

Cơ Sở Dữ Liệu

Chương 6. Các nguyên tắc thiết kế lược đồ quan hệ và phụ thuộc hàm

Trần Hoài Thuận

Ngày 06 tháng 08 năm 2023



DEPARTMENT OF
INFORMATION TECHNOLOGY
GIA DINH UNIVERSITY



GIA DINH
UNIVERSITY



Nội Dung

1. Thế nào là một lượt đồ quan hệ tốt?
2. Các nguyên tắc thiết kế lược đồ quan hệ
3. Phụ thuộc hàm



Thế nào là một lượt
đồ quan hệ tốt?

Thế nào là một lượt đồ quan hệ tốt?

- Tính tốt của một lượt đồ quan hệ định nghĩa ở 2 mức: mức thứ nhất là mức lôgic, mức thứ hai là mức cài đặt.
 - Mức thứ nhất liên quan đến việc các người sử dụng thể hiện các lượt đồ quan hệ và ý nghĩa của các thuộc tính của chúng như thế nào.
 - Mức thứ hai liên quan đến việc các bộ trong một quan hệ cơ sở được lưu trữ và cập nhật như thế nào.



Các giải pháp thiết kế cơ sở dữ liệu

Có hai giải pháp chính:

- Phương pháp thiết kế dưới lên:
 - Điểm xuất phát là quan hệ giữa các thuộc tính riêng rẽ.
 - Sử dụng quan hệ giữa các thuộc tính này để xây dựng các lược đồ quan hệ
 - Đây là phương pháp thiết kế bằng tổng hợp (design by synthesis)
- Phương pháp thiết kế trên xuống, còn gọi là thiết kế bằng phân tích (design by analyse)
 - $\exists [-]$ Bắt đầu từ một số các nhóm thuộc tính trong các quan hệ nhận được từ thiết kế quan niệm.
 - Tách các quan hệ cho đến khi đạt được tính chất mong muốn.





Các nguyên tắc thiết
kế lược đồ quan hệ



Các nguyên tắc thiết kế lược đồ quan hệ

1. Ngữ nghĩa của các thuộc tính quan hệ

Thiết kế một lược đồ quan hệ sao cho dễ giải thích ý nghĩa của nó. Không nên tổ hợp các thuộc tính từ nhiều kiểu thực thể và kiểu liên kết vào một quan hệ.

2. Thông tin dư thừa trong các bộ và sự dị thường cập nhật

Thiết kế các lược đồ quan hệ cơ sở sao cho không sinh ra những dị thường cập nhật trong các quan hệ.

3. Các giá trị không xác định trong các bộ

Tránh việc đặt vào trong các quan hệ cơ sở những thuộc tính mà các giá trị của chúng thường xuyên là null.

4. Sinh ra các bộ giả

Thiết kế các lược đồ quan hệ sao cho chúng có thể được nối với điều kiện bằng trên các thuộc tính là khoá chính hoặc khoá ngoại theo cách đảm bảo không sinh ra các bộ “giả”.



NT1: Các lượt đồ quan hệ cần có ý nghĩa

- Ngữ nghĩa của lược đồ rõ ràng thì dễ thiết kế được lược đồ quan hệ tốt.
- Không tổ hợp các thuộc tính từ nhiều kiểu thực thể và kiểu liên kết vào một quan hệ.
- Nếu một lược đồ quan hệ tương ứng với một kiểu thực thể hoặc một kiểu liên kết thì ý nghĩa của lược đồ đó và các thuộc tính trở nên rõ ràng.



NT2: Tránh các dị thường cập nhật

Giả sử ta có quan hệ NHANVIEN_DONVI như sau:

MaNV	Hodem	Ten	Ngaysinh	Diachi	MaDV	TenDV	MaNQL
NV001	Lê	Vân	1979-02-12	Hà Nội	5	Nghiên cứu	NV002
NV002	Trần Đức	Nam	1976-02-14	Hà Nội	5	Nghiên cứu	NV002
NV010	Hoàng	Thanh	1979-08-05	Nghệ An	4	Hành chính	NV014
NV014	Phạm	Bằng	1952-06-26	Bắc Ninh	4	Hành chính	NV014
NV016	Nguyễn	Sơn	1973-08-14	Hà Nam	5	Nghiên cứu	NV002
NV018	Vũ Hương	Giang	1983-03-26	Nam Định	5	Nghiên cứu	NV002
NV025	Trần Lê	Hoa	1980-03-15	Phú Thọ	4	Hành chính	NV014
NV061	Hoàng	Giang	1967-05-02	Hà Tĩnh	1	Lãnh đạo	NV061

- Có dư thừa thông tin → lãng phí bộ nhớ
- Có thể xảy ra các dị thường cập nhật:
 - Dị thường chèn: Chèn một nhân viên mới chưa làm việc cho đơn vị nào, hoặc chèn một đơn vị vừa thành lập chưa có nhân viên → phải chèn nhiều giá trị null vào quan hệ.
 - Dị thường xóa: Xóa thông tin một nhân viên duy nhất của đơn vị → xóa đơn vị.
 - Dị thường sửa đổi: Đơn vị đổi tên hoặc đổi người quản lý → phải sửa nhiều bộ trong quan hệ.

NT3: Tránh các giá trị null

Nếu một quan hệ chứa quá nhiều giá trị null:

- Tốn không gian lưu trữ
- Kết quả của phép nối bị giảm ý nghĩa
- Các hàm như Count, Sum,... không tính toán được trên giá trị null.



NT4: Tránh sinh các bộ giả

Giả sử chúng ta có hai quan hệ:

NHANVIEN_DIADIEM

Ten	DiadiemDA
Vân	Hà Nội
Vân	Nam Định
Sơn	Bắc Ninh
Giang	Hà Nội

NHANVIEN_DUAN

MaNV	MaDA	Sogio	TenDA	DiadiemDA
NV001	1	32	DA01	Hà Nội
NV001	2	7	DA02	Nam Định
NV016	3	40	DA03	Bắc Ninh
NV018	1	20	DA01	Hà Nội



NT4: Tránh sinh các bộ giả

Nối tự nhiên hai quan hệ trên với nhau, ta có quan hệ:

DiadiemDA	MaNV	MaDA	Sogio	TenDA	Ten
Bắc Ninh	NV016	3	40	DA03	Sơn
Hà Nội	NV001	1	32	DA01	Vân
Hà Nội	NV001	1	32	DA01	Giang
Hà Nội	NV018	1	20	DA01	Vân
Hà Nội	NV018	1	20	DA01	Giang
Nam Định	NV001	2	7	DA02	Vân

- Bộ thứ 3 là bộ giả: Nhân viên có mã số NV001 làm việc ở 2 dự án DA01 và DA02 ở Hà Nội và Nam Định → tên Vân
- Bộ thứ 4 là bộ giả: Nhân viên có mã số NV018 làm việc ở dự án DA01 ở Hà Nội → tên Giang





Phụ thuộc hàm

Định nghĩa phụ thuộc hàm

- Giả sử X và Y là hai tập con của tập thuộc tính R.
- **Một phụ thuộc hàm từ X vào Y** là một ràng buộc trên các bộ quan hệ $r(R)$ sao cho với hai bộ bất kỳ $t_1, t_2 \in r(R)$ nếu $t_1[X] = t_2[X]$ thì $t_1[Y] = t_2[Y]$.
- Phụ thuộc hàm từ X vào Y ký hiệu là $X \rightarrow Y$ với X là vế trái, Y là vế phải của phụ thuộc hàm.
- Các cách diễn đạt khác: **Y phụ thuộc hàm vào X** hoặc **X xác định hàm Y**.
- Phụ thuộc hàm được viết tắt là FD (functional dependency).



Định nghĩa phụ thuộc hàm

- Để biểu diễn các phụ thuộc hàm trong một lược đồ quan hệ, chúng ta sử dụng khái niệm sơ đồ phụ thuộc hàm. Mỗi FD được biểu diễn bằng một đường nằm ngang. Các thuộc tính ở về trái của FD được nối với đường biểu diễn FD bằng các đường thẳng đứng, các thuộc tính ở về phải của FD được nối với đường biểu diễn FD bằng mũi tên chỉ đến các thuộc tính.
- Ví dụ: Ta có lược đồ quan hệ:

MUQN(Sôthẻ, Mãsốsách, Tênngườiimượn, Tênsách, Ngàymượn)

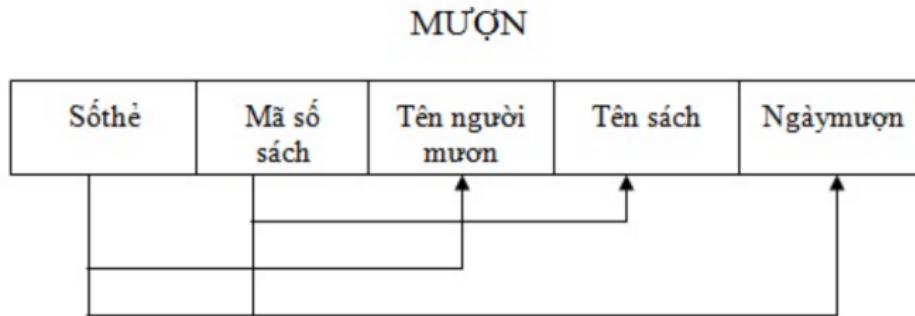
Với các phụ thuộc hàm:

- Sôthẻ → Tênngườiimượn
- Mãsốsách → Tênsách
- Sôthẻ, Mãsốsách → Ngàymượn



Định nghĩa phụ thuộc hàm

Có sơ đồ phụ thuộc hàm như sau:



Hình: Sơ đồ phụ thuộc của quan hệ MUỢN



Phụ thuộc hàm suy diễn được

- Giả sử F là tập các phụ thuộc hàm trên một lược đồ quan hệ R .
- Một phụ thuộc hàm $X \rightarrow Y$ được gọi là **suy diễn được** từ F nếu $X \rightarrow Y$ đúng trong mỗi trạng thái hợp lệ $r(R)$. Điều này có nghĩa là khi $r(R)$ thỏa mãn các phụ thuộc hàm trong F , $r(R)$ cũng thỏa mãn $X \rightarrow Y$.
- Ta sử dụng ký hiệu $F \vDash X \rightarrow Y$ để ký hiệu phụ thuộc hàm $X \rightarrow Y$ được suy diễn từ tập các phụ thuộc hàm F .
- Tập hợp tất cả các phụ thuộc hàm suy diễn được từ F được gọi là bao đóng của F và được ký hiệu là F^+ .

$$F^+ = F \cup \{X \rightarrow Y, F \vDash X \rightarrow Y\}$$

- Để cho tiện, ta viết tắt phụ thuộc hàm có dạng $\{X, Y\} \rightarrow Z$ thành $XY \rightarrow Z$



Các q.tắc suy diễn d.với phụ thuộc hàm

Có 6 quy tắc suy diễn đối với các phụ thuộc hàm:

QT1. (quy tắc phản xạ) : Nếu $X \supseteq Y$ thì $X \rightarrow Y$

QT2. (quy tắc tăng) : $\{ X \rightarrow Y \} \vDash XZ \rightarrow YZ$

QT3. (quy tắc bắc cầu) : $\{ X \rightarrow Y, Y \rightarrow Z \} \vDash X \rightarrow Y$

QT4. (quy tắc chiểu) : $\{ X \rightarrow YZ \} \vDash X \rightarrow Y$ và $X \rightarrow Z$

QT5. (quy tắc hợp) : $\{ X \rightarrow Y, X \rightarrow Z \} \vDash X \rightarrow YZ$

QT6. (quy tắc tựa bắc cầu) : $\{ X \rightarrow Y, WY \rightarrow Z \} \vDash WX \rightarrow Z$



Bao đóng của tập thuộc tính

- Giả sử F là một tập phụ thuộc hàm trên R và X là một tập thuộc tính của R.
- **Bao đóng của tập thuộc tính X dưới F**, ký hiệu là X^+ được định nghĩa như sau:

$$X^+ = \{A, F \vDash X \rightarrow A\}$$

với A là thuộc tính của R.

- Khi cần chỉ rõ tập phụ thuộc hàm, chúng ta ký hiệu bao đóng của X dưới F là X_F^+ .



Tìm bao đóng của tập thuộc tính

Thuật toán 1: Tìm bao đóng X^+ của X dưới \mathcal{F}

Vào: Lược đồ quan hệ R , tập phụ thuộc hàm \mathcal{F} và tập thuộc tính X

Ra: Tập thuộc tính X^+ là bao đóng của X

```
1  $X^+ = X;$ 
2 repeat
3    $OldX^+ = X^+;$ 
4   for mỗi phụ thuộc hàm  $Y \rightarrow Z$  trong  $\mathcal{F}$  do
5     if  $X^+ \supset Y$  then
6       |  $X^+ = X^+ \cup Z;$ 
7     end
8   end
9 until  $OldX^+ = X^+$ ;
```

Hình: Thuật toán tìm bao đóng



Tìm bao đóng của tập thuộc tính

- Ví dụ: Xét lược đồ quan hệ:

$R = \{Mã số NV, Họ tên NV, Mã số DA, Tên DA, Địa điểm DA, Số giờ\}$

và tập phụ thuộc hàm: $F = \{Mã số NV \rightarrow Họ tên NV, Mã số DA \rightarrow Tên DA, Địa điểm DA, \{Mã số NV, Mã số DA\} \rightarrow Số giờ\}$

- Áp dụng thuật toán 1 ta có:

$$\{Mã số NV\}^+ = \{Mã số NV, Họ tên NV\}$$

$$\{Mã số DA\}^+ = \{Mã số DA, Tên DA, Địa điểm DA\}$$

$$\{Mã số NV, Mã số DA\}^+ = \{Mã số NV, Họ tên NV, Mã số DA, Tên DA, Địa điểm DA, Số giờ\}$$



Bao đóng của tập thuộc tính và khóa

- Giả sử ta có lượt đồ quan hệ $R(A_1, A_2, \dots, A_n)$
- Nếu $X^+ = \{ A_1, A_2, \dots, A_n \}$ thì X xác định hàm các thuộc tính còn lại, điều này tương đương với X là siêu khóa.
- Có thể kiểm tra xem một tập thuộc tính X có phải là khóa hay không bằng cách:
 - Kiểm tra X có phải là siêu khóa hay không: $X^+ = \{ A_1, A_2, \dots, A_n \} ?$
 - Nếu có, kiểm tra X có phải là siêu khóa tối thiểu hay không: có tồn tại tập thuộc tính $S \subset X$ sao cho $S^+ = \{ A_1, A_2, \dots, A_n \} ?$



Tìm khóa của lược đồ quan hệ

Thuật toán 2: Tìm một khóa của lược đồ quan hệ

Vào: Lược đồ quan hệ $R(A_1, A_2, \dots, A_n)$ và tập phụ thuộc hàm \mathcal{F}

Ra: Một khóa của lược đồ quan hệ R

```
1  $K = \{A_1, A_2, \dots, A_n\};$ 
2 for mỗi thuộc tính  $A$  của  $K$  do
3   | Xác định  $(K - A)_{\mathcal{F}}^+$ ;           // Thực hiện thuật toán 1
4   | if  $(K - A)_{\mathcal{F}}^+ = \{A_1, A_2, \dots, A_n\}$  then
5     |   |  $K = K - \{A\}$ 
6   | end
7 end
```

Hình: Thuật toán tìm khóa



Tìm khóa của lược đồ quan hệ

Ví dụ: xét lược đồ quan hệ $R(A,B,C,D,E,F)$ và tập phụ thuộc hàm
 $F = \{ A,B \rightarrow F ; A \rightarrow C,D; B \rightarrow E \}$

Ta có $\{A,B\}^+ = \{A,B,C,D,E,F\}$, $A^+ = \{A,C,D\}$, $B^+ = \{B,E\}$, vậy
AB là khóa của quan hệ.



Cơ Sở Dữ Liệu

Chương 6. Các nguyên tắc thiết kế lược đồ quan hệ và phụ thuộc hàm

Trần Hoài Thuận

Ngày 06 tháng 08 năm 2023



GIA DINH
UNIVERSITY



DEPARTMENT OF
INFORMATION TECHNOLOGY
GIA DINH UNIVERSITY