



Quản trị Cơ sở dữ liệu và Phần mềm ứng dụng

Bộ môn CNTT
Khoa Tin học Thương mại



Chương II: *Thiết kế CSDL quan hệ*

1. Giới thiệu chung
 - 1.1. Thiết kế CSDL QH và các cách tiếp cận
 - 1.2. Phụ thuộc hàm
2. Chuẩn hóa lược đồ quan hệ
 - 2.1. Các dạng chuẩn
 - 2.2. Tách lược đồ quan hệ theo chuẩn
3. Ràng buộc toàn vẹn trong CSDL quan hệ
 - 3.1. Khái niệm ràng buộc toàn vẹn
 - 3.2. Ràng buộc toàn vẹn trên thuộc tính
 - 3.3. Ràng buộc toàn vẹn trên quan hệ



1. Giới thiệu chung

1.1. Thiết kế CSDL QH và các cách tiếp cận

- Thiết kế cơ sở dữ liệu quan hệ \Leftrightarrow xây dựng lược đồ CSDL QH gồm một tập các lược đồ quan hệ thỏa mãn hai yêu cầu:
 - Lưu trữ thông tin không dư thừa
 - Tìm kiếm thông tin dễ dàng
- Ví dụ
 - Lược đồ quan hệ
 - CUNG_UNG(**MaNCC**, TenNCC, DiaChi, **SanPham**, Gia)

Dự thừa dữ liệu

○ Quan hệ CUNG_UNG_0

MaNCC	TenNCC	DiaChi	SanPham	Gia
1	Hải Hà	Hà Nội	Kẹo mềm	100
1	Hải Hà	Hà Nội	Kẹo cứng	150
1	Hải Hà	Hà Nội	Bánh	200
2	Kinh đô	Hồ Chí Minh	Kẹo	120
2	Kinh đô	Hồ Chí Minh	Bánh	150

- Một nhà cung cấp cung cấp nhiều mặt hàng.
- Lập các thông tin về nhà cung cấp ứng với mỗi một mặt hàng khác nhau của cùng nhà cung cấp đó.

Không nhất quán

○ Quan hệ CUNG_UNG_0

MaNCC	TenNCC	DiaChi	SanPham	Gia
1	Hải Hà	Đà Nẵng	Kẹo mềm	100
1	Hải Hà	Hà Nội	Kẹo cứng	150
1	Hải Hà	Hà Nội	Bánh	200
2	Kinh đô	Hồ Chí Minh	Kẹo	120
2	Kinh đô	Hồ Chí Minh	Bánh	150

- Dị thường khi cập nhật thông tin về nhà cung cấp như thay đổi địa chỉ.

Dị thường khi thêm bộ

o Quan hệ CUNG_UNG_0

MaNCC	TenNCC	DiaChi	SanPham	Gia
1	Hải Hà	Hà nội	Kẹo mềm	100
1	Hải Hà	Hà Nội	Kẹo cứng	150
1	Hải Hà	Hà Nội	Bánh	200
2	Kinh đô	Hồ Chí Minh	Kẹo	120
2	Kinh đô	Hồ Chí Minh	Bánh	150
3	Bibica	Đà Nẵng	NULL	NULL

- Dị thường khi thêm mới thông tin về nhà cung cấp nhưng nhà cung cấp chưa cung cấp mặt hàng nào.

Dị thường khi xóa bộ

o Quan hệ CUNG_UNG_0

MaNCC	TenNCC	DiaChi	SanPham	Gia
1	Hải Hà	Hà nội	Kẹo mềm	100
1	Hải Hà	Hà Nội	Kẹo cứng	150
1	Hải Hà	Hà Nội	Bánh	200
2	Kinh đô	Hồ Chí Minh	Kẹo	120

- Tồn tại nhà cung cấp chỉ cung cấp một mặt hàng.
- Dị thường khi xóa thông tin về sự cung cấp → xóa luôn thông tin về nhà cung cấp.

Tìm kiếm thông tin

MaNCC	TenNCC	DiaChi
1	Hải Hà	Hà Nội
2	Kinh đô	Hồ Chí Minh

CUNG_UNG_11

MaNCC	SanPham	Gia
1	Kẹo mềm	100
1	Kẹo cứng	150
1	Bánh	200
2	Kẹo	120
2	Bánh	200

CUNG_UNG_12

- Quan hệ CUNG_UNG_0 tách thành 2 quan hệ CUNG_UNG_11 và CUNG_UNG_12
 - Lưu trữ thông tin không dư thừa ???
 - Tìm kiếm thông tin dễ dàng ???



Các cách tiếp cận

- **Từ trên xuống (*Topdown*):**
 - Xây dựng sơ đồ thực thể liên kết ER từ các đặc tả
 - Chuyển đổi sơ đồ ER thành lược đồ CSDL quan hệ.
 - Chuẩn hóa lược đồ CSDL quan hệ (nếu cần)
- **Từ dưới lên (*Bottom Up*):**
 - Xây dựng lược đồ quan hệ ban đầu từ các đặc tả.
 - Chuẩn hóa lược đồ quan hệ.



1.2. Phụ thuộc hàm

a. Khái niệm

- Cho quan hệ R , thuộc tính B của quan hệ R được gọi là phụ thuộc hàm vào thuộc tính A của quan hệ R nếu với mỗi giá trị của A xác định duy nhất một giá trị của B . A được gọi là xác định hàm của B .
- Ký hiệu: $A \rightarrow B$



a. Khái niệm (t)

- Tập các phụ thuộc hàm F của 1 lược đồ quan hệ R là một tập gồm các phụ thuộc hàm xác định trên R .
 - Ví dụ: Tập phụ thuộc hàm $F = \{A \rightarrow B, B \rightarrow C\}$ của $R(A, B, C)$
- Trong quan hệ R , ký hiệu A, B, C dành cho các thuộc tính đơn, X, Y, Z dành cho tập các thuộc tính.

Ví dụ

- Tập tất cả các thuộc tính của quan hệ phải phụ thuộc hàm vào khóa.

- MaNCC → TenNCC
- MaNCC → SoNV
- MaNCC → DiaChi
- MaNCC: Khóa

MaNCC	TenNCC	SoNV	DiaChi
S1	Hải Hà	20	Hà Nội
S2	Kinh Đô	10	Hà Nội
S3	Bibica	30	HCM

- $F = \{ \text{MaNCC} \rightarrow \text{TenNCC}, \text{MaNCC} \rightarrow \text{SoNV}, \text{MaNCC} \rightarrow \text{DiaChi} \}$



Ví dụ

- Một tập thuộc tính là xác định hàm của các thuộc tính khác thì chưa chắc là một khóa.
 - TenNCC → DiaChi
 - TenNCC không phải là khóa

MaNCC	TenNCC	SoNV	DiaChi
S1	Hải Hà	20	Hà Nội
S2	Kinh Đô	10	Hà Nội
S3	Bibica	30	HCM
S4	Hải Hà	10	Hà Nội



b. Hệ tiên đề Amstrong

- Giả thiết
 - Lược đồ quan hệ R.
 - X, Y, Z : tập các thuộc tính thuộc R.
 - $XY = X \cup Y$
- Hệ 3 tiên đề với các phụ thuộc hàm:
 - Phản xạ: $XY \rightarrow X$; $XY \rightarrow Y$
 - Tăng trưởng: $X \rightarrow Y$ thì $XZ \rightarrow YZ$
 - bắc cầu: $X \rightarrow Y, Y \rightarrow Z$ thì $X \rightarrow Z$



Luật suy ra từ hệ tiên đề

- Luật hợp
 - Nếu $X \rightarrow Y, X \rightarrow Z$ thì $X \rightarrow YZ$
- Luật tựa bắc cầu
 - Nếu $X \rightarrow Y, WY \rightarrow Z$ thì $XW \rightarrow Z$
- Luật tách
 - Nếu $X \rightarrow Y, Z$ thuộc Y thì $X \rightarrow Z$



c. Phụ thuộc hàm đầy đủ và phụ thuộc bắc cầu

○ Phụ thuộc hàm đầy đủ

- Y phụ thuộc hàm đầy đủ vào X nếu Y phụ thuộc hàm vào X nhưng không phụ thuộc hàm vào bất kỳ một tập con thực sự nào của X.
- Ví dụ
 - Lược đồ $R(A, B, C, D)$
 - $F = \{AB \rightarrow C; AB \rightarrow D; B \rightarrow D\}$
 - C phụ thuộc hàm đầy đủ vào $\{A, B\}$
 - D không phụ thuộc hàm đầy đủ vào $\{A, B\}$



Phụ thuộc bắc cầu

- Phụ thuộc hàm $X \rightarrow A$, A được gọi là **phụ thuộc bắc cầu** vào X nếu tồn tại Y để cho $X \rightarrow Y$, $Y \rightarrow A$, $Y \not\rightarrow X$ và $A \notin XY$
- Ví dụ:
 - $F = \{A \rightarrow B, B \rightarrow C\}$
 - $A \rightarrow C$: C phụ thuộc bắc cầu vào A



d. Bao đóng và phủ của tập các phụ thuộc hàm

- Cho tập các phụ thuộc hàm F xác định trên R .
 - **Bao đóng F^+** của tập các phụ thuộc hàm F là tập tất cả các phụ thuộc hàm được suy diễn logic từ F .
 - **Phủ G** của tập các phụ thuộc hàm F ($G \approx F$) là tập các phụ thuộc hàm xác định trên R sao cho $G^+ = F^+$.



X^+ ?

- Bao đóng X^+ của thuộc tính X đối với tập phụ thuộc hàm F là tất cả các thuộc tính A mà phụ thuộc hàm $X \rightarrow A$ có thể được suy diễn logic từ F nhờ hệ tiên đề Amstrong.
- Một phụ thuộc hàm $X \rightarrow Y$ thuộc F^+ nếu Y thuộc X^+ : Kiểm tra $X \rightarrow Y$ có thuộc F^+



Ý nghĩa của phụ thuộc hàm

- Chỉ ra các phụ thuộc dữ liệu/ràng buộc có thể xảy ra giữa tập thuộc tính của một lược đồ quan hệ.
 - Giúp xác định khóa tối thiểu, khóa chính của quan hệ.
 - Giúp chuẩn hóa lược đồ quan hệ



2. Chuẩn hóa lược đồ quan hệ

○ Khái niệm

- Là quá trình phân tách các lược đồ quan hệ thành các lược đồ quan hệ nhỏ hơn theo một số **tiêu chuẩn** nhằm loại bỏ việc lưu trữ dư thừa dữ liệu.
- Phép tách thành các lược đồ quan hệ đơn giản hơn, nhỏ hơn phải đảm bảo **không làm mất mát thông tin**.



2.1. Các dạng chuẩn

- Dạng chuẩn 1
- Dạng chuẩn 2
- Dạng chuẩn 3
- Dạng chuẩn Boye-Codd
- *Chuẩn 4 và các dạng chuẩn khác*



a. Dạng chuẩn 1 (1NF)

○ Định nghĩa

- Một lược đồ quan hệ R ở dạng chuẩn 1 nếu và chỉ nếu toàn bộ các miền giá trị của các thuộc tính trong R đều chỉ chứa các giá trị nguyên tố.
- Một quan hệ xác định trên lược đồ quan hệ ở dạng chuẩn 1 được gọi là quan hệ ở dạng chuẩn 1.



Dạng chuẩn 1 (t)

- Một quan hệ thuộc dạng chuẩn 1 là một quan hệ trong đó mỗi miền giá trị của một thuộc tính chỉ chứa những giá trị nguyên tố (không phân chia được nữa).
- Một quan hệ thuộc dạng chuẩn 1 nếu mỗi một ô trong bảng chỉ chứa duy nhất một giá trị



Ví dụ

- Quan hệ CUNG_UNG_0 **chưa thuộc dạng chuẩn 1**

MaNCC	TenNCC	DiaChi	SanPham	Gia
1	Hải Hà	Hà Nội	Kẹo mềm Kẹo cứng Bánh	100 150 200
2	Kinh Đô	Hồ Chí Minh	Kẹo Bánh	120 200

Ví dụ(t)

- Quan hệ CUNG_UNG_1 **đã thuộc dạng chuẩn 1**

MaNCC	TenNCC	DiaChi	SanPham	Gia
1	Hải Hà	Hà Nội	Kẹo mềm	100
1	Hải Hà	Hà Nội	Kẹo cứng	150
1	Hải Hà	Hà Nội	Bánh	200
2	Kinh đô	Hồ Chí Minh	Kẹo	120
2	Kinh đô	Hồ Chí Minh	Bánh	200



b. Dạng chuẩn 2 (2NF)

○ Định nghĩa

- Một lược đồ quan hệ R được gọi là ở dạng chuẩn 2 nếu nó đã ở dạng chuẩn 1 và mọi thuộc tính không khóa đều phụ thuộc hàm đầy đủ vào khóa chính.
- Một quan hệ xác định trên lược đồ quan hệ ở dạng chuẩn 2 được nói là quan hệ ở dạng chuẩn 2.

Ví dụ

- Quan hệ CUNG_UNG_1 **chưa thuộc dạng chuẩn 2.**

Lặp →

MaNCC	TenNCC	DiaChi	SanPham	Gia
1	Hải Hà	Hà Nội	Kẹo mềm	100
1	Hải Hà	Hà Nội	Kẹo cứng	150
1	Hải Hà	Hà Nội	Bánh	200
2	Kinh đô	Hồ Chí Minh	Kẹo	120
2	Kinh đô	Hồ Chí Minh	Bánh	200

Lặp →



Ví dụ(t)

- $R(M, T, D, S, G) = \text{MTDSG} =$ Lược đồ của quan hệ CUNG_UNG_1
- Phụ thuộc hàm
 - $M \rightarrow \text{TD}, MS \rightarrow G$
- MS: Khóa tối thiểu
 - M, S: Thuộc tính khóa
 - T, D, G: Thuộc tính không khóa
- **T, D không phụ thuộc hàm đầy đủ vào MS**
- ➔ Lược đồ quan hệ CUNG_UNG_1 không thuộc dạng chuẩn 2.

Ví dụ

MaNCC	TenNCC	DiaChi
1	Hải Hà	Hà Nội
2	Kinh đô	Hồ Chí Minh

CUNG_UNG_11

MaNCC	SanPham	Gia
1	Kẹo mềm	100
1	Kẹo cứng	150
1	Bánh	200
2	Kẹo	120
2	Bánh	200

CUNG_UNG_12

- Ví dụ 2 thành quan hệ CUNG_UNG_11 và CUNG_UNG_12 tách từ quan hệ CUNG_UNG_1 **đã thuộc dạng chuẩn 2**.
 - TenNCC, DiaChi phụ thuộc hàm đầy đủ vào MaNCC
 - Gia phụ thuộc hàm đầy đủ vào {MaNCC, SanPham}



c. Dạng chuẩn 3

○ **Định nghĩa**

- Một lược đồ quan hệ R được gọi là ở dạng chuẩn 3 nếu nó đã ở dạng chuẩn 2 và mọi thuộc tính không khóa của R đều chỉ phụ thuộc hàm duy nhất vào khóa chính.
- Một quan hệ xác định trên lược đồ quan hệ ở dạng chuẩn ba được nói là quan hệ ở dạng chuẩn 3.

Ví dụ

MaNCC	TenNCC	DiaChi
1	Hải Hà	Hà Nội
2	Kinh đô	Hồ Chí Minh

CUNG_UNG_11

MaNCC	SanPham	Gia
1	Kẹo mềm	100
1	Kẹo cứng	150
1	Bánh	200
2	Kẹo	120
2	Bánh	200

CUNG_UNG_12

- Ví dụ 2 quan hệ CUNG_UNG_11 và CUNG_UNG_12 **đã thuộc dạng chuẩn 3**.
- MTDSG tách thành 1 lược đồ con MTD và MSG:
 - MTD: T, D phụ thuộc chỉ vào khóa M
 - MSG: G phụ thuộc hàm chỉ vào MS



Ví dụ

○ Ví dụ

- Lược đồ quan hệ:
 - $R(\text{Store, Item, Departement, Manager}) \Leftrightarrow R(S, I, D, M)$
- Tập các phụ thuộc hàm:
 - $F = \{SI \rightarrow D, SD \rightarrow M\}$
- Khóa tối thiểu: SI
- Lược đồ thuộc chuẩn 2: D, M phụ thuộc hàm đầy đủ vào SI
- Lược đồ **không thuộc chuẩn 3**: M phụ thuộc bắc cầu vào SI
 - $SI \rightarrow D; SI \rightarrow SD; SD \rightarrow M; SI \rightarrow M(*)$



d. Dạng chuẩn Boye-Codd (BCNF)

- Chuẩn 3 không đáp ứng được những lược đồ quan hệ
 - Có nhiều hơn một khóa tối thiểu
 - Các khóa tối thiểu là khóa kép
 - Các khóa tối thiểu giao nhau
- **Định nghĩa**
 - Một lược đồ quan hệ R thuộc dạng chuẩn Boye-Codd khi và chỉ khi mọi xác định hàm đều là một khóa.



Ví dụ

- Lược đồ quan hệ: R(CITY, STREET, ZIP)
- Phụ thuộc hàm:
 - CITY, STREET \rightarrow ZIP, ZIP \rightarrow CITY
- Khóa tối thiểu:
 - {CITY, STREET}, {STREET, ZIP}
- Thuộc tính khóa:
 - CITY, STREET, ZIP \rightarrow không có thuộc tính không khóa, thuộc dạng chuẩn 3.
- *Xác định* hàm ZIP không phải là một khóa \rightarrow lược đồ không thuộc chuẩn Boye-Codd

Ví dụ

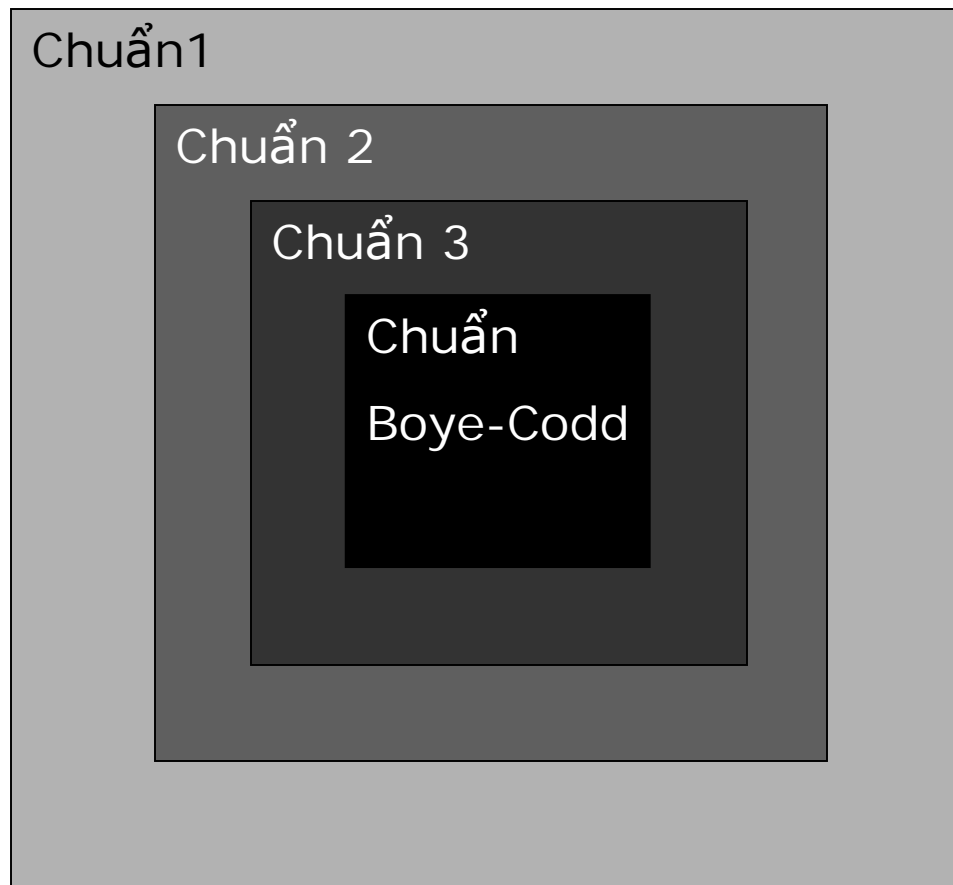
- 2 lược đồ con R1(CITY, ZIP) và R2(STREET, ZIP) thuộc dạng chuẩn Boye-Codd.

CITY	STREET	ZIP
Hà Nội	Láng	0423
Hà Nội	Phạm Hùng	0425
HCM	Võ Thị Sáu	0811

CITY	ZIP
Hà Nội	0423
Hà Nội	0425
HCM	0811

STREET	ZIP
Láng	0423
Phạm Hùng	0425
Võ Thị Sáu	0811

Quan hệ giữa các dạng chuẩn





2.2. Tách lược đồ quan hệ theo chuẩn

2.2.1. Tách bảo toàn thông tin và tách bảo toàn tập phụ thuộc hàm

2.2.2. Các thuật toán tách lược đồ quan hệ

2.2.3. Tách lược đồ quan hệ theo từng bước



2.2.1. Tách bảo toàn thông tin và tách bảo toàn tập phụ thuộc hàm

a. Phép tách bảo toàn thông tin

○ Khái niệm

- Phép tách bảo toàn thông tin là phép tách lược đồ quan hệ sao cho khi **kết nối tự nhiên** các quan hệ xác định trên các lược đồ con, kết quả cho lại quan hệ ban đầu.

Phép kết nối tự nhiên

- Phép ghép các cặp bộ của hai quan hệ trên các thuộc tính bằng nhau của hai quan hệ, một trong hai thuộc tính của phép so sánh "=" được loại bỏ sau khi thực hiện phép ghép.

- $R(A, B, C)$

a1	b1	1
a2	b2	2
a3	b3	3

- $S(D, E)$

1	e1
2	e2
3	e3

- R kết nối tự nhiên với S với $C=D$

a1	b1	1	e1
a2	b2	2	e2
a3	b3	3	e3

Ví dụ

- Lược đồ quan hệ MTDSG = (MaNCC, TenNCC, DiaChi, SanPham, Gia) \leftrightarrow (M, T, D, S, G);
- Tập phụ thuộc hàm $F = \{M \rightarrow TD, MS \rightarrow G\}$
- Quan hệ xác định trên lược đồ MTDSG:

MaNCC	TenNCC	DiaChi	SanPham	Gia
1	Hải Hà	Hà Nội	Kẹo mềm	10
1	Hải Hà	Hà Nội	Kẹo cứng	15
1	Hải Hà	Hà Nội	Bánh ngọt	40
1	Hải Hà	Hà Nội	Bánh mì	16
2	Kinh Đô	Hồ Chí Minh	Kẹo	12
2	Kinh Đô	Hồ Chí Minh	Bánh quy	45
3	Kinh Đô	Đà Nẵng	Bánh mì	10
3	Kinh Đô	Đà Nẵng	Bánh quy	30

Ví dụ(t)

MaNCC	TenNCC	DiaChi
1	Hải Hà	Hà Nội
2	Kinh Đô	Hồ Chí Minh
3	Kinh Đô	Đà Nẵng

MaNCC	SanPham	Gia
1	Kẹo mềm	10
1	Kẹo cứng	15
1	Bánh ngọt	40
1	Bánh mì	16
2	Kẹo	12
2	Bánh quy	45
3	Bánh mì	10
3	Bánh quy	30

- Phép tách bảo toàn thông tin
 - Lược đồ con 1: $MTD = (M, T, D)$
 - Lược đồ con 2: $MSG = (M, S, G)$

Ví dụ(t)

MaNCC	TenNCC	DiaChi
1	Hải Hà	Hà Nội
2	Kinh Đô	Hồ Chí Minh
3	Kinh Đô	Đà Nẵng

TenNCC	SanPham	Gia
Hải Hà	Kẹo mềm	10
Hải Hà	Kẹo cứng	15
Hải Hà	Bánh ngọt	40
Hải Hà	Bánh mỳ	16
Kinh Đô	Kẹo	12
Kinh Đô	Bánh quy	45
Kinh Đô	Bánh mỳ	10
Kinh Đô	Bánh quy	30

- Phép tách mất mát thông tin
 - Lược đồ con 1: $MTD = (M, T, D)$
 - Lược đồ con 2: $MSG = (M, S, G)$



Thuật toán kiểm tra

- *Kiểm tra phép tách không mất mát thông tin của một lược đồ quan hệ thành nhiều lược đồ quan hệ con.*
 - Vào
 - Lược đồ quan hệ: $R = (A_1, A_2, \dots, A_n)$
 - Tập phụ thuộc hàm F trên R
 - Phép tách $\rho = (R_1, R_2, \dots, R_k)$
 - Ra
 - Một khẳng định rằng phép tách ρ có mất mát thông tin hay không?



Thuật toán kiểm tra (t)

- Phương pháp

- Bước 1: Xây dựng một bảng k hàng, n cột
 - Hàng i tương ứng với R_i
 - Cột j tương ứng với thuộc tính A_j
 - Giá trị tại hàng i, cột j
 - a_j : Nếu A_j thuộc R_i
 - b_{ij} : Nếu A_j không thuộc R_i

Ví dụ (bước 1)

- Lược đồ quan hệ MTDSG
- Tập phụ thuộc hàm $F = \{M \rightarrow TD, MS \rightarrow G\}$
- Tách thành hai lược đồ MTD, MSG

	M	T	D	S	G
$R_1 = \text{MTD}$	a1	a2	a3	b14	b15
$R_2 = \text{MSG}$	a1	b22	b23	a4	a5



Phương pháp

- Bước 2: Lặp
 - Áp dụng các phụ thuộc hàm cho bảng vừa được xây dựng: $X \rightarrow Y$: Nếu tồn tại hai hàng có cùng giá trị trên X thì làm bằng nhau các giá trị trên Y.
 - Nếu có giá trị một hàng thuộc Y là a_j thì các giá trị khác thuộc Y gán bằng a_j .
 - Nếu không gán bằng một trong các giá trị b_{ij} .
- Dừng khi các giá trị trong bảng không thể thay đổi được nữa

Ví dụ (bước 2)

	M	T	D	S	G
$R_1 = \text{MTD}$	a1	a2	a3	b14	b15
$R_2 = \text{MSG}$	a1	b22	b23	a4	a5

○ Áp dụng phụ thuộc hàm $M \rightarrow TD$

- Thay $b_{22} = a_2$
- Thay $b_{23} = a_3$



Phương pháp

○ Bước 3: Kiểm tra

- Nếu trong bảng có một hàng gồm toàn ký hiệu $a_1, a_2, a_3, \dots, a_n$ thì phép tách là không mất mát thông tin.
- Ngược lại phép tách là mất mát thông tin.

Ví dụ

	M	T	D	S	G
$R_1 = \text{MTD}$	a1	a2	a3	b14	b15
$R_2 = \text{MSG}$	a1	a2	a3	a4	a5

- Phép tách trên là không mất mát thông tin.



Định lý kiểm tra

- *Kiểm tra phép tách không mất mát thông tin của một lược đồ quan hệ thành hai lược đồ quan hệ con.*
 - Cho
 - lược đồ quan hệ R
 - Tập phụ thuộc hàm F trên R
 - $\rho = (R_1, R_2)$
 - Phép tách là không mất mát thông tin nếu $R_1 \bowtie R_2 \rightarrow R_1 \setminus R_2$ hoặc $R_1 \bowtie R_2 \rightarrow R_2 \setminus R_1$.



Ví dụ

- Lược đồ MTDSG tách thành hai lược đồ con MTD và MSG:
 - $MTD \cap MSG = M$
 - $MTD \setminus MSG = TD$
 - $M \rightarrow TD$ thuộc F^+ : Phép tách là bảo toàn thông tin.
- Lược đồ MTDSG tách thành hai lược đồ con MTD và TSG
 - $MTD \cap TSG = T$
 - $MTD \cap SG = MD$
 - $TSG \setminus MTD = SG$
 - $T \rightarrow MD$ và $T \rightarrow SG$ không thuộc F^+ : Phép tách là mất mát thông tin.



b. Phép tách bảo toàn tập phụ thuộc hàm

○ Khái niệm

- Phép tách lược đồ quan hệ R thành các lược đồ quan hệ con R_i là bảo toàn các tập phụ thuộc hàm nếu hợp của các phụ thuộc hàm là hình chiếu của F trên R_i suy diễn logic được tất cả các phụ thuộc hàm trong F .
- Hình chiếu của F lên R_i là các phụ thuộc hàm $X \rightarrow Y$ thỏa mãn:
 - X, Y thuộc R_i
 - $X \rightarrow Y$ thuộc F^+

Ví dụ

- Cho
 - Lược đồ $R = ACBCD$
 - Tập phụ thuộc hàm $F = \{ A \rightarrow B, C \rightarrow D \}$
 - Phép tách $\rho = (R_1, R_2)$: $R_1 = AB, R_2 = CD$
- Phép tách bảo toàn tập phụ thuộc hàm
 - Hình chiếu của F lên R_1 : $A \rightarrow B, AB \rightarrow A, AB \rightarrow B \dots$
 - Hình chiếu của F lên R_2 : $C \rightarrow D, CD \rightarrow C, CD \rightarrow D \dots$
 - Các phụ thuộc hàm trong F có thể được suy diễn từ các hình chiếu này
- Phép tách không bảo toàn thông tin
 - $AB \bowtie CD = \Phi$
 - $AB \setminus CD = AB$
 - $CD \setminus AB = CD$



Ví dụ

- Cho
 - Lược đồ $R = CSZ$
 - Tập phụ thuộc hàm $F = \{CS \rightarrow Z, Z \rightarrow C\}$
 - Phép tách $\rho = (R_1, R_2)$: $R_1 = CZ, R_2 = SZ$
- Phép tách không bảo toàn tập phụ thuộc hàm
 - Hình chiếu của F lên R_1 : $Z \rightarrow C, CZ \rightarrow C, CZ \rightarrow Z, \dots$
 - Tập phụ thuộc hàm trong R_2 : $SZ \rightarrow S, SZ \rightarrow Z, \dots$
 - Phụ thuộc hàm $CS \rightarrow Z$ không thể được suy ra từ các hình chiếu này.
- Phép tách bảo toàn thông tin
 - $CZ \bowtie SZ \rightarrow CZ \setminus SZ$ hay $Z \rightarrow C$




Ví dụ

- Cho
 - Lược đồ $R = \text{MTDSG}$
 - Tập phụ thuộc hàm $F = \{M \rightarrow \text{TD}, \text{MS} \rightarrow \text{G}\}$
 - Phép tách $\rho = (R_1, R_2)$: $R_1 = \text{MTD}$, $R_2 = \text{MSG}$
- Phép tách bảo toàn tập phụ thuộc hàm
 - Hình chiếu của F lên R_1 : $M \rightarrow \text{TD}, \dots$
 - Hình chiếu của F lên R_2 : $\text{MS} \rightarrow \text{G}, \dots$
 - Các phụ thuộc hàm trong F có thể được suy diễn logic từ các hình chiếu này.
- Phép tách bảo toàn thông tin
 - Đã chứng minh




2.2.2. Các thuật toán tách lược đồ quan hệ

- a. Thuật toán tìm một khóa tối thiểu của lược đồ quan hệ dựa vào tập phụ thuộc hàm
- b. Thuật toán tách không mất mát thông tin và bảo toàn tập phụ thuộc hàm về dạng chuẩn 3
- c. Thuật toán tách không mất mát thông tin về dạng chuẩn Boye-Codd



a. Thuật toán tìm một khóa tối thiểu của lược đồ quan hệ dựa vào tập phụ thuộc hàm

- Khóa của lược đồ quan hệ
 - Giá trị của tập thuộc tính khóa trên mỗi bộ của quan hệ là duy nhất
 - Mọi thuộc tính của quan hệ phải phụ thuộc hàm vào khóa.
- Thuật toán tìm một khóa tối thiểu
 - Loại bỏ dần từng thuộc tính thuộc khóa của quan hệ cho tới khi tập thuộc tính nhỏ nhất còn lại vẫn thỏa mãn là một khóa → khóa tối thiểu của quan hệ.




b. Thuật toán tách không mất mát thông tin và bảo toàn tập phụ thuộc hàm về dạng chuẩn 3

○ Vào:

- Lược đồ quan hệ R
- Tập phụ thuộc hàm F (phủ tối thiểu)

○ Ra:

- Tập sơ đồ con, trong đó mỗi sơ đồ con
 - Thuộc dạng chuẩn 3
 - Phụ thuộc hàm là hình chiếu của F lên nó



b. Thuật toán tách không mất mát thông tin và bảo toàn tập phụ thuộc hàm về dạng chuẩn 3 (t)

○ Phương pháp:

- Nếu tồn tại một thuộc tính thuộc R không có mặt ở vế trái hay vế phải của bất kỳ phụ thuộc hàm nào thì tách thuộc tính này ra khỏi R.
- Nếu tồn tại một phụ thuộc hàm liên quan tới mọi thuộc tính của R thì kết quả là R.
- Nếu các phụ thuộc hàm dạng:
 - $X \rightarrow A$: lược đồ con ở dạng XA .
 - $X \rightarrow A_1, \dots, X \rightarrow A_n$: lược đồ con ở dạng $XA_1 \dots A_n$



Ví dụ

- Cho lược đồ $R = ABCDEG$
- Tập phụ thuộc hàm $F = \{AB \rightarrow C, C \rightarrow A, BC \rightarrow D, ACD \rightarrow B, D \rightarrow EG, BE \rightarrow C, CG \rightarrow BD, CE \rightarrow AG\}$
- Tập phụ thuộc hàm tối thiểu $F' = \{AB \rightarrow C, C \rightarrow A, BC \rightarrow D, D \rightarrow E, D \rightarrow G, BE \rightarrow C, CG \rightarrow B, CE \rightarrow G\}$
- Phép tách $\rho = (ABC, CA, BCD, DEG, BEC, CGB, CEG) = (ABC, BCD, DEG, BEC, CGB, CEG)$ gồm các lược đồ con đã ở dạng chuẩn 3 bảo toàn tập phụ thuộc hàm và bảo toàn thông tin.



C. Thuật toán tách không mất mát thông tin về dạng chuẩn Boye-Codd

○ Vào :

- Lược đồ quan hệ R
- Tập phụ thuộc hàm F

○ Ra:

- Tập sơ đồ con, trong đó mỗi sơ đồ con
 - Thuộc dạng chuẩn Boye-Codd
 - Phụ thuộc hàm là hình chiếu của F lên nó



C. Thuật toán tách không mất mát thông tin về dạng chuẩn Boye-Codd (t)

○ Phương pháp:

- $\rho = (R)$
- Lặp
 - S là một sơ đồ quan hệ trong ρ không ở dạng chuẩn Boye-Codd. Xét một phụ thuộc hàm $X \rightarrow A$ của S. Nếu X không chứa khóa của S, A không thuộc X thì S được tách thành:
 - $S_1 = A \cup \{X\}$
 - $S_2 = S \setminus \{A\}$
- Dừng cho tới khi mọi sơ đồ con trong ρ đã thuộc dạng chuẩn Boye-Codd



2.2.3. Tách lược đồ quan hệ theo từng bước

- Quy tắc

- *Quy tắc 1:* Loại bỏ các thuộc tính phụ thuộc chỉ một phần vào khóa chính.

- ➔ Chuẩn hóa về 2NF

- *Quy tắc 2:* Loại bỏ các thuộc tính không khóa không phụ thuộc vào khóa chính.

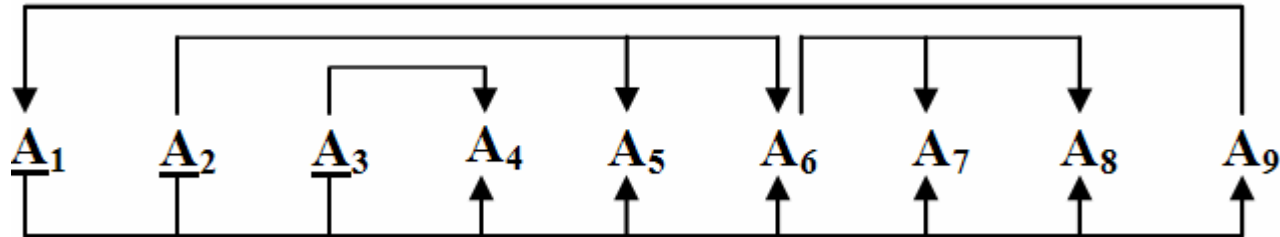
- ➔ Chuẩn hóa về 3NF

- *Quy tắc 3:* Loại bỏ các thuộc tính là giao của các khóa tối thiểu.

- ➔ Chuẩn hóa về BCNF

Ví dụ

- Cho quan hệ $R(A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9)$ trong đó $A_1A_2A_3$ là khóa với sơ đồ phụ thuộc hàm:



- Quan hệ R ở dạng chuẩn nào? Tại sao? Tách R thành các quan hệ ở dạng chuẩn BCNF.



3. Ràng buộc toàn vẹn trong CSDL quan hệ (Integrity Constraint)

3.1. Khái niệm ràng buộc toàn vẹn

- *Ràng buộc toàn vẹn là điều kiện bất biến không được vi phạm trong một CSDL.*
 - Các điều kiện mà mọi thể hiện của quan hệ phải thỏa mãn
 - RBTV xuất phát từ các quy tắc quản lý được áp đặt lên thế giới thực
- Ví dụ:
 - RBTV1: Mỗi bộ của quan hệ DON_DAT_HANG phải ứng với một đơn đặt hàng với mã đơn đặt hàng (MaDDH) duy nhất.
 - RBTV2: Mọi chi tiết về đơn đặt hàng phải có mã hàng (MaHang) thuộc về danh mục hàng.
 - RBTV3: Mã đơn đặt hàng (MaDDH) khác NULL.
 - RBTV4: Tổng các trị giá của các mặt hàng trong CHI_TIET_DON_HANG có cùng MaDDH phải bằng TongGT ghi trong DON_DAT_HANG
- Mục đích của RBTV
 - Đảm bảo tính nhất quán của dữ liệu(RBTV2, RBTV4)
 - Đảm bảo ngữ nghĩa thực tế của dữ liệu(RBTV1, RBTV3)



Khái niệm ràng buộc toàn vẹn

- RBTV được xác định và mô tả trong quá trình thiết kế csdl. RBTV có 3 yếu tố chính:
 - Nội dung
 - Bối cảnh
 - Tâm ảnh hưởng
- RBTV được khai báo thông qua ngôn ngữ định nghĩa và thao tác dữ liệu và được hỗ trợ bởi hqt csdl.
 - Mệnh đề check, arssertion, triger
- RBTV được kiểm tra và xử lý bởi hqt csdl.



Nội dung của RBTV

- Được phát biểu
 - Ngôn ngữ tự nhiên: Đơn giản dễ hiểu
 - Ngôn ngữ hình thức: khó hiểu
 - Đại số quan hệ, phép tính quan hệ, mã giả
 - Biểu thức toán học
- Ví dụ:
 - RBTV5:
 - Mỗi nhân viên có một mã số dùng để phân biệt với những nhân viên khác
 - $$\forall t_1, t_2 \in NHAN_VIEN (t_1 \neq t_2 \Rightarrow t_1.MaNV \neq t_2.MaNV)$$
 - RBTV6:
 - Mỗi nhân viên làm việc trong một phòng ban
 - $$NHAN_VIEN[MaPhong] \in PHONG_BAN[MaPhong]$$
 - RBTV7:
 - Mỗi phòng phải có ít nhất một nhân viên
 - $$\forall s \in PHONG_BAN (\exists t \in NHAN_VIEN (t.MaPhong = s.MaPhong))$$



Bối cảnh của RBTV

- Là những quan hệ mà RBTV có hiệu lực: Một hoặc nhiều quan hệ
- Ví dụ
 - RBTV5 có bối cảnh là quan hệ NHAN_VIEN
 - RBTV6, RBTV7 có bối cảnh là quan hệ NHAN_VIEN, PHONG_BAN



Tầm ảnh hưởng

- RBTV có thể bị vi phạm khi thực hiện các thao tác cập nhật trên bối cảnh: Thêm, xóa, sửa
- Bảng tầm ảnh hưởng dùng để xác định thời điểm cần kiểm tra RBTV

Xây dựng bảng tầm ảnh hưởng cho từng RBTV

- Nội dung mỗi ô
 - +: Cần phải kiểm tra RBTV
 - - : Không cần phải kiểm tra RBTV

Tên RBTV	Thêm	Xóa	Sửa
Quan hệ 1	+	+	-
...
Quan hệ k	+	-	-

Các quan hệ
bối cảnh

Ví dụ

RBTV5	Thêm	Xóa	Sửa
NHAN_VIEN	+	-	-

RBTV6	Thêm	Xóa	Sửa
NHAN_VIEN	+	-	+
PHONG_BAN	-	+	-

RBTV7	Thêm	Xóa	Sửa
NHAN_VIEN	-	-	+
PHONG_BAN	+	-	-

Xây dựng bảng tầm ảnh hưởng tổng hợp

- Xây dựng trên cơ sở bảng tầm ảnh hưởng của các RBTV
- Để xác định thời điểm kiểm tra RBTV khi một thao tác cập nhật trên quan hệ nào đó được thực hiện

	Tên RBTV 1					Tên RBTV r		
	T	X	S	...		T	X	S
Quan hệ 1	+	+	-		+	-	+	
...	
Quan hệ n	-	-	+		+	+	+	



Ví dụ

- Xây dựng bảng tầm ảnh hưởng cho các ràng buộc toàn vẹn RBTV1..., RBTV7 cho lược đồ CSDL quan hệ siêu thị M ?



Phân loại RBTV

- Dựa trên yếu tố bối cảnh của RBTV
 - RBTV có bối cảnh là một quan hệ/RBTV trên thuộc tính
 - RBTV miễn giá trị
 - RBTV liên thuộc tính
 - RBTV liên bộ
 - RBTV có bối cảnh là nhiều quan hệ/RBTV trên quan hệ
 - RBTV tham chiếu
 - RBTV thuộc tính tổng hợp
 - RBTV liên thuộc tính
 - RBTV liên bộ

3.2. Ràng buộc toàn vẹn trên thuộc tính

a. Ràng buộc toàn vẹn về miền giá trị của một thuộc tính

- RBTV đặc tả tập giá trị có thể kết hợp với một thuộc tính.
- Ví dụ: quan hệ NHAN_VIEN
 - RBTV8: Tuổi của nhân viên trong công ty phải lớn hơn 18 và nhỏ hơn 65.

RBTV8	T	X	S
$\forall t \in NHAN_VIEN(18 \leq t.Tuoi \leq 65)$			
NHAN_VIEN	+	-	+

b. RBTV liên thuộc tính

- RBTV có liên quan tới nhiều thuộc tính của một quan hệ. Thông thường đó là các phụ thuộc tính toán, hoặc một suy diễn từ giá trị của một hay nhiều thuộc tính trong cùng một bộ giá trị.
- Ví dụ
 - Quan hệ NHAN_VIEN
 - RBTV9: Nếu nhân viên có giới tính là nữ tuổi của nhân viên trong công ty phải lớn hơn 18 và nhỏ hơn 55.

$\forall t \in NHAN_VIEN (18 \leq t.Tuoi \leq 55 (t.GioiTinh = Nu))$

RBTV9	T	X	S
NHAN_VIEN	+	-	+



c. RBTV liên bộ, liên thuộc tính

- RBTV có liên quan tới nhiều bộ và có thể tới nhiều thuộc tính của (các) bộ giá trị trong một quan hệ.
- Ví dụ: RBTV5



3.3. Ràng buộc toàn vẹn trên quan hệ

a. RBTV tham chiếu/RBTV về phụ thuộc tồn tại (phụ thuộc về khóa ngoại)

- Bộ giá trị của quan hệ này được thêm vào một cách hợp lệ nếu tồn tại một bản ghi tương ứng trong một quan hệ khác.
- Đảm bảo rằng giá trị xuất hiện trong một quan hệ đối với một tập các thuộc tính đã cho cũng xuất hiện đối với một tập các thuộc tính nhất định trong một quan hệ khác.

a. Ràng buộc toàn vẹn về phụ thuộc tồn tại (phụ thuộc về khóa ngoại)

○ Ví dụ:

- RBTV1 : Mỗi bộ của CHI_TIET_DON_HANG phải có một đơn hàng với MaDDH tương ứng

$$\forall t \in CHI_TIET_DON_HANG (\exists u \in DON_HANG (t.MaDDH = u.MaDDH))$$

RVTV4	T	X	S
CHI_TIET_DON_HANG	+	-	-
DON_HANG	-	+	-



Ví dụ

- RBTV2 : Mỗi bộ của CHI_TIET_DON_HANG phải có MaHang thuộc về danh mục hàng trong quan hệ MAT_HANG
- Biểu diễn hình thức ? Bảng xác định tầm ảnh hưởng?



b. *Ràng buộc toàn vẹn tổng hợp*

- Khi có sự hiện diện của 1 thuộc tính mang tính chất tổng hợp (tức là giá trị của thuộc tính có thể được tính toán từ giá trị của các thuộc tính khác trên một hay nhiều bộ giá trị của các quan hệ trong CSDL)
- Ví dụ:
 - DON_DAT_HANG(MaDDH, NgayLap, TenKH, TongGT)
 - CHI_TIET_DON_HANG(MaDDH, TenSanPham, SoLuong, Giatri).
 - RBTV4: Tổng các trị giá của các mặt hàng trong CHI_TIET_DON_HANG có cùng MaDDH phải bằng TongGT ghi trong DON_DAT_HANG .

$\forall t \in DON_DAT_HANG(t.TongGT = \sum u.GiaTri(u \in CHI_TIET_DON_HANG \wedge u.MaDDH = t.MaDDH))$

RVTV4	T	X	S
CHI_TIET_DON_HANG	+	+	+
DON_DAT_HANG	+	-	+



c. RBTV liên thuộc tính

- Mỗi quan hệ giữa các thuộc tính trong nhiều lược đồ quan hệ
- Ví dụ
 - Quan hệ:
 - NHAN_VIEN (TenNV, Luong, NgaySinh, TenPhong)
 - PHONG_BAN (TenPhong, MaPhong, NguoiQuanLy, NgayNhanChuc)
 - RBTV 10:
 - Ngày nhận chức của trưởng phòng phải lớn hơn ngày sinh
 - Biểu diễn hình thức. Bảng xác định tầm ảnh hưởng?



d. Ràng buộc toàn vẹn liên bộ

- Mỗi quan hệ giữa các bộ trên nhiều lược đồ quan hệ
- Ví dụ
 - Quan hệ NHAN_VIEN, PHONG_BAN
 - RBTV11:
 - Lương của nhân viên không được cao hơn lương trưởng phòng
 - Biểu diễn hình thức? Bảng xác định tầm ảnh hưởng?



Chỉ mục

- Functional dependency : Phụ thuộc hàm
- Functional determinant: Xác định hàm
- Normal Form: Dạng chuẩn



Phụ lục

- Thuật toán tìm phủ tối thiểu của một tập phụ thuộc hàm
 - Thuật toán 4.2 (128, Nguyên lý của các hệ csdl)