

Đại số quan hệ

Nguyễn Quốc Tuấn
Bm. Mạng & HTTT

Nội dung

- Tổng quan
- Đại số quan hệ
 - Phép toán một ngôi
 - Phép toán hai ngôi
 - Phép toán khác

Tổng quan

□ Đại số quan hệ

- Là tập 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 cài đặt và tối ưu hóa các câu hỏi trong các HQT CSDL quan hệ
- Được áp dụng trong SQL

Tổng quan

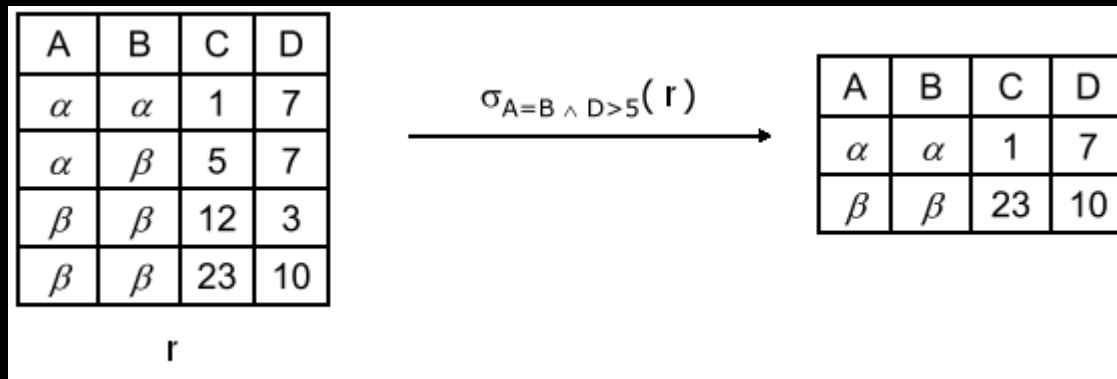
- Toán hạng
 - Các thể hiện quan hệ
 - Các tập hợp
- Toán tử là các phép toán
 - Phép toán tập hợp
 - Hội, giao, hiệu, tích Cartesian
 - Phép toán quan hệ
 - Chọn, chiếu, kết, chia, đổi tên
 - Một số phép toán khác

Phép toán 1 ngôi

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

Phép chọn – Select (σ) - 1

- Để chọn ra các bộ dữ liệu thỏa điều kiện chọn từ một quan hệ
- Ví dụ:



Phép chọn – Select (σ) - 2

- Ký hiệu $\sigma_{\langle DK \rangle}(r)$

- Cú pháp

$$\sigma_{\langle DK \rangle}(r) = \{t | t \in r \text{ và } DK(t) \text{ đúng}\}$$

- DK

- 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: $=, <, \leq, >, \geq, \neq$

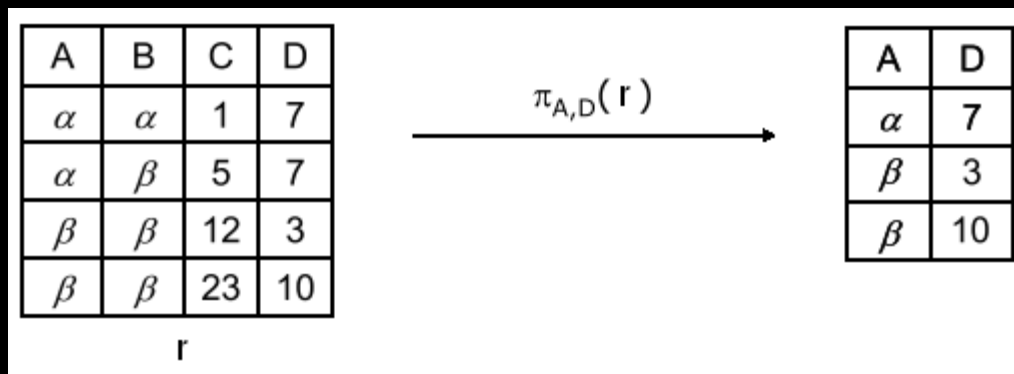
- Các mệnh đề được nối bởi toán tử logic: \wedge, \vee, \neg

Phép chọn – Select (σ) - 3

- Đặc trưng của phép chọn
 - 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 – Project (π) - 1

- Để chọn ra các cột ứng với các thuộc tính nào đó của quan hệ
- Ví dụ:



Phép chiếu – Project (π) - 2

- Ký hiệu: $\pi_{\langle \text{DSTT} \rangle}(\mathbf{r})$

- Cú pháp

$$\pi_{\langle \text{DSTT} \rangle}(\mathbf{r}) = \{t[\text{DSTT}] \mid t \in \mathbf{r}\}$$

- DSTT là danh sách các thuộc tính

Phép chiếu – Project (π) - 3

□ Đặc trưng của phép chiếu

- Phép chiếu không có tính giao hoán

- $\pi_{\langle \text{DSTT1} \rangle}(\pi_{\langle \text{DSTT2} \rangle}(r)) \neq \pi_{\langle \text{DSTT2} \rangle}(\pi_{\langle \text{DSTT1} \rangle}(r))$

- 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.
 - Ví dụ: $\pi_{A,2 * C}(r)$

Chuỗi các phép toán và phép gán

□ Chuỗi các phép toán

- Sử dụng kết quả của phép toán này làm toán hạng của phép toán khác
- Viết các phép toán lồng nhau
- Ví dụ: $\pi_{A,C}(\sigma_{A=B \wedge D>5}(r))$

□ 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
 - $r' \leftarrow \sigma_{A=B \wedge D>5}(r)$
 - $\pi_{A,C}(r')$

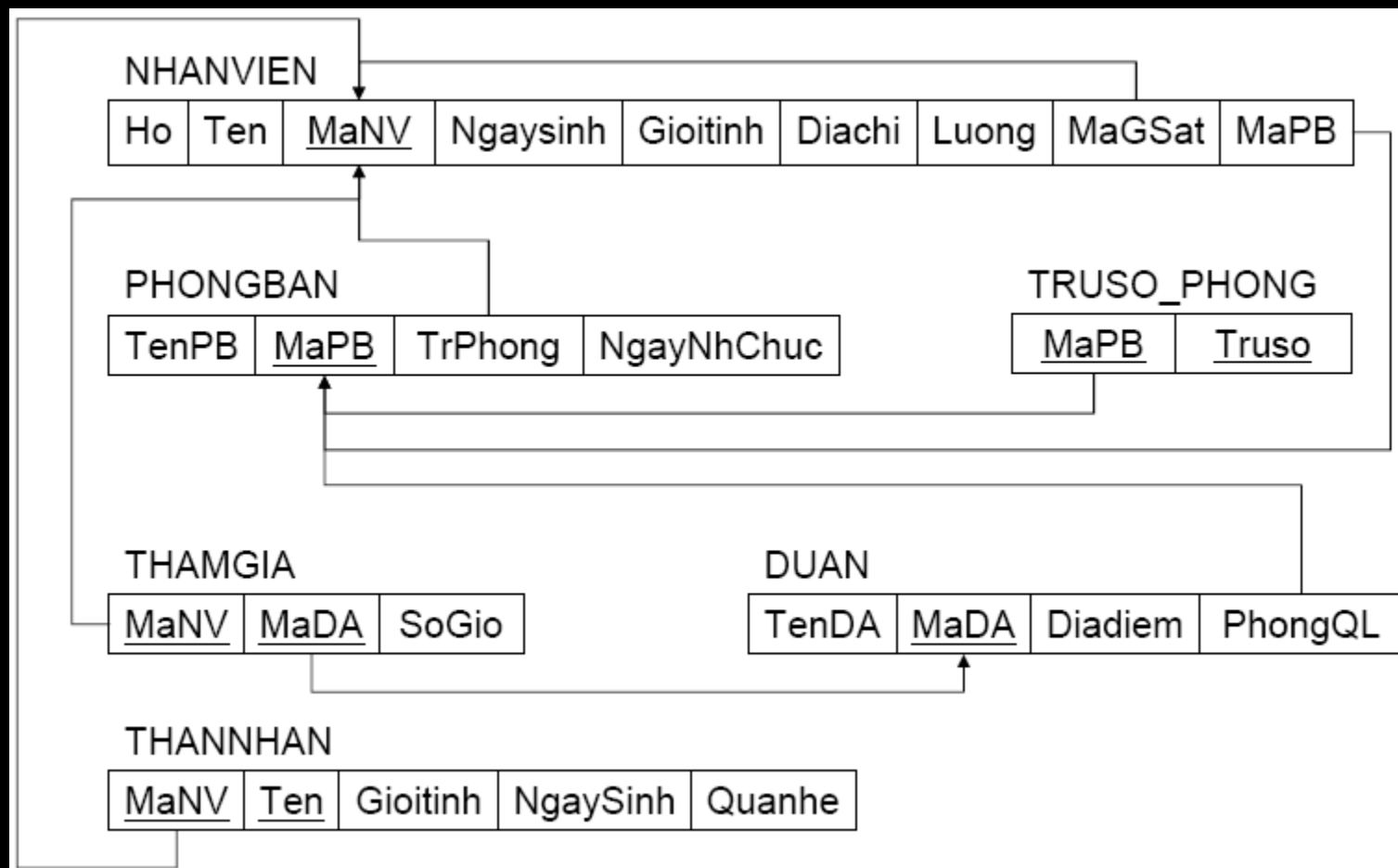
Phép đổi tên – Rename (ρ) -1

- Đề đổi tên quan hệ và tên thuộc tính
- Cú pháp: cho lược đồ 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
 - $\rho_{(B_1, A_2, A_3, \dots, A_n)}(R)$

Đại số quan hệ

Một số ví dụ

□ Cho lược đồ CSDL quan hệ sau:



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.

Một số ví dụ

- $\sigma_{\text{MaPB} = 4}(\text{nhanvien})$
- $\sigma_{\text{MaPB} = 4 \wedge \text{Luong} \geq 25.000 \wedge \text{Luong} \leq 40.000}(\text{nhanvien})$
- $\pi_{\text{Ho, Ten, Gtinh, Luong}}(\text{nhanvien})$
- $\pi_{\text{Ho, Ten, Gtinh, Luong}}(\sigma_{\text{MaPB} = 5}(\text{nhanvien}))$

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 nối (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 và s trên 2 lược đồ 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 xác định $A_i \equiv \text{Miền xác định } B_i$, với $i = 1, \dots, n$

Phép toán tập hợp - 2

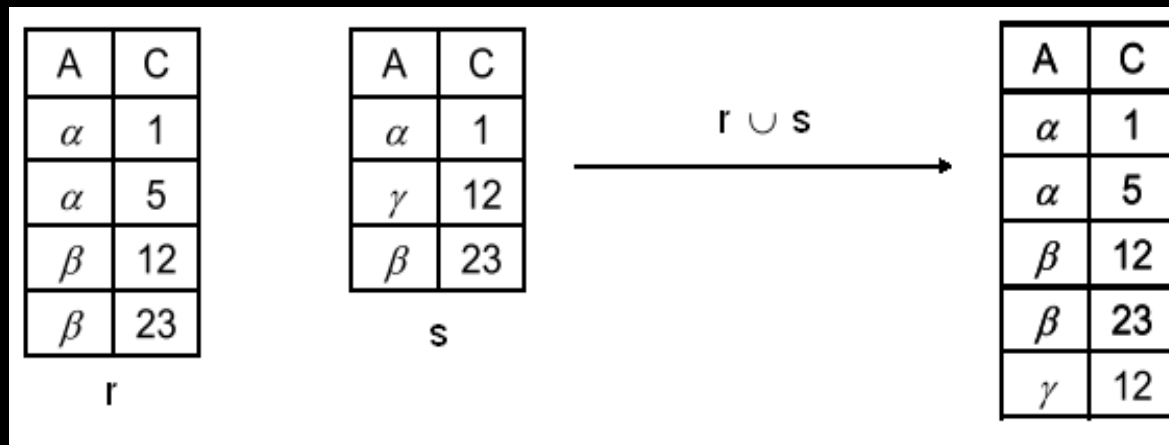
□ Phép hội – Union

- Ký hiệu: $r \cup s$

- Cú pháp

$$r \cup s = \{t \mid t \in r \text{ or } t \in s\}$$

- Ví dụ:



Phép toán tập hợp - 3

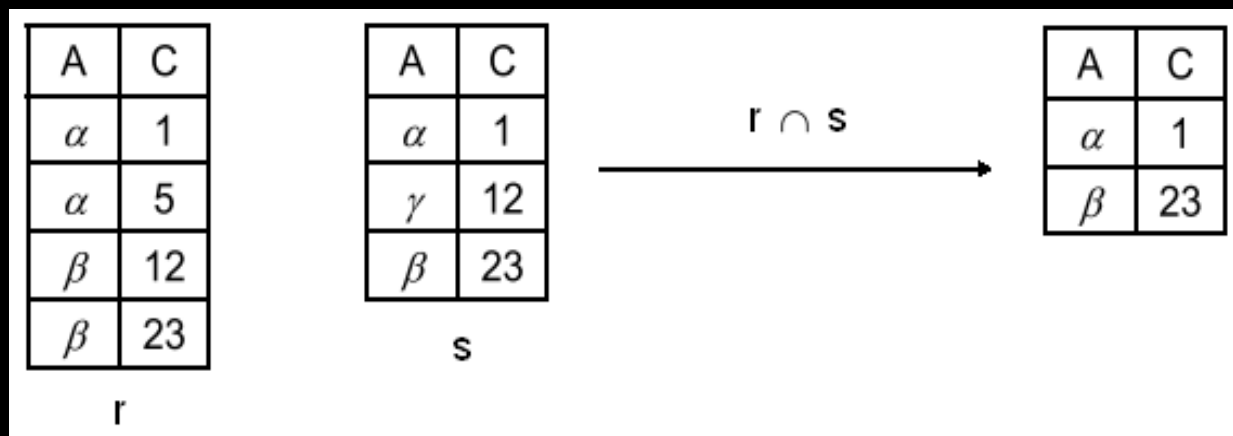
□ Phép giao – Intersection

- Ký hiệu: $r \cap s$

- Cú pháp

$$r \cap s = \{t \mid t \in r \text{ and } t \in s\}$$

- Ví dụ:



Phép toán tập hợp - 4

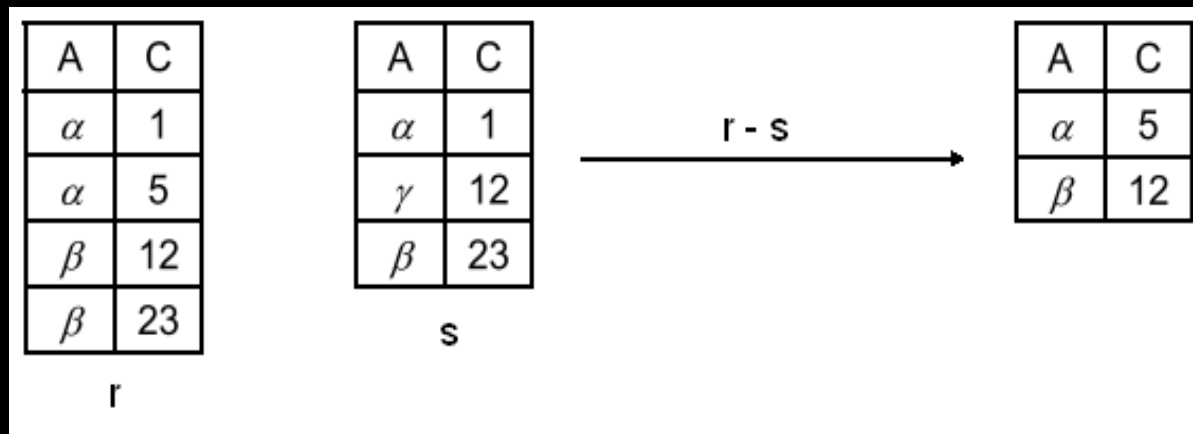
□ Phép hiệu – Difference

- Ký hiệu: $r - s$

- Cú pháp

$$r - s = \{t \mid t \in r \text{ and } t \notin s\}$$

- Ví dụ:



Phép toán tập hợp - 5

□ Đặc trưng

■ Phép hội và giao có tính giao hoán

□ $r \cup s = s \cup r$

□ $r \cap s = s \cap r$

■ Phép hội và giao có tính kết hợp

□ $r \cup (s \cup q) = (r \cup s) \cup q$

□ $r \cap (s \cap q) = (r \cap s) \cap q$

Phép toán tập hợp - 6

□ Phép tích Đề các – Cartesian (không nhất thiết khả hợp)

- Ký hiệu: $r \times s$

- Cú pháp

$$r \times s = \{t \ q \mid t \in r \textbf{ and } q \in s\}$$

- Ví dụ:

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

Một số ví dụ

Xét ví dụ quản lý ở trên. Hãy:

- 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

□ $q_1 \leftarrow \sigma_{\text{MaPB} = 5}(\text{nhanvien})$

$$q_2 \leftarrow \pi_{\text{MaNV}}(q_1)$$

$$q_3 \leftarrow \pi_{\text{MaGS}}(q_1)$$

$$q \leftarrow q_2 \cup q_3$$

□ $q_1 \leftarrow \sigma_{\text{GTinh} = \text{'Nu'}}(\text{nhanvien})$

$$q_2 \leftarrow \rho_{(\text{HoNV}, \text{TenNV}, \text{MaNV1})}(\pi_{\text{Ho}, \text{Ten}, \text{MaNV}}(q_1))$$

$$q_3 \leftarrow q_2 \times \text{thannhan}$$

$$q_4 \leftarrow \sigma_{\text{MaNV1} = \text{MaNV}}(q_3)$$

$$q \leftarrow \pi_{\text{HoNV}, \text{TenNV}, \text{Ten}}(q_4)$$



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 nối theta (Theta Join)
 - $r \bowtie_{\langle \text{ĐK} \rangle} s$
 - $\langle \text{ĐK} \rangle$ là biểu thức logic
 - Kết nối bằng (Equi Join)
 - Kết tự nhiên (Natural Join)
 - $r \bowtie s$ hoặc $r * s$

Phép kết nối theta

□ $r \bowtie_{\langle \text{ĐK} \rangle} s$

□ ĐK: 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

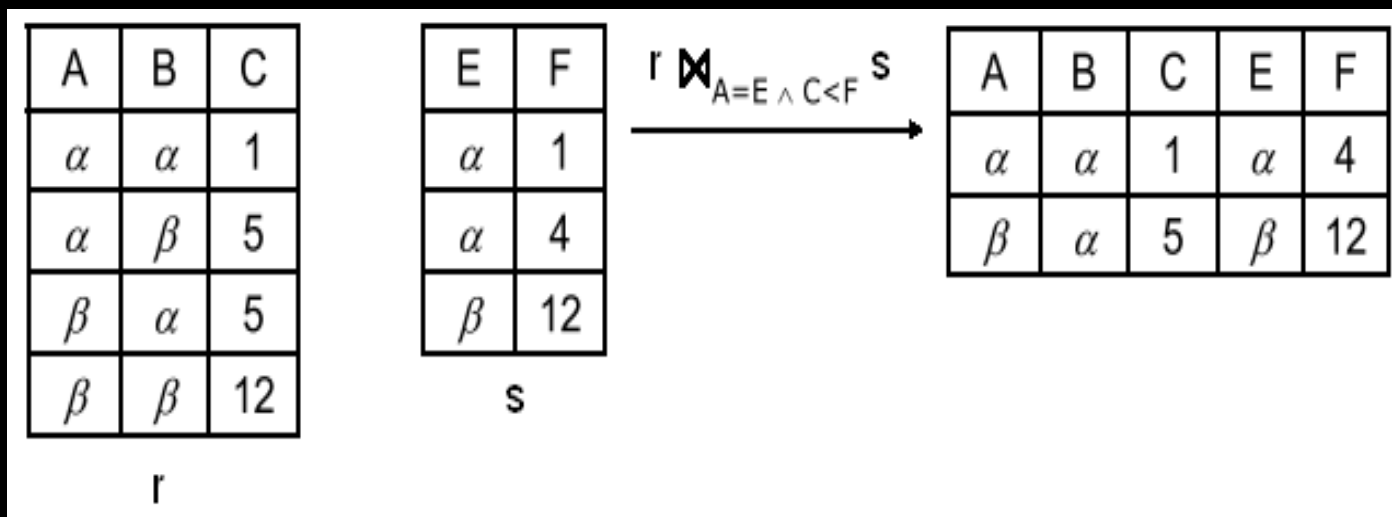
■ B_j là thuộc tính của s

■ Miền xác định $A_i \equiv \text{Miền xác định } B_j$

■ Toán tử so sánh: $=, <, \leq, >, \geq, \neq$

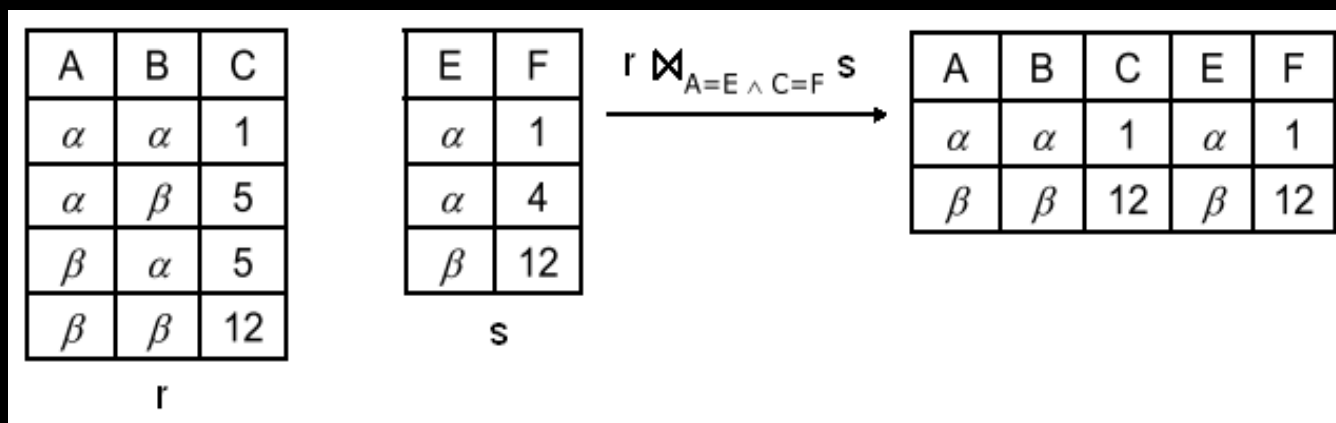
■ Các mệnh đề được nối bởi toán tử logic: \wedge

Phép kết nối theta



Kết nối bằng

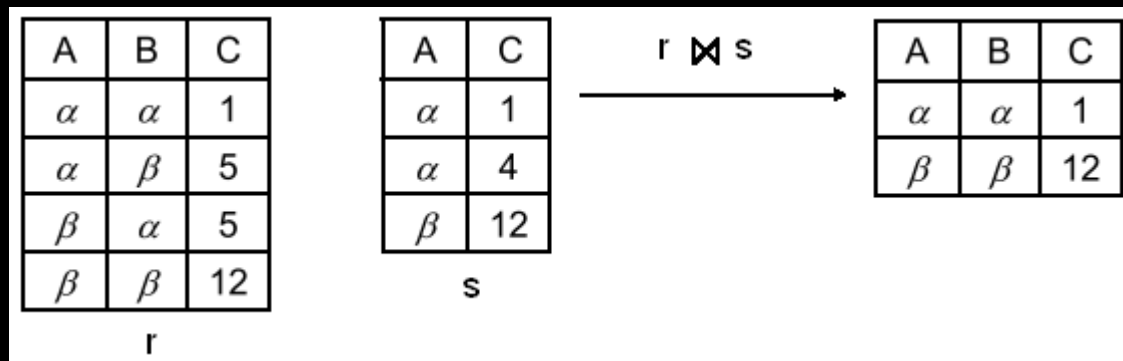
- Tất cả các toán tử so sánh trong biểu thức điều kiện đều là =



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

Kết nối tự nhiên

- Là phép kết bảng và các cặp thuộc tính trong các mệnh đề phải cùng tên và cùng miền xác định



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

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
- Ký hiệu: $r \div s$
- Cho 2 quan hệ r và s trên 2 lược đồ R và S
 - $R = (A_1, \dots, A_m, B_1, \dots, B_n)$
 - $S = (B_1, \dots, B_n)$
 - Kết quả của $r \div s$ là một quan hệ trên lược đồ $R - S = (A_1, \dots, A_m)$

$$r \div s = \{ t \mid t \in \Pi_{R-S}(r) \wedge \forall u \in s ((t, u) \in r) \}$$

Phép chia – Division - 2

<i>A</i>	<i>B</i>
α	1
α	2
α	3
β	1
γ	1
δ	1
δ	3
δ	4
\in	6
\in	1
β	2

r

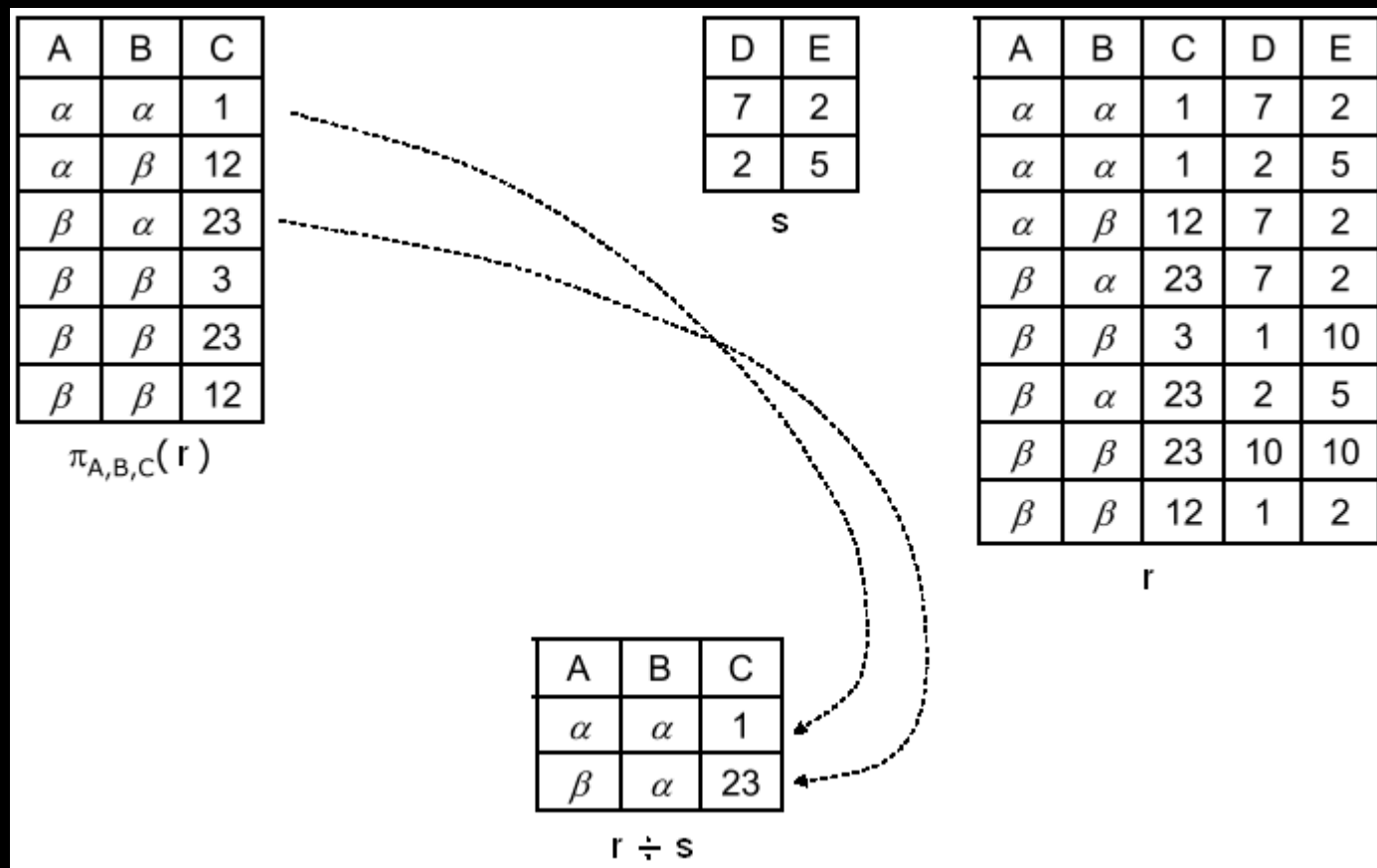
<i>B</i>
1
2

s

<i>A</i>
α
β

$r \div s$

Phép chia – Division - 2



Một số ví dụ

- Cho biết tên, địa chỉ của các nhân viên của phòng Nghiên cứu
- Cho biết tên các nhân viên tham gia tất cả các dự án do phòng số 5 điều phối

Một số ví dụ

$$\square \quad q_1 \leftarrow \sigma_{\text{TenPB} = \text{'Nghien cuu'}}(\text{phongban})$$

$$q_2 \leftarrow q_1 * \text{nhanvien}$$

$$q \leftarrow \pi_{\text{Ho, Ten, Dchi}}(q_2)$$

$$\square \quad q_1 \leftarrow \pi_{\text{MaDA}}(\sigma_{\text{PhongQL} = 5}(\text{duan}))$$

$$q_2 \leftarrow \pi_{\text{MaNV, MaDA}}(\text{thamgia})$$

$$q_3 \leftarrow q_2 \div q_1$$

$$q \leftarrow \pi_{\text{Ho, Ten}}(q_3 * \text{nhanvien})$$

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 thông 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
- Bao gồm
 - Hàm gộp (Aggregate Function)
 - Phép nhóm các bộ dữ liệu (Grouping)
 - Phép kết nối mở rộng (Outer Join)

Hàm gộp và các phép toán nhóm - 1

- Hàm gộp

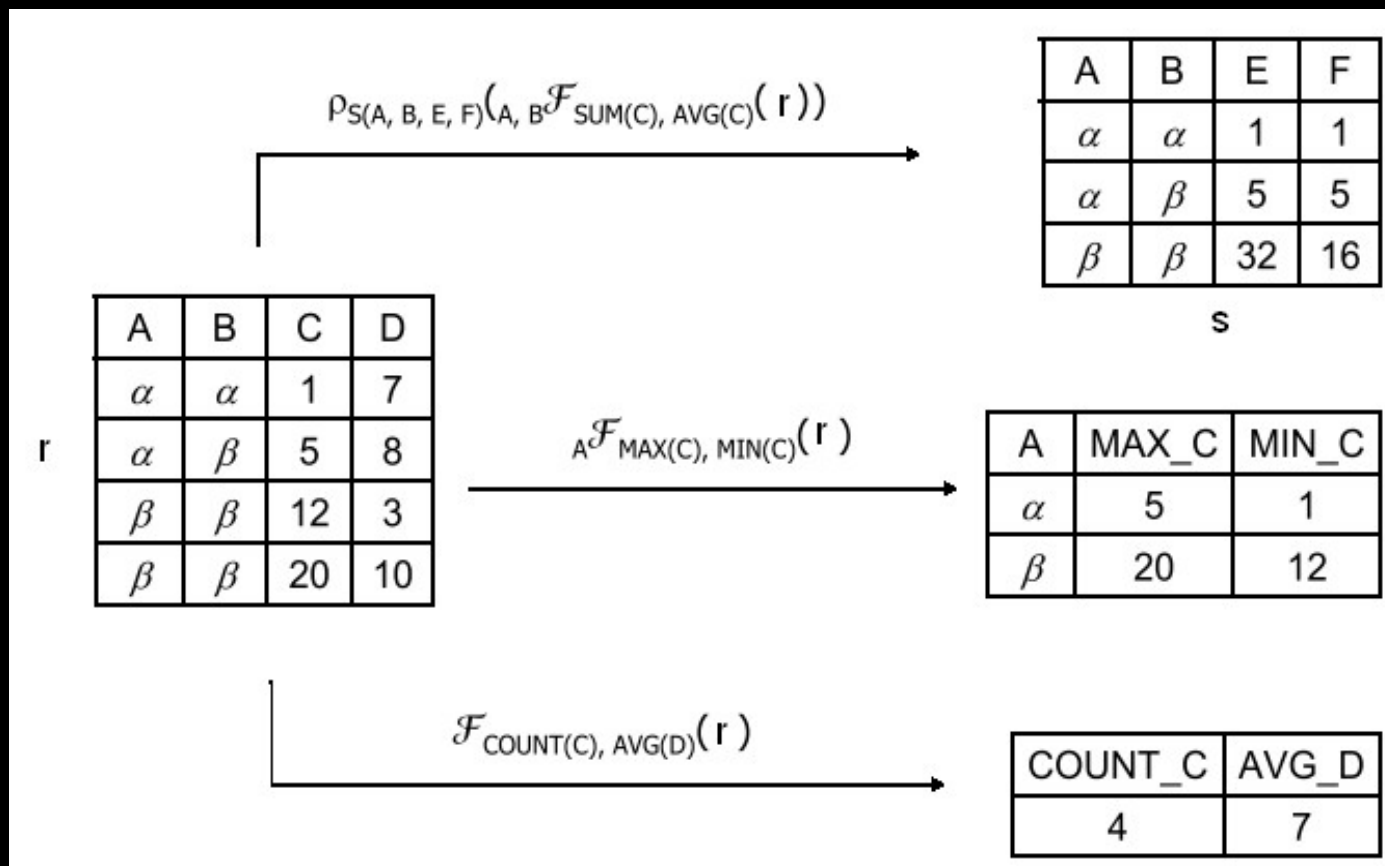
- Avg, sum, max, min, count

- 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 gộp và các phép toán nhóm - 2

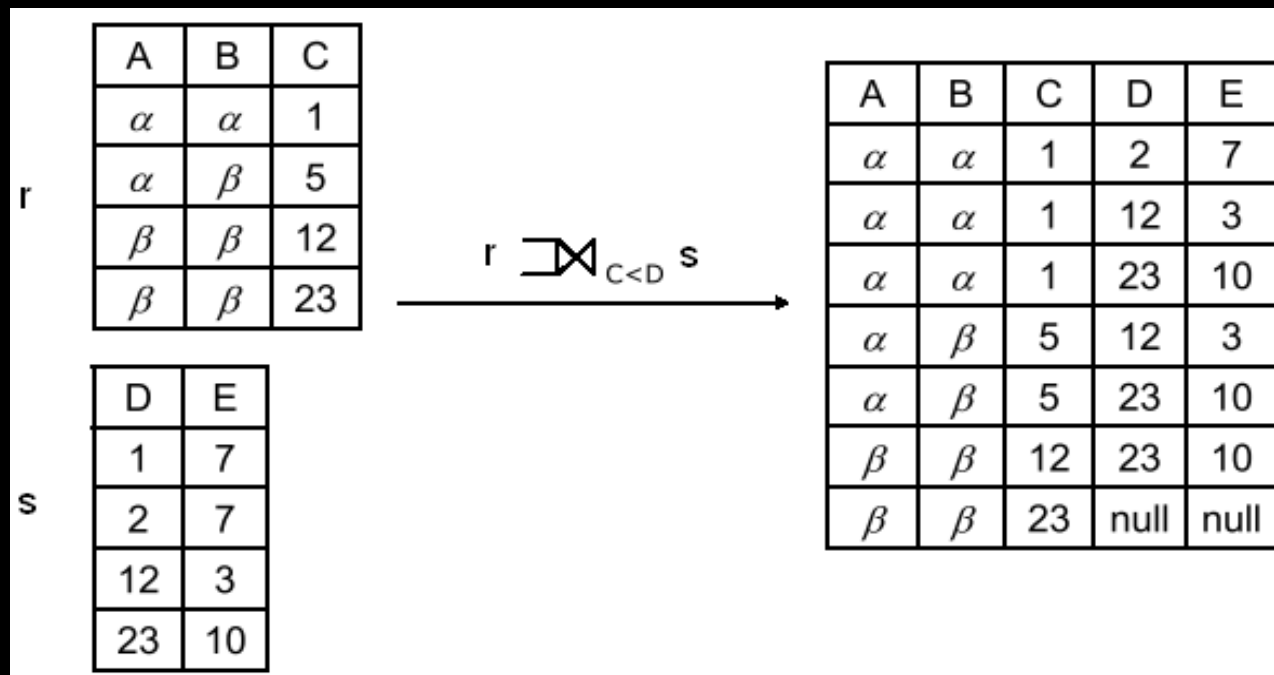


Phép kết nối mở rộng

- 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)
 - $r \bowtie_{\langle DK \rangle} S$

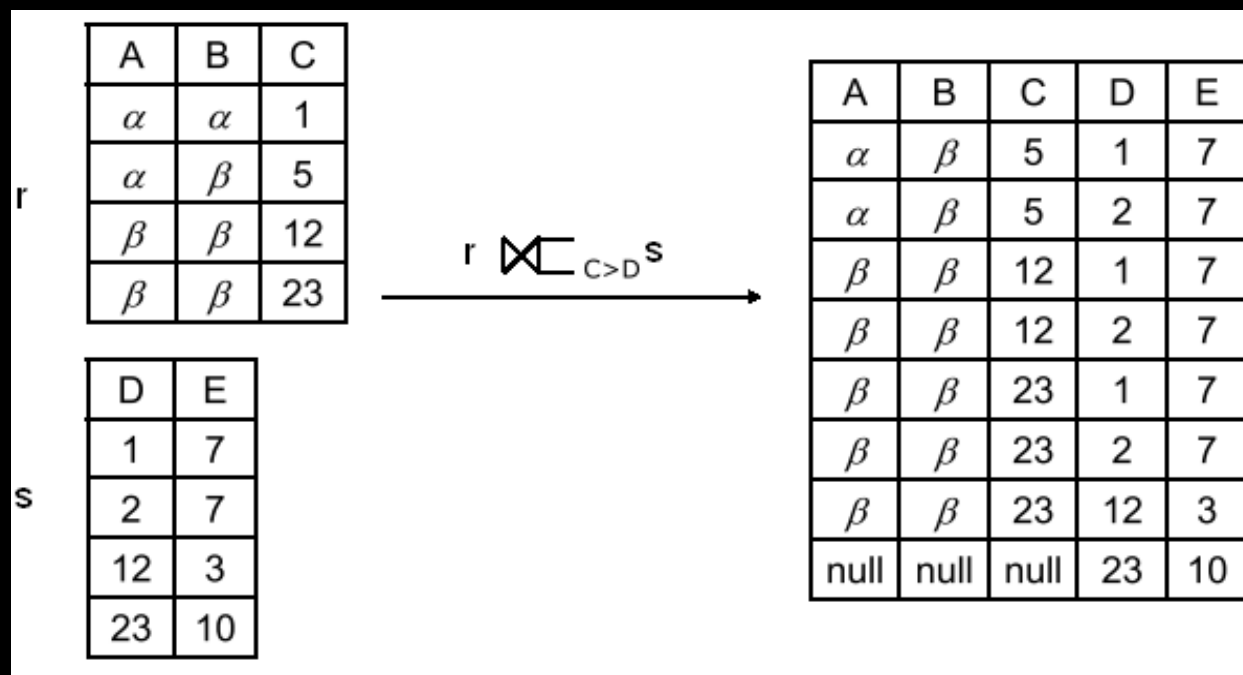
Phép kết nối mở rộng trái

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



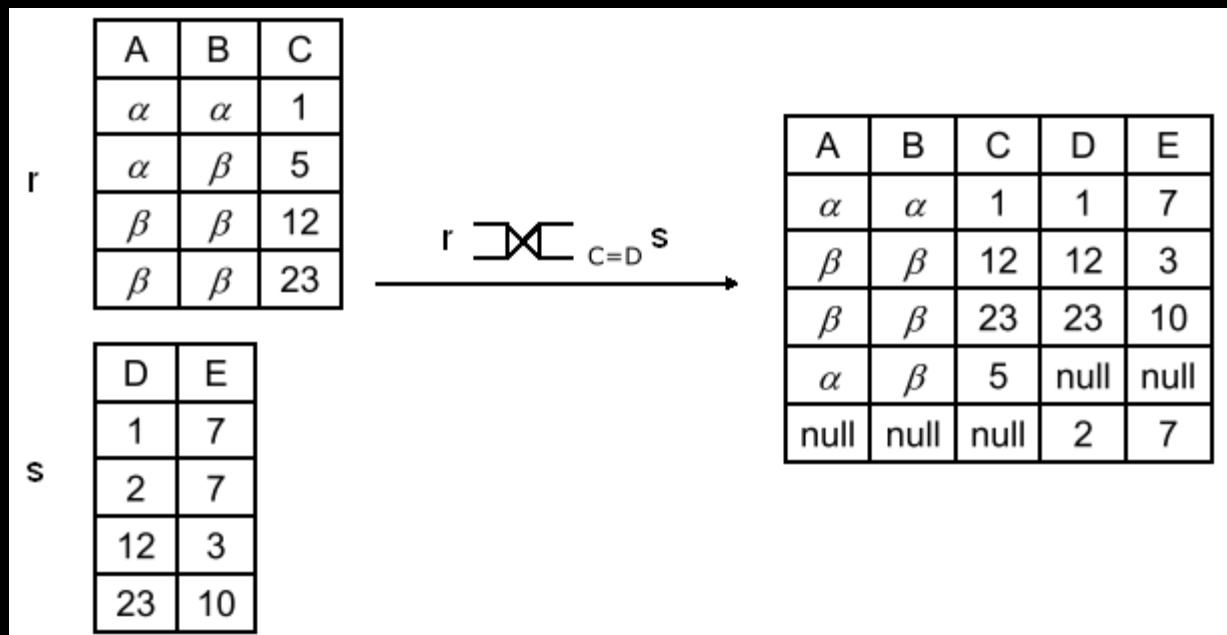
Phép kết nối mở rộng phải

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



Phép kết nối mở rộng 2 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



Một số ví dụ

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

Một số ví dụ

- $\rho_{(\text{MaPB}, \text{SoNV}, \text{LuongTB})}(\text{MaPB} \mathcal{F}_{\text{COUNT}(\text{MaNV}), \text{AVG}(\text{Luong})}(\text{nhanvien}))$
- $q_1 \leftarrow \text{nhanvien} \quad \square \bowtie_{\text{MaNV} = \text{TrPhong}} \text{phongban}$
 $q \leftarrow \pi_{\text{Ho}, \text{Ten}, \text{TenPB}}(q_1)$

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

$account \leftarrow account \cup \{("A-973", "Perryridge", 1200)\}$

$depositor \leftarrow depositor \cup \{("Smith", "A-973")\}$

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

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

■ r là quan hệ

■ F_i là các thuộc tính

□ Ví dụ

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