

Tối ưu hóa câu hỏi

1

Tối ưu hóa câu hỏi

- Biến đổi biểu thức ĐSQH để tìm 1 biểu thức hiệu quả
- Tối ưu dựa trên cấu trúc và nội dung của dữ liệu
- Nâng cao hiệu quả thực hiện câu hỏi trên 1 hay nhiều tiêu chí: thời gian, sử dụng bộ nhớ, ...
- Lưu ý:
 - Không nhất thiết phải tìm biểu thức tối ưu nhất
 - Chú ý tới tài nguyên sử dụng cho tối ưu
- Mục đích của các kỹ thuật tối ưu
 - Giảm số bản ghi
 - Giảm kích thước bản ghi

2

Nguyên tắc tối ưu hóa câu hỏi

- Sáu chiến lược tổng quan của J. D. Ullman
 - 1. Thực hiện phép chọn càng sớm càng tốt
 - 2. Tổ hợp những phép chọn xác định với phép tích Đề-các thành phép kết nối
 - 3. Tổ hợp dãy các phép toán quan hệ một ngôi như các phép chọn và phép chiếu
 - 4. Tìm các biểu thức con chung trong một biểu thức
 - 5. Tiền xử lý các quan hệ / bảng (Table Preprocessing)
 - 6. Đánh giá trước khi thực hiện tính toán

3

Biểu thức tương đương

- Sử dụng các phép biến đổi tương đương để tìm ra biểu thức ĐSQH tốt
- Biểu thức trong ngôn ngữ ĐSQH có các hạng thức là biến quan hệ R_1, \dots, R_n ; các quan hệ hằng, được xác định như là một ánh xạ từ các k-bộ của các quan hệ (r_1, \dots, r_k) trong đó r_i là quan hệ trên lược đồ R_i và thay thế r_i vào R_i khi đánh giá biểu thức.
- Hai biểu thức E_1 và E_2 được gọi là tương đương (Equivalent), viết tắt là $E_1 \equiv E_2$, nếu chúng biểu diễn cùng một ánh xạ, nghĩa là, nếu chúng ta thay thế cùng các quan hệ cho tên các lược đồ tương ứng ở hai biểu thức E_1 và E_2 , thì chúng sẽ cho ra cùng một kết quả.

4

Quy tắc biến đổi tương đương

- 1. Quy tắc giao hoán của phép kết nối và tích Đề-các
 - E_1, E_2 là các biểu thức quan hệ
 - $E_1 \bowtie E_2 \equiv E_2 \bowtie E_1$ // Tính giao hoán của kết nối
 - $E_1 * E_2 \equiv E_1 * E_2$ // Tính giao hoán của kết bằng
 - $E_1 \times E_2 \equiv E_1 \times E_2$ // Tính giao hoán của tích Đề-các.
- 2. Quy tắc kết hợp của phép kết nối và tích Đề-các
 - Nếu E_1, E_2 và E_3 là các biểu thức quan hệ: F_1, F_2 là điều kiện thì:
 - $(E_1 \bowtie E_2) \bowtie E_3 \equiv E_1 \bowtie (E_2 \bowtie E_3)$
 - $(E_1 * E_2) * E_3 \equiv E_1 * (E_2 * E_3)$
 - $(E_1 \times E_2) \times E_3 \equiv E_1 \times (E_2 \times E_3)$

5

Quy tắc biến đổi tương đương

3. Dãy các phép chọn

$$\sigma_{F_1}(\sigma_{F_2}(E)) = \sigma_{F_1 \wedge F_2}(E)$$

4. Dãy các phép chiếu

$$\pi_x(\pi_y(E)) = \pi_{x,y}(E)$$

6

Quy tắc biến đổi tương đương

- 5. Giao hoán phép chọn và phép chiếu

$$\sigma_F(\pi_X(E)) = \pi_X(\sigma_F(E))$$

- 6. Giao hoán phép chọn và tích Đề-các

$$\sigma_F(E1 \times E2) = \sigma_F(E1) \times E2$$

- 7. Giao hoán phép chọn và một phép hợp

$$\sigma_F(E1 \cup E2) = \sigma_F(E1) \cup \sigma_F(E2)$$

7

Quy tắc biến đổi tương đương

- 8. Phân phối phép chọn và một phép hiệu tập hợp

$$\sigma_F(E1 - E2) = \sigma_F(E1) - \sigma_F(E2)$$

- 9. Giao hoán một phép chiếu với một phép tích Đề-các

$$\pi_X(E1 \times E2) = \pi_Y(E1) \times \pi_Z(E2), X = YZ$$

- 10. Phân phối một phép chiếu với một phép hợp

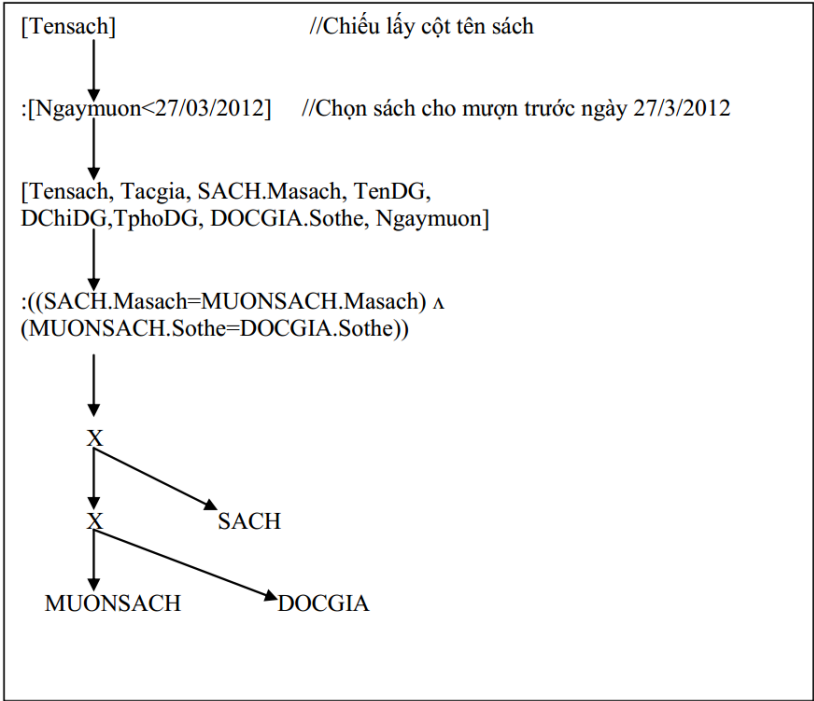
$$\pi_X(E1 \cup E2) = \pi_X(E1) \cup \pi_X(E2)$$

8

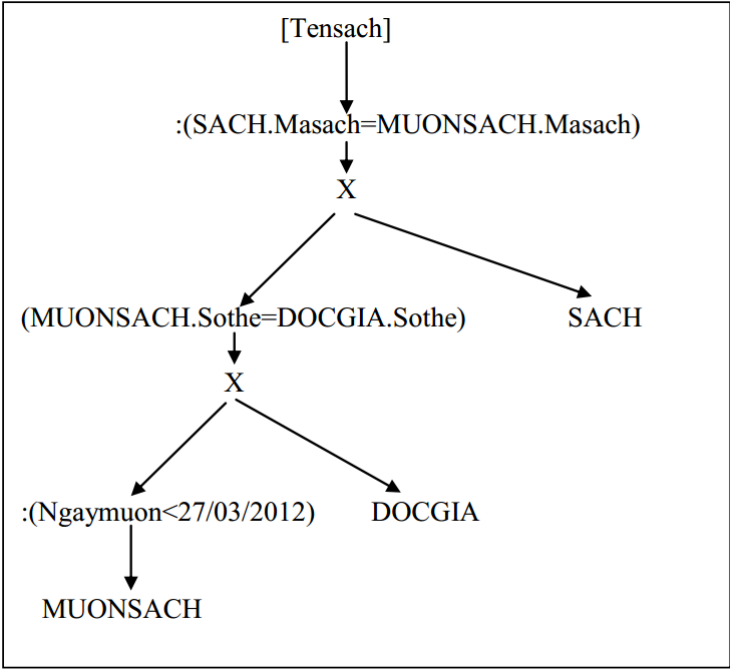
Ví dụ

- Hãy xét một CSDL quản lý thư viện bao gồm các quan hệ sau đây:
- 1. SACH (Tensach, Tacgia, NhaXB, Masach).
- 2. NHAXUATBAN (NhaXB, Diachi, Thanhpho)
- 3. DOCGIA (TenDG, DchiDG, TphoDG, Sothe)
- 4. MUONSACH (Sothe, Masach, Ngaymuon)
- **Câu hỏi:** Cho danh sách những cuốn sách đã cho mượn trước ngày 27/03/2012.
- $Q1 \leftarrow \pi_{Masach} (\sigma_{Ngaymuon < '27/03/2012'} (MUONSACH))$
- $Q2 \leftarrow SACH * Q1$
- $Q3 \leftarrow \pi_{Tensach, Tacgia, NhaXB} (Q2)$

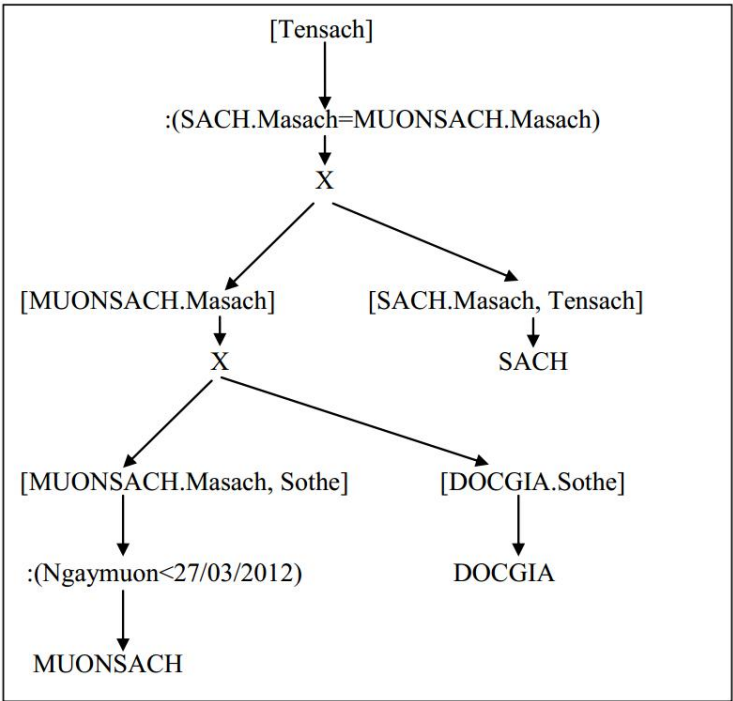
9



10



11



12