

Cơ Sở Dữ Liệu

Chương 7. Chuẩn hóa lược đồ quan hệ

Trần Hoài Thuận

Ngày 06 tháng 08 năm 2023



GIA DINH
UNIVERSITY



Nội Dung

1. Giới thiệu về chuẩn hóa

2. Các dạng chuẩn dựa trên khóa chính

- Dạng chuẩn 1 (1NF)
- Dạng chuẩn 2 (2NF)
- Dạng chuẩn 3 (3NF)
- Dạng chuẩn Boyce-Codd (BCNF)



Giới thiệu về chuẩn hóa

Giới thiệu về chuẩn hóa

- Quá trình chuẩn hóa (Codd đề nghị năm 1972): lấy một lược đồ quan hệ và thực hiện các phép kiểm tra để xác nhận lược đồ có thỏa mãn một số điều kiện nhất định (dạng chuẩn) hay không; nếu không các lược đồ quan hệ sẽ được tách ra.
- Codd đề xuất ba dạng chuẩn 1, 2 và 3 (1NF, 2NF, 3NF).
- Sau đó Boyce và Codd đề nghị dạng chuẩn Boyce-Codd (BCNF) mạnh hơn 3NF.
- Các dạng chuẩn này dựa trên các phụ thuộc hàm giữa các thuộc tính của một lược đồ quan hệ.
- Sau đó dạng chuẩn 4 (4NF) và 5 (5NF) được đề xuất dựa trên phụ thuộc hàm đa trị và phụ thuộc hàm nối.



Giới thiệu về chuẩn hóa

- Chuẩn hóa dữ liệu có thể được xem là quá trình phân tích và tách các lược đồ quan hệ để đạt được hai mục tiêu:
 - Cực tiểu hóa sự dư thừa
 - Cực tiểu hóa các phép cập nhật bất thường
- Các lược đồ quan hệ không thỏa mãn các kiểm tra dạng chuẩn sẽ được tách ra thành các lược đồ quan hệ nhỏ hơn thỏa mãn các kiểm tra và có các tính chất mong muốn.
- Phép tách các lược đồ quan hệ cần có hai tính chất:
 - Tính chất nối không mất mát (phải đạt được)
 - Tính chất bảo toàn phụ thuộc





Các dạng chuẩn dựa
trên khóa chính

Dạng chuẩn 1 (1NF)

- Một lược đồ quan hệ $R(A_1, A_2, \dots, A_n)$ là ở dạng chuẩn 1 (1NF - first normal form) nếu $\text{Dom}(A_i)$ chỉ chứa các giá trị nguyên tố (không phân chia được) $\forall i$ và
- $\forall t = < t_1, t_2, \dots, t_n > \in r(R), t_i \in \text{Dom}(A_i)$, t_i nhận giá trị đơn.
- Như vậy 1NF không cho phép có các thuộc tính đa trị hoặc các nhóm thuộc tính lặp.



Thuộc tính đa trị và nhóm thuộc tính lặp

Ví dụ: xét quan hệ DONVI và quan hệ NHANVIEN_DUAN sau:

MaDV	TenDV	MaNQL	Diadiem
5	Nghiên cứu	NV002	Nam Định,Hà Nội,Bắc Ninh
4	Hành chính	NV014	Hà Nội
1	Tài vụ	NV061	Hà Nội

MaDA	TenDA	TenNV	Sogio
1	DA01	Vân,Nam	15,20
2	DA02	Nam,Thanh,Bằng	10,12,28
3	DA03	Thanh	20

Hình: Quan hệ chứa thuộc tính đa trị



Chuẩn hóa 1NF

- Giả sử K là khóa của lược đồ quan hệ R và A là thuộc tính đa trị hoặc là nhóm thuộc tính lặp.
- Để chuẩn hóa R về 1NF:
 - Loại A ra khỏi R
 - Tạo một lược đồ quan hệ mới S(K, A) với khóa chính của S là:
 - { A, K } nếu A là thuộc tính đa trị
 - { K, KA } nếu A là nhóm thuộc tính lặp, KA là khóa bộ phận của A



Ví dụ chuẩn hóa 1NF

Áp dụng: Lược đồ quan hệ DONVI ở trên sẽ được tách thành hai:

- DONVI (MADV, TENDV, MANQL)
- DONVI_ĐIADIỂM (MADV, Địađiểm)

Lược đồ quan hệ NHANVIEN_DUAN cũng được tách thành hai:

- DUAN (MADA, TENDA)
- NHANVIEN_DUAN(MADA, TENNV, SOGIO)



Dạng chuẩn 2 (2NF)

- Một phụ thuộc hàm $X \rightarrow Y$ là một **phụ thuộc hàm đầy đủ** nếu bỏ đi bất kỳ một thuộc tính $A \in X$ thì phụ thuộc hàm đó không còn đúng nữa.
- Nếu tồn tại một thuộc tính B sao cho $(X-B) \rightarrow Y$, chúng ta nói rằng Y **phụ thuộc hàm bộ phận** vào X .
- Một lược đồ quan hệ R là ở dạng chuẩn 2 (2NF - second normal form) nếu:
 - R là 1NF, và
 - Mỗi thuộc tính không khóa A trong R phụ thuộc hàm đầy đủ vào khóa chính của R



Chuẩn hóa 2NF

- Giả sử R là lược đồ quan hệ ở dạng chuẩn 1NF và không là 2NF, có khóa K.
- Để chuẩn hóa R về 2NF, đối với mỗi nhóm thuộc tính không khóa X phụ thuộc hàm bộ phận vào K:
 - Loại X ra khỏi R
 - Gọi khóa bộ phận của K xác định hàm X là $K_X : K_X \rightarrow X$
 - Tạo lược đồ quan hệ mới $S(\underline{K_X}, X)$ với K_X là khóa chính





Ví dụ chuẩn hóa 2NF

Ví dụ, xét lược đồ quan hệ:

NHANVIEN_DUAN(MANV, MADA, SOGIO, TENNV, TENDA,
DIADIEM)

$\{MANV, MADA\} \rightarrow \{SOGIO, TENNV, TENDA, DIADIEM\}$ là
phụ thuộc hàm đầy đủ.

MANV \rightarrow TENNV là phụ thuộc hàm bộ phận.

MADA \rightarrow {TENDA, DIADIEM} là phụ thuộc hàm bộ phận.

- NHANVIEN_DUAN không là 2NF do TENNV, TENDA và
DIADIEM phụ thuộc hàm bộ phận vào khóa chính {MANV,
MADA}
- Chuẩn hóa về 2NF:
 - Loại TENNV, TENDA, DIADIEM ra khỏi
NHANVIEN_DUAN(MANV, MADA, SOGIO)
 - Tạo lược đồ quan hệ mới 1: N1(MANV, TENNV)
 - Tạo lược đồ quan hệ mới 2: N2(MADA, TENDA, DIADIEM)

Dạng chuẩn 3 (3NF)

- Một phụ thuộc hàm $X \rightarrow Y$ trong một lược đồ quan hệ R là một **phụ thuộc hàm bắc cầu** nếu tồn tại một tập hợp thuộc tính Z của R không phải là khóa và không phải là tập con của khóa sao cho $X \rightarrow Z$ và $Z \rightarrow Y$ đều đúng.
- Một lược đồ quan hệ R là ở dạng chuẩn 3 (3NF - third normal form) nếu:
 - R là 2NF, và
 - Không có thuộc tính không khóa nào của R phụ thuộc hàm bắc cầu vào khóa chính



Chuẩn hóa 3NF

- Giả sử R là lược đồ quan hệ có khóa chính K; R ở dạng chuẩn 2NF và không ở dạng chuẩn 3NF.
- Để chuẩn hóa R về 3NF, với mỗi thuộc tính X phụ thuộc bắc cầu vào K:
 - Loại X ra khỏi R
 - Gọi Y là thuộc tính bắc cầu, ta có: $K \rightarrow Y$ và $Y \rightarrow X$
 - Tạo lược đồ quan hệ mới S(Y, X) với Y là khóa chính



Ví dụ chuẩn hóa 3NF

Ví dụ: Xét lược đồ quan hệ

NHANVIEN_DONVI(MANV, TENNV, NGAYSINH, DIACHI,
MADV, TENDV, MANQL)

Với các phụ thuộc hàm:

$\text{MANV} \rightarrow \{\text{TENVN}, \text{NGAYSINH}, \text{DIACHI}, \text{MADV}, \text{TENDV}, \text{MANQL}\}$

$\text{MADV} \rightarrow \{\text{TENDV}, \text{MANQL}\}$

- Các thuộc tính TENDV, MANQL phụ thuộc bắc cầu vào khóa chính, do đó lược đồ quan hệ không thoả mãn điều kiện 3NF.
- Chuẩn hóa về 3NF:
 - Loại TENDV, MANQL ra khỏi NHANVIEN_DONVI(TENVN, MANV, NGAYSINH, DIACHI, MADV)
 - Tạo lược đồ quan hệ mới 1: N1(MADV, TENDV, MANQL)



Dạng chuẩn Boyce-Codd (BCNF)

- Một lược đồ quan hệ R là ở dạng chuẩn Boyce-Codd (BCNF - Boyce-Codd normal form) nếu:
 - R là 3NF, và
 - Không có thuộc tính khóa phụ thuộc hàm vào thuộc tính không khóa trong R



Chuẩn hóa BCNF

- Giả sử R là lược đồ quan hệ có khóa K , R là 3NF và không là BCNF.
- Để chuẩn hóa R về BCNF, đối với mỗi thuộc tính không khóa A xác định hàm thuộc tính khóa $A \rightarrow K_S$, với $K_S \subset K$:
 - Loại K_S ra khỏi R và bổ sung A vào khóa chính của R. Khóa chính mới của R là: $K - K_S \cup \{A\}$
 - Tạo lược đồ mới $S(\underline{A}, K_S)$ với A là khóa chính





Ví dụ chuẩn hóa BCNF

Ví dụ: Xét lược đồ quan hệ

R(A₁, A₂, A₃, A₄, A₅, A₆) với các phụ thuộc hàm:

$$\{A_1, A_2\} \rightarrow \{A_3, A_4, A_5, A_6\}.$$

$$A_4 \rightarrow A_2$$

$$A_6 \rightarrow A_1$$

- R là 2NF (không có phụ thuộc hàm bộ phận) và 3NF (không có phụ thuộc hàm bắc cầu từ khóa chính).
- R không là BCNF vì có thuộc tính khóa phụ thuộc hàm vào thuộc tính không khóa.
- Chuẩn hóa về BCNF:
 - Loại A₂ ra khỏi R và bổ sung A₄ vào khóa chính: R(A₁, A₄, A₃, A₅, A₆)
 - Tạo quan hệ mới S₁(A₄, A₂) với A₄ là khóa chính
 - Loại A₁ ra khỏi R và bổ sung A₆ vào khóa chính: R(A₆, A₄, A₃, A₅)
 - Tạo quan hệ mới S₂(A₆, A₁) với A₆ là khóa chính

Cơ Sở Dữ Liệu

Chương 7. Chuẩn hóa lược đồ quan hệ

Trần Hoài Thuận

Ngày 06 tháng 08 năm 2023



DEPARTMENT OF
INFORMATION TECHNOLOGY
GIA DINH UNIVERSITY



GIA DINH
UNIVERSITY