CƠ SỞ DỮ LIỆU

GIÁO VIÊN: ĐỖ THỊ MAI HƯỜNG

BỘ MÔN: CÁC HỆ THỐNG THÔNG TIN

KHOA: CÔNG NGHỆ THÔNG TIN

CHƯƠNG 3

Lý thuyết thiết kế cơ sở dữ liệu quan hệ

Nội dung chi tiết

- Giới hạn của ER
- Sự dư thừa
- Phụ thuộc hàm
- Hệ suy diễn Amstrong
- Thuật toán tìm bao đóng
- Thuật toán tìm khóa
- Các dạng chuẩn
- Kiểm tra kết nối không mất thông tin

Sự dư thừa

- Sự phụ thuộc giữa các thuộc tính gây ra sự dư thừa

  Ví dụ: Điểm các môn học → Điểm trung bình → xếp loại
- Thuộc tính đa trị trong lược đồ ER → nhiều bộ số liệu trong lược đồ quan hệ
- Ví dụ:

  NHANVIEN(TENNV, HONV, NS,DCHI,GT,LUONG, BANGCAP)

TENNV	HONV	NS	DCHI	GT	LUONG	BANGCAP
Tung	Nguyen	12/08/1955	638 NVC Q5	Nam	40000	Trung học
Nhu	Le	06/20/1951	291 HVH QPN	Nu	43000	Trung học
Nhu	Le	06/20/1951	291 HVH QPN	Nu	43000	Đại học
Hung	Nguyen	09/15/1962	Ba Ria VT	Nam	38000	Thạc sỹ

Sự dư thừa (tt)

- Sự dư thừa → sự dị thường
  - Thao tác sửa đổi: cập nhật tất cả các giá trị liên quan
  - Thao tác xóa: người cuối cùng của đơn vị → mất thông tin về đơn vị
  - Thao tác chèn

TENPB	MAPB	MaTP	NG_NHANCHUC	MANV	TENNV	HONV	...
Nghien cuu	5	NV01	05/22/1988	NV01	Tung	Nguyen	...
Dieu hanh	4	NV02	01/01/1995	NV02	Hung	Nguyen	...
Quan ly	1	NV03	06/19/1981	NV03	Vinh	Pham	...

Sự dư thừa (tt)

- Các giá trị không xác định
  - Đặt thuộc tính Trưởng phòng vào quan hệ NHANVIEN thay vì vào quan hệ PHONGBAN
- Các bộ giả
  - Sử dụng các phép nối

Sự dư thừa (tt)

- Một số quy tắc

  1.Rõ ràng về mặt ngữ nghĩa, tránh các phụ thuộc giữa các thuộc tính với nhau

  2.Tránh sự trùng lặp về nội dung →đảm bảo tránh được các dị thường khi thao tác cập nhật dữ liệu

  3. Tránh đặt các thuộc tính có nhiều giá trị Null
     - Khó thực hiện các phép nối và kết hợp

  4. 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"

=> Lý thuyết về chuẩn hóa: (dựa trên phụ thuộc hàm, …) sẽ là nền tảng cơ sở để thực hiện việc phân tích và chuẩn hóa lược đồ quan hệ

Phụ thuộc hàm

Phụ thuộc hàm trong quan hệ r

- Cho lược đồ quan hệ R và X, Y là các tập con của R. r là một quan hệ trên R.
- Ta nói X xác định phụ thuộc hàm Y ký hiệu $X \rightarrow Y$ trong r nếu với mọi t và t' của r mà t, t' bằng nhau trên tập X thì chúng cũng bằng nhau trên tập Y, tức là $\forall$ t, t' $\in$ r nếu $t.X = t'.X \Rightarrow t.Y = t'.Y$
- Ví dụ:
  - X={MaNV}, Y={Hoten,NS} thỏa mãn $X \rightarrow Y$
  - X={Hoten}, Y={DC, GT} không thỏa mãn $X \rightarrow Y$
- Phụ thuộc hàm trên r là trường hợp riêng của phụ thuộc hàm trên R.

Phụ thuộc hàm(tt)

- **Phụ thuộc hàm trong quan hệ r**

Ví dụ: trong lược đồ quan hệ sau nếu giả thiết Hoten nhập vào là khác nhau thì từ Hoten có thể suy diễn ra tất cả các thuộc tính khác. Nhưng nếu thêm vào bộ có Hoten giống với bộ đã có thì phụ thuộc hàm không còn đúng nữa.

HoTen	NgaySinh	MaPB	TenPB
Nguyễn Văn A	1/1/1980	PB01	Hành chính
Nguyễn Văn B	20/2/1981	PB02	Tổng hợp
Trần C	13/6/1981	PB03	Dự án

Trần C	10/2/1982	PB01	Hành chính

Phụ thuộc hàm(tt)

Bài tập:

Cho bảng quan hệ r như sau:

A	B	C	D
x	u	x	y
y	x	z	x
z	y	y	y
y	z	w	z

Trong các phụ thuộc hàm sau PTH nào không thỏa mãn r

$A \rightarrow B, A \rightarrow C, B \rightarrow A, C \rightarrow D, D \rightarrow C, D \rightarrow A$

Phụ thuộc hàm(tt)

Phụ thuộc hàm trên lược đồ quan hệ R

- Cho lược đồ quan hệ R và X, Y là các tập con của R. Ta nói X xác định phụ thuộc hàm Y ký hiệu X → Y trên lược đồ quan hệ R. Nếu với mọi r trên R xác định X → Y .

Phụ thuộc hàm(tt)

Các tính chất của phụ thuộc hàm

- **A1. Tính phản xạ**

  $X \rightarrow X$, hay tổng quát hơn nếu $Y \subset X$ thì $X \rightarrow Y$
- **A2. Tính mở rộng hai vế**

  $X \rightarrow Y$ thì $XZ \rightarrow YZ$. (Mở rộng hai vế Z)
- **A3. Tính bắc cầu**: $X \rightarrow Y$ và $Y \rightarrow Z$ thì $X \rightarrow Z$.
- Hệ A bao gồm các tính chất {A1, A2, A3} của phụ thuộc hàm được gọi là hệ tiên đề Armstrong của lớp các phụ thuộc hàm.

Phụ thuộc hàm(tt)

Giả sử $t, t' \in r$

1. **Tính phản xạ**: hiển nhiên vì t và t’ đã bằng nhau trong tập X thì chúng phải bằng nhau trong mọi tập con của X, nói cách khác $t.X=t'.X \Rightarrow t.X=t'.X$ & $t.Y=t'.Y$ với $Y \subset X \Rightarrow X \rightarrow Y$

2. **Tính mở rộng 2 vế**: giả sử $t.XZ=t'.XZ$ ta phải chứng minh $t.YZ=t'.YZ$

Thật vậy từ $t.XZ=t'.XZ$ ta có $t.X=t'.X$ và $t.Z=t'.Z$. Theo giả thiết $t.X=t'.X$ thì $t.Y=t'.Y$. Như vậy ta có $t.Y=t'.Y$ và $t.Z=t'.Z$ thì $t.YZ=t'.YZ$. Suy ra $XZ \rightarrow YZ$

3. Tính bắc cầu:

$t.X = t'.X \Rightarrow t.Y = t'.Y$

$t.Y = t'.Y \Rightarrow t.Z = t'.Z$

$\Rightarrow$ $t.X = t'.X$ thì $t.Z = t'.Z \Leftrightarrow X \rightarrow Z$

Các tính chất bổ sung từ Hệ Amstrong

Các tính chất sau đều được suy ra từ hệ tiên đề Armstrong.

Tính tựa bắc cầu:

$X \rightarrow Y$ và $YZ \rightarrow W$ thì $XZ \rightarrow W$

Tính chất chiếu:

$X \rightarrow YZ$ thì $X \rightarrow Y$ và $X \rightarrow Z$

Tính cộng đầy đủ:

$X \rightarrow Y$ và $Z \rightarrow W$ thì $XZ \rightarrow YW$

Hệ tiên đề Armstrong

- Chứng minh:
- Tính chất tựa bắc cầu: X → Y và YZ → W thì XZ → W

X → Y theo tính mở rộng hai vế

XZ → YZ

Và YZ → W

Theo tính bắc cầu

XZ → W

Hệ tiên đề Amstrong

Bài tập: Chứng minh các tính chất còn lại

- **Tính chất chiếu:**

  $X \rightarrow YZ$ thì $X \rightarrow Y$ và $X \rightarrow Z$

- **Tính cộng đầy đủ:**

  $X \rightarrow Y$ và $Z \rightarrow W$ thì $XZ \rightarrow YW$

Hệ tiên đề Armstrong

- **Phép suy dẫn theo hệ tiên đề Armstrong**

PTH f được suy dẫn theo hệ tiên đề Armstrong là f có thể nhận được từ F sau một số hữu hạn bước áp dụng các luật của tiên đề Armstrong. Ký hiệu F |= f.

- **Phép suy dẫn theo quan hệ**

PTH f suy dẫn được từ tập PTH F theo quan hệ (hoặc PTH f được suy dẫn theo quan hệ từ tập PTH F) ký hiệu F |- f, nếu với mọi quan hệ r trên lược đồ R mà F thỏa mãn thì f cũng thỏa mãn.

Hệ tiên đề Armstrong

Bổ đề: Giả sử $X \subseteq R$, nếu gọi X^+ là tập các thuộc tính A của R mà $F \models X \rightarrow A$ thì với mọi tập $Y \subseteq R$, $F \models X \rightarrow Y \Leftrightarrow Y \subseteq X^+$.

a. Chứng minh chiều thuận

Ta có $F \models X \rightarrow Y$. Giả sử Y={A, B, C, ...} theo tính mở rộng trái thu hẹp phải:

$F \models X \rightarrow A$, nên theo định nghĩa X^+ ta có $A \in X^+$.

$F \models X \rightarrow B$, nên theo định nghĩa X+ ta có $B \in X^+$.

$F \models X \rightarrow C$, nên theo định nghĩa X+ ta có $C \in X^+$.

..., vậy $\{A, B, C, ...\} = Y \subset X^+$.

b. Chứng minh điều ngược lại

$Y \subset X^+$. Theo định nghĩa tập X^+ thì mọi $A \in Y$ ta có $F \models X \rightarrow A$, vậy theo tính chất cộng đầy đủ ta có $F \models X \rightarrow Y$.

Hệ tiên đề Armstrong

Định lý:

Cho tập PTH F và một PTH f trên R khi đó ta có

F |- f khi và chỉ khi F |= f.

Chứng minh:

- Giả sử có F $|= X \rightarrow Y$ cần chứng minh F $|- X \rightarrow Y$
- Theo bổ đề ta có $Y \subseteq X^+$. Để chứng minh F $|- X \rightarrow Y$, ta lấy một quan hệ R tuỳ ý thoả mãn tất cả các fds của F và ta phải chứng minh R thoả $X \rightarrow Y$.
- Ta lấy 2 thực thể bất kỳ t, t' của R mà t[X]=t'[X], ta phải chứng tỏ t[Y]=t'[Y] mà $Y \subseteq X^+$ nên t[Y]=t'[Y] (đpcm)

Hệ tiên đề Armstrong

Giả sử có F $|-X \rightarrow Y$ chứng minh F $|=X \rightarrow Y$, hay chỉ cần chứng minh $Y \subseteq X^+$

Nhận xét: Nếu $X' \subseteq X^+$ thì $(X')^+ \subseteq X^+$.

Xét một quan hệ r trên tập thuộc tính R

r	*X*	*R-X*
t1	111…111	111…111
t2	111…111	000…000

Hệ tiên đề Armstrong

- Ta thấy R thoả mãn tất cả các phụ thuộc hàm của F. Vì lấy một pth $P \rightarrow Q$ của F thì R thoả $P \rightarrow Q$. Thật vậy:
  - *TH1*: P không là tập con của $X^+ \Rightarrow$R thoả $P \rightarrow Q$ vì t[P]=t'[P] thì $t \equiv t'$ và t[Q]=t'[Q]
  - *TH2:* $P \subseteq X^+ \Rightarrow P^+ \subseteq X^+$
    - Nếu $t \equiv t'$ thì t[Q]=t'[Q]
    - Nếu $t \neq t'$ ta có thể giả thiết t=t1 và t'=t2. Do $P \rightarrow Q$ thuộc F nên $Q \subseteq P^+$ hay t[Q]=t'[Q]
    - Vậy trong mọi trường hợp R thoả các phụ thuộc hàm của F.
    - Do giả thiết F $|$-$X \rightarrow Y$, mà R thoả tất cả các fds của F, R cũng thoả fd $X \rightarrow Y$.
    - Do t1[X]=t2[X] nên t1[Y]=t2[Y] suy ra $Y \subseteq X^+ \rightarrow$ đpcm.
- KL: Hai phương pháp suy dẫn là tương đương nhau.

Hệ tiên đề Armstrong

Bài tập:

Cho lược đồ quan hệ <R,F>

- R={A, B, C, D, E, I} và
- F={BC$\rightarrow$DE, BE $\rightarrow$C, BI$\rightarrow$A, CE $\rightarrow$I}

a. Chứng minh F|- BC$\rightarrow$I

b. Chứng minh F|= BC$\rightarrow$I

Bao đóng F^+ của tập PTH F

4. Bao đóng F^+ của tập PTH F

Tập PTH được suy dẫn từ F được gọi là bao đóng của tập PTH F, ký hiệu F^+.

Ví dụ:

R={A, B, C, D}

$F=\{A \rightarrow B, B \rightarrow C, A \rightarrow D, B \rightarrow D\}$

$F^+ = \{A \rightarrow B, B \rightarrow C, A \rightarrow D, B \rightarrow D, A \rightarrow C, A \rightarrow BD, A \rightarrow BCD, A \rightarrow BC, A \rightarrow CD, B \rightarrow CD\}$

Các tính chất của F^+

a. Tính phản xạ: $F \subseteq F^+$

b. Tính đơn điệu: $F \subseteq G \Rightarrow F^+ \subseteq G^+$

c. Tính lũy đẳng: $F^{++}=F^+$

Bao đóng X^+

- Định nghĩa bao đóng X^+

Cho lược đồ quan hệ R = $\{A_1, ..., A_n\}$. Giả sử F là tập PTH trên R. X là tập con của tập thuộc tính R.

Bao đóng X đối với F, ký hiệu X^+ (X^+_F để chỉ bao đóng lấy theo tập F) là tập thuộc tính A của R mà $X \rightarrow A$ được suy dẫn từ tập F.

$X^+ = \{A: A \in R \text{ và } X \rightarrow A \in F^+\}$

hoặc $X^+ = \{X \cup A: A \in R, A \notin X \text{ và } X \rightarrow A \in F^+\}$

Ví dụ:

$R = \{A, B, C, D, E, G\}$

$F = \{A \rightarrow C, A \rightarrow EG, B \rightarrow D, G \rightarrow E\}$

$X = \{A, B\}$

$Y = \{C, D, G\}$

$X^+ = \{A, B, C, D, E, G\}$

$Y^+ = \{C, D, E, G\}$

Tính chất của bao đóng X+

1. Tính phản xạ: $X \subseteq X^+$
2. Tính đơn điệu: $X \subseteq Y \Rightarrow X^+ \subseteq Y^+$.
3. Tính lũy đẳng: $X^{++} = X^+$
4. Bao đóng tổng chứa tổng các bao đóng: $X^+Y^+ \subseteq (XY)^+$
5. $(X^+Y)^+ = (XY^+)^+ = (X^+Y^+)^+ = (XY)^+$
6. $X \rightarrow Y \Rightarrow Y \subseteq X^+$
7. $X \rightarrow Y \Rightarrow Y^+ \subseteq X^+$
8. $X \rightarrow X^+$ và $X^+ \rightarrow X$
9. $X^+ = Y^+ \Leftrightarrow X \rightarrow Y$ và $Y \rightarrow X$.

Thuật toán tìm bao đóng X+

- **Bài toán thành viên**

Vấn đề được đưa ra ở đây là: Cho trước một tập PTH F có hay không một khẳng định f $\in F^+$. Để giải quyết bài toán này người ta sử dụng tính chất 6 của tập bao đóng hay bổ đề: $X \rightarrow Y \in F^+ \Leftrightarrow Y \subset X^+$.

Do vậy chỉ cần tìm được X^+ ta sẽ giải quyết được bài toán $X \rightarrow Y$ có thuộc F^+.

Thuật toán tìm bao đóng X^+

- **Thuật toán tìm bao đóng X^+**

Thuật toán tìm bao đóng X^+ của Beeri và Bernstein

Cho $R = \{A_1, ..., A_n\}$. F là tập PTH trên R. X là tập thuộc tính.

Ta xây dựng tập $X^0, .., X^k$ như sau:

$X^0=X$

$X^{(i+1)} = X^i Z^i$ với $Z^i=\{A: A \in R \setminus X^i$ và $X^i \rightarrow A \in F^+\}$ i=0, 1, 2, ...

Tập $X^0, X^1, ...$ là tập tăng dần và tập R là hữu hạn nên sau hữu hạn bước thuật toán phải kết thúc. Tồn tại $X^k = X^{k+1} = ...$ Chính X^k là tập X^+.

Thuật toán tìm bao đóng X^+

Thuật toán

Input: Lược đồ quan hệ R

Tập PTH F, Tập thuộc tính X

Output: Tập X^+

Begin

Y:=X

Repeat

Z:=∅

For each A in R do

If ($A \notin Y$ and $Y \rightarrow A \in F^+$) then $Z= Z \cup A$;

$Y := Y \cup Z$;

Until $Z = \emptyset$;

$X^+ = Y$

End;

Thuật toán tìm bao đóng X^+

Ví dụ

Cho R = {A, B, C, D, E, G}

Cho tập PTH F ={$AB \rightarrow C$, $BC \rightarrow D$, $ACD \rightarrow B$, $D \rightarrow EG$, $BE \rightarrow C$, $CG \rightarrow DB$, $CE \rightarrow AG$}

X = {B, D}

$X^0 = \{B, D\}$, $Z^0 = \{E, G\}$ ($D \rightarrow EG \in F$)

$X^1 = \{B, D, E, G\}$, $Z^1 = \{C\}$ ($BE \rightarrow C \in F$)

$X^2 = \{B, C, D, E, G\}$, $Z^2 = \{A\}$ ($CE \rightarrow A \in F^+$ Hoặc $CE \rightarrow AG \in F$)

$X^3 = \{A, B, C, D, E, G\}$ $Z^3 = \varnothing$

$X^+ = X^3$

Thuật toán tìm bao đóng X^+

Chứng minh tính đúng đắn của thuật toán

Chứng minh: $X^+ \subset X^k$ và $X^k \subset X^+$.

a. $X^+ \subset X^k$

Thật vậy lấy $A \in X^+$. Như trên ta thấy $X^+=XZ$ với $Z = \{A: A \notin X \text{ và } X \rightarrow A \in F^+\}$

Nếu $A \in X$ thì $A \in X^k$ vì $X \subset X^k$.

Nếu $A \in Z$ thì theo định nghĩa các tập Z^i, tồn tại một chỉ số i để $A \in Z^i$ vậy $A \in X^k \Rightarrow X^+ \subset X^k$.

b. $X^k \subset X^+$

$X^0 \rightarrow X^1 \rightarrow \ldots \rightarrow X^k \Rightarrow X \rightarrow X^k \Rightarrow X^k \subset X^+$

Bài tập về phụ thuộc hàm và bao đóng

1. Cho lược đồ quan hệ <R,F>

- R={A, B, C, D, E, I} và
- $F=\{BC \rightarrow DE, BE \rightarrow C, BI \rightarrow A, CE \rightarrow I\}$

a. Chứng minh $F \vdash BC \rightarrow I$

b. Chứng minh $F \models BC \rightarrow I$

c. Tìm bao đóng của BC, BE,BI,CE

d. Chứng minh $BC \rightarrow A \in F^+$

2. Cho $F = \{AB \rightarrow E, AG \rightarrow I, E \rightarrow G, GI \rightarrow H\}$.

- Chứng minh rằng $AB \rightarrow GH \in F^+$

Bài tập

3. Cho $F = \{AB \rightarrow C, B \rightarrow D, CD \rightarrow E, CE \rightarrow GH, G \rightarrow A\}$.

a. Chứng minh rằng $AB \rightarrow E \in F^+$

b. Chứng minh rằng $AB \rightarrow G \in F^+$?

4. Cho $F = \{XY \rightarrow W, Y \rightarrow Z, WZ \rightarrow P, WP \rightarrow QR, Q \rightarrow X\}$.

Chứng minh rằng $XY \rightarrow P \in F^+$

Bài tập

5. Cho bảng quan hệ r như sau:

A	B	C	D
x	u	x	y
y	x	z	x
z	y	y	y
y	z	w	z

Trong các phụ thuộc hàm sau PTH nào không thỏa mãn r

$A \rightarrow B, A \rightarrow C, B \rightarrow A, C \rightarrow D, D \rightarrow C, D \rightarrow A$

Khóa của sơ đồ quan hệ

Trong lược đồ quan hệ một số thuộc tính đóng vai trò quan trọng, và từ các thuộc tính này có thể xác định các thuộc tính khác.

Khái niệm sơ đồ quan hệ:

Cho lược đồ quan hệ: R ={A_1, A_2, ..., A_n}, F là tập các phụ thuộc hàm trên R. Sơ đồ quan hệ là cặp R, F như trên, ký hiệu SĐQH là W=<R,F>

Sơ đồ quan hệ là một lược đồ quan hệ và tập phụ thuộc hàm trên nó

Khóa của sơ đồ quan hệ

- **Định nghĩa khóa sơ đồ quan hệ**

$k \subset R$ được gọi là khóa tối thiểu của sơ đồ quan hệ trên W nếu k là tập tối thiểu kéo theo R, tức là k là khóa tối thiểu nếu: $k^+ = R$ ($k \rightarrow R$) và nếu bớt khỏi k dù một phần tử thì bao đóng của tập còn lại khác R.

$k \subset R$ gọi là khóa tối thiểu nếu thỏa mãn:

1. $k^+ = R$
2. $(k-A)^+ \neq R$, với mọi $A \in k$.

Những thuộc tính thuộc khóa gọi là thuộc tính khóa, những thuộc tính không thuộc khóa gọi là thuộc tính không khóa (thuộc tính thứ cấp).

<u>Siêu khóa</u>:

X là siêu khóa nếu X thỏa mãn điều kiện 1.

Khóa của sơ đồ quan hệ

- **Định lý**

a. Nếu k là khóa của sơ đồ quan hệ W = <R, F>, r là quan hệ trên R thì mọi cặp phần tử khác nhau t_1, t_2 của r ta luôn có $t_1.k \neq t_2.k$.

b. Ngược lại nếu k là tập tối thiểu và với mọi quan hệ r trên R mà mọi cặp t_1, t_2 của r mà $t_1.k \neq t_2.k$ thì k là khóa của sơ đồ quan hệ W = <R, F>

Thuật toán tìm khóa

Do bao đóng của k là R nên ban đầu ta coi k = R.

Ta loại trừ dần các phần tử và kiểm tra xem $k^+ = R$.

Input: W = <R, F>

Ouput: k – khóa của W

Thuật toán:

Bước 1: Đặt k = R

Bước 2: Lặp quá trình loại khỏi k phần tử A mà $(k\text{-}A)^+ = R$

Thuật toán tìm khóa

Mô tả thuật toán bằng giả mã

```
Begin
          k = R
          For each A in k do
                    if (k-A)+ = R then
                              k = k –A
End.
```

Thuật toán sẽ tìm được một khóa k cho R

Nếu muốn tìm được khóa khác ta có thể thay đổi thứ tự loại trừ các phần tử A ra khỏi k.

Thuật toán tìm khóa

- Ví dụ 1: Tìm khóa

Cho W=<R,F>, R={A, B, C}, F={A → B, A → C, B→A, B →C, AC→B }

- <u>Ví dụ 2</u>:
- Cho W=<R,F>, R={A, B, C, D, E, H}
- F={A → B, A → C, B→A, B →C, AC→B, E → C, C→H, B→H }

Tìm một khóa

Các tính chất của khóa

1. Các thuộc tính không xuất hiện trong cả hai vế trái hoặc phải của tập F phải có trong mọi khóa k.

2. Các thuộc tính chỉ xuất hiện bên trái của các PTH trong F cũng phải thuộc mọi khóa k.

3. Những thuộc tính xuất hiện, và chỉ xuất hiện bên vế phải của tập PTH sẽ không thuộc bất kỳ khóa nào.

4. Thuật toán khẳng định mọi SĐQH W đều có khóa, tuy nhiên thuật toán không khẳng định có bao nhiêu khóa và số lượng các phần tử trong khóa có như nhau hay không.

5. Họ tất cả các khóa của một SĐQH W là hệ Sperner (tức là không có hai khóa bao nhau).

Thuật toán tìm một khóa

Ví dụ:

Cho W=<R,F> với R={A,B,C,D,E,G,H,I},

F={AC->B,BI->ACD,ABC->D,H->I,ACE->BCG,CG->AE}. Tìm K?

Bước 1: Đặt K=R\{D}={A,B,C,E,G,H,I}

Bước 2: Lần lượt loại các thuộc tính có trong K:

-Loại A:

-....

K={C,G,H}

Thuật toán tìm mọi khóa

- **Ý tưởng**

Cho W=<R,F>,

1. Tìm tất cả tập con khác rỗng của R
2. Loại tập con có bao đóng khác R
3. Loại tập con bao tập con khác
4. Những tập còn lại là khóa của W

Thuật toán tìm mọi khóa…

- **Thuật toán**

Cho W=<R,F>, R={A, B, C}, F={A → B, A → C, B→A, B →C, AC→B }

	Bao đóng	Tập con có bao đóng bằng R	Khóa
A	ABC	A	A
B	ABC	B	B
C	C		
AB	ABC	AB	
AC	ABC	AC	
BC	ABC	BC	
ABC	ABC	ABC	

Thuật toán tìm mọi khóa…

- **Một số cải tiến**

- Theo tính chất của khóa chúng ta sẽ có một số thuộc tính luôn thuộc khóa. Trong thuật toán tìm khóa sẽ không xét nó và thêm vào khóa

- Một số thuộc tính không thuộc khóa nào cả. Ta loại bỏ nó trong quá trình tìm kiếm khóa

Thuật toán tìm mọi khóa…

- **Thuật toán**

Cho W=<R,F>, R={A, B, C, D, E, H}

F={A → B, A → C, B→A, B →C, AC→B, E → C, C→H, B→H }

- Chắc chắn D, E tham gia mọi khóa

- H sẽ không tham gia vào khóa nào cả

- Thuật toán tìm khóa: sẽ không cố gắng loại trừ D, E ra khỏi tập. Tập khởi tạo ban đầu có thể là K=R\{H}.
- Thuật toán tìm mọi khóa: Thêm một cột mới luôn chứa D, E. Trong các tập con của thuộc tính còn lại không xem xét đến H và xét cả tập con bằng rỗng .

Thuật toán tìm mọi khóa…

- **Thuật toán**

		Bao đóng	Siêu khóa	Khóa
DE		DEHC		
DE	A	DEHABC	DEA	DEA
DE	B	DEHABC	DEB	DEB
DE	C	DEHC		
DE	AB	DEHABC	DEAB	
DE	AC	DEHABC	DEAC	
DE	BC	DEHABC	DEBC	
DE	ABC	DEHABC	DEABC	

Bài tập

1. R(A,B,C,D) với F=[AB → C; D → B; C → ABD]
2. R(ABCDEG); với F=[A → BC, C → DE, E → G]
3. R(ABCDEGH); với F=[C → AB, D → E, B → G]
4. R(ABCDEGH); với F=[A → BC. D → E, H → G]
5. R(ABCDEG); với F=[AB → C; C → B; ABD → E;G → A]
6. R(ABCDEGHI);

F=[AC → B; BI → ACD; ABC → D; H → I; ACE → BCG, CG → AE]

Bài tập về nhà: Các bài tập cuối chương 2, Lý thuyết CSDL, Nguyễn Bá Tường

Các dạng chuẩn

- Mỗi một dạng chuẩn là một tập các điều kiện trên lược đồ nhằm đảm bảo các tính chất của nó (liên quan tới dư thừa và bất thường trong cập nhật)
- Chuẩn hóa dữ liệu: quá trình phân tích lược đồ quan hệ dựa trên các FD và các khóa chính để đạt được
  - Cực tiểu sự dư thừa
  - Cực tiểu các phép cập nhật bất thường

Các dạng chuẩn (tt)

- Thủ tục chuẩn hoá cung cấp
  - Một cơ cấu hình thức để phân tích các lược đồ quan hệ dựa trên các khoá của nó và *các phụ thuộc hàm giữa các thuộc tính* của nó.
  - Một loạt các kiểm tra dạng chuẩn có thể thực hiện trên các lược đồ quan hệ riêng rẽ sao cho cơ sở dữ liệu quan hệ có thể được chuẩn hoá đến một mức cần thiết.
- Tính chất
  - Nối không mất mát (hoặc nối không phụ thêm)
  - Bảo toàn sự phụ thuộc
    - nó đảm bảo rằng từng phụ thuộc hàm sẽ được biểu hiện trong các quan hệ riêng rẽ nhận được sau khi tách.

Các dạng chuẩn (tt)

- Phân loại
  - Boyce Codd đề nghị 3 dạng
    - 1NF (first normal form): tương đương với định nghĩa của lược đồ quan hệ (quan hệ và bộ)
    - 2NF: ko có giá trị trong thực tiễn
    - 3NF → BCNF: thường sử dụng nhiều nhất
  - 4NF, 5NF do tính đa trị và phụ thuộc hàm nối

Dạng chuẩn 1

- Đn: gọi là 1NF nếu miền giá trị của một thuộc tính chỉ chứa giá trị nguyên tử (đơn, ko phân chia được) và giá trị của mỗi thuộc tính cũng là một giá trị đơn lấy từ miền giá trị của nó hoặc không chứa nhóm thuộc tính lặp.
- Trong 1 NF:
    - Không có thuộc tính đa trị
- Ví dụ

PHONGBAN(MaPHG, TenPHG, DDIEM)

Thuộc tính đa trị

PHONGBAN(MaPHG, TenPHG)

DDIEM_PHG(MaPHG, MADDIEM,TENDD)

Dạng chuẩn 1 (tt)

- Vấn đề còn tồn tại trong 1NF
- Xét lược đồ DDIEM_PHG(MaPHG, DDIEM)
  - Vẫn bị lặp lại
  - Ẩn chứa các phụ thuộc hàm bộ phận
  -

MAPHG	DIADIEM
1	TP HCM
4	HA NOI
5	VUNGTAU
5	NHATRANG
5	TP HCM

Dạng chuẩn 2

- Phụ thuộc hàm đầy đủ: Một phụ thuộc hàm $X \rightarrow Y$ là một phụ thuộc hàm đầy đủ nếu loại bỏ bất kỳ thuộc tính A nào ra khỏi X thì phụ thuộc hàm không còn đúng nữa.

$$A, A \in X, (X - \{A\}) \rightarrow Y : \text{là sai.}$$

- Phụ thuộc hàm bộ phận: Một phụ thuộc hàm $X \rightarrow Y$ là phụ thuộc bộ phận nếu có thể bỏ một thuộc tính $A \in X$, ra khỏi X phụ thuộc hàm vẫn đúng, điều đó có nghĩa là với

$$\exists A \in X, (X - \{A\}) \rightarrow Y$$

Dạng chuẩn 2 (tt)

- 2NF:
  - Thỏa mãn 1NF
  - Phụ thuộc hàm đầy đủ vào khóa chính (không tồn tại phụ thuộc hàm vào một phần của khóa)

- Với các quan hệ có thuộc tính khóa đơn thì ko phải xét
- Chỉ kiểm tra các lược đồ có chứa phụ thuộc hàm bộ phận

Dạng chuẩn 2 (tt)

- Ví dụ

Phụ thuộc vào cả 2 MaNV, MaDA

NV_DA(MaNV, MaDA, Sogio, TenDA, DDiemDA)

Chỉ phụ thuộc vào MaDA

Dạng chuẩn 2 (tt)

- Ví dụ

Phụ thuộc vào cả 2 MaNV, MaDA

NV_DA(<u>MaNV</u>, <u>MaDA</u>, Sogio, TenDA, DDiemDA)

Chỉ phụ thuộc vào MaDA

NV_DA(<u>MaNV</u>, <u>MaDA</u>, Sogio)

DUAN(<u>MaDA</u>, TenDA, DDiemDA)

Dạng chuẩn 3

- 3NF dựa trên khái niệm phụ thuộc bắc cầu.
- ĐN: Một lược đồ quan hệ R là ở 3NF nếu nó thoả mãn (theo Codd)
  - Thỏa mãn 2NF
  - Không có thuộc tính không khoá nào của R là phụ thuộc bắc cầu vào khoá chính(không có phụ thuộc hàm ngoài khóa tức là **không có phụ thuộc hàm có nguồn là thuộc tính không khóa, đích là thuộc tính không khóa**)

Dạng chuẩn 3 (tt)

Phụ thuộc vào MaNV

NV_DV(MaNV, TenNV, NS, DCHI, MaDV, TenDV, TruongPHG)

Phụ thuộc vào MaDV

- Tất cả các thuộc tính phải phụ thuộc vào thuộc tính khóa
  - Một vài thuộc tính phụ thuộc vào thuộc tính ko phải là khóa
  - Chuẩn hóa → Tách nhóm các thuộc tính đó thành quan hệ mới

Dạng chuẩn 3 (tt)

Phụ thuộc vào MaNV

NV_DV(MaNV, TenNV, NS, DCHI, MaDV, TenDV, TruongPHG)

Phụ thuộc vào MaDV

NHANVIEN(MaNV, TenNV, NS, DCHI, MaDV)

DONVI(MaDV, TenDV, TruongPHG)

Tóm tắt 3 dạng chuẩn 1-3

NF	Nhận biết	Cách chuẩn hóa
1	Quan hệ có thuộc tính đa trị hoặc quan hệ lặp	Loại bỏ các thuộc tính vi phạm dạng chuẩn 1 và đặt chúng vào một bảng riêng cùng với khoá chính của quan hệ ban đầu. Khoá chính của bảng này là một tổ hợp của khoá chính của quan hệ ban đầu và thuộc tính đa trị hoặc khoá bộ phận của nhóm lặp. Các thuộc tính còn lại lập thành một quan hệ với khóa chính là khóa chính ban đầu.
2	Phụ thuộc 1 phần vào thuộc tính khóa	Loại bỏ các thuộc tính không khoá phụ thuộc vào một bộ phận khoá chính và tách thành ra một bảng riêng, khoá chính của bảng là bộ phận khoá mà chúng phụ thuộc vào. Các thuộc tính còn lại lập thành một quan hệ, khóa chính của nó là khóa chính ban đầu.
3	Phụ thuộc ẩn, tồn tại phụ thuộc hàm giữa các thuộc tính ko phải là khóa	Loại bỏ các thuộc tính phụ thuộc bắc cầu ra khỏi quan hệ và tách chúng thành một quan hệ riêng có khoá chính là thuộc tính bắc cầu. Các thuộc tính còn lại lập thành một quan hệ có khóa chính là quan hệ ban đầu

Ví dụ kiểm tra dạng chuẩn (tt)

- *Ví dụ 1:* Cho lược đồ W=<R={MaSV, TenSV, MaMT, ĐiemThi}; F={ MaSV→TenSV ; MaSV, MaMT→ĐiemThi}>

  Kiểm tra W có ở dạng chuẩn 3NF không?
- *Ví dụ 2:* Cho lược đồ quan hệ

  W=<R= ABC; F={AB → C; BC → A}>
- Hỏi W có ở dạng chuẩn 3NF hay không?

Ví dụ kiểm tra dạng chuẩn

- *<u>Ví dụ 3:</u>* Cho sơ đồ quan hệ W như sau:
- $W=<U=CTRHSG;F=\{C \rightarrow T,HR \rightarrow C,HT\rightarrow R,CS\rightarrow G,HS \rightarrow R\}>$

- Yêu cầu: Kiểm tra W có ở dạng chuẩn 3NF hay không? Nếu không W ở dạng chuẩn nào?

Ví dụ kiểm tra dạng chuẩn (tt)

- Xác định tập K các khoá của W và tập N các thuộc tính không khoá.
- *Nhận xét:* Thuộc tính HS tham gia khoá vì chúng chỉ xuất hiện ở vế trái. Kiểm tra HS có phải là khoá không.
- HS^+=HSRCTG=U.
- Ta có khoá K={HS} là khoá duy nhất, N={RCTG} là các thuộc tính không khoá.
- *Ta có:* C $\rightarrow$ T là phụ thuộc hàm ngoài khóa. Suy ra W không ở dạng chuẩn 3NF.
- *Hoặc:*
  - *R là thuộc tính không khoá:* HS $\rightarrow$ HT $\rightarrow$ R (HT không $\rightarrow$ HS).
  - Như vậy R phụ thuộc bắc cầu vào khoá HS thông qua cầu HT. Suy ra W không ở dạng chuẩn 3NF.

Ví dụ kiểm tra dạng chuẩn (tt)

- Kiểm tra N=RCTG có phụ thuộc đầy đủ vào HS hay không?

$$\begin{cases} H \to C : H^{+} = H \Rightarrow C \not\subset H^{+} \\ S \to C : S^{+} = S \Rightarrow C \not\subset S^{+} \end{cases}$$

Vậy C phụ thuộc đầy đủ vào HS

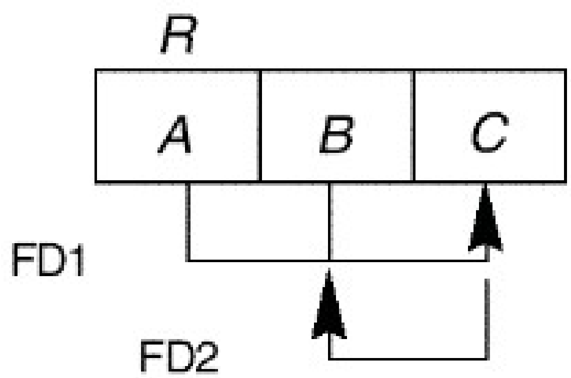
» Tương tự RTG

- KL: W ở dạng chuẩn 2NF.

Dạng chuẩn Boyce-Codd

- Một lược đồ quan hệ R được gọi là ở dạng chuẩn Boyce-Codd (BCNF) nếu nó
  – Thỏa mãn dạng chuẩn 3NF
  – Không có các thuộc tính khóa phụ thuộc hàm vào thuộc tính không khóa.

- Ví dụ



Dạng chuẩn Boyce-Codd(tt)

- Ví dụ:

  R (A1,A2,A3,A4,A5)

  Với các phụ thuộc hàm:

  – A1,A2 $\rightarrow$ A3,A4,A5

  – Giả sử A4 $\rightarrow$ A2

Dạng chuẩn Boyce-Codd(tt)

- Nếu một lược đồ quan hệ không thoả mãn điều kiện BCNF, thủ tục chuẩn hóa bao gồm:
  - tách các thuộc tính khóa phụ thuộc hàm vào thuộc tính không khóa ra thành một quan hệ, bổ sung thêm thuộc tính gây ra sự phụ thuộc vào làm khóa của quan hệ này.
  - Các thuộc tính còn lại tạo thành một LĐQH, khóa là một phần của khóa ban đầu và thuộc tính gây ra sự phụ thuộc.

- Ví dụ trên: R (<u>A1</u>,<u>A2</u>,A3,A4,A5)

  Với các phụ thuộc hàm:
  - A1,A2 $\rightarrow$ A3,A4,A5
  - G/s A4 $\rightarrow$ A2
- lược đồ được tách ra như sau:
  - R1(<u>A4</u>, A2)
  - R2(<u>A1</u>, <u>A4</u>, A3, A5)

Ví dụ về chuẩn hóa quan hệ

Ví dụ: Cho lược đồ quan hệ ở dạng chuẩn 1, R = {$\underline{A}$,$\underline{B}$,C,D,E,F,G,H,I,J} có khóa chính là A,B

Với tập các phụ thuộc hàm :

- A,B → C,D,E,F,G,H,I,J
- A→ E,F,G,H,I,J
- F → I, J
- D →B
- Kiểm tra dạng chuẩn của R, nếu chưa thuộc 3NF hoặc BCNF thì tách R đưa về 3NF hoặc BCNF?

Ví dụ về chuẩn hóa quan hệ (tt)

- Do có phụ thuộc hàm A→ E,F,G,H,I,J mà A là một bộ phận của khóa chính nên quan hệ R là vi phạm 2NF.
- Ta tách R thành R1(A,E,F,G,H,I,J) và R2(A,B,C,D).

  Trong R1, do có phụ thuộc hàm F→ I, J, nên ta có I,J phụ thuộc bắc cầu vào khóa chính, R1 là quan hệ vi phạm 3NF. Trong R2 ta có phụ thuộc hàm D → B trong đó B là một thuộc tính khóa, R2 vi phạm BCNF.

  Tách R1 và R2 ta có:
- R11(F,I,J) , R12(A,E,F,G,H), R21(D,B), R22(A,D,C)

Tách kết nối không mất thông tin

- ***Định nghĩa:*** Cho lược đồ quan hệ R={A_1,A_2,…,A_n}, R_1, R_2,.., R_k là các tập con của R và chúng là một phép tách của R tức là: R= $R_1 \cup R_2 \cup \ldots \cup R_k$,F là phụ thuộc hàm trên R.

  Ta nói phép tách R thành các lược đồ con ρ= (R_1, R_2,.., R_k) không mất thông tin đối với F nếu với mọi quan hệ r trên R mà F thỏa mãn thì:

$$r = r.R_1 \ |><| \ r.R_2 \ldots |><| \ r.R_k$$

Tách kết nối không mất thông tin(tt)

- *Ví dụ:* Cho lược đồ

  α=<U={MaSV, TenSV, MaMT, TenMon, ĐiemThi}; F={MaSV$\rightarrow$TenSV; MaMT$\rightarrow$TenMon; MaSV, MaMT$\rightarrow$ĐiemThi}>
- Tách lược đồ α thành các lược đồ sau:
- α_1=<U_1={MaSV, TenSV}; F_1={MaSV$\rightarrow$TenSV}>
- α_2=<U_2={MaMT, TenMon}; F_2={MaMT$\rightarrow$TenMon}>
- α_3=<U_3={MaSV, MaMT, DiemThi}; F_3={MaSV, MaMT$\rightarrow$ĐiemThi}>

Tách kết nối không mất thông tin(tt)

- Lấy quan hệ R là quan hệ KETQUA(MaSV, TenSV, MaMT, TenMon, ĐiemThi) khi đó ta có:
- R_1= KETQUA[MaSV, TenSV] =SINHVIEN$\in\alpha_1$
- R_2= KETQUA[MaMT, TenMon] =MONTHI$\in\alpha_2$
- R_3= KETQUA[MaSV, MaMT, DiemThi] =KQUA$\in\alpha_3$
- Khi cần ta có thể khôi phục lại thông tin:

KETQUA=SINHVIEN |><| KQUA |><| MONTHI

Tách kết nối không mất thông tin(tt)

Ví dụ mô tả phép tách tổn thất thông tin và không tổn thất thông tin:

- Cho R = { X, Y , Z} tập các thuộc tính và
- Giả sử quan hệ gốc: R và các quan hệ tách: R1 và R2

R				R_1			R_2	
X	Y	Z		X	Y		Y	Z
x_1	y_1	z_1		x_1	y_1		y_1	z_1
x_2	y_2	z_2		x_2	y_2		y_2	z_2
x_3	y_2	z_3		x_3	y_2		y_2	z_3
x_4	y_3	z_4		x_4	y_3		y_3	Z_4

Ta thấy $R \subseteq R1 \mid\times\mid R2$

Tách kết nối không mất thông tin(tt)

R				R1			R2	
X	Y	Z		X	Y		X	Z
x_1	y_1	z_1		x_1	y_1		x_1	z_1
x_2	y_2	z_2		x_2	y_2		x_2	z_2
x_3	y_2	z_3		x_3	y_2		x_3	z_3
x_4	y_3	z_4		x_4	y_3		x_4	z_4

- Ta thấy R = R_1 |><| R_2
- Nếu một phép tách không có tính chất nối không mất mát thông tin thì chúng ta có thể nhận được các bộ phụ thêm (các bộ giả) sau khi áp dụng các phép chiếu và nối tự nhiên. Nghĩa của từ mất mát ở đây là mất mát thông tin chưa không phải mất các bộ giá trị. Vì vậy, với tính chất này ta nên gọi chính xác hơn là tính chất nối không phụ thêm.

Kiểm tra tính tách kết nối không mất thông tin…

- Định lý: Nếu R={R1, R2} là một phép tách của R, F là tập PTH thì R là phép tách không mất thông tin đối với F khi và chỉ khi:

  R1 ∩ R2 -> R1-R2 hoặc

  R1 ∩ R2 -> R2-R1

- Ví dụ: R(MaSV, MaMT, DiemThi)
  - R_1 (MaSV, TenSV)
  - R_2 (MaSV, MaMT, DiemThi)
  - R1 ∩ R2 ={MaSV}
  - R1-R2={Hoten}

Ta có R1 ∩ R2 -> R1-R2 nên phép tách không mất thông tin

Kiểm tra tính tách kết nối không mất thông tin

Thuật toán Chase

- *Input Sơ đồ quan hệ: R*={A_1,A_2,…,A_n} là lược đồ quan hệ. F là tập các PTH; phép tách ρ= (R_1, R_2,.., R_k)
- *Output*: Khẳng định phép tách kết nối không mất hay không?

Thuật toán:

- Xây dựng bảng

- Xây dựng một bảng gồm n cột và m hàng, cột j tương ứng với thuộc tính A_j , hàng i tương ứng với R_i.
- Ở vị trí hàng i cột j, ta ký hiệu là a_j nếu A_j thuộc R_i , Ngược lại ta ký hiệu là b_{ij}
- Ta có bảng T như sau:

T	A_1	A_2	…	A_n
R_1				
R_2				
…				
R_k				

Kiểm tra tính tách kết nối không mất thông tin…

Bước lặp: Áp dụng các PTH trong F cho bảng vừa xây dựng:

- Xét nhiều lần mỗi phụ thuộc hàm $X \rightarrow Y$ thuộc F cho đến khi không xét được nữa. Với mỗi FTH $X \rightarrow Y$ mà trong bảng có giá trị giống nhau trên tập X thì ta cho chúng bằng nhau trên tập Y với lưu ý là nếu một trong 2 ký hiệu là a_j thì chọn a_j, ngược lại làm bằng chúng bằng một trong các kí hiệu b_{ij}.
- Thuật toán dừng khi trong bảng có một dòng toàn là ký tự a hoặc cho đến khi không xét dược nữa.

- *Kết luận*: Nếu có một dòng toàn a thì phép tách kết nối đó không mất thông tin. Ngược lại phép tách có mất thông tin.

Kiểm tra tính tách kết nối không mất thông tin…

Ví dụ: Cho R= {A , B, C, D, E, F}; R1 := {A , B, D, E}

- R2 := {A , C, D, F }
- R3 := {B, C, E, F } và F := {A → B, F → E}.

Bước 1: Thành lập bảng ban đầu gồm 3 hàng và 6 cột:

	A	B	C	D	E	F
R_1	a_1	a_2	b_{13}	a_4	a_5	b_{16}
R_2	a_1	b_{22}	a_3	a_4	b_{25}	a_6
R_3	b_{31}	a_2	a_3	b_{34}	a_5	a_6

Bước 2: Áp dụng $A \rightarrow B$ suy ra $b_{22} = a_2$

	A	B	C	D	E	F
R_1	a_1	a_2	b_{13}	a_4	a_5	b_{16}
R_2	a_1	a_2	a_3	a_4	b_{25}	a_6
R_3	b_{31}	a_2	a_3	b_{34}	a_5	a_6

Kiểm tra tính tách kết nối không mất thông tin…

Bước 3: Áp dụng $F \rightarrow E$ suy ra $b_{25} = a_5$

	A	B	C	D	E	F
R_1	a_1	a_2	b_{13}	a_4	a_5	b_{16}
R_2	a_1	a_2	a_3	a_4	a_5	a_6
R_3	b_{31}	a_2	a_3	b_{34}	a_5	a_6

Như vậy tồn tại hàng thứ 2 R_2 chứa các ký tự $\{a_1, a_2, a_3, a_4, a_5, a_6\}$.

- Suy ra phép tách không làm mất thông tin

Kiểm tra tính tách kết nối không mất thông tin…

- Ví dụ: $F=\{A\rightarrow B, C\rightarrow D, D\rightarrow EG\}$
- $R_1=\{A,B,C,D\}$, $R_2=\{D,E,G\}$

	A	B	C	D	E	G
R_1	a_1	a_2	a_3	a_4	b_{15}	b_{16}
R_2	b_{21}	b_{22}	b_{23}	a_4	a_5	a_6

- Ta thấy trong bảng trên 2 bộ bằng nhau ở thuộc tính D. Sử dụng PTH $D\rightarrow EG$ ta có $b_{15}=a_5$ và $b_{16}=a_6$. Như vậy dòng đầu của bảng chứa toàn ai, nên phép tách không mất thông tin.

Kiểm tra tính tách kết nối không mất thông tin…

Ví dụ:

- Cho lược đồ quan hệ R=ABCDE
- Tách R thành các lược đồ sau:
- R1 = AD, R2=AB, R3= BE, R4= CDE, R5= AE
- Tập phụ thuộc hàm $F=(A\rightarrow C, B\rightarrow C, C\rightarrow D, DE\rightarrow C, CE\rightarrow A)$
- Xác định phép tách trên có mất thông tin hay không?

Kiểm tra tính tách kết nối không mất thông tin…

$F=(A\rightarrow C, B\rightarrow C, C\rightarrow D, DE\rightarrow C, CE\rightarrow A$

lập bảng:

	A	B	C	D	E
AD	a1	b12	b13	a4	b15
AB	a1	a2	b23 ***(b13)***	b24 ***(a4)***	b25
BE	b31 ***(a1)***	a2	b33 ***(b13)(a3)***	b34 ***(a4)***	a5
CDE	b41	b42	a3	a4	a5
AE	a1	b52	b53 ***(b13)(a3)***	b54 ***(a4)***	a5

Phép tách trên không mất thông tin vì có dòng BE toàn ký hiệu a

Tách quan hệ thành các quan hệ BCNF với tính chất nối không mất thông tin.

Input: Lược đồ quan hệ R và tập 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.
- Mỗi lần đi vào vòng lặp trong thuật toán trên, chúng ta tách một quan hệ Q không phải BCNF thành hai lược đồ quan hệ. Kết thúc thuật toán, tất cả các quan hệ trong D sẽ ở BCNF.

Tách quan hệ thành các quan hệ BCNF với tính chất nối không mất thông tin.

- Trong bước 2 của thuật toán trên, cần xác định xem một lược đồ quan hệ Q có ở BCNF hay không. Một phương pháp để làm điều đó là kiểm tra. Với mỗi phụ thuộc hàm $X \rightarrow Y$ trong Q, ta tính X^+. Nếu X^+ không chứa tất cả các thuộc tính trong Q thì $X \rightarrow Y$ vi phạm BCNF bởi vì X không phải là một siêu khóa.

Ví dụ

- Xét lược đồ quan hệ R = { A, B, C, D, E, F)

Với các phụ thuộc hàm: $A \rightarrow BCDEF$, $BC \rightarrow ADEF$, $B \rightarrow F$, $D \rightarrow E$, $D \rightarrow B$

Lược đồ quan hệ này có hai khóa là A và BC.

- Ta có $B \rightarrow F$ vi phạm BCNF vì $B^+ \neq R$ nên B không phải là siêu khóa, R được tách thành: $R_1(B,F)$ với phụ thuộc hàm $B \rightarrow F$, $R_2(A,B,C,D,E)$ với các phụ thuộc hàm $A \rightarrow BCDE$, $BC \rightarrow ADE$, $D \rightarrow E$, $D \rightarrow B$
- Xét R_2 : Do $D \rightarrow E$ vi phạm BCNF (D là một thuộc tính không khóa), R_2 được tách thành: $R_{21}(D,E)$ với phụ thuộc hàm $D \rightarrow E$, $R_{22}(ABCD)$ với các phụ thuộc hàm $A \rightarrow BCD$, $BC \rightarrow AD$, $D \rightarrow B$
- Xét R_{22} **:** Do $D \rightarrow B$ vi phạm BCNF (Dkhông phải là thuộc tính khóa), R_{22} được tách thành: $R_{221}(D,B)$, $R_{222}(A,C,D)$ với phụ thuộc hàm $A \rightarrow CD$ (phụ thuộc hàm $BC \rightarrow AD$ bị mất)
- Tóm lại, ta có phép tách $D = \{R_1, R_{21}, R_{221}, R_{222}\}$.

THẢO LUẬN

Nội dung 1: Trả lời câu hỏi

- Trình khái niệm về phụ thuộc hàm, Hệ suy diễn Amstrong?
- Các khái niệm về phụ thuộc hàm, bao đóng và khóa?
- Phân biệt các dạng chuẩn?
- Các cách kiểm tra tách kết nối có làm mất thông tin?
- Trình bày các phương pháp chuẩn hóa dữ liệu?

THẢO LUẬN

Nội dung 2: Thảo luận theo nhóm

Thực hiện chuẩn hóa dữ liệu với đề tài nhóm đã chọn trong chương 2 theo lý thuyết các dạng chuẩn 1, 2, 3.