

Nội dung

- Xử lý truy vấn phân tán
 - Phân rã và cục bộ hóa truy vấn
 - Tối ưu hóa truy vấn phân tán
 - Trình tự kết nối
 - Xử lý truy vấn thích ứng

18

Bước 1 – Phân rã truy vấn

Tương tự như xử lý truy vấn tập trung.

Đầu vào: Truy vấn tính toán trên quan hệ toàn cục

- Chuẩn hóa: Biến đổi câu truy vấn thành một dạng chuẩn để xử lý tiếp.
 - Thao tác lượng từ hóa truy vấn và xác định đủ điều kiện
- Phân tích
 - Phát hiện và từ chối các truy vấn "không đúng"
- Đơn giản hóa
 - Loại bỏ các vị từ dư thừa
- Tái cấu trúc truy vấn
 - Truy vấn tính toán → Truy vấn đại số
 - Sử dụng các quy tắc chuyển đổi

19

Bước 2 – Cục bộ hóa dữ liệu

Đầu vào: Truy vấn đại số trên các quan hệ phân tán

- Xác định những mảnh nào có liên quan
- Chương trình cục bộ hóa
 - Thay thế mỗi truy vấn toàn cục bằng chương trình cụ thể hóa của nó
 - Tối ưu hóa

20

Ví dụ

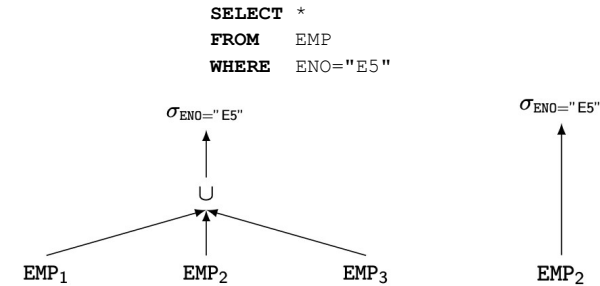
- Giả thiết
 - EMP được phân thành các mảnh như sau:
 - $EMP_1 = \sigma_{ENO \leq 'E3'}(EMP)$
 - $EMP_2 = \sigma_{'E3' < ENO \leq 'E6'}(EMP)$
 - $EMP_3 = \sigma_{ENO > 'E6'}(EMP)$
 - ASG được phân mảnh như sau:
 - $ASG_1 = \sigma_{ENO \leq 'E3'}(ASG)$
 - $ASG_2 = \sigma_{ENO > 'E3'}(ASG)$
- Trong bất kỳ truy vấn nào: Chương trình cục bộ hóa cho quan hệ phân mảnh ngang là hợp của các mảnh.
 - Thay thế EMP bởi $(EMP_1 \cup EMP_2 \cup EMP_3)$
 - Thay thế ASG bởi $(ASG_1 \cup ASG_2)$

21

Rút gọn phân mảnh ngang cơ sở

Rút gọn bằng phép chọn

- Quan hệ R và $F_R = \{R_1, R_2, \dots, R_w\}$ trong đó $R_i = \sigma_{p_i}(R)$
 $\sigma_{p_i}(R) = \emptyset$ nếu $\forall x$ trong $R: \neg(p_i(x) \wedge p_j(x))$



22

Rút gọn phân mảnh ngang cơ sở

Rút gọn với phép kết nối

- Có thể nếu việc phân mảnh được thực hiện trên thuộc tính kết nối
- Phân tán phép kết nối trên phép hợp

$$(R_1 \cup R_2) \bowtie S \Leftrightarrow (R_1 \bowtie S) \cup (R_2 \bowtie S)$$

- Cho $R_i = \sigma_{p_i}(R)$ và $R_j = \sigma_{p_j}(R)$

$$R_i \bowtie R_j = \emptyset \text{ nếu } \forall x \text{ trong } R_i, \forall y \text{ trong } R_j: \neg(p_i(x) \wedge p_j(y))$$

23

Rút gọn phân mảnh ngang cơ sở

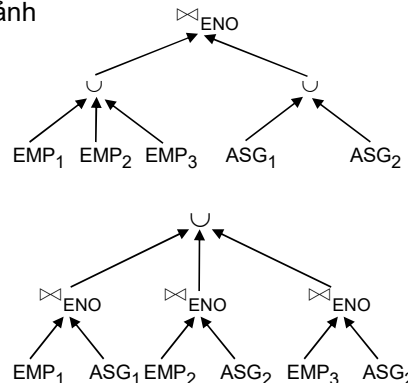
Giả sử EMP được phân mảnh như trước và

- ASG₁: $\sigma_{ENO \leq "E3"}(ASG)$
- ASG₂: $\sigma_{ENO > "E3"}(ASG)$

Xét câu truy vấn

SELECT *
FROM EMP
NATURAL JOIN ASG

- Phân tán phép kết nối trên phép hợp
- Áp dụng quy tắc rút gọn



24

Rút gọn phân mảnh dọc

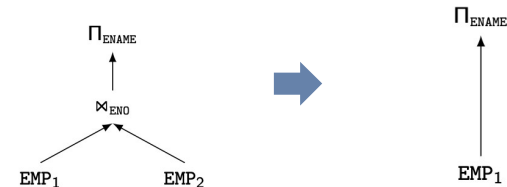
Các quan hệ trung gian (không rỗng) không được sử dụng

Quan hệ R được xác định trên các thuộc tính $A = \{A_1, \dots, A_n\}$ được phân mảnh dọc như sau $R_i = \Pi_{A'}(R)$ trong đó $A' \subseteq A$:

$\Pi_{D,K}(R_i)$ là không được sử dụng khi tập các thuộc tính chiếu D không nằm trong A'

Ví dụ: $EMP_1 = \Pi_{ENO, ENAME}(EMP)$; $EMP_2 = \Pi_{ENO, TITLE}(EMP)$

SELECT ENAME
FROM EMP



25

Rút gọn phân mảnh ngang dẫn xuất

■ Quy tắc:

- Phân tán các phép kết nối trên phép hợp
- Áp dụng rút gọn kết nối cho phân mảnh ngang

■ Ví dụ

ASG₁: ASG \bowtie_{ENO} EMP₁

ASG₂: ASG \bowtie_{ENO} EMP₂

EMP₁: $\sigma_{\text{TITLE}=\text{"Programmer"}}(\text{EMP})$

EMP₂: $\sigma_{\text{TITLE}\neq\text{"Programmer"}}(\text{EMP})$

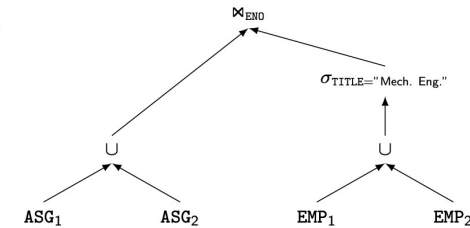
■ Truy vấn

```
SELECT *
FROM EMP NATURAL JOIN ASG
WHERE EMP.TITLE = "Mech. Eng."
```

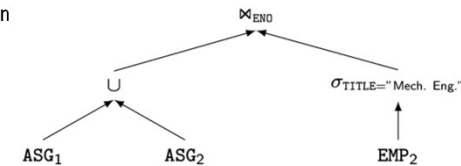
26

Rút gọn phân mảnh ngang dẫn xuất

Truy vấn chung



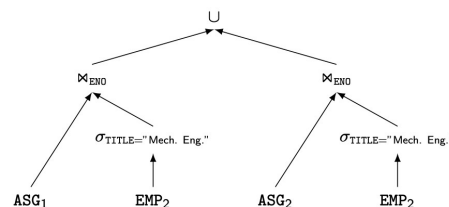
Các lựa chọn
đầu tiên



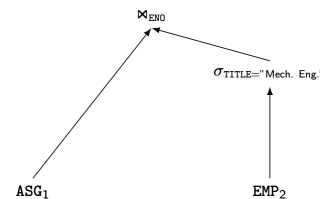
27

Rút gọn phân mảnh ngang dẫn xuất

Phép kết nối trên
phép hợp



Loại bỏ các quan
hệ trung gian rỗng
(cây con bên trái)



28

Rút gọn phân mảnh lại

■ Kết hợp các quy tắc đã được xác định:

- Loại bỏ các quan hệ rỗng được tạo ra bởi các phép chọn mâu thuẫn trên các mảnh ngang;
- Loại bỏ các quan hệ không được sử dụng được tạo ra bởi các phép chiếu trên các mảnh dọc;
- Phân tán các phép kết nối trên các phép hợp để cô lập và loại bỏ các phép kết nối không được sử dụng.

29

Rút gọn phân mảnh lại

Ví dụ:

Xét phân mảnh lại như sau:

$EMP_1 = \sigma_{ENO \leq "E4"} (\Pi_{ENO, ENAME} (EMP))$

$EMP_2 = \sigma_{ENO > "E4"} (\Pi_{ENO, ENAME} (EMP))$

$EMP_3 = \sigma_{ENO, TITLE} (EMP)$

và truy vấn:

```
SELECT  ENAME
FROM    EMP
WHERE   ENO = "E5"
```

