



BÀI GIẢNG

CƠ SỞ DỮ LIỆU

8. Nguyên tắc thiết kế lược đồ quan hệ

Nguyễn Hải Châu

Khoa Công nghệ Thông tin
Trường Đại học Công nghệ, ĐHQGHN

Các giải pháp thiết kế cơ sở dữ liệu quan hệ

Có hai giải pháp chính:

- 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)
- Thiết kế trên xuống, hay thiết kế bằng phân tích (design by analyse):
 - Xuất phát từ các lược đồ quan hệ có được từ thiết kế quan niệm
 - Tách các quan hệ cho đến khi đạt mục đích mong muốn

Các nguyên tắc thiết kế lược đồ quan hệ

- NT1. (Ngữ nghĩa của các thuộc tính): 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 lược đồ quan hệ
- NT2. (Thông tin dư thừa trong các bộ và 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
- NT3. (Các giá trị không xác định trong các bộ): Tránh sử dụng các thuộc tính của các lược đồ quan hệ mà giá trị của các thuộc tính này thường xuyên là *null*
- NT4. (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 ngoài để đảm bảo không sinh ra các bộ "giả"

NT1: Các lược đồ quan hệ cần có ý nghĩa rõ ràng

- 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 trong 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 trong quan hệ
- 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

Nếu một quan hệ chứa nhiều giá trị *null*

- Tổn không gian lưu trữ
- Kết quả của các phép nối bị giảm ý nghĩa
- Các hàm nhóm 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 |

và *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

Kết quả $NHANVIEN_DIADIEM$ * $NHANVIEN_DUAN$:
DiadiemDA

| 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 là Vân
- Bộ thứ 4 là bộ giả: Nhân viên có mã số NV018 làm việc cho dự án DA01 ở Hà Nội → tên là Giang
- Nên thiết kế các lược đồ để thực hiện các phép nối với điều kiện bằng trên khóa chính và/hoặc khóa ngoài để tránh các bộ giả

Cơ sở lý thuyết cho thiết kế

- Phụ thuộc hàm (functional dependency)
- Chuẩn hóa (normalization)
- Các thuật toán thiết kế cơ sở dữ liệu quan hệ