



Chương 7

Chuẩn hóa cơ sở dữ liệu

Các vấn đề khi dư thừa dữ liệu

- Dư thừa trong một lược đồ quan hệ làm cho cơ sở dữ liệu quan hệ không tối ưu vì sẽ gây ra những bất thường khi thực hiện các thao tác:
 - Chèn
 - Xóa
 - Cập nhật

Các vấn đề khi dư thừa dữ liệu

- Ví dụ: lược đồ quan hệ Thi được tổ chức như sau:
Thi (Masv, Hoten, Monhoc, Diemthi)

MASV	HOTEN	MONHOC	DIEMTHI
00CDTH189	Nguyễn Văn Thành	Cấu Trúc Dữ Liệu	7
00CDTH189	Nguyễn Văn Thành	Cơ Sở Dữ Liệu	9
00CDTH211	Trần Thu Hà	Kỹ Thuật Lập Trình	5
00CDTH189	Nguyễn Văn Thành	Kỹ Thuật Lập Trình	8

Các vấn đề khi dư thừa dữ liệu

➤ Bất thường khi cập nhật

- Do dư thừa nên khi cập nhật họ tên của một sinh viên trong một bộ nào đó nhưng vẫn để lại họ tên cũ trong những bộ khác

➤ Bất thường khi chèn

- Không thể biết họ tên của một sinh viên nếu hiện tại sinh viên đó không dự thi môn nào

➤ Bất thường khi xóa

- Ngược lại, khi xóa tất cả các môn thi của một sinh viên, vô ý làm mất dấu vết để tìm ra họ tên của sinh viên này

Chuẩn hóa cơ sở dữ liệu

- **Chuẩn hóa:** Là quá trình phân rã những quan hệ chưa đạt chuẩn bằng cách phân rã thành những quan hệ nhỏ hơn và đạt chuẩn
- Ví dụ: Phân rã lược đồ quan hệ Thi thành ba lược đồ quan hệ:
 - Sinhvien(masv,hoten)
 - Monhoc(mamh, Tenmon)
 - Ketqua(masv, Mamh, Diemthi)

Chuẩn hóa cơ sở dữ liệu

MASV	HOTEN
00CDTH189	Nguyễn Văn Thành
00CDTH211	Trần Thu Hà

MAMH	TENMON
M1	Cơ sở dữ liệu
M2	Cấu trúc dữ liệu
M3	Kỹ thuật lập trình

MASV	MAMH	DIEMTHI
00CDTH189	M2	7
00CDTH189	M2	9
00CDTH211	M3	5
00CDTH189	M3	8

Dạng chuẩn 1 (First Normal Form)

“Lược đồ quan hệ R ở dạng chuẩn 1(1NF) nếu thỏa:

- Mọi thuộc tính của R đều chứa các giá trị nguyên tố (atomic value).
- Có khóa chính – mỗi bộ là duy nhất
- Tất cả thuộc tính còn lại phải phụ thuộc vào khóa chính

Dạng chuẩn 1 (First Normal Form)

Ví dụ: Xét sát lược đồ quan hệ

NVien_DuAn(MaNV, TenNV, MaDuAn, Sogio, TenDuAn, Diadiem

NVien_DuAn

Các thuộc tính vi phạm chuẩn 1

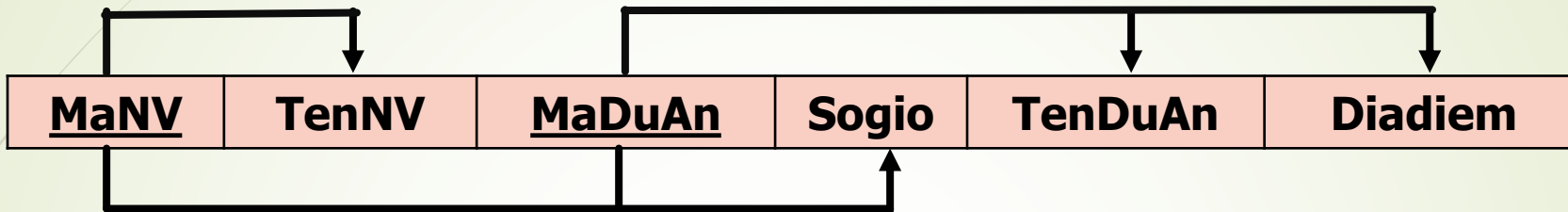
MaNV	TenNV	MaDuAn	Sogio	TenDuAn	Diadiem
NV01	Tuấn	1 2	32 7	Dự án A Dự án B	Bình Thạnh Gò Vấp
NV02	Hoàng	3	40	Dự án C	Thủ Đức
NV03	Phong	1 2	30 20	Dự án A Dự án B	Bình Thạnh Gò Vấp

Lược đồ quan hệ PhongBan không đạt chuẩn 1

Chuẩn hóa về dạng chuẩn 1

- Loại bỏ thuộc tính đa trị
- Xác định khóa chính
- Xác định tất cả các phụ thuộc hàm

Ví dụ: chuyển lược đồ *NVien_DuAn* thành lược đồ quan hệ đạt chuẩn 1



NVien_DuAn

MaNV	TenNV	MaDuAn	Sogio	TenDuAn	Diadiem
NV01	Tuấn	1	32	Dự án A	Bình Thạnh
NV01	Tuấn	2	7	Dự án B	Gò Vấp
NV02	Hoàng	3	40	Dự án C	Thủ Đức
NV03	Phong	1	30	Dự án A	Bình Thạnh
NV03	Phong	2	20	Dự án B	Gò Vấp

Lược đồ quan hệ NVien_DuAn đạt chuẩn 1

Dạng chuẩn 2 (Second Normal Form)

➤ Thuộc tính khóa – thuộc tính không khóa

➤ Xét lược đồ quan hệ $R(A, B, C, D)$, có khóa là $\{A, B\}$

➤ Thuộc tính khóa: **A, B**

➤ Thuộc tính không khóa: **C, D**

Dạng chuẩn 2 (Second Normal Form)

➤ Phụ thuộc hàm đầy đủ

- Phụ thuộc hàm $X \rightarrow Y$ là *phụ thuộc hàm đầy đủ* nếu với bất kỳ thuộc tính A , với $A \subseteq X$, thì $(X - \{A\})$ không xác định được Y .

Ví dụ: $\{\text{MaNV}, \text{MaDuAn}\} \rightarrow \text{Sogio}$

Dạng chuẩn 2 (Second Normal Form)

□ Phụ thuộc hàm từng phần

- Phụ thuộc hàm $X \rightarrow Y$ là *phụ thuộc hàm từng phần* nếu loại bỏ một số thuộc A , với $A \in X$, thì $(X - \{A\}) \rightarrow Y$.

Ví dụ: $\{\text{MaNV}, \text{MaDuAn}\} \rightarrow \text{TenNV}$

Dạng chuẩn 2 (Second Normal Form)

“Lược đồ quan hệ R đạt dạng chuẩn 2 (2NF) đối với tập phụ thuộc hàm F nếu thỏa hai điều kiện sau

- *R đạt dạng chuẩn 1*
- *Mọi thuộc tính không khóa đều phụ thuộc hàm đầy đủ vào mọi khóa của R ”*

Ví dụ: Xét lược đồ quan hệ **NVien_DuAn**

NVien_DuAn(MaNV, MaDuAn, Sogio, TenNV, TenDuAn, Diadiem)

Tập phụ thuộc hàm trên **NVien_DuAn**

MaNV \rightarrow TenNV

MaDuAn \rightarrow {TenDuAn, Diadiem}

{MaNV, MaDuAn} \rightarrow Sogio

Lược đồ quan
hệ **NVien_DuAn** có
đạt chuẩn 2 ?

MaNV	MaDuAn	Sogio	TenNV	TenDuAn	Diadiem
NV01	1	32	Tuấn	Dự án A	Bình Thạnh
NV01	2	7	Tuấn	Dự án B	Gò Vấp
NV02	3	40	Hoàng	Dự án C	Thủ Đức
NV03	1	30	Phong	Dự án A	Bình Thạnh
NV03	2	20	Phong	Dự án B	Gò Vấp

Lược đồ quan hệ **NVien_DuAn** không đạt chuẩn 2 vì tồn tại thuộc tính không khóa phụ thuộc không đầy đủ vào khóa

Chuẩn hóa về dạng chuẩn 2

Loại bỏ phụ thuộc hàm từng phần

Ví dụ: Chuyển lược đồ quan hệ **NVien_DuAn** thành các lược đồ đạt chuẩn 2

NVien(MaNV, TenNV)

DuAn(MaDuAn, TenDuAn, Diadiem)

NV_DA(MaNV, MaDuAn, Sogio)

NhanVien

MaNV	TenNV
NV01	Tuấn
NV02	Hoàng
NV03	Phong

NV_DA

MaNV	MaDuAn	Sogio
NV01	1	32
NV01	2	7
NV02	3	40
NV03	1	30
NV03	2	20

DuAn

MaDuAn	TenDuAn	Diadiem
1	Dự án A	Bình Thạnh
2	Dự án B	Gò Vấp
3	Dự án C	Thủ Đức

Ví dụ: Lược đồ quan hệ **NhanVien**

<u>MaNV</u>	TenNV	Ngaysinh	Diachi	MaPB	TenPB	Ma_NguoiQL
-------------	-------	----------	--------	------	-------	------------

Thể hiện của lược đồ quan hệ **NhanVien**

NhanVien

*Lược đồ quan hệ
NhanVien có đạt
chuẩn 2 ?*

MaNV	TenNV	Ngaysinh	Diachi	MaPB	TenPB	Ma_NguoiQL
NV01	Tuấn	10/11/90	Bình Thạnh	TC	Tổ chức	NV01
NV02	Duy	20/10/89	Gò Vấp	KT	Kỹ thuật	NV02
NV03	Hoàng	03/11/88	Thủ Đức	QT	Quản trị	NV03
NV04	Nam	10/12/89	Bình Thạnh	TC	Tổ chức	NV01
NV05	Phong	12/11/87	Gò Vấp	KT	Kỹ thuật	NV02

Một lược đồ quan hệ có khóa chính chứa một thuộc tính thì lược đồ quan hệ đạt dạng chuẩn 2

Giải thuật kiểm tra lược đồ quan hệ đạt 2NF

- Cho lược đồ quan hệ R và tập phụ thuộc hàm F
 - B1: Tìm tất cả khóa K của R
 - B2: Với mỗi khóa K , tìm bao đóng của tất cả tập con thực sự S của K .
 - B3: Nếu có $S+$ chứa thuộc tính không khóa thì R không đạt 2NF, ngược lại thì R đạt 2NF

➤ Ví dụ:

Cho $Q(A,B,C,D)$ và $F=\{ AB \rightarrow CD; B \rightarrow D; C \rightarrow A \}$

Xác định dạng chuẩn của PTH

- Khoá là $\{A,B\}$ và $\{B,C\}$.
- D là thuộc tính không khoá; $AB \rightarrow D$ không là phụ thuộc hàm đầy đủ vì có $B \rightarrow D$.
- Vậy Q đạt chuẩn 1

- Ví dụ: Xác định dạng chuẩn của lược đồ quan hệ sau. $Q(GMVNHP)$ với $F=\{G\rightarrow N; G\rightarrow H; G\rightarrow P; M\rightarrow V; NHP\rightarrow M\}$
- **Xác định dạng chuẩn của PTH**
- Khoá của Q là G .
- Thuộc tính không khoá là M, V, N, H, P .
- Do các phụ thuộc hàm $G \rightarrow M; G \rightarrow V; G \rightarrow N; G \rightarrow H; G \rightarrow P$ là các phụ thuộc hàm đầy đủ, nên lược đồ quan hệ Q đạt dạng chuẩn 2

Giải thuật kiểm tra lược đồ quan hệ đạt 2NF

- Ví dụ: cho lược đồ quan hệ $R(A,B,C,D)$ và tập phụ thuộc hàm $F=\{AB \rightarrow C; B \rightarrow D; BC \rightarrow A\}$. R đạt 2 NF?

$$TN=\{B\}, TG=\{AC\}$$

X_i	$(TN \cup X_i)$	$(TN \cup X_i)^+$	Super key	Key
ϕ	B	BD		
A	AB	ABCD	AB	AB
C	BC	ABCD	BC	BC
AC	ABC	ABCD	ABC	

- Khóa của lược đồ: AB và BC
- Phụ thuộc hàm $B \rightarrow D$, thuộc tính không khóa D, phụ thuộc không đầy đủ vào khóa, do đó R không đạt 2 NF

Giải thuật kiểm tra lược đồ quan hệ đạt 2NF

➤ Hệ quả:

- R đạt 2NF nếu R đạt 1NF và tập thuộc tính không khoá của R bằng rỗng.
- Nếu lược đồ quan hệ có 1 khóa và khóa có một thuộc tính thì quan hệ đó ít nhất đạt chuẩn 2.
- Ví dụ: $R(ABCDEH)$ $F=\{A \rightarrow EB; D \rightarrow C; E \rightarrow DH\}$
 - Khoá của R là $K=\{A\}$
 - R đạt 2NF

Bài tập

- Cho lược đồ quan hệ $R(A,B,C,D,E)$ và tập phụ thuộc hàm $F = \{AB \rightarrow C, CD \rightarrow E, DE \rightarrow B\}$.
 - a) Xác định khóa của R
 - b) Lược đồ quan hệ đạt chuẩn 2 không?

Dạng chuẩn 3 (Third Normal Form)

Định nghĩa 1

Một lược đồ quan hệ R đạt dạng chuẩn 3 nếu R đạt dạng chuẩn 2 và không có thuộc tính không khóa phụ thuộc bắc cầu vào khóa chính

Dạng chuẩn 3 (Third Normal Form)

Ví dụ: xét lược đồ quan hệ

Nvien_Pban(MaNV, TenNV, Ngaysinh, Diachi, MaPB, TenPB, Ma_NguoiQL)

MaNV	TenNV	Ngaysinh	Diachi	MaPB	TenPB	Ma_NguoiQL
NV01	Tuấn	10/11/90	Bình Thạnh	NS	Nhân sự	NV01
NV02	Tuấn	20/10/89	Gò Vấp	KT	Kỹ thuật	NV02
NV03	Hoàng	03/11/88	Thủ Đức	NS	Nhân sự	NV01
NV04	Phong	10/12/89	Bình Thạnh	NS	Nhân sự	NV01
NV05	Phong	12/11/87	Gò Vấp	KT	Kỹ thuật	NV02

Dạng chuẩn 3 (Third Normal Form)

➤ Nhận xét:

- Lược đồ quan hệ **Nvien_Pban** tồn tại thuộc tính không khóa **Ma_nguoiQL** phụ thuộc bắc cầu vào khóa chính **MaNV** thông qua **MaPB**

$MaNV \rightarrow MaPB$

$MaPB \rightarrow Ma_NguoiQL$

Dạng chuẩn 3 (Third Normal Form)

- Có thể tách lược đồ quan hệ ***Nvien_Pban*** thành hai quan hệ đạt chuẩn 3

NhanVien

MaNV	TenNV	Ngaysinh	Diachi	MaPB
NV01	Tuấn	10/11/90	Bình Thạnh	NS
NV02	Tuấn	20/10/89	Gò Vấp	KT
NV03	Hoàng	03/11/88	Thủ Đức	NS
NV04	Phong	10/12/89	Bình Thạnh	NS
NV05	Phong	12/11/87	Gò Vấp	KT

PhongBan

MaPB	TenPB	Ma_NguoiQL
NS	Nhân sự	NV01
KT	Kỹ thuật	NV02

Dạng chuẩn 3 (Third Normal Form)

► Định nghĩa 2

Một lược đồ quan hệ R đạt dạng chuẩn 3 nếu mọi phụ thuộc hàm $X \rightarrow A \in F^+$, với $A \notin X$, đều thỏa:

- *Hoặc X là siêu khóa*
- *Hoặc A là thuộc tính khóa*

► *Nhận xét: Nếu R đạt chuẩn 3 thì R đạt chuẩn 2*

Dạng chuẩn 3 (Third Normal Form)

➤ Ví dụ: Xét lược đồ quan hệ $R(A,B,C,D)$ và tập phụ thuộc hàm $F=\{AB \rightarrow C; D \rightarrow B; C \rightarrow ABD\}$

Lược đồ quan hệ có 3 khóa

➤ $K1 = \{AB\};$

➤ $K2 = \{AD\};$

➤ $K3 = \{C\}$

R đạt chuẩn 3

Dạng chuẩn 3 (Third Normal Form)

- ▶ Ví dụ: cho $R(C,S,Z)$, $F = \{CS \rightarrow Z, Z \rightarrow C\}$
 - ▶ Khóa : CS và SZ
 - ▶ Tất cả các thuộc tính đều là thuộc tính khóa
 - ▶ R đạt 3NF
- ▶ Ví dụ: $R(A,S,I,P)$, $F=\{SI \rightarrow P, S \rightarrow A\}$,
 - ▶ Khóa: SI
 - ▶ $S \rightarrow A$ nhưng S là tập con của khóa và A là thuộc tính không khóa \rightarrow R không đạt 2NF \rightarrow R không đạt 3NF

Giải thuật kiểm tra lược đồ quan hệ đạt 3NF

- Cho lược đồ quan hệ R và tập phụ thuộc hàm F
 - B1: Tìm tất cả khóa của R
 - B2: Tạo một tập phụ thuộc hàm tương đương (F_{1tt}) với vế phải của các phụ thuộc hàm chỉ chứa một thuộc tính.
 - B3: Nếu mọi phụ thuộc hàm $X \rightarrow A \in F_{1tt}$ với $A \notin X$, X là siêu khóa hoặc A là thuộc tính khóa thì R đạt 3NF, ngược lại thì R không đạt 3NF

Giải thuật kiểm tra lược đồ quan hệ đạt 3NF

- Ví dụ: Cho $R(ABCD)$ và tập phụ thuộc hàm
$$F=(AB \rightarrow C ; D \rightarrow B; C \rightarrow ABD)$$
 - $K1=[AB]$; $K2=[AD]$; $K3=[C]$ là các khoá, vậy R không có thuộc tính không khoá nên R đạt chuẩn 3
- Hệ quả: Nếu lược đồ quan hệ R, F mà R không có thuộc tính không khoá thì R đạt chuẩn 3.

Giải thuật kiểm tra lược đồ quan hệ đạt 3NF

- Ví dụ: Xác định dạng chuẩn của lược đồ quan hệ $R(NGPM)$, tập phụ thuộc hàm $F=\{NGP \rightarrow M; M \rightarrow P\}$
 - Khoá của R là $\{NGP\}, \{NGM\}$
 - $NGP \rightarrow M$ có vế trái là siêu khoá
 - $M \rightarrow P$ có vế phải là thuộc tính khoá.
- ➔ R đạt chuẩn 3.

Giải thuật kiểm tra lược đồ quan hệ đạt 3NF

➡ Bài tập:

1. Cho lược đồ quan hệ $R(A,B,C,D)$ và tập phụ thuộc hàm $F=\{AB \rightarrow C; D \rightarrow B; C \rightarrow ABD\}$, Q đạt 3NF?
2. Cho lược đồ quan hệ $R(A,B,C,D,E)$ và tập phụ thuộc hàm $F=\{A \rightarrow C; AD \rightarrow B; CB \rightarrow D\}$, Q đạt chuẩn mấy? Nếu không đạt chuẩn, thì chỉ ra phụ thuộc hàm nào vi phạm chuẩn tương ứng?

Dạng chuẩn BCNF(Boyce-Codd)

- *Một lược đồ quan hệ R ở dạng chuẩn BCNF nếu với mỗi phụ thuộc hàm không hiển nhiên $X \rightarrow A \in F$ thì X là một siêu khoá của R .*
- *Nhận xét. Nếu R đạt BCNF thì R đạt 3NF*

Dạng chuẩn BCNF(Boyce-Codd)

- ▶ Ví dụ: Xác định dạng chuẩn của lược đồ quan hệ sau. $R(ACDEIB) F=\{ACD \rightarrow EBI; CE \rightarrow AD\}$
 - ▶ R có hai khoá là: ACD và CE.
 - ▶ Các phụ thuộc hàm của F đều có vế trái là siêu khoá, nên R đạt dạng chuẩn BC.

Giải thuật kiểm tra BCNF

- Cho lược đồ quan hệ R và tập phụ thuộc hàm F
 - $B1$: Tìm tất cả khóa của Q
 - $B2$: Tạo tập phụ thuộc hàm tương đương ($F1_{tt}$) mà vế phải của các phụ thuộc hàm chỉ chứa một thuộc tính.
 - $B3$: Nếu mọi phụ thuộc hàm $X \rightarrow A \in F1_{tt}$ với $A \notin X$, X là siêu khóa, thì Q đạt chuẩn BCNF, ngược lại thì không đạt

Giải thuật kiểm tra BCNF

- VD: Cho $R(A,B,C,D,E,I)$ $F=\{ACD \rightarrow EBI; CE \rightarrow AD\}$. R đạt chuẩn BCNF?
- TN={C}, TG={ADE}

X_i	$(TN \cup X_i)$	$(TN \cup X_i)^+$	Super key	Key
ϕ	C	C		
A	AC	AC		
D	CD	CD		
AD	ACD	ABCDEI	ACD	ACD
E	CE	ABCDEI	CE	CE
AE	ACE	ABCDEI	ACE	
DE	CDE	ABCDEI	CDE	
ADE	ACDE	ABCDEI	ACDE	

$F \equiv F_{tt} = \{ACD \rightarrow E, ACD \rightarrow B, ACD \rightarrow I, CE \rightarrow A, CE \rightarrow D\}$

Vế trái của F_{tt} đều là siêu khóa \rightarrow Q đạt BCNF

Bài tập

1. Cho biết dạng chuẩn cao nhất của các LDQH sau:

- a) $Q(ABCDEG) \ F = \{A \rightarrow BC, C \rightarrow DE, E \rightarrow G\}$
- b) $Q(ABCDEFGH) \ F = \{C \rightarrow AB, D \rightarrow E, B \rightarrow G\}$
- c) $Q(ABCDEFGH) \ F = \{A \rightarrow BC, D \rightarrow E, H \rightarrow G\}$
- d) $Q(ABCDEG) \ F = \{AB \rightarrow C, C \rightarrow B, ABD \rightarrow E, G \rightarrow A\}$
- e) $Q(ABCDEFGHI) \ F = \{AC \rightarrow B, BI \rightarrow ACD, ABC \rightarrow D$
 $H \rightarrow I, ACE \rightarrow BCG, CG \rightarrow A\}$

2. Cho $Q(CDEGHK)$ và $F = \{CK \rightarrow H, C \rightarrow D, E \rightarrow C, E \rightarrow G, CK \rightarrow E\}$

- a) Chứng minh $EK \rightarrow DH$
- b) Tìm tất cả các khóa của Q
- c) Xác định dạng chuẩn cao nhất của Q

Phân rã (Decompositions)

➡ Phân rã một lược đồ quan hệ R : là thay thế lược đồ quan hệ R bằng các lược đồ quan hệ R_1, R_2, \dots, R_n , mà mỗi lược đồ quan hệ này chứa tập con thuộc tính của R , sao cho:

➡ $R = R_1 \cup R_2 \dots \cup R_n$

➡ Mỗi R_i là tập con của R (với $i = 1, 2, \dots, n$)

Ví dụ: Lược đồ quan hệ $R(x, y, z)$ có thể có 2 tập con:

$R_1(x, z)$ và $R_2(y, z)$, nếu hội R_1 và R_2 , ta được
 $R = R_1 \cup R_2$

Phân rã (Decompositions)

➤ Mục đích của phân rã:

- Giảm sự dư thừa bằng cách phân rã một quan hệ thành nhiều quan hệ ở dạng chuẩn cao hơn.

➤ Vấn đề với việc phân rã:

- Nếu $R \neq R_1 \cup R_2 \dots \cup R_n$ thì phép phân rã là mất mát thông tin

Phân rã (Decompositions)

➡ Ví dụ

T

Employee	Project	Branch
Brown	Mars	L.A.
Green	Jupiter	San Jose
Green	Venus	San Jose
Hoskins	Saturn	San Jose
Hoskins	Venus	San Jose

Employee → Branch, Project → Branch

T1

Employee	Branch
Brown	L.A.
Green	San Jose
Hoskins	San Jose

T2

Project	Branch
Mars	L.A.
Jupiter	San Jose
Saturn	San Jose
Venus	San Jose

Phân rã (Decompositions)

- Ví dụ: sau khi phân rã, dùng phép kết tự nhiên để kết 2 lược đồ T1 và T2 → không bảo toàn thông tin

Employee	Project	Branch
Brown	Mars	L.A.
Green	Jupiter	San Jose
Green	Venus	San Jose
Hoskins	Saturn	San Jose
Hoskins	Venus	San Jose
Green	Saturn	San Jose
Hoskins	Jupiter	San Jose

Employee	Project	Branch
Brown	Mars	L.A.
Green	Jupiter	San Jose
Green	Venus	San Jose
Hoskins	Saturn	San Jose
Hoskins	Venus	San Jose

Phép phân rã bảo toàn thông tin

(LossLess-join decomposition)

► Cho lược đồ quan hệ R và tập phụ thuộc hàm F .

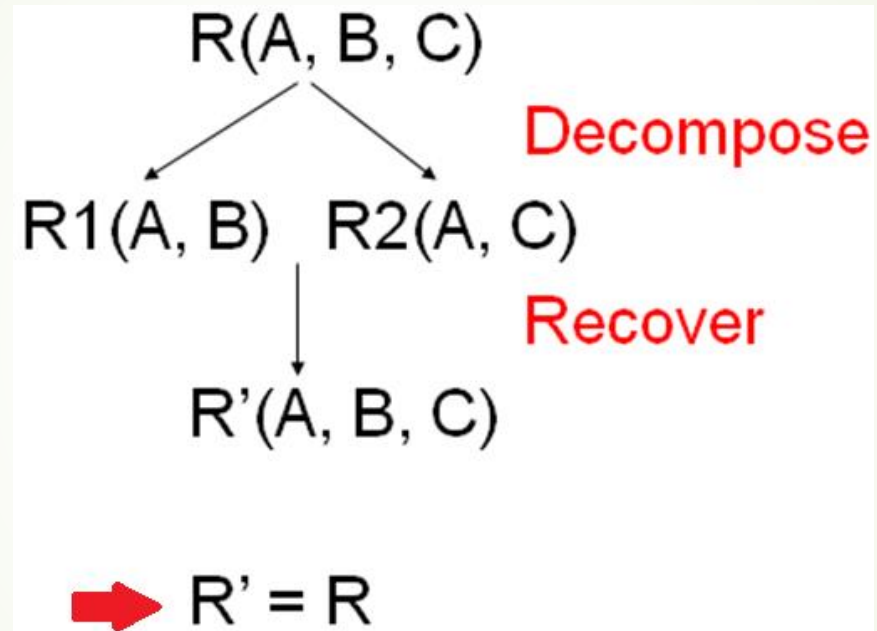
Phép phân rã R thành hai lược đồ với tập thuộc tính X và Y được gọi là phép phân rã bảo toàn thông tin đối với F nếu mọi thể hiện r của R thỏa mãn F ta luôn có:

$$\pi_X(r) \bowtie \pi_Y(r) = r$$

Phép phân rã bảo toàn thông tin

(LossLess-join decomposition)

➡ Ví dụ



Thuật toán kiểm tra phân rã bảo toàn thông tin

- Cho lược đồ quan hệ R , tập phụ thuộc hàm F và phép phân rã của R : $D=\{R_1, R_2, \dots, R_m\}$
- **Bước 1:** *Lập ma trận m hàng ứng với m lược đồ con R_i và n cột ứng với n thuộc tính của R . Phần tử (i,j) của ma trận được xác định như sau:*
 - $(i,j) = a_j$ nếu $a_j \in R_i$
 - $(i,j)=b_{ij}$ nếu $a_j \notin R_i$ trong đó, $a_j, b_{ij} \in \text{Dom}(A_j)$

Thuật toán kiểm tra phân rã bảo toàn thông tin

➤ Bước 2: Biến đổi bảng

- Với mỗi $FD X \rightarrow Y$ nếu có hai hàng giống nhau trên X và khác nhau trên Y thì làm cho hai hàng đó cũng giống nhau trên Y . (ưu tiên dùng aj).
- Tiếp tục với các phụ thuộc hàm trong bảng cho đến khi không còn áp dụng được nữa.
- Mục đích của việc biến đổi bảng là để thu được bảng cuối cùng, xem như một quan hệ thỏa tập F .

Thuật toán kiểm tra phân rã bảo toàn thông tin

- **Bước 3:** Xem bảng kết quả
 - Nếu trong bảng cuối cùng có chứa hàng gồm toàn ký hiệu a thì kết luận phép phân rã $D = (R_1, R_2, \dots, R_m)$ là bảo toàn thông tin, ngược lại thì D là phân rã mất thông tin

Ví dụ

- Phân rã lược đồ quan hệ $Q(ABCDE)$ thành $Q_1(AD)$, $Q_2(AB)$, $Q_3(BE)$, $Q_4(CDE)$, $Q_5(AE)$ và tập $F=\{A\rightarrow C, B\rightarrow C, A\rightarrow D, DE\rightarrow C, CE\rightarrow A\}$
- Phép phân rã trên có bảo toàn thông tin không?
- Bước 1

	A	B	C	D	E
$Q_1(AD)$	a1	b1	b2	a4	b3
$Q_2(AB)$	a1	a2	b4	b5	b6
$Q_3(BE)$	b7	a2	b8	b9	a5
$Q_4(CDE)$	b10	b11	a3	a4	a5
$Q_5(AE)$	a1	b12	b13	b14	a5

➡ Xét $A \rightarrow C$

$A \rightarrow C$	A	B	C	D	E
$Q_1(AD)$	a1	b1	b2	a4	b3
$Q_2(AB)$	a1	a2	b4 b2	b5	b6
$Q_3(BE)$	b7	a2	b8	b9	a5
$Q_4(CDE)$	b10	b11	a3	a4	a5
$Q_5(AE)$	a1	b12	b13 b2	b14	a5

Xét $B \rightarrow C$

$B \rightarrow C$	A	B	C	D	E
$Q_1(AD)$	a1	b1	b2	a4	b3
$Q_2(AB)$	a1	a2	b4 b2	b5	b6
$Q_3(BE)$	b7	a2	b8 b2	b9	a5
$Q_4(CDE)$	b10	b11	a3	a4	a5
$Q_5(AE)$	a1	b12	b4 b2	b14	a5

Xét $A \rightarrow D$

$A \rightarrow D$	A	B	C	D	E
$Q_1(AD)$	a1	b1	b2	a4	b3
$Q_2(AB)$	a1	a2	b4 b2	b5 a4	b6
$Q_3(BE)$	b7	a2	b8 b2	b9	a5
$Q_4(CDE)$	b10	b11	a3	a4	a5
$Q_5(AE)$	a1	b12	b4 b2	b14 a4	a5

Xét $DE \rightarrow C$

$DE \rightarrow C$	A	B	C	D	E
$Q_1(AD)$	a1	b1	b2 a3	a4	b3
$Q_2(AB)$	a1	a2	b4 b2 a3	b5 a4	b6
$Q_3(BE)$	b7	a2	b8 b2 a3	b9	a5
$Q_4(CDE)$	b10	b11	a3	a4	a5
$Q_5(AE)$	a1	B12	b4 b2 a3	b14 a4	a5

Xét $CE \rightarrow A$

$CE \rightarrow A$	A	B	C	D	E
$Q_1(AD)$	a1	b1	b2 a3	a4	b3
$Q_2(AB)$	a1	a2	b4 b2 a3	b5 a4	b6
$Q_3(BE)$	b7 a1	a2	b8 b2 a3	b9	a5
$Q_4(CDE)$	b10 a1	b11	a3	a4	a5
$Q_5(AE)$	a1	b12	b4 b2 a3	b14 a4	a5

Lần lượt xét lần 2, Xét $A \rightarrow D$

$A \rightarrow D$	A	B	C	D	E
$Q_1(AD)$	a1	b1	b2 a3	a4	b3
$Q_2(AB)$	a1	a2	b4 b2 a3	b5 a4	b6
$Q_3(BE)$	b7 a1	a2	b8 b2 a3	b9 a4	a5
$Q_4(CDE)$	b10 a1	b11	a3	a4	a5
$Q_5(AE)$	a1	b12	b4 b2 a3	b14 a4	a5

Dòng $Q_3(BE)$ chứa toàn giá trị aj \rightarrow Phép phân
rã trên là bảo toàn thông tin

Bài tập

- Kiểm tra phép tách sau có bảo toàn thông tin?
 1. Cho $R(A,B,C,D)$ và $F=\{A \rightarrow B, B \rightarrow C, A \rightarrow D, D \rightarrow C\}$. Phân rã thành: $R1(A,B)$, $R2(A,C)$, $R3(B,D)$.
 2. Cho $R(A,B,C,D,E)$ $F=\{A \rightarrow C,E; E \rightarrow A,D; C,D \rightarrow B\}$. Phân rã thành: $R1(A,B,C,E)$, $R2(A,C,D)$, $R3(D,E)$.

Hệ quả

- Gọi $\rho(Q1, Q2)$ là phép phân rã của lược đồ quan hệ Q và F là tập phụ thuộc hàm, ρ là phép phân rã không mất mát thông tin nếu và chỉ nếu

$$Q1 \cap Q2 \rightarrow Q1-Q2 \text{ hoặc } Q1 \cap Q2 \rightarrow Q2-Q1$$

Hệ quả

Ví dụ:

$Q(\text{SAIP}), Q_1(\text{SA}), Q_2(\text{SIP}), F=\{S \rightarrow A, SI \rightarrow P\}$

➡ $Q_1 \cup Q_2 = \text{SAIP} = Q^+$

➡ $Q_1 \cap Q_2 = S \rightarrow Q_1 - Q_2 = A$

=> Phép phân rã bảo toàn thông tin

Thuật toán phân rã dạng BCNF/3NF bảo toàn PHỤ THUỘC HÀM

- Bước 1: Tìm tất cả khóa của R
- Bước 2: Tìm phụ thuộc hàm $X \rightarrow A \in F$ có X không là siêu khóa và $A \in X$ (vi phạm BCNF). Nếu tìm thấy thì tách R thành R1 và R2 theo quy tắc sau:
 - R1(XA). Tìm bao đóng của tất cả tập con của XY để suy ra $F1 = \Pi_{R1}(F)$
 - R2($R^+ - A$). Tìm bao đóng của tất cả tập con của $R^+ - A$ để suy ra $F2 = \Pi_{R2}(F)$
 - Thực hiện thuật toán phân rã (R1, F1) và (R2, F2)

Thuật toán phân rã dạng BCNF/3NF bảo toàn thông tin

- Nếu không tìm thấy thì xét dạng chuẩn Qi:
 - Nếu mọi phụ thuộc hàm trong F_i đều có vế trái là siêu khóa thì Qi đạt dạng chuẩn BC
 - Nếu có phụ thuộc hàm trong F_i có vế trái là siêu khóa hoặc vế phải là thuộc tính khóa thì Qi đạt dạng chuẩn 3

Thuật toán phân rã dạng BCNF/3NF bảo toàn thông tin

➤ Ví dụ: Cho $Q(S,D,I,M)$ $F=\{SI \rightarrow D; SD \rightarrow M\}$ hãy phân rã thành các lược đồ con đạt chuẩn BCNF.

➤ Bước 1: Tìm tất cả khóa của Q

X_i	$TN \cup X_i$	$(TN \cup X_i)^+$	Siêu khóa	Khóa
\emptyset	SI	SDIM	SI	SI
D	SID	SDIM	SID	

➤ Bước 2: $SD \rightarrow M$ có SD không là siêu khóa \rightarrow tách Q thành $Q_1(SDM)$ và $Q_2(SDI)$.

Thuật toán phân rã dạng BCNF/3NF bảo toàn thông tin

- Tính $F_1 = \Pi_{Q_1}(F)$ của $Q_1(SDM)$ dựa trên F

$$\begin{array}{llll}
 S^+ = S & D^+ = D & M^+ = M & \\
 & SD^+ = SDM & SM^+ = SM & \\
 & & DM^+ = DM & \\
 & & SDM^+ = SDM &
 \end{array}$$

- Ta được

$$F_1^+ = \Pi_{Q_1}(F) = \{SD \rightarrow M, SD \rightarrow SM, SD \rightarrow DM, SD \rightarrow SDM\}$$

- $\rightarrow F_1 = \{SD \rightarrow M\}$

Thuật toán phân rã dạng BCNF/3NF bảo toàn thông tin

- Tính $F_2 = \Pi_{Q_2}(F)$ của $Q_2(SID)$ dựa trên F

$$\begin{array}{llll}
 S^+ = S & D^+ = D & I^+ = I \\
 & SD^+ = SDM & SI^+ = SDIM \\
 & & DI^+ = DI \\
 & & SDI^+ = SDIM
 \end{array}$$

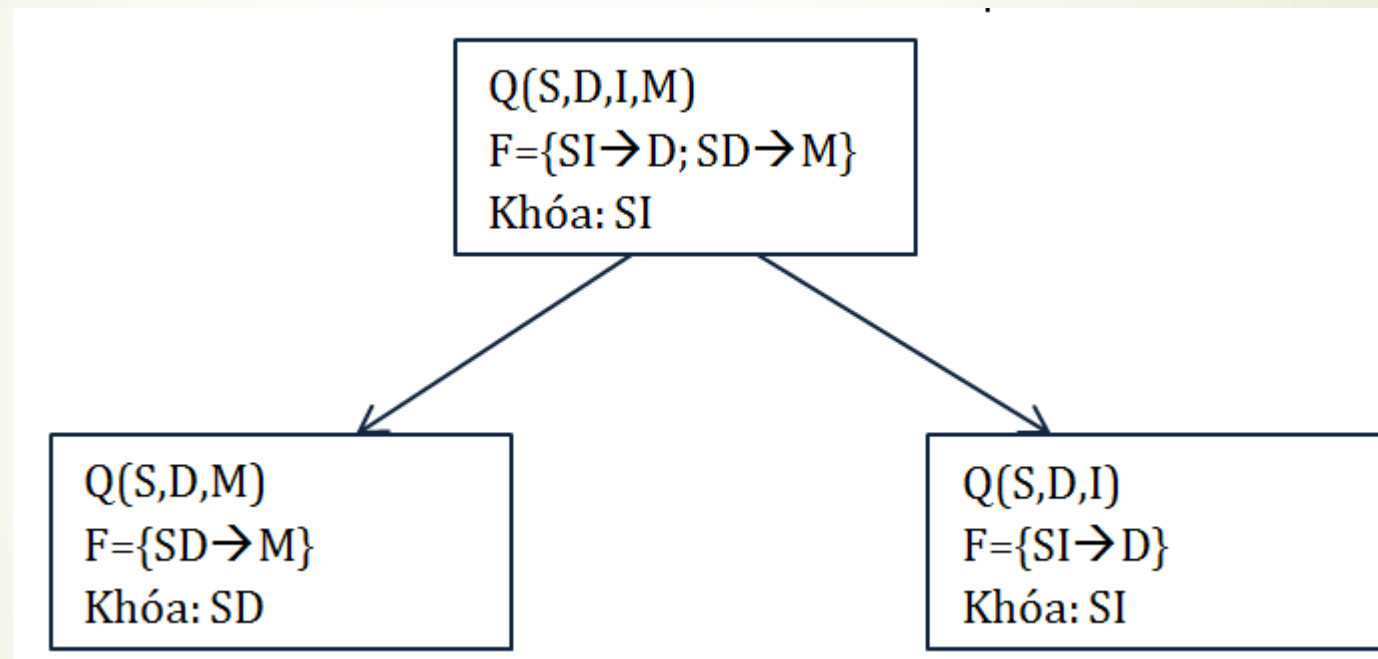
- Ta được:

$$F_2^+ = \Pi_{Q_2}(F) = \{SI \rightarrow D, SI \rightarrow SD, SI \rightarrow DI, SI \rightarrow SDI\}$$

- $\rightarrow F_2 = \{SI \rightarrow D\}$

Thuật toán phân rã dạng BCNF/3NF bảo toàn thông tin

➤ Kết quả:



Thuật toán phân rã dạng 3NF bảo toàn thông tin và bảo toàn phụ thuộc hàm

➡ Bước 1 : Tìm phủ tối thiểu của F

➡ Bước 2 :

Nếu có một PTH 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)

Nếu có những thuộc tính của Q không nằm trong một PTH nào của Ftt, dù ở vế trái hay vế phải của F, thì chúng tạo thành lược đồ cần tìm

Cứ mỗi PTH $X \rightarrow A$ thuộc Ftt thì XA là lược đồ cần tìm

Nếu có một lược đồ con chứa Khoá của Q thì kết thúc thuật toán.

Ngược lại tạo một lược đồ là K

Thuật toán phân rã dạng 3NF bảo toàn thông tin và bảo toàn phụ thuộc hàm

Cho $Q(CTHRSG)$ và

$F=\{C\rightarrow T, HR\rightarrow C, TH\rightarrow R, CS\rightarrow G, HS\rightarrow R\}$

Hãy phân rã Q thành các lược đồ con đạt dạng chuẩn 3 vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm

$Q1(CT)$ $Q2(HRC)$ $Q3(THR)$ $Q4(CSG)$ $Q5(HSR)$

Bài tập

1. Kiểm tra phân rã có bảo toàn thông tin hay không?
 - a) $R(CSZ)$, $F=\{CS \rightarrow Z, Z \rightarrow C\}$ có $\rho(SZ,CZ)$ là phép tách của lược đồ quan hệ R .
 - b) $R(ABCD)$, $F=\{A \rightarrow B, C \rightarrow D\}$ có $\rho(AB,CD)$ là phép tách của lược đồ quan hệ R .

Bài tập

2. Cho $R(ABCDEFGG)$, $F=\{B \rightarrow A, D \rightarrow C, D \rightarrow EB, DF \rightarrow G\}$, phân rã về dạng chuẩn BC bảo toàn thông tin.
3. Cho $Q(ABCDEG)$, $F=\{AE \rightarrow C, CG \rightarrow A, BD \rightarrow G, GA \rightarrow E\}$.
 - a) Tìm tất cả các khóa của Q .
 - b) Xác định dạng chuẩn của Q .
 - c) Nếu Q chưa đạt chuẩn 3. Hãy phân rã Q thành các quan hệ đạt chuẩn và bảo toàn thông tin.

Bài tập

4. Cho $R(A,B,C,D,E)$ và $F=\{CDE \rightarrow B, AE \rightarrow BD, DE \rightarrow C, CE \rightarrow D, B \rightarrow CD\}$

Hãy phân rã R thành lược đồ quan hệ đạt tối thiểu dạng chuẩn 3 vừa bảo toàn thông tin vừa bảo toàn phụ thuộc hàm.

Bài tập

5. Cho lược đồ quan hệ $R(U, F)$ với $U = ABCDEHIKJ$
 $F = \{ C \rightarrow EHI, HI \rightarrow ABC, AC \rightarrow DJ, EC \rightarrow AB \}$
- a) Tìm tất cả các khóa của lược đồ quan hệ trên
 - b) Lược đồ quan hệ trên đã thỏa 2NF chưa? Tại sao?
 - c) Dùng phép tách bảo toàn phụ thuộc hàm để tách R thành các LĐQH thỏa dạng chuẩn 3NF
 - d) Dùng phép tách có nối kết không mất thông tin để tách R thành các LĐQH thỏa 3NF hoặc BCNF

Bài tập

6. Cho $Q(SDIBQO)$ với các phụ thuộc hàm: $F = \{ S \rightarrow D, I \rightarrow B, IS \rightarrow Q, B \rightarrow O \}$.
- Lược đồ trên đang tồn tại ở dạng chuẩn cao nhất là bao nhiêu?
 - Hãy tách lược đồ trên thành dạng chuẩn 3NF có nối không mất thông tin và bảo toàn tập phụ thuộc hàm.

Bài tập

7. Cho $Q(ABCD)$ và $F=\{A \rightarrow B, B \rightarrow C, A \rightarrow D, D \rightarrow C\}$.
Tách Q thành $Q1(AB)$, $Q2(AC)$, $Q3(BD)$
- a) Có bảo toàn thông tin
 - b) Có bảo toàn phụ thuộc hàm

➤ **CÂU 1 (3 điểm)**

➤ Cho hai tập phụ thuộc hàm F và G

➤ $F = \{AB \rightarrow C, B \rightarrow C, CD \rightarrow E, D \rightarrow E\}$ và $G = \{AB \rightarrow C, B \rightarrow C, C \rightarrow B, BD \rightarrow E\}$

➤ Cho biết G và F có tương đương với nhau không (có diễn giải) ?

CÂU 2 (4 điểm)

➤ Cho quan hệ $R(ABCDEH)$ và tập các phụ thuộc hàm

➤ $F = \{ AB \rightarrow ED, BD \rightarrow CH, AC \rightarrow BD, ABC \rightarrow DH \}$

➤ Tìm một phủ tối thiểu của F (có diễn giải) ?

CÂU 3 (3 điểm)

➤ Cho tập phụ thuộc hàm:

➤ $F = \{AB \rightarrow E, AG \rightarrow I, BE \rightarrow I, E \rightarrow G, GI \rightarrow H\}$

➤ Chứng minh rằng $(AB \rightarrow GH)$ bằng tiên đề Armstrong