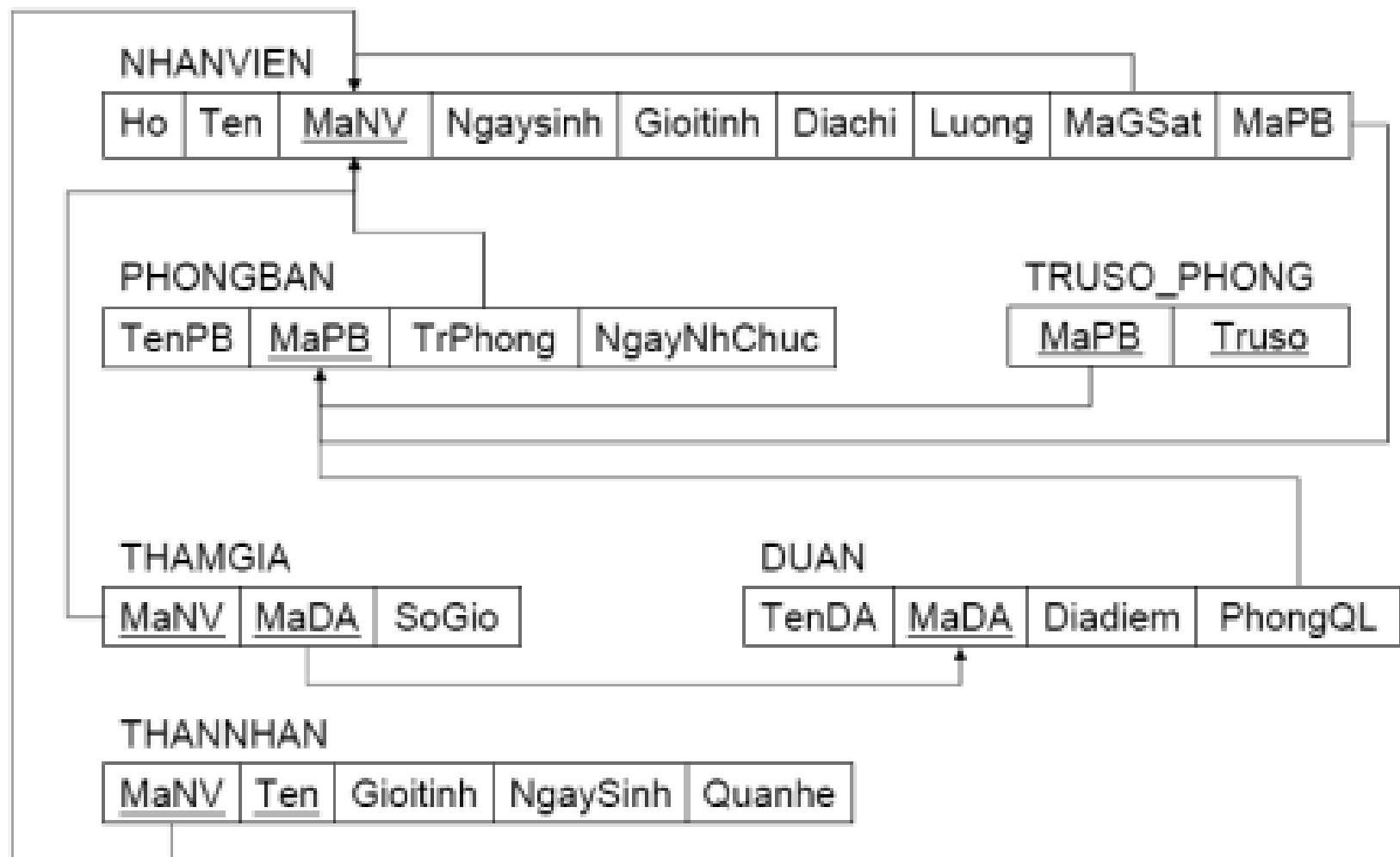
$\pi_{A,B,C}(R)$

S	D	E
	7	2
	2	5

A	B	C	D	E	$\in R$
α	α	1	7	2	Có
α	α	1	2	5	Có
α	β	12	7	2	Có
α	β	12	2	5	k0
β	α	23	7	2	Có
β	α	23	2	5	Có
β	β	3	7	2	ko
β	β	23	7	2	Ko
...

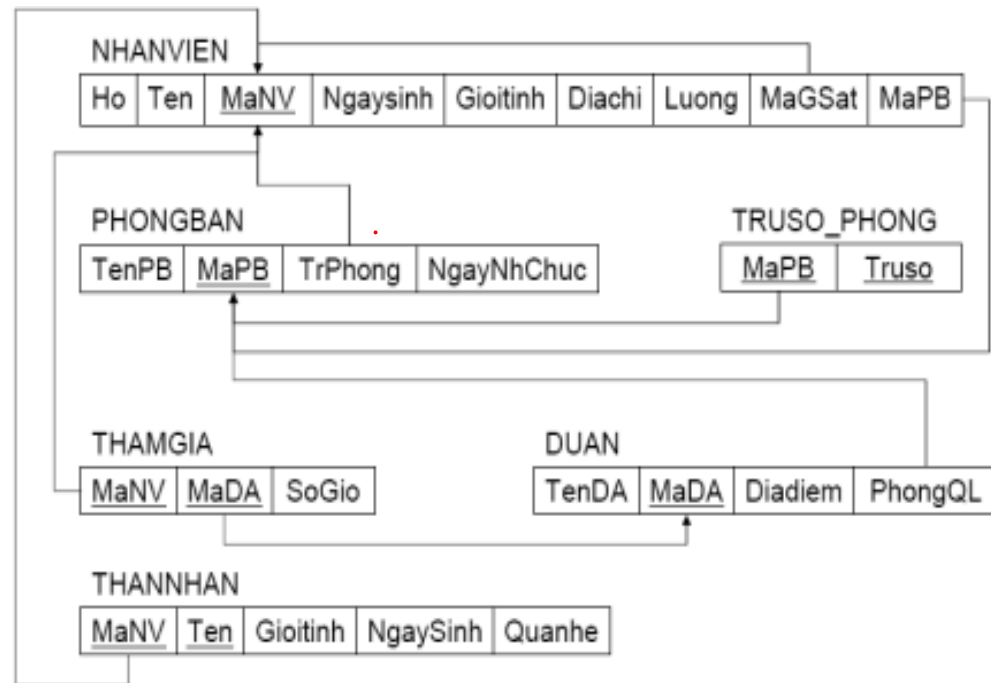
R	A	B	C	D	E
	α	α	1	7	2
	α	α	1	2	5
	α	β	12	7	2
	β	α	23	7	2
	β	β	3	1	10
	β	α	23	2	5
	β	β	23	10	10
	β	β	12	1	2

$R \div S$	A	B	C
	α	α	1
	β	α	23



Một số ví dụ

- Cho biết tên các nhân viên tham gia *tất cả* các dự án do phòng số 5 điều phối.



Các phép toán khác

- Để biểu diễn các truy vấn mà không thể thực hiện với các phép toán đại số quan hệ cơ sở
 - Các truy vấn mang tính chất thống kê đơn giản trên một tập hợp các giá trị hoặc các nhóm tập hợp giá trị dữ liệu.
 - Các truy vấn dùng để tạo các báo cáo.
- Gồm
 - Hàm tập hợp (Aggregate Function).
 - Phép gom nhóm các bộ dữ liệu (Grouping).
 - Phép kết mở rộng (Outer Join).

Hàm gộp và các phép toán nhóm (1)

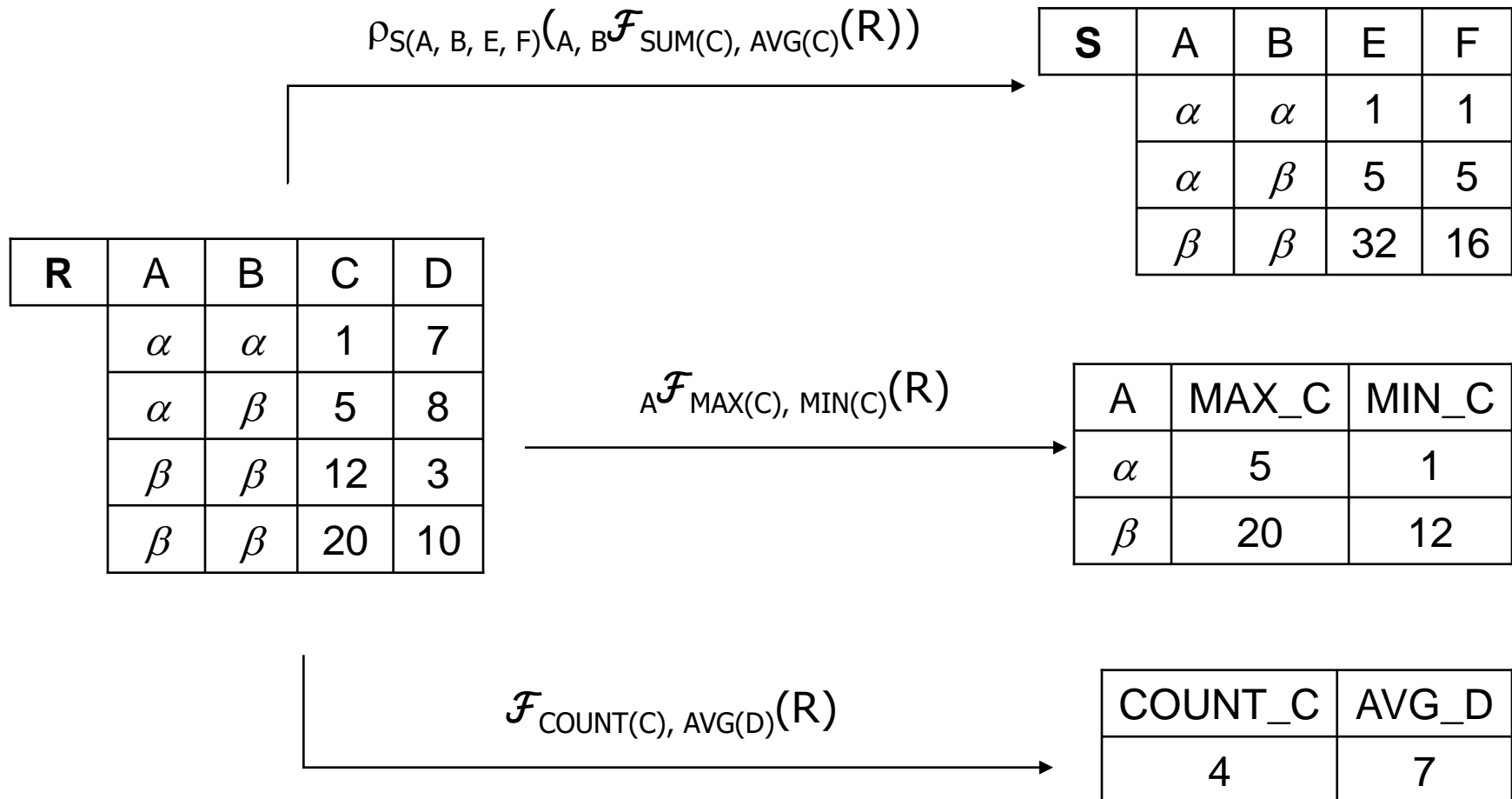
- Hàm gộp

- SUM - Tính tổng của các giá trị trong tập hợp.
- AVG - Tính giá trị trung bình của các giá trị trong tập hợp.
- MAX, MIN - Tìm giá trị lớn nhất, nhỏ nhất của các giá trị trong tập hợp.
- COUNT - Đếm số bộ của một quan hệ hoặc số các giá trị của một thuộc tính.

- Cú pháp

- $\langle \text{DSTT} \rangle \mathcal{F}_{\langle \text{DSH} \rangle}(\mathbf{R})$
- $\langle \text{DSTT} \rangle$ là danh sách các thuộc tính thuộc \mathbf{R} .
- $\langle \text{DSH} \rangle$ là danh sách các cặp (hàm tập hợp, thuộc tính) áp dụng trên các nhóm.

Hàm tập hợp và gom nhóm (2)



$A, B \mathcal{F}_{\text{SUM}(C), \text{AVG}(C)}(R)$

STT	A	B	C	D
1	α	β	1	3
2	α	α	3	5
3	α	β	2	6
4	α	α	3	3
5	α	β	4	5
6	α	α	7	8
7	β	β	1	11
8	α	α	1	3
9	β	β	2	9
10	α	α	4	4

A	B	SUM(C)	AVG(C)
α	β	7	$\frac{7}{3}$
α	α	18	$\frac{18}{5}$
β	β	3	$\frac{3}{2}$

$A, B \mathcal{F} \text{SUM}(C), \text{AVG}(C)(R)$

Tên trường	A	B	C	D
	X	Y	3	5
	X	Y	5	6
	X	X	7	8
	Y	Y	5	6
	Y	Y	8	9

A	B	Sum(C)	AVG(C)
X	Y	8	4
X	X	7	7
Y	Y	13	6.5

Phép kết mở rộng (1)

- Để giữ lại tất cả các bộ trong một quan hệ bất chấp chúng có được liên kết với các bộ trong quan hệ còn lại hay không nhằm tránh mất thông tin hoặc tạo các báo cáo.
- Có 3 dạng
 - Mở rộng trái (Left Outer Join)
 - $R \bowtie_{\langle DK \rangle} S$.
 - Mở rộng phải (Right Outer Join)
 - $R \bowtie_{\langle DK \rangle} S$.
 - Mở rộng hai phía (Full Outer Join)
 - $R \bowtie_{\langle DK \rangle} S$.

Phép kết mở rộng trái

- Giữ lại tất cả các bộ của quan hệ ở bên trái phép toán kết mà không liên kết được với bộ nào của quan hệ bên phải.

R	A	B	C
	α	α	1
	α	β	5
	β	β	12
	β	β	23

S	D	E
	1	7
	2	7
	12	3
	23	10

$R \bowtie_{C < D} S$

A	B	C	D	E
α	α	1	2	7
α	α	1	12	3
α	α	1	23	10
α	β	5	12	3
α	β	5	23	10
β	β	12	23	10
β	β	23	null	null

Phép kết mở rộng phải

- Giữ lại tất cả các bộ của quan hệ ở bên phải phép toán kết mà không liên kết được với bộ nào của quan hệ bên trái.

R	A	B	C
	α	α	1
	α	β	5
	β	β	12
	β	β	23

S	D	E
	1	7
	2	7
	12	3
	23	10

$R \bowtie_{C > D} S$

A	B	C	D	E
α	β	5	1	7
α	β	5	2	7
β	β	12	1	7
β	β	12	2	7
β	β	23	1	7
β	β	23	2	7
β	β	23	12	3
null	null	null	23	10

Phép kết nối mở rộng hai phía

- Giữ lại tất cả các bộ của từng quan hệ ở hai bên phép toán kết mà không liên kết được với bộ nào của quan hệ còn lại.

R	A	B	C
	α	α	1
	α	β	5
	β	β	12
	β	β	23

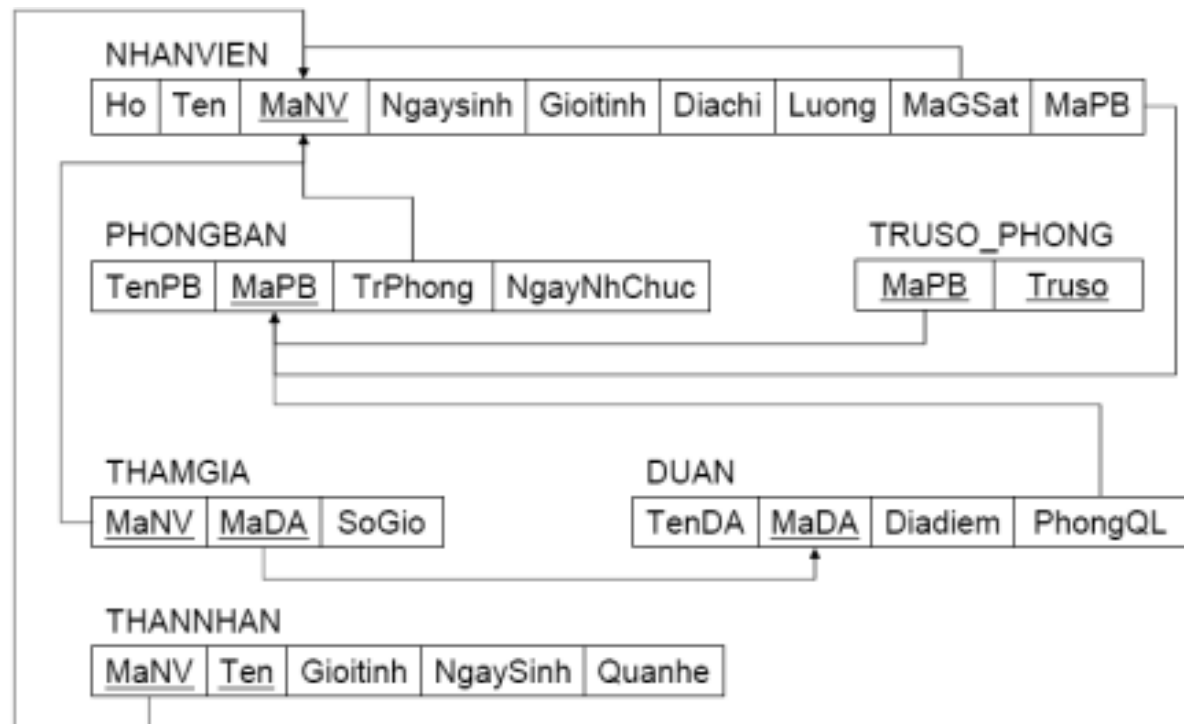
S	D	E
	1	7
	2	7
	12	3
	23	10



A	B	C	D	E
α	α	1	1	7
β	β	12	12	3
β	β	23	23	10
α	β	5	null	null
null	null	null	2	7

Một số ví dụ

- Với mỗi phòng ban cho biết mã số, tổng số nhân viên và mức lương trung bình.
- Với mỗi nhân viên cho biết họ, tên và tên phòng nếu họ là trưởng phòng.



Các phép sửa đổi CSDL

- Thêm (Insert)
Cập nhật (Update)
Xóa (Delete)
Các phép sửa đổi dữ liệu đều sử dụng phép gán

Thêm

- Cú pháp

$$r \leftarrow r \cup E$$

- r là quan hệ
- E là một biểu thức đại số quan hệ

- Ví dụ

$$\begin{aligned} \text{account} &\leftarrow \text{account} \cup \{(\text{"A-973"}, \text{"Perryridge"}, 1200)\} \\ \text{depositor} &\leftarrow \text{depositor} \cup \{(\text{"Smith"}, \text{"A-973"})\} \end{aligned}$$

Xóa

- Cú pháp

$$r \leftarrow r - E$$

- r là quan hệ
- E là một biểu thức đại số quan hệ

- Ví dụ

$$account \leftarrow account - \sigma_{branch_name = "Perryridge"}(account)$$

$$loan \leftarrow loan - \sigma_{amount \geq 0 \text{ and } amount \leq 50}(loan)$$

Cập nhật

- Cú pháp

$$r \leftarrow \Pi_{F1, F2, \dots, Fn}(r)$$

- r là quan hệ
- F_i là các hàm tính toán trên các thuộc tính

- Ví dụ

$$account \leftarrow \Pi_{account_number, branch_name, balance * 1.05}(account)$$