

BÀI GIẢNG CƠ SỞ DỮ LIỆU

Phần 1: Các khái niệm cơ bản - Mô hình thực thể-liên kết - Mô hình quan hệ - Phụ thuộc hàm

Nguyễn Hải Châu
Khoa Công nghệ Thông tin
Trường Đại học Công nghệ, ĐHQGHN

- [1] Ramez A. Elmasri, Shamkant Navathe, *Fundamentals of Database systems*, 3rd edition, John Wiley & Sons, Inc., 2000.
- [2] Raghu Ramakrishnan, Johannes Gehrke, *Database Management Systems*, 2nd edition
- [3] Nguyễn Tuệ, *Giáo trình cơ sở dữ liệu*, Nhà xuất bản Đại học Quốc gia Hà Nội, 2008.
- [4] J. Date, *An introduction to database systems*, 7th edition, Addison Wesley Longman Inn., 2000
- [5] Hector Garcia-Molina et. al., *Database Systems: The complete book*, Prentice Hall, 2002

Tài liệu tham khảo

Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Phần1: Các khái niệm cơ bản của cơ sở dữ liệu (CSDL):

- Các khái niệm cơ bản
- Mô hình thực thể-liên kết (ER)
- Mô hình quan hệ, đại số quan hệ
- Phụ thuộc hàm, chuẩn hóa và thiết kế cơ sở dữ liệu

Tài liệu tham khảo

Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Khái niệm cơ bản

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Khái niệm cơ bản

Hệ QT CSDL

Mô hình CSDL

Lược đồ CSDL

Con người và CSDL

Ngôn ngữ/Giao diện

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Các khái niệm cơ bản của cơ sở dữ liệu

- **Dữ liệu** là những sự kiện có thể ghi lại được và có ý nghĩa.
- Một **cơ sở dữ liệu** là một tập hợp *dữ liệu có liên quan với nhau*, được lưu trữ trên máy tính, có nhiều người sử dụng và được tổ chức theo một mô hình.
 - ◆ Một CSDL biểu thị một khía cạnh nào đó của thế giới thực. Thông tin được đưa vào trong CSDL tạo thành một không gian CSDL hoặc một “thế giới nhỏ” (miniworld)
 - ◆ Một cơ sở dữ liệu là một tập hợp dữ liệu liên kết với nhau một cách logic và mang một ý nghĩa nào đó
 - ◆ Một cơ sở dữ liệu được thiết kế và được phổ biến cho một mục đích riêng
- Một **hệ quản trị cơ sở dữ liệu** là một tập chương trình giúp cho người sử dụng tạo ra, duy trì và khai thác CSDL.
- Người ta gọi cơ sở dữ liệu và hệ quản trị cơ sở dữ liệu bằng một thuật ngữ chung là **hệ cơ sở dữ liệu**.

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Khái niệm cơ bản

Hệ QT CSDL

Mô hình CSDL

Lược đồ CSDL

Con người và CSDL

Ngôn ngữ/Giao diện

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Chức năng của một hệ quản trị cơ sở dữ liệu

1. Lưu trữ các định nghĩa, các mối liên kết dữ liệu (gọi là siêu dữ liệu) vào một từ điển dữ liệu
2. Tạo ra các cấu trúc phức tạp theo yêu cầu để lưu trữ dữ liệu
3. Biến đổi các dữ liệu được nhập vào để phù hợp với các cấu trúc dữ liệu ở điểm trên
4. Tạo ra một hệ thống bảo mật và áp đặt tính bảo mật và riêng tư trong cơ sở dữ liệu
5. Tạo ra các cấu trúc phức tạp cho phép nhiều người sử dụng truy cập đến dữ liệu
6. Cung cấp các thủ tục sao lưu và phục hồi dữ liệu để đảm bảo sự an toàn và toàn vẹn dữ liệu
7. Xúc tiến và áp đặt các quy tắc an toàn để loại bỏ vấn đề toàn vẹn dữ liệu
8. Cung cấp việc truy cập dữ liệu thông qua một ngôn ngữ truy vấn

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Khái niệm cơ bản

Hệ QT CSDL

Mô hình CSDL

Lược đồ CSDL

Con người và CSDL

Ngôn ngữ/Giao diện

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Mô hình cơ sở dữ liệu

- Một **mô hình cơ sở dữ liệu** là một tập hợp các khái niệm dùng để biểu diễn các cấu trúc của cơ sở dữ liệu
- Các loại mô hình cơ sở dữ liệu
 - ◆ Các **mô hình dữ liệu bậc cao** hoặc **mô hình dữ liệu mức quan niệm** cung cấp các khái niệm gắn liền với cách cảm nhận dữ liệu của nhiều người sử dụng
 - ◆ Các **mô hình dữ liệu bậc thấp** hoặc các **mô hình dữ liệu vật lý** cung cấp các khái niệm mô tả chi tiết về việc dữ liệu được lưu trữ trong máy tính như thế nào
 - ◆ Các **mô hình dữ liệu thể hiện (mô hình dữ liệu mức logic)**, chúng cung cấp những khái niệm mà người sử dụng có thể hiểu được và không xa với cách tổ chức dữ liệu bên trong máy tính
- Các mô hình dữ liệu thể hiện là các mô hình được sử dụng nhiều nhất. Ba mô hình nổi tiếng thuộc loại này là **mô hình quan hệ**, **mô hình mạng** và **mô hình phân cấp**

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Khái niệm cơ bản

Hệ QT CSDL

Mô hình CSDL

Lược đồ CSDL

Con người và CSDL

Ngôn ngữ/Giao diện

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Lược đồ và trạng thái cơ sở dữ liệu

- Trong một mô hình dữ liệu cần phải phân biệt rõ giữa *mô tả của cơ sở dữ liệu* và *bản thân cơ sở dữ liệu*
- Mô tả của một cơ sở dữ liệu được gọi là **lược đồ cơ sở dữ liệu**, nó được xác định rõ trong quá trình thiết kế cơ sở dữ liệu và không bị thay đổi thường xuyên
- Các dữ liệu trong một cơ sở dữ liệu tại một thời điểm cụ thể được gọi là một **trạng thái cơ sở dữ liệu** hoặc là ảnh (snapshot) của cơ sở dữ liệu
- Việc phân biệt giữa lược đồ cơ sở dữ liệu và trạng thái cơ sở dữ liệu là rất quan trọng

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Khái niệm cơ bản

Hệ QT CSDL

Mô hình CSDL

Lược đồ CSDL

Con người và CSDL

Ngôn ngữ/Giao diện

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Con người và cơ sở dữ liệu

- Người quản trị hệ cơ sở dữ liệu (Database Administrator – DBA)
- Người thiết kế cơ sở dữ liệu (Database Designer)
- Người sử dụng (End User)
- Người phân tích hệ thống và lập trình ứng dụng
- Người thiết kế và cài đặt hệ quản trị dữ liệu
- Những người phát triển công cụ
- Các thao tác viên và những người bảo trì

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Khái niệm cơ bản

Hệ QT CSDL

Mô hình CSDL

Lược đồ CSDL

Con người và CSDL

Ngôn ngữ/Giao diện

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Ngôn ngữ và giao diện cơ sở dữ liệu

- Các ngôn ngữ hệ quản trị cơ sở dữ liệu:
 - ◆ **Ngôn ngữ định nghĩa dữ liệu** (data definition language – DDL): Dùng để định nghĩa các lược đồ
 - ◆ **Ngôn ngữ thao tác dữ liệu** (data manipulation language – DML): Dùng để thao tác cơ sở dữ liệu
- Các loại giao diện hệ quản trị cơ sở dữ liệu
 - ◆ Giao diện dựa trên bảng chọn (Menu)
 - ◆ Giao diện dựa trên mẫu biểu
 - ◆ Giao diện đồ họa (Graphic User Interface - GUI)
 - ◆ Giao diện cho người quản trị hệ thống

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Khái niệm cơ bản

Hệ QT CSDL

Mô hình CSDL

Lược đồ CSDL

Con người và CSDL

Ngôn ngữ/Giao diện

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Mô hình thực thể-liên kết và thực thể liên kết mở rộng

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Thiết kế CSDL

Mô hình ER

Kiểu & tập thực thể

Khóa & tập giá trị

Kiểu & tập liên kết

Cấp liên kết...

Các ràng buộc...

Mô hình EER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Sử dụng mô hình quan niệm cho thiết kế cơ sở dữ liệu

- Tập hợp các yêu cầu và phân tích
- Thiết kế quan niệm
- Thiết kế logic hoặc là ánh xạ mô hình dữ liệu
- Thiết kế vật lý

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Thiết kế CSDL

Mô hình ER

Kiểu & tập thực thể

Khóa & tập giá trị

Kiểu & tập liên kết

Cấp liên kết...

Các ràng buộc...

Mô hình EER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Mô hình thực thể-liên kết

- **Mô hình thực thể-liên kết (mô hình ER)** là một mô hình dữ liệu mức quan niệm phổ biến, tập trung vào các cấu trúc dữ liệu và các ràng buộc
- **Thực thể** là một “vật” trong thế giới thực, có sự tồn tại độc lập
- Mỗi một thực thể có các **thuộc tính**, đó là các đặc trưng cụ thể mô tả thực thể đó
 - ◆ **Thuộc tính đơn** là thuộc tính không thể phân chia ra được thành các thành phần nhỏ hơn
 - ◆ **Thuộc tính phức hợp** là thuộc tính có thể phân chia được thành các thành phần nhỏ hơn, biểu diễn các thuộc tính cơ bản hơn với các ý nghĩa độc lập
 - ◆ Những thuộc tính có giá trị duy nhất cho một thực thể cụ thể gọi là các **thuộc tính đơn trị**
 - ◆ Một thuộc tính có thể có một tập giá trị cho cùng một thực thể: Đó là **thuộc tính đa trị**

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Thiết kế CSDL

Mô hình ER

Kiểu & tập thực thể

Khóa & tập giá trị

Kiểu & tập liên kết

Cấp liên kết...

Các ràng buộc...

Mô hình EER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Mô hình thực thể-liên kết (tiếp)

- **Thuộc tính được lưu trữ** là các thuộc tính mà giá trị của nó được nhập vào khi cài đặt cơ sở dữ liệu
- **Thực thể** là một “vật” trong thế giới thực, có sự tồn tại độc lập
- Thuộc tính mà giá trị của nó có thể tính được thông qua giá trị của các thuộc tính khác gọi là **thuộc tính suy diễn được**
- Trong một số trường hợp, một thực thể cụ thể có thể không có các giá trị áp dụng được cho một thuộc tính. Trong trường hợp như vậy, ta phải tạo ra một giá trị đặc biệt gọi là giá trị không xác định (*null*): **Thuộc tính null**
- **Thuộc tính phức tạp**: Là sự kết hợp của các thuộc tính phức hợp và đa trị

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Thiết kế CSDL

Mô hình ER

Kiểu & tập thực thể

Khóa & tập giá trị

Kiểu & tập liên kết

Cấp liên kết...

Các ràng buộc...

Mô hình EER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Kiểu thực thể và tập thực thể

- Một **kiểu thực thể** là một tập hợp các thực thể có các thuộc tính như nhau được mô tả bằng tên và các thuộc tính. Ví dụ: NHÂNVIÊN (Họ tên, Tuổi, Lương) là một kiểu thực thể
- Một kiểu thực thể mô tả một lược đồ cho một tập các thực thể có cùng một cấu trúc
- Một kiểu thực thể được biểu diễn trong lược đồ ER bởi một hộp hình chữ nhật có chứa tên kiểu thực thể
- Một tập hợp các thực thể của một kiểu thực thể cụ thể trong cơ sở dữ liệu tại một thời điểm được gọi là một **tập thực thể**

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Thiết kế CSDL

Mô hình ER

Kiểu & tập thực thể

Khóa & tập giá trị

Kiểu & tập liên kết

Cấp liên kết...

Các ràng buộc...

Mô hình EER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

- Thuộc tính mà các giá trị của nó là khác nhau đối với mỗi thực thể riêng biệt trong một tập thực thể gọi là **thuộc tính khóa**
- Nhiều thuộc tính kết hợp với nhau tạo thành một khóa: ta có **khóa phức hợp**. Khóa phức hợp phải *tối thiểu*
- Khóa dùng để phân biệt hai thực thể
- Một kiểu thực thể có thể có nhiều hơn một khóa
- Kiểu thực thể không có khóa gọi là **kiểu thực thể yếu**
- Mỗi thuộc tính đơn của một kiểu thực thể được kết hợp với một **miền giá trị**
- Một thuộc tính A của kiểu thực thể E có tập giá trị V có thể được định nghĩa là một hàm từ E vào tập hợp lực lượng $P(V)$ của V : $A : E \rightarrow P(V)$
- Ký hiệu giá trị của thuộc tính A với thực thể e là $A(e)$
 - ◆ Nếu A là thuộc tính phức hợp, tập giá trị V là tích Đề-các $P(V_1) \times P(V_2) \times \dots \times P(V_n)$, với V_1, V_2, \dots, V_n là tập các giá trị cho các thành phần đơn của A .

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Thiết kế CSDL

Mô hình ER

Kiểu & tập thực thể

Khóa & tập giá trị

Kiểu & tập liên kết

Cấp liên kết...

Các ràng buộc...

Mô hình EER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Kiểu liên kết, tập liên kết và các thể hiện

- Một **kiểu liên kết** R là một quan hệ toán học trên E_1, E_2, \dots, E_n hoặc có thể định nghĩa như là một tập con của tích Đề-các $E_1 \times E_2 \times \dots \times E_n$
- Một kiểu liên kết R giữa n kiểu thực thể E_1, E_2, \dots, E_n xác định một **tập liên kết** giữa các thực thể của các kiểu đó
- **Tập liên kết** R là một tập hợp các thể hiện liên kết $r_i, i = 1, 2, \dots$ trong đó mỗi r_i liên kết n thực thể riêng biệt e_1, e_2, \dots, e_n và mỗi một thực thể e_j trong r_i là một thành phần của kiểu thực thể $E_j, 1 \leq j \leq n$
- Cũng như các kiểu thực thể và tập thực thể, một kiểu liên kết và tập liên kết tương ứng cũng có tên chung là R
- Mỗi kiểu thực thể E_1, E_2, \dots, E_n được gọi là tham gia vào kiểu liên kết R , và tương tự, mỗi thực thể riêng biệt e_1, e_2, \dots, e_n được gọi là tham gia vào thể hiện liên kết $r_i = (e_1, e_2, \dots, e_n)$

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Thiết kế CSDL

Mô hình ER

Kiểu & tập thực thể

Khóa & tập giá trị

Kiểu & tập liên kết

Cấp liên kết...

Các ràng buộc...

Mô hình EER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Cấp liên kết, tên vai trò và kiểu liên kết đệ quy

- **Cấp của một kiểu liên kết** là số các kiểu thực thể tham gia vào kiểu liên kết đó.
- **Tên vai trò** dùng để chỉ rõ vai trò của các thực thể của kiểu thực thể tham gia liên kết, nó giúp đỡ việc giải thích ý nghĩa của liên kết
- Trong một số trường hợp, một kiểu thực thể có thể tham gia vào một kiểu liên kết với các vai trò khác nhau. Trong những trường hợp như vậy, tên vai trò trở nên cần thiết để phân biệt ý nghĩa của mỗi sự tham gia. Các kiểu liên kết như vậy gọi là **kiểu liên kết đệ quy**

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Thiết kế CSDL

Mô hình ER

Kiểu & tập thực thể

Khóa & tập giá trị

Kiểu & tập liên kết

Cấp liên kết...

Các ràng buộc...

Mô hình EER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Các ràng buộc trên các kiểu liên kết

Các kiểu liên kết thường có một số ràng buộc để hạn chế số các tổ hợp có thể của các thực thể có thể tham gia trong tập hợp liên kết tương ứng. Các ràng buộc này được xác định từ tình trạng của thế giới thực mà kiểu liên kết biểu diễn. Hai loại ràng buộc chính:

- **Tỷ số lực lượng:** Tỷ số lực lượng cho một kiểu liên kết chỉ ra số các thể hiện liên kết mà một thực thể có thể tham gia. Với các kiểu liên kết cấp 2, có thể có các tỷ số lực lượng $1 : 1$, $1 : N$, và $M : N$
- **Các ràng buộc tham gia và sự phụ thuộc tồn tại:** Ràng buộc tham gia chỉ ra rằng có phải sự tồn tại của một kiểu thực thể phụ thuộc vào một kiểu thực thể khác thông qua một kiểu liên kết hay không. Có hai kiểu ràng buộc tham gia:
 - ◆ **Ràng buộc tham gia toàn bộ (phụ thuộc tồn tại):**
 - ◆ **Ràng buộc tham gia bộ phận:**

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Thiết kế CSDL

Mô hình ER

Kiểu & tập thực thể

Khóa & tập giá trị

Kiểu & tập liên kết

Cấp liên kết...

Các ràng buộc...

Mô hình EER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Mô hình thực thể-liên kết mở rộng (EER)

- Mô hình EER bao gồm tất cả các khái niệm của mô hình ER, ngoài ra còn có các khái niệm lớp, kiểu liên kết lớp cha lớp con, tính thừa kế, chuyên biệt, tổng quát, phạm trù
- Lớp cha, lớp con và sự thừa kế
- Chuyên biệt hoá, tổng quát hoá
 - ◆ Chuyên biệt hóa
 - ◆ Tổng quát hóa
 - ◆ Phân cấp chuyên biệt và lưới chuyên biệt
 - ◆ Các ràng buộc và các đặc trưng của chuyên biệt hoá, tổng quát hoá
- Sơ đồ mô hình ERR

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Thiết kế CSDL

Mô hình ER

Kiểu & tập thực thể

Khóa & tập giá trị

Kiểu & tập liên kết

Cấp liên kết...

Các ràng buộc...

Mô hình EER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Mô hình quan hệ, các ràng buộc quan hệ và đại số quan hệ

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Miền, thuộc tính, bộ, quan hệ

- Một miền D là một tập hợp các giá trị nguyên tố, nghĩa là mỗi giá trị trong miền là không thể phân chia được trong phạm vi mô hình quan hệ. Để đặc tả một miền, người ta chỉ ra một tên, một kiểu dữ liệu và khuôn dạng dữ liệu
- Một lược đồ quan hệ R , ký hiệu là $R(A_1, A_2, \dots, A_n)$, được tạo nên từ một tên quan hệ R và một danh sách các thuộc tính A_1, A_2, \dots, A_n
- $Dom(A_i)$ là miền giá trị của A_i
- Cấp của một quan hệ là số các thuộc tính của lược đồ quan hệ của nó
- Một quan hệ r (hoặc trạng thái quan hệ) của lược đồ quan hệ $R(A_1, A_2, \dots, A_n)$ được ký hiệu là $r(R)$, là tập hợp các n -bộ $r = t_1, t_2, \dots, t_n$. Mỗi n -bộ t là một danh sách có thứ tự của n giá trị, $t = \langle v_1, v_2, \dots, v_n \rangle$, trong đó mỗi v_i , $1 \leq i \leq n$, là một phần tử của $Dom(A_i)$ hoặc là một giá trị không xác định (null value). Giá trị thứ i của bộ t được ký hiệu là $t[A_i]$

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Các ràng buộc quan hệ, lược đồ cơ sở dữ liệu quan hệ

- Các ràng buộc miền: Các ràng buộc miền chỉ ra rằng giá trị của mỗi thuộc tính A phải là một giá trị nguyên tử thuộc miền giá trị $Dom(A)$
- Một **siêu khoá** SK xác định rõ một ràng buộc về tính duy nhất, phát biểu rằng không có hai bộ khác nhau trong một trạng thái r của R có cùng một giá trị cho SK
- Ràng buộc khoá và ràng buộc trên các giá trị không xác định ($null$): với hai bộ khác nhau bất kỳ t_1 và t_2 trong một trạng thái quan hệ r của R chúng ta có ràng buộc là $t_1[SK] \neq t_2[SK]$

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Cơ sở dữ liệu quan hệ và lược đồ cơ sở dữ liệu quan hệ

- Một **lược đồ cơ sở dữ liệu quan hệ** S là một tập hợp các lược đồ quan hệ $S = \{R_1, R_2, \dots, R_n\}$ và một tập các ràng buộc toàn vẹn
- Một **trạng thái cơ sở dữ liệu quan hệ** (hoặc **một cơ sở dữ liệu quan hệ**) DB của S là một tập hợp các trạng thái quan hệ $DB = \{r_1, r_2, \dots, r_n\}$ sao cho mỗi r_i là một trạng thái của R_i và sao cho các trạng thái quan hệ r_i thoả mãn các ràng buộc toàn vẹn chỉ ra trong tập các ràng buộc toàn vẹn

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Toàn vẹn thực thể, toàn vẹn tham chiếu và khóa ngoài

- **Ràng buộc toàn vẹn thực thể** được phát biểu là: khoá chính phải luôn luôn có giá trị xác định, nghĩa là không được phép có giá trị *null*
- **Ràng buộc toàn vẹn tham chiếu** được phát biểu là: một bộ giá trị trong một quan hệ có liên kết đến một quan hệ khác phải liên kết đến một bộ giá trị tồn tại trong quan hệ đó

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Toàn vẹn thực thể, toàn vẹn tham chiếu và khóa ngoài (tiếp)

- Một tập hợp các thuộc tính FK trong một lược đồ quan hệ R_1 là một **khóa ngoài** của R_1 tham chiếu đến quan hệ R_2 nếu nó thoả mãn hai quy tắc sau:
 1. Các thuộc tính trong FK có cùng miền giá trị như các thuộc tính của khóa chính PK của R_2 . Các thuộc tính FK được gọi là tham chiếu đến (hoặc là liên hệ đến) quan hệ R_2 .
 2. Một giá trị của FK trong một bộ t_1 của trạng thái hiện tại $r_1(R_1)$ hoặc có mặt như một giá trị của khóa chính của một bộ t_2 nào đó trong trạng thái hiện tại $r_2(R_2)$, hoặc là $null$. Trong trường hợp này ta có $t_1[FK] = t_2[PK]$ và ta nói rằng bộ t_1 liên hệ (tham chiếu) đến bộ t_2 . R_1 được gọi là quan hệ tham chiếu và R_2 được gọi là quan hệ bị tham chiếu

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Các phép toán trên mô hình quan hệ

- **Phép chèn (insert)** cung cấp một danh sách các giá trị cho một bộ mới t được chèn vào trong một quan hệ R . Phép chèn có thể vi phạm các kiểu ràng buộc được mô tả ở trên
- **Phép xoá (delete)** được sử dụng để xoá một hoặc nhiều bộ giá trị của một quan hệ. Phép xoá chỉ có thể vi phạm ràng buộc tham chiếu trong trường hợp bộ bị xoá được tham chiếu bởi một khoá ngoài từ các bộ khác trong CSDL
- **Phép sửa đổi (update)** được dùng để thay đổi các giá trị của một hoặc nhiều thuộc tính trong một (hoặc nhiều) bộ của một quan hệ R nào đây
- **Phép chọn được (select)** sử dụng để chọn một tập hợp các bộ thoả mãn điều kiện chọn từ một quan hệ. Có thể xem phép chọn như một bộ lọc, nó chỉ giữ lại các bộ thoả mãn điều kiện đặt ra. Phép chọn được ký hiệu là $\sigma_{\langle cond \rangle}(R)$, trong đó $\langle cond \rangle$ là điều kiện chọn.

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Các phép toán trên mô hình quan hệ (tiếp)

- Nếu ta coi một quan hệ như một bảng thì **phép chiếu (project)** là phép toán chọn một số cột của bảng. Phép chiếu được ký hiệu là: $\pi_{\langle attr \rangle}(R)$ trong đó $\langle attr \rangle$ là danh sách con các thuộc tính của R .
- Các phép toán lý thuyết tập hợp:
 - ◆ **Phép hợp:** Hợp của hai quan hệ R và S , được ký hiệu là $R \cup S$, cho kết quả là một quan hệ chứa tất cả các bộ có trong R hoặc ở trong S hoặc ở trong cả hai. Các bộ trùng lặp bị loại bỏ
 - ◆ **Phép giao:** Giao của hai quan hệ R và S , được ký hiệu là $R \cap S$, cho kết quả là một quan hệ chứa tất cả các bộ có trong cả hai quan hệ R và S
 - ◆ **Phép trừ quan hệ:** Phép trừ quan hệ R và S , được ký hiệu là $R - S$, cho kết quả là một quan hệ chứa tất cả các bộ có trong R nhưng không có trong S

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Các phép toán trên mô hình quan hệ (tiếp)

- **Phép nối (join)** Phép nối được ký hiệu là \bowtie và được dùng để kết hợp các bộ có liên hệ với nhau từ hai quan hệ thành một bộ. Phép toán này rất quan trọng đối với cơ sở dữ liệu quan hệ có nhiều bảng bởi vì nó cho phép ta xử lý các mối liên kết giữa các quan hệ
- **Dạng tổng quát của phép nối trên hai quan hệ**
 $R(A_1, A_2, \dots, A_n)$ và $S(B_1, B_2, \dots, B_m)$ là $R \bowtie^{<cond>} S$ trong đó $<cond>$ là điều kiện nối. Kết quả của phép nối là một quan hệ $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$ có $n + m$ thuộc tính. Mỗi bộ của Q là một sự kết nối giữa một bộ của R và một bộ của S khi chúng thoả mãn điều kiện nối
- **Phép nối tự nhiên (*)** nhằm loại bỏ thuộc tính thứ hai (thuộc tính thừa) trong điều kiện nối bằng. Định nghĩa chuẩn của nối tự nhiên đòi hỏi hai thuộc tính nối (hoặc mỗi cặp thuộc tính nối) phải có tên như nhau trong cả hai quan hệ

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Các phép toán trên mô hình quan hệ (tiếp)

- Có ba phép nối ngoài gọi là **nối ngoài trái (left outer join)**, **nối ngoài phải (right outer join)** và **nối ngoài đầy đủ (full outer join)**, được ký hiệu tương ứng là: $\sqcup\bowtie$, $\bowtie\sqcup$ và $\sqcup\bowtie\sqcup$
- **Phép nối ngoài trái** giữ lại mọi bộ trong quan hệ bên trái R trong phép nối. Nếu không có bộ liên kết nào được tìm thấy trong S thì các thuộc tính của S trong kết quả phép nối được “làm đầy” bằng các giá trị *null*. Tương tự như vậy đối với các **phép nối ngoài phải** và các **phép nối ngoài đầy đủ**
- **Phép toán hợp ngoài** được mở rộng để lấy hợp của các bộ từ các quan hệ nếu các bộ không tương thích đồng nhất. Phép toán này chỉ lấy hợp của các quan hệ mà chúng chỉ tương thích bộ phận, nghĩa là chỉ một vài thuộc tính của chúng là tương thích phép hợp

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Chuyển đổi mô hình ER thành mô hình quan hệ

Bước 1. Với mỗi kiểu thực thể thông thường E trong lược đồ ER, hãy tạo một quan hệ R chứa mọi thuộc tính đơn của E . Với các thuộc tính phức hợp, chỉ lấy các thuộc tính thành phần đơn của nó. Chọn một trong các thuộc tính khoá của E làm khoá chính cho R . Nếu khoá được chọn của E là phức hợp (gồm nhiều thuộc tính) thì tập các thuộc tính đơn đó sẽ cùng nhau tạo nên khoá chính của R

Bước 2. Với mỗi kiểu thực thể yếu W trong lược đồ ER cùng với kiểu thực thể chủ E , hãy tạo một quan hệ R chứa tất cả các thành phần đơn (hoặc các thành phần đơn của các thuộc tính phức hợp) của W như là các thuộc tính của R . Đưa các thuộc tính khoá chính của các quan hệ tương ứng với kiểu thực thể chủ làm khoá ngoài của R . Các thuộc tính này sẽ xác định kiểu liên kết của W . Khoá chính của R là một tổ hợp của khoá chính của các quan hệ tương ứng với kiểu thực thể chủ và khoá bộ phận của kiểu thực thể yếu W nếu có

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Chuyển đổi mô hình ER thành mô hình quan hệ (tiếp)

Bước 3. Với mỗi kiểu liên kết 1:1 R trong lược đồ ER, hãy xác định các quan hệ S và T tương ứng với các kiểu thực thể tham gia trong R . Hãy chọn một trong các quan hệ, chẳng hạn S , và đưa khoá chính của T vào làm khoá ngoài trong S . Tốt nhất là chọn S là một kiểu thực thể tham gia toàn bộ vào R . Đưa tất cả các thuộc tính đơn (hoặc các thành phần đơn của các thuộc tính phức hợp) của kiểu liên kết 1:1 R vào làm các thuộc tính của S

Bước 4. Với mỗi kiểu liên kết hai ngôi R kiểu 1 : N , hãy xác định quan hệ S biểu diễn kiểu thực thể tham gia ở phía N của kiểu liên kết. Đưa khoá chính của quan hệ T biểu diễn kiểu thực thể tham gia vào R ở phía 1 vào làm khoá ngoài trong S . Làm như vậy là vì mỗi thực thể cụ thể của phía N được liên kết với nhiều nhất là một thực thể cụ thể của phía 1 của kiểu liên kết. Đưa các thuộc tính đơn (hoặc các thành phần đơn của các thuộc tính phức hợp) của kiểu liên kết 1 : N vào làm các thuộc tính của S

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Chuyển đổi mô hình ER thành mô hình quan hệ (tiếp)

Bước 5. Với mỗi kiểu liên kết $N : M$ hai ngôi R , hãy tạo ra một quan hệ mới S để biểu diễn R . Đưa các khoá chính của các quan hệ biểu diễn các kiểu thực thể tham gia vào làm khoá ngoài của SA . Tổ hợp các khoá chính đó sẽ tạo nên khoá chính của S . Đưa tất cả các thuộc tính đơn (hoặc các thành phần đơn của các thuộc tính phức hợp) của kiểu liên kết $N : M$ vào làm các thuộc tính của S . Chú ý rằng ta không thể biểu diễn một kiểu liên kết $N : M$ bằng một thuộc tính khoá ngoài đơn giản trong một trong các quan hệ tham gia (như đã làm với các kiểu liên kết $1 : 1$ và $1 : N$) vì tỷ số lực lượng $N : M$

Bước 6. Với mỗi thuộc tính đa trị A , hãy tạo ra một quan hệ mới R . Quan hệ R này sẽ chứa một thuộc tính tương ứng với A cộng với thuộc tính khoá K của quan hệ biểu diễn kiểu thực thể hoặc kiểu liên kết có thuộc tính là A làm khoá ngoài của R . Khoá chính của R là một tổ hợp của A và K . Nếu thuộc tính đa trị là phức hợp thì chúng ta chỉ đưa vào R các thành phần đơn của nó

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Chuyển đổi mô hình ER thành mô hình quan hệ (tiếp)

Bước 7. Với mỗi kiểu liên kết n ngôi R , trong đó $n > 2$, hãy tạo ra một quan hệ S để biểu diễn R . Đưa các khoá chính của các quan hệ biểu diễn các kiểu thực thể tham gia vào làm khoá ngoài của S . Đưa tất cả các thuộc tính đơn (hoặc các thành phần đơn của các thuộc tính phức hợp) của kiểu liên kết n -ngôi vào làm thuộc tính của S . Khoá chính của S thường là một tổ hợp các khoá chính của các quan hệ biểu diễn các kiểu thực thể tham gia. Tuy nhiên, nếu ràng buộc lực lượng trên một kiểu thực thể E nào đó tham gia vào R là 1 thì khoá chính của S không được chứa thuộc tính khoá ngoài tham chiếu đến quan hệ E tương ứng với kiểu thực thể E

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Miền-thuộc tính...

Ràng buộc quan hệ

CSDL quan hệ

Các loại ràng buộc

Phép toán quan hệ

Chuyển đổi ER-quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Phụ thuộc hàm và chuẩn hóa cơ sở dữ liệu quan hệ

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Nguyên tắc thiết kế

Phụ thuộc hàm

Qui tắc suy diễn

Bao đóng

Phụ thuộc hàm tương đương

Phụ thuộc hàm tối thiểu

Các dạng chuẩn

Thiết kế CSDL

Các nguyên tắc thiết kế lược đồ quan hệ

- Nguyên tắc 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 đơn*
- Nguyên tắc 2. (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*
- Nguyên tắc 3. (Các giá trị không xác định trong các bộ): *Tránh càng xa càng tốt việc đặt vào trong các quan hệ cơ sở những thuộc tính mà các giá trị của chúng thường xuyên là null*
- Nguyên tắc 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 ngoài theo cách đảm bảo không sinh ra các bộ “giả”*

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Nguyên tắc thiết kế

Phụ thuộc hàm

Qui tắc suy diễn

Bao đóng

Phụ thuộc hàm
tương đương

Phụ thuộc hàm tối
thiểu

Các dạng chuẩn

Thiết kế CSDL

- Một **phụ thuộc hàm** (viết tắt là FD) ký hiệu là $X \rightarrow Y$, giữa hai tập thuộc tính X và Y chỉ ra một ràng buộc trên các bộ có thể có tạo nên một trạng thái quan hệ r của R . Ràng buộc đó là: với hai bộ bất kỳ t_1 và t_2 trong r , nếu có $t_1[X] = t_2[X]$ thì cũng phải có $t_1[Y] = t_2[Y]$
- Chúng ta nói rằng có một phụ thuộc hàm từ X vào Y hoặc Y phụ thuộc hàm vào X
- Một phụ thuộc hàm là một tính chất của lược đồ quan hệ R chứ không phải là tính chất của một trạng thái hợp pháp r của R . Vì vậy, một phụ thuộc hàm không thể được phát hiện một cách tự động từ một trạng thái r mà phải do một người hiểu biết ngữ nghĩa của các thuộc tính xác định một cách rõ ràng

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Nguyên tắc thiết kế

Phụ thuộc hàm

Qui tắc suy diễn

Bao đóng

Phụ thuộc hàm
tương đương

Phụ thuộc hàm tối
thiểu

Các dạng chuẩn

Thiết kế CSDL

Các qui tắc suy diễn đối với phụ thuộc hàm

- QT1 (phản xạ): Nếu $X \supset Y$ thì $X \rightarrow Y$
QT2 (tăng): $\{X \rightarrow Y\} \models XZ \rightarrow YZ$
QT3 (bắc cầu): $\{X \rightarrow Y, Y \rightarrow Z\} \models X \rightarrow Z$
QT4 (chiều): $\{X \rightarrow YZ\} \models X \rightarrow Y$ và $X \rightarrow Z$
QT5 (hợp): $\{X \rightarrow Y, X \rightarrow Z\} \models X \rightarrow YZ$
QT6 (tựa bắc cầu): $\{X \rightarrow Y, WY \rightarrow Z\} \models WX \rightarrow Z$

Amstrong đã chứng minh rằng các quy tắc suy diễn từ QT1 đến QT3 là đúng và đầy đủ:

- Đúng: cho trước một tập phụ thuộc hàm F trên một lược đồ quan hệ R , bất kỳ một phụ thuộc hàm nào suy diễn được bằng cách áp dụng các quy tắc từ QT1 đến QT3 cũng đúng trong mỗi trạng thái quan hệ r của R thoả mãn các phụ thuộc hàm trong F
- Đầy đủ: việc sử dụng các quy tắc từ QT1 đến QT3 lặp lại nhiều lần để suy diễn các phụ thuộc hàm cho đến khi không còn suy diễn được nữa sẽ cho kết quả là một tập

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Nguyên tắc thiết kế

Phụ thuộc hàm

Quy tắc suy diễn

Bao đóng

Phụ thuộc hàm
tương đương

Phụ thuộc hàm tối
thiểu

Các dạng chuẩn

Thiết kế CSDL

Bao đóng của tập thuộc tính

- Thông thường khi thiết kế cơ sở dữ liệu, đầu tiên ta chỉ ra một tập các phụ thuộc hàm để xác định được nhờ ngữ nghĩa của các thuộc tính của R . Sau đó, sử dụng các quy tắc Armstrong để suy diễn các phụ thuộc hàm bổ sung
- Với mỗi tập thuộc tính X , chúng ta xác định tập X^+ các thuộc tính phụ thuộc hàm vào X dựa trên F . X^+ được gọi là bao đóng của X dưới F

Thuật toán 4.1: (xác định X^+ , bao đóng của X dưới F)

$X^+ = X$;

repeat

$OldX^+ = X^+$;

với mỗi phụ thuộc hàm $Y \rightarrow Z$ trong F thực hiện

nếu $X^+ \supset Y$ thì $X^+ = X^+ \cup Z$;

until ($X^+ = OldX^+$);

Tài liệu tham khảo

Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Nguyên tắc thiết kế

Phụ thuộc hàm

Quy tắc suy diễn

Bao đóng

Phụ thuộc hàm

tương đương

Phụ thuộc hàm tối thiểu

Các dạng chuẩn

Thiết kế CSDL

Sự tương đương của các tập phụ thuộc hàm

- Một tập hợp các phụ thuộc hàm E được phủ bởi một tập các phụ thuộc hàm F - hoặc F phủ E - nếu mỗi một phụ thuộc hàm trong E đều ở trong F^+ , điều đó có nghĩa là mỗi phụ thuộc hàm trong E có thể suy diễn được từ F
- Hai tập phụ thuộc hàm E và F là tương đương nếu $E^+ = F^+$
- Một tập phụ thuộc hàm là tối thiểu nếu nó thoả mãn các điều kiện sau đây:
 - ◆ Vê phải của các phụ thuộc hàm trong F chỉ có một thuộc tính.
 - ◆ Chúng ta không thể thay thế bất kỳ một phụ thuộc hàm $X \rightarrow A$ trong F bằng phụ thuộc hàm $Y \rightarrow A$, trong đó Y là tập con đúng của X mà vẫn còn là một tập phụ thuộc hàm tương đương với F .
 - ◆ Chúng ta không thể bỏ đi bất kỳ phụ thuộc hàm nào ra khỏi F mà vẫn có một tập phụ thuộc hàm tương đương với F

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Nguyên tắc thiết kế

Phụ thuộc hàm

Quy tắc suy diễn

Bao đóng

Phụ thuộc hàm tương đương

Phụ thuộc hàm tối thiểu

Các dạng chuẩn

Thiết kế CSDL

Thuật toán tìm phụ thuộc hàm tối thiểu

Thuật toán 4.2 (Tìm phủ tối thiểu G cho F)

1. Đặt $G := F$;
2. Thay thế mỗi phụ thuộc hàm $X \rightarrow \{A_1, A_2, \dots, A_n\}$ trong G bằng n phụ thuộc hàm $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_n$.
3. Với mỗi phụ thuộc hàm $X \rightarrow A$ trong G ,
với mỗi thuộc tính B là một phần tử của X
nếu $((G - (X \rightarrow A)) \cup ((X - \{B\}) \rightarrow A))$ là tương đương với G
thì thay thế $X \rightarrow A$ bằng $(X - \{B\}) \rightarrow A$ ở trong G
4. Với mỗi phụ thuộc hàm $X \rightarrow A$ còn lại trong G
nếu $(G - \{X \rightarrow A\})$ là tương đương với G
thì loại bỏ $X \rightarrow A$ ra khỏi G

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Nguyên tắc thiết kế

Phụ thuộc hàm

Qui tắc suy diễn

Bao đóng

Phụ thuộc hàm
tương đương

**Phụ thuộc hàm tối
thiểu**

Các dạng chuẩn

Thiết kế CSDL

Các dạng chuẩn dựa trên khóa chính

- Một lược đồ quan hệ R là ở **dạng chuẩn 1 (1NF)** nếu miền giá trị của các thuộc tính của nó chỉ chứa các giá trị nguyên tử (đơn, không phân chia được) và giá trị của một thuộc tính bất kỳ trong một bộ giá trị phải là một giá trị đơn thuộc miền giá trị của thuộc tính đó.
- Một lược đồ quan hệ R là ở **dạng chuẩn 2 (2NF)** nếu mỗi thuộc tính không khóa A trong R không phụ thuộc bộ phận vào một khóa bất kỳ của R .
- Một lược đồ quan hệ R là ở **dạng chuẩn 3 (3NF)** nếu khi một phụ thuộc hàm $X \rightarrow A$ thỏa mãn trong R , thì:
 - ◆ Hoặc X là một siêu khóa của R .
 - ◆ Hoặc A là một thuộc tính khóa của R .
- Một lược đồ quan hệ là ở **dạng chuẩn Boyce-Codd (BCNF)** nếu khi một phụ thuộc hàm $X \rightarrow A$ thỏa mãn trong R thì X là một siêu khóa của R .

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Nguyên tắc thiết kế

Phụ thuộc hàm

Quy tắc suy diễn

Bao đóng

Phụ thuộc hàm
tương đương

Phụ thuộc hàm tối
thiểu

Các dạng chuẩn

Thiết kế CSDL

Các thuật toán thiết kế lược đồ cơ sở dữ liệu quan hệ

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nối không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

Tách quan hệ và điều kiện bảo toàn

- Tách quan hệ: Lược đồ quan hệ vũ trụ đơn $R = \{A_1, A_2, \dots, A_n\}$ được tách thành một tập hợp các lược đồ quan hệ $D = \{R_1, R_2, \dots, R_m\}$. Một cách hình thức, ta có điều kiện bảo toàn thuộc tính: $\cup R_i = R$
- Tính không đầy đủ của các dạng chuẩn: Mục đích của chúng ta là mỗi quan hệ riêng rẽ R_i trong phép tách D là ở dạng chuẩn BCNF hoặc 3NF. Tuy nhiên, điều đó không đủ để đảm bảo một thiết kế CSDL tốt. Bên cạnh việc xem xét từng quan hệ riêng rẽ, chúng ta cần xem xét toàn bộ phép tách.
- Việc mỗi phụ thuộc hàm $X \rightarrow Y$ trong F hoặc được xuất hiện trực tiếp trong một trong các lược đồ quan hệ R_i trong phép tách D hoặc có thể được suy diễn từ các phụ thuộc hàm có trong R_i là rất có lợi. Ta gọi đó là **điều kiện bảo toàn phụ thuộc**
- *Định lý*: Luôn luôn tìm được một phép tách bảo toàn phụ thuộc D đối với F sao cho mỗi quan hệ R_i trong D là 3NF

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nối không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

Thuật toán tách bảo toàn phụ thuộc

Thuật toán 5.1: Tạo một phép tách bảo toàn phụ thuộc $D = \{R_1, R_2, \dots, R_m\}$ của một quan hệ dữ liệu R dựa trên một tập phụ thuộc hàm F sao cho mỗi R_i trong D là ở 3NF. Thuật toán này chỉ đảm bảo tính chất bảo toàn phụ thuộc, không đảm bảo tính chất nội không mất mát.

Input: Một quan hệ dữ liệu R và một tập phụ thuộc hàm F trên các thuộc tính của R .

1. Tìm phủ tối thiểu G của F .
2. Với mỗi vế trái X của một phụ thuộc hàm xuất hiện trong G , hãy tạo một lược đồ trong D với các thuộc tính $\{X \cup \{A_1\} \cup \{A_2\} \cup \dots \cup \{A_k\}\}$ trong đó $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_k$ chỉ là các phụ thuộc hàm trong G với X là vế trái (X là khóa của quan hệ này).
3. Đặt các thuộc tính còn lại (những thuộc tính chưa được đặt vào quan hệ nào) vào một quan hệ đơn để đảm bảo tính chất bảo toàn thuộc tính.

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nội không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

Phép tách và kết nối không mất mát

Thuật toán 5.2: Kiểm tra tính chất nối không mất mát

Input: Một quan hệ vũ trụ $R(A_1, A_2, \dots, A_n)$, một phép tách $D = \{R_1, R_2, \dots, R_m\}$ của R và một tập phụ thuộc hàm F .

1. Tạo một ma trận S có m hàng, n cột. Mỗi cột của ma trận ứng với một thuộc tính, mỗi hàng ứng với mỗi quan hệ R_i
2. Đặt $S(i, j) = 1$ nếu thuộc tính A_j thuộc về quan hệ R_i và bằng 0 trong trường hợp ngược lại.
3. Lặp lại vòng lặp sau đây cho đến khi nào việc thực hiện vòng lặp không làm thay đổi S : Với mỗi phụ thuộc hàm $X \rightarrow Y$ trong F , xác định các hàng trong S có các ký hiệu 1 như nhau trong các cột ứng với các thuộc tính trong X . Nếu có một hàng trong số đó chứa 1 trong các cột ứng với thuộc tính Y thì hãy làm cho các làm cho các cột tương ứng của các hàng khác cũng chứa 1.
4. Nếu có một hàng chứa toàn ký hiệu 1 thì phép tách có tính chất nối không mất mát, ngược lại, phép tách không có tính chất đó.

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nối không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

Tách quan hệ với tính chất nổi không mất mát

Thuật toán 5.3: Tách quan hệ thành các quan hệ BCNF với tính chất nổi không mất mát

Input: Một quan hệ vũ trụ R và một tập hợp các phụ thuộc hàm F trên các thuộc tính của R .

1. Đặt $D := \{R\}$
2. Khi có một lược đồ quan hệ Q trong D không phải ở BCNF, thực hiện vòng lặp: Với mỗi một lược đồ quan hệ Q trong D không ở BCNF hãy tìm một phụ thuộc hàm $X \rightarrow Y$ trong Q vi phạm BCNF và thay thế Q trong D bằng hai lược đồ quan hệ $(Q - Y)$ và $(X \cup Y)$. Quá trình lặp dừng khi không còn quan hệ nào trong D vi phạm BCNF.

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nổi không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

Thuật toán tổng hợp quan hệ bảo toàn phụ thuộc và nội không mất mát

Thuật toán 5.4: Thuật toán tổng hợp quan hệ với tính chất bảo toàn phụ thuộc và nội không mất mát.

Input: Một quan hệ vũ trụ R và một tập các phụ thuộc hàm F trên các thuộc tính của R .

1. Tìm phủ tối thiểu G cho F .
2. Với mỗi vế trái X của một phụ thuộc hàm xuất hiện trong G hãy tạo ra một lược đồ quan hệ trong D với các thuộc tính $\{X \cup \{A_1\} \cup \{A_2\} \cup \dots \cup \{A_k\}\}$, trong đó $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_k$ chỉ là các phụ thuộc hàm ở trong G với X là vế trái (X là khóa của quan hệ này).
3. Nếu không có lược đồ quan hệ nào trong D chứa một khóa của R thì hãy tạo ra thêm một lược đồ quan hệ trong D chứa các thuộc tính tạo nên một khóa của R .

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nội không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

Thuật toán xác định khóa

Thuật toán xác định khóa: Tìm một khóa K của R dựa trên tập F các phụ thuộc hàm (để thực hiện bước 3 trong thuật toán 5.4).

1. Đặt $K := R$
2. Với mỗi thuộc tính A trong K {
 tính $(K - A)^+$ đối với F ;
 Nếu $(K - A)^+$ chứa tất cả các thuộc tính trong R thì
 đặt $K := K - \{A\}$
}

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nối không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

Giả thiết có một lược đồ quan hệ R , X và Y là hai tập con của R . Một **phụ thuộc đa trị (MVD)**, ký hiệu là $X \twoheadrightarrow Y$, chỉ ra ràng buộc sau đây trên một trạng thái quan hệ bất kỳ của R : Nếu hai bộ t_1 và t_2 tồn tại trong R sao cho $t_1[X] = t_2[X]$ thì hai bộ t_3 và t_4 cũng tồn tại trong R với các tính chất sau:

- $t_3[X] = t_4[X] = t_1[X] = t_2[X]$
- $t_3[Y] = t_1[Y]$ và $t_4[Y] = t_2[Y]$
- $t_3[Z] = t_2[Z]$ và $t_4[Z] = t_1[Z]$ với $Z = (R - (X \cup Y))$

Khi $X \twoheadrightarrow Y$ thỏa mãn, ta nói rằng X đa xác định Y . Bởi vì tính đối xứng trong định nghĩa, khi $X \twoheadrightarrow Y$ thỏa mãn trong R , $X \twoheadrightarrow Z$ cũng thỏa mãn trong R . Như vậy $X \twoheadrightarrow Y$ kéo theo $X \twoheadrightarrow Z$ và vì thế đôi khi nó được viết là $X \twoheadrightarrow Y|Z$

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nổi không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

Các qui tắc suy diễn với các phụ thuộc hàm và phụ thuộc đa trị

- QT1 (phản xạ cho FD): Nếu $X \supseteq Y$ thì $X \rightarrow Y$
- QT2 (tăng cho FD): $\{X \rightarrow Y\} \models XZ \rightarrow YZ$
- QT3 (bắc cầu cho FD): $\{X \rightarrow Y, Y \rightarrow Z\} \models X \rightarrow Z$
- QT4 (bù cho MVD): $\{X \twoheadrightarrow Y\} \models \{X \twoheadrightarrow (R - (X \cup Y))\}$
- QT5 (tăng cho MVD): Nếu $X \twoheadrightarrow Y$ và $W \supseteq Z$ thì $WX \twoheadrightarrow YZ$
- QT6 (bắc cầu cho MVD):
 $X \twoheadrightarrow Y, Y \twoheadrightarrow Z \models X \twoheadrightarrow (Z \sim Y)$
- QT7 (tái tạo cho FD và MVD): $X \rightarrow Y \mid = X \twoheadrightarrow Y$
- QT8 (liên hợp cho FD và MVD): Nếu $X \twoheadrightarrow Y$ và có tồn tại W với các tính chất a) $W \cap Y = \emptyset$, b) $W \rightarrow Z$ và c) $Y \supseteq Z$ thì $X \rightarrow Z$

QT1 đến QT3 là các quy tắc suy diễn Armstrong đối với các phụ thuộc hàm. QT4 đến QT6 là các quy tắc suy diễn chỉ liên quan đến các phụ thuộc đa trị. QT7 và QT8 liên kết các phụ thuộc hàm và các phụ thuộc đa trị.

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nối không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

- Một lược đồ quan hệ R là ở **dạng chuẩn 4 (4NF)** đối với một tập hợp các phụ thuộc F (gồm các phụ thuộc hàm và phụ thuộc đa trị) nếu với mỗi phụ thuộc đa trị không tầm thường $X \twoheadrightarrow Y$ trong F^+ , X là một siêu khóa của R
- Tách có tính chất nối không mất mát thành các quan hệ 4NF: Các lược đồ quan hệ R_1 và R_2 tạo thành một phép tách có tính chất nối không mất mát của R khi và chỉ khi $(R_1 \cap R_2) \twoheadrightarrow (R_1 - R_2)$ (hoặc $(R_1 \cap R_2) \twoheadrightarrow (R_2 - R_1)$).

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nối không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa
trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

Thuật toán tách quan hệ không mất mát thành các quan hệ 4NF

Thuật toán 5.5: Tách quan hệ thành các quan hệ 4NF với tính chất nối không mất mát.

Input: Một quan hệ vũ trụ R và một tập phụ thuộc hàm và phụ thuộc đa trị F .

1. Đặt $D := \{R\}$
2. Khi có một lược đồ quan hệ Q trong D không ở 4NF, thực hiện:
 - { Chọn một lược đồ quan hệ Q trong D không ở 4NF;
 - Tìm một phụ thuộc đa trị không tầm thường $X \twoheadrightarrow Y$ trong Q vi phạm 4NF;
 - Thay thế Q trong D bằng hai lược đồ quan hệ $(Q - Y)$ và $(X \cup Y)$ }

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nối không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5

Các phụ thuộc nối và dạng chuẩn 5

- Một **phụ thuộc nối** (JD), ký hiệu là $JD(R_1, R_2, \dots, R_n)$ trên lược đồ quan hệ R chỉ ra một ràng buộc trên các trạng thái r của R . Ràng buộc đó tuyên bố rằng mỗi trạng thái hợp pháp r của R phải có phép tách có tính chất nối không mất mát thành R_1, R_2, \dots, R_n . Điều đó nghĩa là:
$$*(\pi_{R_1}(r), \pi_{R_2}(r), \dots, \pi_{R_n}(r)) = r$$
- Một phụ thuộc nối $JD(R_1, R_2, \dots, R_n)$ là một phụ thuộc nối tầm thường nếu một trong các lược đồ quan hệ R_i ở trong $JD(R_1, R_2, \dots, R_n)$ là bằng R
- Một lược đồ quan hệ R là ở **dạng chuẩn 5 (5NF)** (hoặc **dạng chuẩn nối chiếu PJNF – Project-Join normal form**) đối với một tập F các phụ thuộc hàm, phụ thuộc đa trị và phụ thuộc nối nếu với mỗi phụ thuộc nối không tầm thường $JD(R_1, R_2, \dots, R_n)$ trong F^+ , mỗi R_i là một siêu khóa của R .

Tài liệu tham khảo
Mở đầu

Khái niệm cơ bản

Mô hình ER

Mô hình quan hệ

Phụ thuộc hàm

Thiết kế CSDL

Tách quan hệ

Thuật toán 5.1

Nối không mất mát

Tổng hợp quan hệ

Xác định khóa

Phụ thuộc hàm đa trị

Các qui tắc suy diễn

Dạng chuẩn 4

Tách quan hệ

Dạng chuẩn 5