

Chương 10

Phân rã lược đồ (Decomposition)

Nội dung

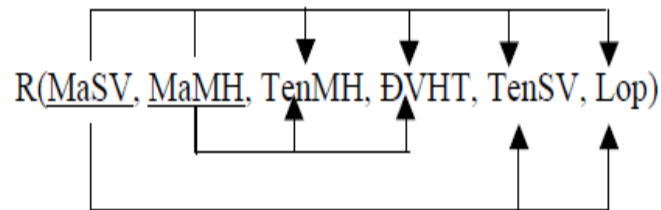
- Mục đích phân rã
- Định nghĩa phân rã
 - Phân rã không mất thông tin
 - Phân rã bảo toàn phụ thuộc
- Phân rã thành BCNF
- Phân rã thành 3NF
 - Phân rã thông thường
 - Tổng hợp

Mục đích của phân rã lược đồ quan hệ

Định nghĩa

- Phép phân rã các lược đồ quan hệ $R = \{A_1, A_2, \dots, A_n\}$ là việc thay thế lược đồ quan hệ R thành các lược đồ con $\{R_1, \dots, R_k\}$, trong đó $R_i \subseteq R$ và $R = R_1 \cup R_2 \dots \cup R_k$
- Ví dụ: Cho quan hệ R với các phụ thuộc hàm như sau:

Ví dụ: Cho quan hệ R với các phụ thuộc hàm như sau:



Ta có thể phân rã thành 3 lược đồ $R_1(\text{MaSV}, \text{TenSV}, \text{Lop})$ và $R_2(\text{MaMH}, \text{TenMH}, \text{ĐVHT})$ và $R_3(\text{MaSV}, \text{MaMH})$.

Mục đích của phân rã lược đồ quan hệ

- Được xem như 1 công cụ bổ sung vào phương pháp ER để loại trừ dư thừa dữ liệu
 - Phụ thuộc hàm được xem như là sự khái quát hóa các ràng buộc chính (key constraint). Các FD được dùng để xác định các dạng chuẩn (normal form). Việc phân rã lược đồ sẽ dựa theo các dạng chuẩn này
- Lý thuyết phân rã còn được gọi là lý thuyết chuẩn hóa.

Tính chất của phân rã lược đồ

- Vì chuẩn BCNF không có dư thừa và chuẩn 3NF tuy có dư thừa nhưng cũng hạn chế, nên việc phân rã lược đồ sẽ chỉ tập trung vào hai dạng này.

Phân rã lược đồ – Decomposition

- Phân rã 1 lược đồ $R = (U, F)$ với U là tập các thuộc tính, F là tập phụ thuộc hàm sẽ cho ra 1 tập hợp các lược đồ

$$R_1 = (U_1, F_1) \quad R_2 = (U_2, F_2) \dots$$

$$R_n = (U_n, F_n)$$

Sao cho thỏa mãn điều kiện sau:

- $U = \bigcup_{i=1}^n U_i$
- F suy dẫn F_i với $i = 1, \dots, n$

Phân rã lược đồ – Decomposition

- Phân rã lược đồ sẽ dẫn đến việc **phân rã quan hệ**.
- Phân rã 1 quan hệ r trên lược đồ R , cho ra 1 tập hợp các quan hệ

$$r_1 = \pi_{U_1}(r) \quad r_2 = \pi_{U_2}(r), \dots$$

$$r_n = \pi_{U_n}(r)$$

Phân rã không mất mát thông tin

(Lossless decomposition)

- Khảo sát quan hệ r và các phân rã của nó r_1, \dots, r_n
- Sau phân rã, CSDL không còn lưu trữ quan hệ r nữa mà chỉ lưu lại các quan hệ chiếu của nó r_1, \dots, r_n . CSDL phải có khả năng khôi phục lại quan hệ gốc r từ các quan hệ chiếu này.
- Nếu không khôi phục lại được quan hệ r thì việc phân rã không biểu diễn cùng 1 thông tin với CSDL gốc → Phân rã mất mát thông tin (lossy decomposition)

Phân rã kết nối không mất mát thông tin

(Lossless-join decomposition)

- Phân rã lược đồ $R = (U, F)$ thành 1 tập hợp các lược đồ

$$R_1 = (U_1, F_1) \quad R_2 = (U_2, F_2) \dots$$

$$R_n = (U_n, F_n)$$

- ✧ Không mất mát (lossless) nếu với mỗi điển hình (instance) hợp lệ r của lược đồ R thì

$$r = r_1 \bowtie r_2 \bowtie \dots \bowtie r_n$$

$$\text{Với } r_1 = \pi_{U_1}(r) \quad r_2 = \pi_{U_2}(r), \dots$$

$$r_n = \pi_{U_n}(r)$$

Phân rã mất mát thông tin

(Lossless decomposition)

- Cho lược đồ quan hệ $Q(\text{TENNCC}, \text{DIACHI}, \text{SANPHAM}, \text{DONGIA})$ có quan hệ tương ứng là r
- Đặt r_1 là quan hệ có được bằng cách chiếu r lên $Q_1(\text{TENNCC}, \text{SANPHAM}, \text{DONGIA})$,
- Đặt r_2 là quan hệ có được bằng cách chiếu r lên $Q_2(\text{TENNCC}, \text{DIACHI})$
- Đặt r' là quan hệ có được bằng cách kết tự nhiên giữa r_1 và r_2 qua TENNCC .

Phân rã mất mát thông tin (Lossless decomposition)

$$r$$

TENNCC	DIACHI	SANPHAM	DONGIA
Hung	12 Nguyễn Kiệm	Gạch ống	200
Hung	12 Nguyễn Kiệm	Gạch thẻ	250
Hung	40 Nguyễn Oanh	Gạch ống	200

$$r_2 = r.Q_2^+$$

TENNCC	DIACHI
Hung	12 Nguyễn Kiệm
Hung	40 Nguyễn Oanh

$$r_1 = r.Q_1^+$$

TENNCC	SANPHAM	DONGIA
Hung	Gạch ống	200
Hung	Gạch thẻ	250

$$r' = r_1 |><| r_2$$

TENNCC	DIACHI	SANPHAM	DONGIA
Hung	12 Nguyễn Kiệm	Gạch ống	200
Hung	12 Nguyễn Kiệm	Gạch thẻ	250
Hung	40 Nguyễn Oanh	Gạch ống	200
Hung	40 Nguyễn Oanh	Gạch thẻ	250

Kết quả là $r \neq r'$ hay $r \neq r.Q_1 |><| r.Q_2$.

→ phép tách $\rho(Q_1, Q_2)$ tách Q thành Q_1, Q_2 là tách-kết nối (phân rã) mất mát thông tin.

Phân rã mất mát thông tin

(Lossless-join decomposition)

- Thực tế sẽ nhận được nhiều bộ (tuple) từ phép kết các r_1, r_2, \dots, r_n hơn là các bộ gốc ban đầu \rightarrow Vậy tại sao lại gọi là mất mát (lossy)??
- Tuy nhiều bộ hơn nhưng lại thiếu thông tin và không có cách nào biết được bộ nào là đúng, bộ nào là không đúng với bộ gốc.
- Nhiều bộ hơn nhưng không đúng thông tin thì sẽ đồng nghĩa với mất mát thông tin

Ví dụ phân rã kết nối mất mát thông tin

Cho quan hệ r

S	P	D
s1	p1	d1
s2	p2	d2
s3	p1	d3

Instance r

Kết nối tự nhiên 2 quan hệ phân rã này:

Phân rã thành 2 quan hệ

S	P
s1	p1
s2	p2
s3	p1

$\pi_{SP}(r)$

P	D
p1	d1
p2	d2
p1	d3

$\pi_{PD}(r)$

S	P	D
s1	p1	d1
s2	p2	d2
s3	p1	d3
s1	p1	d3
s3	p1	d1

$\pi_{SP}(r) \bowtie \pi_{PD}(r)$

→ Có những bộ không thuộc quan hệ gốc ban đầu

Phân rã không mất mát thông tin (Lossless-join decomposition)



Tính chất

- Nếu Q là một lược đồ quan hệ, $Q1, Q2$ là hai lược đồ quan hệ con có

$$Q1^+ \cap Q2^+ = X$$

$$Q1^+ \cup Q2^+ = Q^+$$

$$X \rightarrow Q2^+ \quad \times$$

$$\text{Thì } r = r.Q1 \bowtie r.Q2$$

Thuật toán kiểm tra không mất mát thông tin (Lossless-join decomposition)

- **Input:**
 - Lược đồ quan hệ $R = \{A_1, A_2, \dots, A_n\}$
 - Tập các phụ thuộc hàm F
 - Phép tách $\rho(R_1, R_2, \dots, R_k)$
- **Output:** Kết luận phép tách ρ không mất mát thông tin.

Phân rã không mất mát thông tin (Lossless-join decomposition)



Ví dụ 10: cho $Q(\text{SAIP})$, $Q1 = (\text{SA})$, $Q2 = (\text{SIP})$
 $F = \{S \rightarrow A, SI \rightarrow P\}$. Hỏi việc tách Q thành $Q1$ và $Q2$ có gây ra mất mát thông tin không?

- Áp dụng tính chất trên, ta có

$$Q1^+ \cap Q2^+ = S$$

$$Q1^+ \cup Q2^+ = \text{SAIP} = Q^+$$

$$S \rightarrow SA = Q1^+$$

- Theo tính chất trên, với mọi quan hệ r của Q ta luôn có

$$r = r.Q1 \overset{S}{\bowtie} r.Q2.$$

- Suy ra phép tách trên là phép tách kết nối bảo toàn thông tin.

Thuật toán kiểm tra không mất mát thông tin (Lossless-join decomposition)

Các bước của thuật toán kiểm tra phép tách kết nối bảo toàn thông tin:

- Dữ liệu vào: lược đồ quan hệ $Q(A_1, A_2, \dots, A_n)$, tập phụ thuộc hàm F , phép tách $\rho = (Q_1, Q_2, \dots, Q_k)$.
- Dữ liệu ra: kết luận phép tách ρ có phải là phép tách bảo toàn thông tin ?
- Bước 1:
 - Thiết lập một bảng với n cột (tương ứng với n thuộc tính) và k dòng (tương ứng với k quan hệ), trong đó cột thứ j ứng với thuộc tính A_j , dòng thứ i ứng với lược đồ R_i .
 - Tại dòng i và cột j , ta điền ký hiệu a_j nếu thuộc tính $A_j \in R_i$. Ngược lại ta điền ký hiệu b_{ij} .

Thuật toán kiểm tra không mất mát thông tin (Lossless-join decomposition)

Các bước của thuật toán:

- Bước 2:
 - Xét các phụ thuộc hàm trong F và áp dụng cho bảng trên.
 - Giả sử ta có phụ thuộc hàm $X \rightarrow Y \in F$, xét các dòng có giá trị bằng nhau trên thuộc tính X **thì làm bằng** các giá trị của chúng trên Y .
 - Ngược lại làm bằng chúng bằng ký hiệu bij. Tiếp tục áp dụng các pth cho bảng (kể cả việc lặp lại các phụ thuộc hàm đã áp dụng) cho tới khi không còn áp dụng được nữa.

Thuật toán kiểm tra không mất mát thông tin (Lossless-join decomposition)

Các bước của thuật toán:

- Bước 3:
 - Xem xét bảng kết quả. Nếu xuất hiện một dòng chứa toàn giá trị **a1, a2 ,...,an** thì kết luận phép tách ρ không mất mát thông tin.

Thuật toán kiểm tra không mất mát thông tin (Lossless-join decomposition)

Ví dụ:

Ví dụ: Cho quan hệ:
EMP_DEPT

ENAME	<u>SSN</u>	BDATE	ADDRESS	DNUMBER	DNAME	DMGRSSN
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Research	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Research	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administration	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administration	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 Fire Oak, Humble, TX	5	Research	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Research	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administration	987654321
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Headquarters	888665555

Hình 6.15. Minh họa dữ liệu của quan hệ EMP_DEPT

Thuật toán kiểm tra không mất mát thông tin (Lossless-join decomposition)

Ví dụ:

- Tách thành 2 quan hệ:

EMPLOYEE

ENAME	<u>SSN</u>	BDATE	ADDRESS	DNUMBER
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4
Narayan, Remesh K.	666884444	1962-09-15	975 Fire Oak, Humble, TX	5
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1

DEPARTMENT

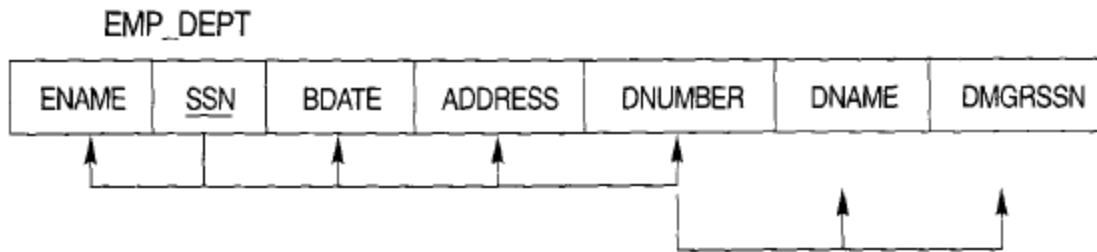
DNAME	<u>DNUMBER</u>	DMGRSSN
Research	5	333445555
Administration	4	987654321
Headquarters	1	888665555

Hình 6.16. Quan hệ EMPLOYEE được phân rã (tách) thành 2 quan hệ

Thuật toán kiểm tra không mất mát thông tin (Lossless-join decomposition)

Ví dụ:

Tập phụ thuộc hàm F:



Kiểm tra phép tách trên là không mất mát thông tin:

Bước 1:

	ENAME	SSN	BDate	Address	DNumber	DName	DMgrSsn
EMPLOYEE	a1	a2	a3	a4	a5	b16	b17
DEPARTMENT	b21	b22	b23	b24	a5	a6	a7

Thuật toán kiểm tra không mất mát thông tin (Lossless-join decomposition)

Ví dụ:

- Bước 2: Xét phụ thuộc hàm $Dnumber \rightarrow DName, DMgrSsn$. Ta nhận thấy có giá trị a5 ở dòng thứ 2, nên ta sẽ làm bằng giá trị a6, a7 cho dòng thứ 1.
- Bước 3: Tồn tại một dòng chứa giá trị a1, a2,...a7. Kết luận, phép phân rã trên không mất mát thông tin.

	EName	SSN	BDate	Address	DNumber	DName	DMgrSsn
EMPLOYEE	a1	a2	a3	a4	a5	a6	a7
DEPARTMENT	b21	b22	b23	b24	a5	a6	a7

Ghi chú: Sinh viên thực hiện phép nối tự nhiên 2 quan hệ EMPLOYEE và DEPARTMENT trên để kiểm tra có bằng quan hệ ban đầu EMP DEPT

Thuật toán kiểm tra không mất mát thông tin (Lossless-join decomposition)

Bài tập:

Cho quan hệ $R(W, A, Z, Y, Q, P)$. Hỏi phép tách R thành các quan hệ sau có mất mát thông tin không?

- $R_1(A, Z);$
- $R_2(W, Y, Q, P)$
- $R_3(Y, Q, P, A)$
- $F = \{W \rightarrow AYQP, A \rightarrow Z, YQP \rightarrow A\}$

Phân rã nhị phân

(Binary Decomposition)

- Cho lược đồ $R = (U, F)$ và $R_1 = (U_1, F_1)$, $R_2 = (U_2, F_2)$ là một phân rã nhị phân của R .
- Sự phân rã này không mất thông tin nếu và chỉ nếu thỏa mãn một trong các điều kiện sau:
 - $(U_1 \cap U_2) \rightarrow U_1 \in F^+$
 - $(U_1 \cap U_2) \rightarrow U_2 \in F^+$

Phân rã nhị phân

(Binary Decomposition)



Các thuộc tính chung của U1 và U2 phải chứa khóa của hoặc R1 hoặc R2.

- Kiểm tra này là cần thiết để bảo đảm phân rã có kết nối không bị mất.
- **Ví dụ:** Cho $R(SNLRWH)$ có $FD R \rightarrow W$ vi phạm chuẩn 3NF, nên tách thành $SNLRH$ and RW . Phân rã này có bị mất kết nối không???
- ❖ *Không, vì R là thuộc tính chung của cả 2 lược đồ $R1$, $R2$ nên phân rã này kết nối không mất..*

Phân rã nhị phân

(Binary Decomposition)

Ví dụ

- Xét lược đồ quan hệ

PERSON(SSN, Name, Address, Hobby)

SSN	Name	Address	Hobby
1111111	John	123 Main St.	Stamps
1111111	John	123 Main St.	Coins
5556667	Mary	7 Lake Dr.	Hiking
5556667	Mary	7 Lake Dr.	Skating
9876543	Simpson	Fox 5 TV	Acting

Phân rã nhị phân

(Binary Decomposition)

- Nếu phân rã lược đồ trên thành 2 lược đồ sau:

PERSON1(SSN, Name, Address)

HOBBY(SSN, Hobby)

Việc phân rã này có mất thông tin không??

- Vì $\text{PERSON1} \cap \text{HOBBY} = \{\text{SSN}\}$ mà SSN là khóa chính của PERSON1, do đó

$\text{PERSON1} \cap \text{HOBBY} \rightarrow \text{PERSON1}$

Phân rã này không mất thông tin

Phân rã bảo toàn phụ thuộc hàm

- Cho lược đồ $R = (U, F)$ và $R_1 = (U_1, F_1)$, $R_2 = (U_2, F_2), \dots, R_n = (U_n, F_n)$ là phân rã của R .
- Phân rã được gọi là bảo toàn phụ thuộc hàm nếu và chỉ nếu F và $\bigcup_{i=1}^n F_i$ là tương đương nhau.

Phân rã bảo toàn phụ thuộc hàm (Dependency-Preserving Decomposition)

- Khảo sát lược đồ quan hệ sau:
HASACCOUNT(ClientId, OfficeId, AccountNumber)
Với các FD sau:
 - ClientId, OfficeId \rightarrow AccountNumber
 - AccountNumber \rightarrow OfficeId

Nếu phân rã lược đồ trên thành 2 lược đồ sau:

ACCTOFFICE (AccountNumber, OfficeId)

ACCTCLIENT (AccountNumber, ClientId)

Phân rã trên có mất mát thông tin không???

Phân rã bảo toàn phụ thuộc hàm

- Phân rã trên không mất mát thông tin vì:

$$ACCTOFFICE \cap ACCTCLIENT = \{AccountNumber\}$$

Mà AccountNumber là khóa chính của ACCTOFFICE, nên

$$ACCTOFFICE \cap ACCTCLIENT \rightarrow ACCTOFFICE$$

- Nhưng phân rã này không bảo toàn phụ thuộc hàm
- Phụ thuộc hàm gốc $ClientId, OfficeId \rightarrow AccountNumber$ (1) không tồn tại trong các phụ thuộc hàm của các lược đồ phân rã vì:
 - Cả hai phân rã đều không chứa đủ các thuộc tính khóa của phụ thuộc hàm gốc (1) nên không thể suy diễn lại được phụ thuộc hàm này

Phân rã bảo toàn phụ thuộc hàm

- Nếu 1 phụ thuộc hàm $f \in F$ nhưng không thuộc bất kỳ F_i nào không có nghĩa là phân rã không bảo toàn phụ thuộc hàm nếu f có thể được suy diễn từ

$$\bigcup_{i=1}^n F_i$$

- Chỉ khi nào f không suy diễn được từ $\bigcup_{i=1}^n F_i$ thì lúc đó phân rã mới không bảo toàn phụ thuộc \rightarrow để duy trì f đòi hỏi phải có kết nối các lược đồ phân rã trước, kiểm tra phụ thuộc hàm sau

Ví dụ

- Phân rã quan hệ HASACCOUNT

AccountNumber	ClientId	OfficeId
B123	111111	SB01
A908	123456	MN08

AccountNumber	OfficeId	Account Number	ClientId
B123	SB01	B123	111111
A908	MN08	A908	123456

Ví dụ

- HASACCOUNT và phân rã của nó sau khi chèn thêm 1 hàng

AccountNumber	ClientId	OfficeId
B123	111111	SB01
B567	111111	SB01
A908	123456	MN08

AccountNumber	OfficeId
B123	SB01
B567	SB01
A908	MN08

Account Number	ClientId
B123	111111
B567	111111
A908	123456

Sau khi join 2 lược đồ phân rã lại, phụ thuộc hàm
ClientId, OfficeId → AccountNumber bị vi phạm

Phép chiếu của tập phụ thuộc hàm

- Khảo sát lược đồ $R=(U,F)$, một quan hệ r trên R và 1 tập thuộc tính $S \subseteq U$
- Phép chiếu của tập F lên tập các thuộc tính S được định nghĩa như sau:

$$\pi_S(F)=\{X \rightarrow Y | X \rightarrow Y \in F^+ \text{ and } X \cup Y \subseteq S\}$$

Phép chiếu này được ký hiệu F_S

Ví dụ về phép chiếu tập FD

- Cho $R=(U,F)$ với $U=\{A,B,C\}, F=\{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$
- Giả sử R được phân rã thành 2 lược đồ AB và BC
 - $F_{AB} = \{A \rightarrow B\}$
 - $F_{BC} = \{B \rightarrow C\}$
 - Còn $C \rightarrow A$ thì sao??

➔ Phép phân rã R có bảo toàn phụ thuộc hàm không??

Ví dụ về phép chiếu tập FD

- $F^+ = F \cup \{A \rightarrow C, B \rightarrow A, \text{ and } C \rightarrow B\}$
 - $F_{AB} = \{A \rightarrow B; B \rightarrow A\}$, $F_{BC} = \{B \rightarrow C; C \rightarrow B\}$
 - $F_{AB} \cup F_{BC} = \{A \rightarrow B, B \rightarrow C, B \rightarrow A, C \rightarrow B\}$
 - $(F_{AB} \cup F_{BC})^+$ có chứa FD $C \rightarrow A$
- Phân rã này bảo toàn phụ thuộc hàm

Các ví dụ trong giáo trình CSDL trang 64

Thuật toán kiểm tra bảo toàn phụ thuộc hàm

Thuật toán tìm bao đóng của tập thuộc tính X đối với $G = \cup \Pi Q_i(F)$

- Vào: $\rho = (Q_1, Q_2, \dots, Q_k), F, X$
- Ra: XG^+
- **Bước 1**: Với mỗi phụ thuộc hàm $X \rightarrow Y \in F$ ta thực hiện từ bước 2 đến bước 4
- **Bước 2**: đặt $Z' = X$
- **Bước 3**: thế $Z' = Z' \cup ((Z' \cap Q^+)^+ \cap Q^+)$
- **Bước 4**: nếu ở Q_i , Z' thay đổi thì thực hiện lại bước 3 cho $Q_{\text{đầu tiên}}$
Ngược lại kết thúc thuật toán và trả về Z' (là bao đóng XG^+)

Thuật toán kiểm tra bảo toàn phụ thuộc hàm

Thuật toán kiểm tra bảo toàn phụ thuộc hàm

- Vào: $\rho = (Q_1, Q_2, \dots, Q_k), F$
- Ra: kết luận phép tách ρ bảo toàn hay không bảo toàn phụ thuộc hàm
- **Bước 1:** Với mỗi phụ thuộc hàm $X \rightarrow Y \in F$ ta thực hiện từ bước 2 đến bước 3:
- **Bước 2:** Tìm bao đóng X_{G^+} với $G = \cup \Pi Q_i(F)$
- **Bước 3:** Nếu $Y \subseteq X_{G^+}$ thì $X \rightarrow Y \in \cup \Pi Q_i(F)^+$
- **Bước 4:** Nếu tất cả phụ thuộc $X \rightarrow Y \in F$ đều thuộc $\cup \Pi Q_i(F)^+$ thì ta kết luận phân rã ρ bảo toàn phụ thuộc hàm ngược lại ρ không bảo toàn phụ hàm

Thuật toán kiểm tra bảo toàn phụ thuộc hàm

Thuật toán kiểm tra bảo toàn phụ thuộc hàm

- Bài tập: Cho lược đồ quan hệ $Q(C, S, Z)$ và $F = \{CS \rightarrow Z, Z \rightarrow C\}$.
- Phép tách $\rho = (Q_1, Q_2)$ tách Q thành hai lược đồ $Q_1(S, Z)$ và $Q_2(C, Z)$.
- Hỏi phép tách có bảo toàn phụ thuộc hàm không?

Thuật toán kiểm tra bảo toàn phụ thuộc hàm

Ví dụ 1: thực hiện lại ví dụ 13, nghĩa là kiểm tra phép tách có bảo toàn phụ thuộc hàm không?

- Vào: $Q(C,S,Z), F = \{CS \rightarrow Z, Z \rightarrow C\}, Q1(S,Z)$ và $Q2(C,Z)$
- Đương nhiên $Z \rightarrow C \in G = \Pi Q1(F) \cup \Pi Q2(F) \Rightarrow Z \rightarrow C \in (\Pi Q1(F) \cup \Pi Q2(F))^+$

1. $Z' = CS$

2. Gán $Z' = Z' \cup ((Z' \cap)^+ \cap)$: $Z' = CS \cup (S \cap SZ) = CS$

Bước 1 và 2 có Z' không thay đổi, ta sang lược đồ $Q2$ và tính tiếp Z'

3. Gán $Z' = Z' \cup ((Z' \cap)^+ \cap)$: $Z' = CS \cup (C \cap CZ) = CS$

Z' không thay đổi và hết lược đồ quan hệ \Rightarrow ngưng không tính tiếp Z'

4. Vậy $= CS \Rightarrow CS \rightarrow Z \notin (\Pi Q1(F) \cup \Pi Q2(F))^+$ phép phân rã không bảo toàn phụ thuộc hàm.

Thuật toán kiểm tra bảo toàn phụ thuộc hàm

Ví dụ 2: thực hiện lại ví dụ 12 với nội dung kết luận phép tách ρ có bảo toàn phụ thuộc hàm không (không tính F^+)

- Vào: $Q(A,B,C), F = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}, Q1(A,B)$ và $Q2(B,C)$
- Hiển nhiên $G = \Pi Q1(F) \cup \Pi Q2(F) \supseteq \{A \rightarrow B, B \rightarrow C\}$
- Ta xác định $C \rightarrow A$ có thuộc $(\Pi Q1(F) \cup \Pi Q2(F))^+ Z' = CS$

1. $Z' = C$

2. Gán $Z' = Z' \cup ((Z' \cap)^+ \cap)$: $Z' = C \cup (\emptyset \cap AB) = C$

Bước 1 và 2 có Z' không thay đổi, ta sang lược đồ $Q2$ và tính tiếp Z'

3. Gán $Z' = Z' \cup ((Z' \cap)^+ \cap)$: $Z' = C \cup (ABC \cap BC) = BC$

Z' thay đổi \Rightarrow tính tiếp Z' bắt đầu từ lược đồ $Q1$

3. Gán $Z' = Z' \cup ((Z' \cap)^+ \cap)$: $Z' = BC \cup (ABC \cap AB) = ABC$

do $Z' = Q^+ \Rightarrow Z'$ sẽ không bao giờ thay đổi.

5. vậy $= ABC \Rightarrow C \rightarrow A \in (\Pi Q1(F) \cup \Pi Q2(F))^+$ phép phân rã bảo toàn phụ thuộc hàm.

Phân rã lược đồ quan hệ

- Hai tính chất của phân rã:
 - Lossless (không mất thông tin)
 - Dependency-preserving (bảo toàn phụ thuộc hàm)
 - Tính chất nào quan trọng hơn???
- ➔ Lossless là bắt buộc (mandatory) trong khi dependency-preserving là tùy chọn (optional)

Phân rã lược đồ quan hệ

Ví dụ: Cho lược đồ quan hệ $Q(C,S,Z)$ và $F=\{CS\rightarrow Z, Z\rightarrow C\}$.

Phép tách $\rho=(Q_1, Q_2)$ tách Q thành hai lược đồ $Q_1(S,Z)$ và $Q_2(C,Z)$.

- Hỏi phép tách có bảo toàn phụ thuộc hàm không?

Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

1. Cách thông thường

- Thuật toán phân rã Q,F thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin
- Bước 1: Tìm tất cả khóa của Q
- Bước 2: Tìm phụ thuộc hàm $X \rightarrow Y \in F$ có X không là siêu khóa và Y không chứa thuộc tính khóa.

Nếu tìm thấy thì tách Q thành Q1 và Q2 theo quy tắc sau:

$Q1 = Q[XY]; F1 \equiv \Pi Q1(F)$ tìm bao đóng của tất cả tập con của XY để suy ra $\Pi Q1(F) \Rightarrow F1$

$Q2 = Q[Q^+ - Y]; F2 \equiv \Pi Q2(F)$ tìm bao đóng của tất cả tập con của $Q^+ - Y$ để suy ra $\Pi Q2(F) \Rightarrow F2$

thực hiện thuật toán phân rã (Q1,F1)

thực hiện thuật toán phân rã (Q2,F2)

Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

1. Cách thông thường

- Thuật toán phân rã Q,F thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

Ngược lại nếu không tìm thấy thì có hai trường hợp:

Trường hợp 1: mọi phụ thuộc hàm trong F_i đều có vế trái là siêu khóa thì Q_i đạt chuẩn BC

Trường hợp 2: nếu có phụ thuộc hàm có vế trái không là siêu khóa và vế phải là thuộc tính khóa thì Q_i đạt chuẩn 3.

Tính chất: Theo thuật toán trên, khi phân rã Q thành $Q_1(XY)$ với $X \rightarrow Y$ và Q_2 thì tập khóa SQ của Q luôn luôn bằng với tập khóa SQ_2 của Q_2 .

Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

2. Thuật toán phân rã Q,F thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

- Bước 1: Tìm tập tất cả khóa SK của Q
- Bước 2: Tìm phụ thuộc hàm $X \rightarrow Y \in F$ có X không là siêu khóa và Y không chứa thuộc tính khóa. Nếu tìm thấy thì tách Q thành Q1 và Q2 theo quy tắc sau:
 - $Q1 = Q[XY]$; Tính F1 bằng cách tính bao đóng tất cả tập con của XY
 - $Q2 = Q[Q^+ - Y]$ SK cũng là tập khóa của Q2
 - thực hiện bước 1 cho Q1
 - thực hiện bước 2 cho Q2

Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

2. Thuật toán phân rã Q,F thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

- Ngược lại nếu không tìm thấy thì có hai trường hợp:
 - Trường hợp 1: mọi phụ thuộc hàm trong F_i đều có vế trái là siêu khóa thì Q_i đạt chuẩn BC
 - Trường hợp 2: nếu có phụ thuộc hàm có vế trái không là siêu khóa và vế phải là thuộc tính khóa thì Q_i đạt chuẩn 3.

Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

3. Thuật Toán phân rã Q, F thành dạng chuẩn BC bảo toàn thông tin

- Bước 1: $Z' = Q+$
- Bước 2: phân rã Z' theo thuật toán chi tiết để được 2 lược đồ $Z'-A$ và XA trong đó XA ở dạng chuẩn BC và $X \rightarrow A$
 - Nếu thuật toán chi tiết cho kết quả thì qua bước 3
 - Ngược lại kết thúc thuật toán
- Bước 3: nhận XA là một lược đồ con của các lược đồ kết quả Q_1, \dots, Q_k
- Bước 4: thực hiện phân rã $Z'-A, F$

Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

3. Thuật Toán phân rã Q, F thành dạng chuẩn BC bảo toàn thông tin

- Thuật toán chi tiết
- Bước 1:
 - Nếu Z' không chứa AB sao cho $(Z'-AB) \rightarrow A$. thì báo không phân rã được.
 - Ngược lại qua bước 2
- Bước 2: đặt $Y' = Z'$
- Bước 3: nếu Y' chứa AB sao cho $(Y'-AB) \rightarrow A$. thì gán $Y' = Y' - B$ thực hiện lại bước 2
- Bước 4: bước 3 cho kết quả $Y' = XA$ với XA ở dạng chuẩn BC và $X \rightarrow A$. Trả về XA

Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

Bài tập: cho $Q(CTHRSG)$,

$F = \{C \rightarrow T; HR \rightarrow C; HT \rightarrow R; CS \rightarrow G; HS \rightarrow R\}$

hãy phân rã Q thành các lược đồ con đạt chuẩn BC bảo toàn thông tin.

Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

- Bài tập:

- Cho $R(A, B, C, D, E)$

$$F = \{DE \rightarrow A; B \rightarrow C; E \rightarrow AD; AB \rightarrow D\}$$

- 1) Xác định dạng chuẩn của Q
- 2) Hãy phân rã Q thành các lược đồ con đạt chuẩn BC bảo toàn thông tin.

Phân rã thành dạng chuẩn BC (hay chuẩn 3) bảo toàn thông tin

Cho lược đồ quan hệ $R(ABCDEFG)$ và tập phụ thuộc hàm

$F = \{ \mathbf{AB} \rightarrow \mathbf{C} ; \mathbf{AD} \rightarrow \mathbf{E} ; \mathbf{C} \rightarrow \mathbf{DG} ; \mathbf{AB} \rightarrow \mathbf{DE} \}$

Câu 1 : Chứng minh phụ thuộc hàm $\mathbf{AB} \rightarrow \mathbf{G}$ được suy dẫn từ F bằng 2 phương pháp ?

- Sử dụng luật dẫn Amstrong
- Sử dụng bao đóng của tập thuộc tính

Câu 2 : Tìm tất cả các khóa của R ?

Câu 3 : Tìm dạng chuẩn cao nhất của R ?

Câu 4 : Tìm phủ tối thiểu của F ?

Câu 5 : Nếu R chưa đạt 3NF , hãy phân rã R đạt 3NF bảo toàn thông tin và bảo toàn pth ?

Phân rã thành dạng chuẩn 3 vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm

1. Thuật Toán phân rã Q, F thành dạng chuẩn 3, bảo toàn thông tin, bảo toàn phụ thuộc hàm

- Dữ liệu vào: lược đồ quan hệ Q và tập phụ thuộc hàm F.
 - Dữ liệu ra: một phân rã sao cho mỗi lược đồ quan hệ con đều đạt chuẩn 3 vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm.
1. Tìm phủ tối thiểu Ftt của F
 2. Nếu có một phụ thuộc hàm nào của Ftt mà liên quan đến tất cả các thuộc tính của Q thì kết quả phân rã chính là Q (Q không thể phân rã)
 3. Nếu có những thuộc tính của Q không nằm trong một phụ thuộc nào của Ftt - dù ở vế phải hay vế trái của F thì chúng tạo thành một lược đồ cần tìm.
 4. Cứ mỗi phụ thuộc hàm $X \rightarrow A \in Ftt$ thì XA là một lược đồ cần tìm
 5. Nếu có một lược đồ con chứa khóa K của Q thì kết thúc thuật toán
Ngược lại tạo một lược đồ con K

Giải thuật phân rã BCNF

- $R=(U,F)$ là 1 lược đồ quan hệ không ở chuẩn BCNF.
- Giải thuật: thực hiện lặp lại việc phân chia R thành những lược đồ nhỏ hơn sao cho các lược đồ mới có ít FD vi phạm BCNF hơn. Giải thuật kết thúc khi tất cả lược đồ kết quả đều ở dạng BCNF

Giải thuật phân rã BCNF

Input $R = (U, F)$

Decomposition = R

While có lược đồ $S = (V, F')$ trong Decomposition không phải BCNF

/*Nếu có $X \rightarrow Y \in F$ sao cho $X \cup Y \subseteq S$ và vi phạm BCNF, dùng FD này để phân rã*/

- Thay S trong Decomposition với $S_1 = (XY, F_1)$
- $S_2 = ((S - Y) \cup X, F_2)$ với F_1, F_2 là tất cả các FD của F'

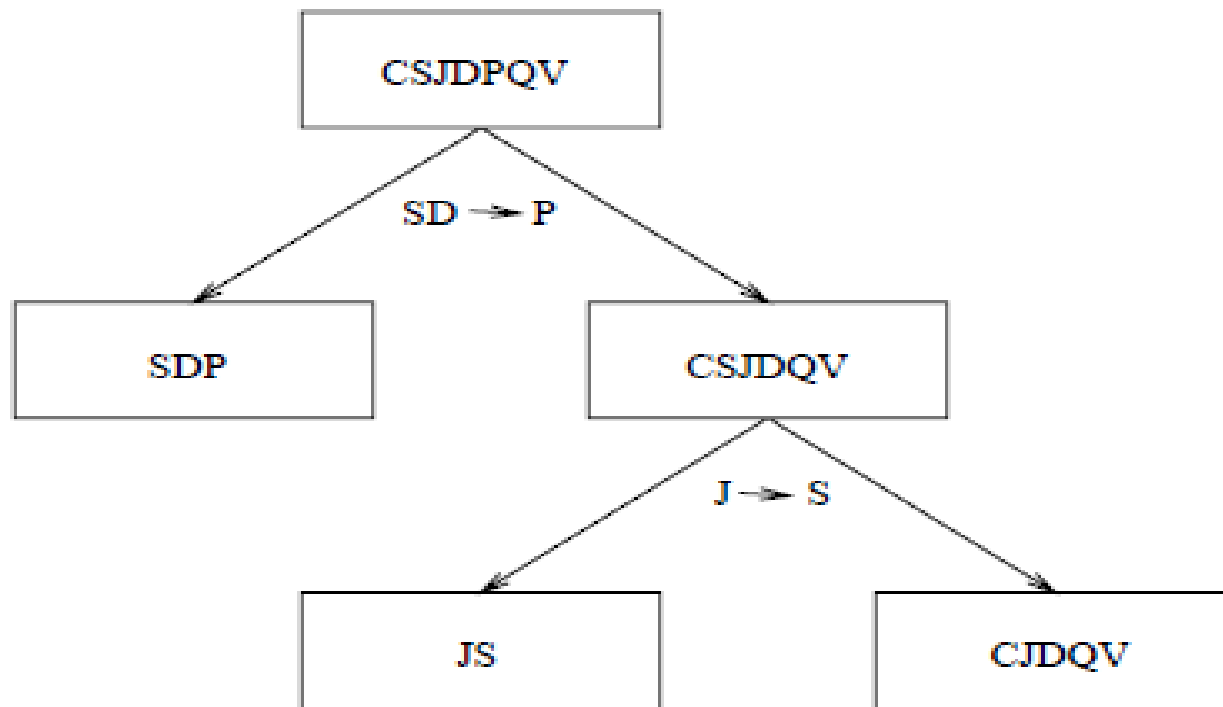
End

Return Decomposition

Ví dụ 1

- Cho lược đồ $R(CSJDPQV)$ và C là khóa.
- Tập phụ thuộc hàm $\{C \rightarrow SJDPQV; JP \rightarrow C; SD \rightarrow P; J \rightarrow S\}$.
- Do $SD \rightarrow P$ vi phạm chuẩn BCNF nên tách thành 2 lược đồ.
 - $R1(SDP)$ ở dạng chuẩn BCNF.
 - $R2(CSJDPQV)$: vì $J \rightarrow S$ vi phạm chuẩn BCNF nên tiếp tục phân rã $R2$ thành 2 lược đồ $R21(JS)$ và $R22(CJDQV)$

Ví dụ 1



Phân rã này có bảo toàn phụ thuộc hàm không???

Không (do $JP \rightarrow C$ đòi hỏi phải thực hiện phép kết join mới có được)

Ví dụ 2

- Cho $R = (U, F)$, $U = \{ABCDEFGH\}$, $F = \{ABH \rightarrow C, A \rightarrow DE, BGH \rightarrow F, F \rightarrow ADH, BH \rightarrow GE\}$
- Tìm FD vi phạm BCNF
 - $(ABH)^+ = U$, ABH là siêu khóa, $ABH \rightarrow C$ không vi phạm BCNF
 - $A^+ \neq U$, $A \rightarrow DE$ vi phạm BCNF
- Chia R thành
 - $R_1 = (ADE, \{A \rightarrow DE\})$
 - $R_2 = (ABCFGH, \{ABH \rightarrow C, BGH \rightarrow F, F \rightarrow AH, BH \rightarrow G\})$

Ví dụ 2

- Sau khi phân rã, chú ý đến 2 phụ thuộc hàm gốc $F \rightarrow ADH$, $BH \rightarrow GE$
 - Chia $F \rightarrow ADH$ thành $\{F \rightarrow AH, F \rightarrow D\}$
 - Chia $BH \rightarrow GE$ thành $\{BH \rightarrow G, BH \rightarrow E\}$
 - $F \rightarrow D$, $BH \rightarrow E$ không có chỗ trong các phân rã mới (vì không có ràng buộc nào có đủ thuộc tính cho các FD này)
 - Nhưng
 - $F \rightarrow D$ có thể suy diễn từ $F \rightarrow AH \in R_2$ và $A \rightarrow DE \in R_1$
 - $BH \rightarrow E$ có thể suy diễn được dựa vào $(BH)^+$ từ R_1, R_2
- ➔ Phân rã R_1, R_2 bảo toàn phụ thuộc hàm

Ví dụ 2

- R1 là BCNF
- Với R2:
 - $ABH \rightarrow C$, $BGH \rightarrow F$ không vi phạm BCNF (ABH , BGH đều là siêu khóa)
 - $F \rightarrow AH$ vi phạm BCNF
- ➔ Phân rã R2 thành
 - $R21 = (FAH, \{F \rightarrow AH\})$
 - $R22 = (FBCG, \{\})$
- ➔ R21, R22 đều là BCNF nhưng khi đó các FD $ABH \rightarrow C$, $BGH \rightarrow F$ và $BH \rightarrow G$ không có mặt nữa và cũng không thể suy dẫn được từ các FD của R21, R22 và R1
- ➔ Phân rã R2 không bảo toàn phụ thuộc hàm

Nhận xét

- Việc phân rã R thành R1, R21, R22 không phải là duy nhất.
- Nếu bắt đầu từ FD $F \rightarrow ADH$ thì sẽ có

$$R1 = (FADH; \{F \rightarrow ADH\})$$

$$R2 = (FBCEG, \{\})$$

R1, R2 cũng ở chuẩn BCNF và 1 số FD gốc cũng bị mất, không thể suy diễn được

Tính chất của giải thuật phân rã BCNF

- Không mất mát thông tin
- Nhưng có thể không bảo toàn phụ thuộc hàm
- Là giải thuật không xác định (nondeterministic), phụ thuộc vào thứ tự các FD được chọn để xét phân rã

Phân rã thành chuẩn 3NF

- Hai cách thực hiện:

- **Cách 1:** dùng giải thuật phân rã như của BCNF nhưng có bổ sung để phân rã bảo toàn phụ thuộc.
- **Cách 2:** dùng phương pháp tổng hợp (**synthesis**), bắt đầu từ các thuộc tính riêng rẽ, nhóm chúng lại thành các lược đồ

Cách 1: Phân rã thành chuẩn 3NF

- Cho $R(U, F)$ với F là phủ tối thiểu. Giả sử đã phân rã không mất mát R thành $R_1; R_2; \dots; R_n$. Mỗi R_i đều đã ở dạng chuẩn 3NF. Để bảo đảm phân rã này bảo toàn phụ thuộc hàm, cần thực hiện thêm 2 bước sau:
 - Xác định tập N chứa các FD không được bảo toàn (*not **preserved***) nghĩa là không có trong bao đóng của hợp các F_i $(\bigcup F_i)^+$
 - Với mỗi FD trong N , tạo thêm 1 lược đồ quan hệ mới XA và bổ sung nó vào tập phân rã.

Cách 1: Phân rã thành chuẩn 3NF (tt)

- Để tối ưu, nếu tập N chứa nhiều FD có cùng vế trái $X \rightarrow A_1; X \rightarrow A_2; \dots X \rightarrow A_n$, có thể thay thế chúng thành FD đơn tương đương $X \rightarrow A_1A_2..A_n$. Vì vậy chỉ cần tạo thêm 1 lược đồ quan hệ mới $(X \ A_1A_2..A_n)$ thay vì phải tạo nhiều lược đồ (XA_i)

Ví dụ: phân rã lược đồ thành 3NF

- Xét lược đồ $(CSJDPQV)$ và tập $FD \{C \rightarrow SJDPQV; JP \rightarrow C, SD \rightarrow P, J \rightarrow S\}$
- Nếu ta phân rã thành SDP và $CSJDQV$, thì SDP ở chuẩn BCNF nhưng $CSJDQV$ không đạt được chuẩn 3NF. Vì vậy tiếp tục phân rã thành JS và $CJDQV$.
- Các lược đồ quan hệ SDP , JS và $CJDQV$ đều đạt chuẩn 3. Phân rã kết nối không mất (lossless join) nhưng $JP \rightarrow C$ không được bảo toàn (not preserved) \Rightarrow bổ sung thêm lược đồ CJP nữa.

Cách 2: dùng phương pháp synthesis

- Xuất phát từ phủ tối thiểu của FD F , tạo 1 lược đồ quan hệ mới R_i cho mỗi FD trong F . Tất cả lược đồ này đều ở 3NF và bảo toàn phụ thuộc.
- Nếu phân rã này làm kết nối bị mất (lossless-join) thì bổ sung thêm 1 lược đồ chỉ chứa các thuộc tính khóa.

Ví dụ dùng cách synthesis

- Cho $R(ABC)$ với phủ tối thiểu (*minimal cover*) $F = \{A \rightarrow B, C \rightarrow B, A \rightarrow C\}$.
- **Bước 1:** phân rã thành 2 lược đồ $R1(AB)$ và $R2(BC)$.
- **Bước 2:** phân rã bị mất kết nối vì $A \rightarrow C$ không thể suy diễn được từ hợp các FD của $R1, R2$. Bổ sung thêm lược đồ $R3(AC)$.

Tính chất của giải thuật phân rã 3NF

- Bảo toàn phụ thuộc hàm
- Không mất thông tin

Ví dụ

- Phủ tối thiểu G của tập F ví dụ trước:
 $G = \{ABH \rightarrow C, A \rightarrow D, C \rightarrow E, F \rightarrow A, E \rightarrow F\}$
- Phân rã thành 6 lược đồ:
 - $(ABHC; \{ABH \rightarrow C\})$
 - $(AD; \{A \rightarrow D\})$
 - $(CE; \{C \rightarrow E\})$
 - $(FA; \{F \rightarrow A\})$
 - $(EF; \{E \rightarrow F\})$
 - $(BHE; \{BH \rightarrow E\})$
- Không có lược đồ phân rã nào là siêu khóa BCGH, nên bổ sung thêm lược đồ thứ 6
 - $(BCGH; \{\})$

Phân rã BCNF thông qua phân rã 3NF

- Vì giải thuật phân rã BCNF có thể không bảo toàn phụ thuộc hàm \rightarrow nên phân rã BCNF thông qua phân rã 3NF. Nếu lược đồ sau phân rã là BCNF thì dừng, nếu không thì dùng lúc đó mới dùng giải thuật BCNF để phân rã tiếp

Thuật toán phân rã một lược đồ quan hệ thành các lược đồ con ở 3NF.

● **Input:**

- Lược đồ quan hệ R
- Tập các phụ thuộc hàm F, không làm mất tính tổng quát giả sử đó là phủ tối thiểu.

● **Output:**

- Phép tách không mất mát thông tin trên R thành các lược đồ con ở dạng chuẩn 3 sao cho vẫn bảo toàn các phụ thuộc hàm.

Thuật toán phân rã một lược đồ quan hệ thành các lược đồ con ở 3NF.

Các bước của thuật toán:

- **Bước 1:** Loại bỏ các thuộc tính của R nếu thuộc tính đó không liên quan đến phụ thuộc hàm nào của F. (không có mặt ở cả hai vế của phụ thuộc hàm).
- **Bước 2:** Nếu có một phụ thuộc hàm của F liên quan đến tất cả các thuộc tính của R thì kết quả chính là R.
- **Bước 3:** Ngoài ra, phép tách ρ đưa ra các lược đồ gồm các thuộc tính XA ứng với phụ thuộc hàm $X \rightarrow A \in F$. Nếu tồn tại các phụ thuộc hàm $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_n$ thuộc F thì thay thế XA_i ($1 \leq i \leq n$) bằng $XA_1A_2 \dots A_n$. Quá trình tiếp tục.

Thuật toán phân rã một lược đồ quan hệ thành các lược đồ con ở 3NF.

Ví dụ:

- Cho lược đồ quan hệ $R(C, T, H, R, S, G)$ với tập phụ thuộc hàm tối thiểu F :
 - $C \rightarrow T, HR \rightarrow C, HT \rightarrow R, CS \rightarrow G, HS \rightarrow R$.
- Yêu cầu: Phân rã lược đồ quan hệ trên thành các quan hệ con đều ở dạng 3NF.
- Sử dụng thuật toán tìm khoá \rightarrow Khoá chính của R là HS .

Thực hiện thuật toán:

- Bước 1: Không có thuộc tính bị loại bỏ
- Bước 2: Không có phụ thuộc hàm nào liên quan tới tất cả các thuộc tính

Thuật toán phân rã một lược đồ quan hệ thành các lược đồ con ở 3NF.

Ví dụ:

Bước 3:

- Phụ thuộc hàm $C \rightarrow T$ vi phạm 3NF (phụ thuộc bắc cầu vào khoá), vì vậy tách R thành $R1(C,T)$ và $R2(C,H,R,S,G)$.
- Phụ thuộc hàm $CS \rightarrow G$ vi phạm 3NF (phụ thuộc bộ phận vào khoá), tách $R2$ thành $R21(C,S,G)$ và $R22(C,H,R,S)$.
- Phụ thuộc hàm $HR \rightarrow C$ vi phạm 3NF, tách $R22$ thành
- $R221(H,R,C)$ và $R222(H,S,R)$
- Như vậy, quan hệ R được tách thành các quan hệ sau: $R1$, $R21$, $R221$, $R222$

Thuật toán phân rã một lược đồ quan hệ thành các lược đồ con ở 3NF.

- **Lưu ý:**
- Kết quả của phép tách có thể khác nhau phụ thuộc vào thứ tự áp dụng các phụ thuộc hàm khi thực hiện thuật toán.
- Sinh viên tự kiểm tra xem việc tách quan hệ như trên có mất mát thông tin không.

Bài tập

1. Cho một quan hệ $R = \{A, B, C, D, E, F, G, H, I, J\}$ và tập phụ thuộc hàm

● $F = \{ A, B \rightarrow C, A \rightarrow D, E, B \rightarrow F, F \rightarrow G, H, D \rightarrow I, J \}$

● **Yêu cầu:**

- Tìm $\{A\}^+ = \{D, E, I, J\}$
- Tìm khóa của quan hệ R.
- Tách quan hệ R thành BCNF.
- Kiểm tra xem việc tách trên có mất mát thông tin không?

Bài tập

2. Lập lại yêu cầu ở bài 1 với tập phụ thuộc hàm sau:

- $G = \{A, B \rightarrow C$
 $B, D \rightarrow E, F$
 $A, D \rightarrow G, H$
 $A \rightarrow I$
 $H \rightarrow J\}$

Bài tập

3. Cho một quan hệ $R = \{\text{CourseNo}, \text{SecNo}, \text{OfferingDept}, \text{Credit_Hours}, \text{CourseLevel}, \text{InstructorSSN}, \text{Semester}, \text{Year}, \text{Days_Hours}, \text{RoomNo}, \text{NoOfStudents}\}$ và tập phụ thuộc hàm:
- $F = \{ \text{CourseNo} \rightarrow \text{OfferingDept}, \text{Credit_Hours}, \text{CourseLevel};$
 $\text{CourseNo}, \text{SecNo}, \text{Semester}, \text{Year} \rightarrow \text{Days_Hours},$
 $\text{RoomNo}, \text{NoOfStudents}, \text{InstructorSSN};$
 $\text{RoomNo}, \text{Days_Hours}, \text{Semester}, \text{Year} \rightarrow \text{InstructorSSN},$
 $\text{CourseNo}, \text{SecNo} \}$
 - Tìm khóa của quan hệ R.
 - Quan hệ trên thuộc dạng chuẩn mấy?
 - Tách quan hệ về dạng 3NF.
 - Kiểm tra xem việc tách trên có mất mát thông tin không?

Bài tập

3. Cho một quan hệ $R = \{\text{CourseNo}, \text{SecNo}, \text{OfferingDept}, \text{Credit_Hours}, \text{CourseLevel}, \text{InstructorSSN}, \text{Semester}, \text{Year}, \text{Days_Hours}, \text{RoomNo}, \text{NoOfStudents}\}$ và tập phụ thuộc hàm:
- $F = \{ \text{CourseNo} \rightarrow \text{OfferingDept}, \text{Credit_Hours}, \text{CourseLevel};$
 $\text{CourseNo}, \text{SecNo}, \text{Semester}, \text{Year} \rightarrow \text{Days_Hours},$
 $\text{RoomNo}, \text{NoOfStudents}, \text{InstructorSSN};$
 $\text{RoomNo}, \text{Days_Hours}, \text{Semester}, \text{Year} \rightarrow \text{InstructorSSN},$
 $\text{CourseNo}, \text{SecNo} \}$

Bài tập

Cho quan hệ $R(A, B, C, D, E, F)$ và tập phụ thuộc hàm $F = \{A \rightarrow BD; F \rightarrow AB; AC \rightarrow DF; BC \rightarrow E; \}$

1. Chứng minh rằng $AC \rightarrow E$ là thành viên của F bằng tiên đề Armstrong và bao đóng (0.5đ)
2. Tìm phủ tối thiểu của F (2 đ)
3. Xác định khóa cho quan hệ R . (1đ)
4. Xác định dạng chuẩn cao nhất cho quan hệ R . (1đ)
5. Bạn đưa ra giải pháp để sửa quan hệ trên (tách quan hệ) đạt chuẩn cao hơn. (0.5đ)

Bài tập

- **CHUYENNGANH(MaCN, TenCN)**
- **SINHVIEN(MaSV, HoTen, Email, SoDT, ChuyenNganh)**
- **GIANGVIEN(MaGV, HoTenGV, Email, ChuyenNganh)**
- **DETAI(MaDT, TenDT, LoaiDT, YeuCau, MucTieu, MaGV)**
- **PHIEUDANGKY(MaSV, MaDT, HocKy, NamHoc)**

Bài tập

- **Dựa trên lược đồ cơ sở dữ liệu của câu 1, bạn hãy viết các lệnh SQL sau:**
- Liệt kê danh sách sinh viên (MaSV, HoTen, chuyenNganh) thực hiện đề tài do các giảng viên khác chuyên ngành hướng dẫn trong học kỳ 2 năm học '2017-2018', được sắp xếp theo chuyên ngành của sinh viên.
- Liệt kê danh sách các đề tài (MaDT, TenDT) do giảng viên chuyên ngành 'Hệ thống thông tin' hướng dẫn trong học kỳ 2 năm học '2017-2018'
- Liệt kê danh sách giảng viên (MaGV, HoTenGV) hướng dẫn nhiều hơn 10 sinh viên.
- Liệt kê danh sách tên những đề tài không được sinh viên đăng ký.