



Trường Đại học Nha Trang
Khoa Công nghệ Thông tin

CƠ SỞ DỮ LIỆU

Chủ đề 6+8: Chuẩn hoá CSDL, Phân rã bảo toàn thông tin

TS. Phạm Thị Thu Thúy
thuthuy@ntu.edu.vn

Nội dung

- ❖ Tại sao
- ❖ Chuẩn hóa dữ liệu
- ❖ Các dạng chuẩn
 - Dạng chuẩn 1 (1NF)
 - Dạng chuẩn 2 (2NF)
 - Dạng chuẩn 3 (3NF)
 - Dạng chuẩn Boyce codd (BCNF)
 - Dạng chuẩn 4
- ❖ Các bước chuẩn hóa lược đồ quan hệ
- ❖ Phân rã lược đồ quan hệ
 - Kiểm tra phân rã bảo toàn thông tin'
 - Kiểm tra phân rã bảo toàn phụ thuộc hàm
 - Phân rã 3NF
 - Phân rã BCNF
- ❖ Tổng kết về chuẩn hóa CSDL
- ❖ Bài tập

Tại sao?

- CungCapSP(**maNCC**, **tenNCC**, **diaChi**, **sanPham**, **gia**)

maNCC	tenNCC	diaChi	sanPham	gia
S01	Sanyo	Tokyo, Nhật Bản	Tủ lạnh	250
S01	Sanyo	Tokyo, Nhật Bản	Máy giặt	270
S01	Sanyo	Tokyo, Nhật Bản	Máy lạnh	200
S02	Sharp	Hiroshima, Nhật Bản	TiV	140
S03	Samsung	Hồng Kông, Đài Loan	LCD	200
S03	Samsung	Hồng Kông, Đài Loan	TiV	120
M01	Maytag	New York, Mỹ	Tủ lạnh	260

- Trùng lặp thông tin làm tăng chi phí lưu trữ thông tin và chi phí kiểm tra RBTV.
- **Cách giải quyết:** tách thành hai lược đồ như sau:
 - *CungCap(**maNCC**, **sanPham**, **gia**) và NhaCC(**maNCC**, **tenNCC**, **diaChi**) → quan hệ tương ứng:*

Tại sao?

Khóa chính là gì?

→ {maNV, tenKH}

maNV	tenNV	donV	luong	tenKH	ngayKTKH
100	Nguyễn Văn An	P. Tiếp thị	480	Anh văn	31/12/2009
100	Nguyễn Văn An	P. Tiếp thị	480	NC thị trường	10/03/2010
140	Trần Thị Mỹ	P. Kế toán	550	Luật thuế	15/04/2010
110	Lê Văn Minh	P. Hệ thống thông tin	500	Anh văn	31/12/2009
110	Lê Văn Minh	P. Hệ thống thông tin	500	C++	31/01/2010
190	Trần Văn Ban	P. Tài chính	520		
150	Hồ Minh Nhật	P. Tiếp thị	450	Anh văn	31/12/2009
150	Hồ Minh Nhật	P. Tiếp thị	450	Java	10/01/2010

Sự bất thường ở bảng này?

- **Thêm vào** – không thể thêm vào 1 nhân viên mới mà không tham gia lớp nào
- **Xoá đi** – nếu xoá bỏ nhân viên 140, ta sẽ mất thông tin về lớp **Luật thuế**
- **Hiệu chỉnh** – để tăng lương cho nhân viên 100, đòi hỏi phải cập nhật 2 hàng

Tại sao?

- ❖ **Sự trùng lặp thông tin là nguyên nhân làm cho CSDL có chất lượng kém.**
- ❖ **⇒ Cần đưa ra các tiêu chuẩn để có thể đánh giá chất lượng thiết kế của một lược đồ CSDL cũng như cần có phương pháp để làm cho một lược đồ CSDL trở nên tốt hơn.**

Chuẩn hóa dữ liệu

❖ Là một công cụ cơ bản để:

- **kiểm tra và cải tiến** một thiết kế CSDL luận lý, thỏa mãn các ràng buộc toàn vẹn → giúp tránh sự trùng lặp dữ liệu (data duplication) không cần thiết.

❖ Là quá trình phân rã các quan hệ không bình thường (anomaly) thành các quan hệ có cấu trúc tốt (well-structured) nhỏ hơn

- Quan hệ có cấu trúc tốt
 - Có sự dư thừa dữ liệu tối thiểu
 - Cho phép người dùng thêm vào, xóa đi và cập nhật các hàng (row) mà không gây ra sự mâu thuẫn dữ liệu nào

Chuẩn hóa dữ liệu (tt)

❖ Mục tiêu:

- Giảm dư thừa dữ liệu
- Tránh sự dị thường
 - Dị thường khi thêm vào (Insertion anomaly) – thêm các hàng mới đòi hỏi người dùng tạo ra các dữ liệu trùng lặp
 - Dị thường khi xóa (Deletion anomaly) – xóa các hàng đã có gây ra sự mất dữ liệu cần thiết cho các hàng khác
 - Dị thường khi hiệu chỉnh (Modification anomaly) – thay đổi dữ liệu ở một hàng đòi hỏi thay đổi dữ liệu của các hàng khác vì trùng lặp

❖ Để chuẩn hóa dữ liệu dựa vào các dạng chuẩn

Chuẩn hóa dữ liệu (tt)

❖ Có các dạng chuẩn sau:

- Dạng chuẩn 1 (First Normal Form) – 1NF
- Dạng chuẩn 2 (Second Normal Form) – 2NF
- Dạng chuẩn 3 (Third Normal Form) – 3NF
- Dạng chuẩn Boyce codd (Boyce codd Normal Form) – BCNF

1. Là gì?

2. Cách kiểm tra?

3. Cách chuẩn hóa?

Các dạng chuẩn – Dạng chuẩn 1 (1NF)

❖ **Khái niệm thuộc tính đơn (thuộc tính nguyên tố):**

- Một thuộc tính A của R là một thuộc tính đơn nếu nó không phải là sự tích hợp của nhiều thuộc tính khác.
- Ngược lại, nếu người sử dụng có thể truy xuất đến từng phần của thuộc tính A thì A được gọi là thuộc tính kép (thuộc tính không nguyên tố).

Các dạng chuẩn – Dạng chuẩn 1 (1NF)

- ❖ Khái niệm thuộc tính đơn (thuộc tính nguyên tố): (tt)
- ❖ Ví dụ: LĐQH Chuyen_Mon (MaGV, Mon) và quan hệ $r_{\text{chuyen_Mon}}$ như sau:

MaGV	Mon
GV01	Pascal, CTDL, TRR
GV02	CSDL
GV03	CSDI, CTDL

Mon là thuộc tính kép

- LĐQH *VatTu*(*maVT*, *tenVT*, *dvTinh*) trong đó *tenVT* bao gồm cả tên của vật tư và quy cách tính của nó \rightarrow *tenVT* là thuộc tính kép.

❖ Định nghĩa 1NF:

- Một lược đồ quan hệ được gọi là ở dạng chuẩn 1 (1NF) nếu mọi thuộc tính của R đều là thuộc tính đơn.
- Một lược đồ CSDL được gọi là ở dạng chuẩn 1 nếu mọi lược đồ quan hệ con R_i của nó đều ở dạng chuẩn 1.

❖ Cách gọi:

- Lược đồ quan hệ ở dạng chuẩn 1 (1NF) hay còn được gọi là lược đồ quan hệ 1NF.
- Lược đồ quan hệ 1NF còn được gọi là lược đồ quan hệ chuẩn hóa.
- Lược đồ quan hệ không ở dạng chuẩn 1 được gọi là lược đồ quan hệ không chuẩn hay phi chuẩn.

❖ Kiểm tra xem một lược đồ qh có ở dạng chuẩn 1 không?

→ Kiểm tra xem các thuộc tính của lược đồ qh có phải là thuộc tính đơn ko?

❖ Đưa lược đồ qh về dạng chuẩn 1:

→ Tách thuộc tính kép thành các thuộc tính đơn (tách dọc) hoặc tách thành các dòng (tách ngang)

Các dạng chuẩn – Dạng chuẩn 1 (1NF) (tt)

❖ Ví dụ:

- LĐQH VatTu(**maVT**, **tenVT**, dvTinh) được chuyển thành: VatTu(**maVT**, **tenVT**, **quyCach**, dvTinh)
- quan hệ $T_{\text{Chuyen_Mon}}$ ở trên không ở dạng chuẩn 1 vì thuộc tính Mon là thuộc tính kép \rightarrow tách thành các dòng để loại bỏ thuộc tính kép

maVT	tenVT	quyCach	dvTinh


maGV	Mon
GV01	Pasc
GV01	CTDL
GV01	TRR
GV02	CSDL
GV03	CSDL
GV03	CTDL

Các dạng chuẩn – Dạng chuẩn 1 (1NF) (tt)

❖ Ví dụ:

- Cho LĐQH CungCap(**maNCC**, **tenNCC**, **diaChi**, **sanPham**) trong đó thuộc tính **sanPham** bao gồm tên sản phẩm và giá. Giả sử ta có quan hệ r_{CungCap} sau:

maNCC	tenNCC	diaChi	sanPham
1	Hải Hà	Hà Nội	Kẹo mềm 100, Kẹo cứng 150, Bánh 200
2	Kinh Đô	Hồ Chí Minh	Kẹo 120, Bánh 150



maNCC	tenNCC	diaChi	tenSP	gia

Phi chuẩn

Các dạng chuẩn – Dạng chuẩn 1 (1NF) (tt)

❖ Ví dụ:

- Cho LĐQH CungCap(maNCC, tenNCC, diaChi, sanPham) trong đó thuộc tính sanPham bao gồm tên sản phẩm và giá. Giả sử ta có quan hệ r_{CungCap} sau:

maNCC	tenNCC	diaChi	sanPham
1	Hải Hà	Hà Nội	Kẹo mềm 100, Kẹo cứng 150, Bánh 200
2	Kinh Đô	Hồ Chí Minh	Kẹo 120, Bánh 150

Thông tin trùng lặp!!!

maNCC	tenNCC	diaChi	tenSP	gia
1	Hải Hà	Hà Nội	Kẹo mềm	100
1	Hải Hà	Hà Nội	Kẹo cứng	150
1	Hải Hà	Hà Nội	Bánh	200
2	Kinh Đô	Hồ Chí Minh	Kẹo	120
2	Kinh Đô	Hồ Chí Minh	Bánh	150

Phi chuẩn

chuẩn hóa

Nhận xét Dạng chuẩn 1

- Mọi lược đồ quan hệ đều thuộc dạng chuẩn 1.
- Dạng chuẩn 1 có thể dẫn đến sự trùng lặp dữ liệu. Do đó gây ra các dị thường về cập nhật dữ liệu.

Các dạng chuẩn – Dạng chuẩn 2 (2NF)

❖ Khái niệm PTH đầy đủ và PTH riêng phần

- Phụ thuộc hàm đầy đủ (Full FD)

PTH đầy đủ là PTH $X \rightarrow Y$ mà nếu loại bỏ bất kỳ thuộc tính A nào trong X thì $(X - A) \nrightarrow Y$

- Phụ thuộc hàm riêng phần (partial FD)

PTH riêng phần là PTH mà không phải là PTH đầy đủ.

Các dạng chuẩn – Dạng chuẩn 2 (2NF) (tt)

❖ K.n thuộc tính khóa, thuộc tính không khóa

- Thuộc tính khóa (thuộc tính nguyên tố):
 - Là một thuộc tính của lược đồ quan hệ R mà là thành phần của ít nhất một khóa dự tuyển trong R.
- Thuộc tính không khóa (thuộc tính không nguyên tố):
 - Là thuộc tính không phải là thuộc tính nguyên tố (không tham gia vào bất kỳ một khóa nào của LĐQH)
- Vd: $R(U, F)$, $U = ABCDE$, $K_1 = AB$, $K_2 = CB$
 - Tập thuộc tính khóa là: ABC
 - Tập thuộc tính không khóa là: DE

❖ Định nghĩa dạng chuẩn 2 (2NF):

- LĐQH R được gọi là ở dạng chuẩn 2 nếu:
 - R đạt 1NF
 - Mọi thuộc tính không khóa A trong R phụ thuộc hàm đầy đủ vào khóa của R

Các dạng chuẩn – Dạng chuẩn 2 (2NF) (tt)

❖ Ví dụ:

- LĐQH Dat_hang(soDH, maHH, ngayDH, maKH, soluongdat) với quan hệ như sau:

<u>soDH</u>	<u>maHH</u>	soluongdat	ngayDH	maKH
DH001	PCDX48	3	05/01/06	CTCN01
DH002	HDQT05	5	05/01/06	CTVN01
DH002	HDQT10	2	05/01/06	CTCN01
DH001	KBMS01	5	05/01/06	CTCN01
DH004	PCDX48	5	15/01/07	CTHV01

- *FD: soDH → ngayDH, maKH*
soDH, maHH → soluongdat, ngayDH, maKH
- LĐQH trên không đạt 2NF vì thuộc tính không khóa **ngayDH, maKH** không phụ thuộc đầy đủ vào khóa
K = {soDH, maHH} (chỉ phụ thuộc vào soDH)

Các dạng chuẩn – Dạng chuẩn 2 (2NF) (tt)

- ❖ LĐQH Ctiet_HD(**soHD**, **maHH**, giaban, soluongban)
 - FD: soHD, maHH \rightarrow giaban, soluongban
 - Khóa: K = {SoHD, maHH}
 - Hai thuộc tính không khóa **giaban** và **soluongban** đều phụ thuộc đầy đủ vào khóa nên LĐQH này đạt 2NF.

- ❖ **Nhận xét:**
 - Nếu LĐQH R chỉ có một khóa K và $\text{card}(K)=1$ thì R ở dạng chuẩn 2
 - Một LĐQH ở dạng chuẩn 2 vẫn có thể chứa đựng sự trùng lặp thông tin.

Các dạng chuẩn – Dạng chuẩn 2 (2NF) (tt)

❖ **Định nghĩa dạng chuẩn 2 (2NF):**

- LĐQH R được gọi là ở dạng chuẩn 2 nếu:
 - R đạt 1NF
 - Mọi thuộc tính không khóa A trong R phụ thuộc hàm đầy đủ vào khóa của R

❖ **Định nghĩa khác:**

- Lược đồ quan hệ đạt 2NF nếu với mọi FD $X \rightarrow A$ đúng trong R thì:
 - (1) $A \in X$ hoặc
 - (2) A là thuộc tính khóa trong R hoặc
 - (3) X không là tập con của một khóa nào, hoặc
 - (4) X là siêu khóa của R

Các dạng chuẩn – Dạng chuẩn 2 (2NF)

- Kiểm tra LĐQH đạt 2NF

❖ Cách kiểm tra lược đồ quan hệ ở dạng chuẩn 2:

- Tìm tất cả các khóa của lược đồ quan hệ
- Xác định tập thuộc tính khóa, tập thuộc tính không khóa
- Kiểm tra lần lượt từng PTH: $X \rightarrow A$ trong tập F
 - Nếu có PTH $X \rightarrow A$ nào mà:
 - X là tập con của khóa và
 - A là thuộc tính không khóathì kết luận lược đồ này không ở dạng chuẩn 2
 - Ngược lại, nếu tất cả các PTH $X \rightarrow A$ trong F có:
 - vế trái không phải là tập con của khóa hoặc
 - vế phải là thuộc tính khóathì kết luận lược đồ ở dạng chuẩn 2

Các dạng chuẩn – Dạng chuẩn 2 (2NF)

- Kiểm tra LĐQH đạt 2NF (tt)

❖ Ví dụ:

- Cho $R(ABCD)$
- $F = \{AB \rightarrow C, A \rightarrow D, BD \rightarrow C\}$

Hỏi: Lược đồ R có ở dạng chuẩn 2 không?

❖ Cách làm:

- Khóa là: $K = AB$
- Tập thuộc tính khóa là $\{A, B\}$
- Tập thuộc tính không khóa là $\{C, D\}$
- Xét fd: $A \rightarrow D$

- A là tập con của khóa, D là thuộc tính không khóa

→ Rõ ràng thuộc tính không khóa D phụ thuộc không đầy đủ vào khóa → R không đạt dạng chuẩn 2 (không là 2NF)

Các dạng chuẩn – Dạng chuẩn 2 (2NF)

- Kiểm tra LĐQH đạt 2NF (tt)

❖ Ví dụ:

Cho $R(CSZ)$ $F = \{CS \rightarrow Z, Z \rightarrow C\}$

Hỏi: Lược đồ R có ở dạng chuẩn 2 không?

❖ Cách làm:

- Khóa là: $K1 = CS, K2 = SZ$
 - Tập thuộc tính khóa: CSZ
 - Không có thuộc tính không khóa
- Không thể nào có trường hợp thuộc tính không khóa phụ thuộc không đầy đủ vào khóa → Lược đồ R ở dạng chuẩn 2.

❖ Nhận xét:

- **Nếu tất cả các thuộc tính của lược đồ quan hệ đều là thuộc tính khóa thì lược đồ là 2NF.**

Các dạng chuẩn – Dạng chuẩn 2 (2NF)

- Kiểm tra LĐQH đạt 2NF (tt)

❖ Ví dụ:

- Cho $R(ABCD)$
- $F = \{A \rightarrow BC, B \rightarrow D\}$

Hỏi: Lược đồ R có ở dạng chuẩn 2 không?

❖ Cách làm:

- Lược đồ quan hệ trên có một khóa duy nhất là: $K = A$
- \rightarrow Lược đồ R là 2NF vì không tồn tại tập con thực sự của khóa nên không tồn tại phụ thuộc hàm không đầy đủ của thuộc tính không khóa vào khóa.

❖ Nhận xét:

- *Nếu tất cả các khóa của lược đồ quan hệ chỉ có một thuộc tính thì lược đồ quan hệ là 2NF*

Các dạng chuẩn – Dạng chuẩn 2 (2NF) - Kiểm tra LĐQH đạt 2NF (tt)

❖ Ví dụ: cho LĐQH sau:

CungCap(maNCC, tenNCC, diaChi, sanPham, gia)

F = {maNCC → tenNCC, diaChi
maNCC, sanPham → gia}

- Khóa: K = {maNCC, sanPham}
 - Tập thuộc tính không khóa: {tenNCC, diaChi, gia}
 - Ta thấy với PTH: **maNCC → tenNCC, diaChi**
 - tenNCC, diaChi là thuộc tính không khóa, phụ thuộc vào maNCC là tập con thực sự của khóa K = {maNCC, sanPham}
- Lược đồ CungCap không ở dạng chuẩn 2

→ Đưa về dạng chuẩn 2

Các dạng chuẩn – Dạng chuẩn 2 (2NF)

- Chuẩn hóa lược đồ về 2NF

❖ **Cách làm:**

- Xác định phụ thuộc hàm vi phạm dạng chuẩn 2
- Nhóm các thuộc tính có mặt trong phụ thuộc hàm đó, bổ sung thêm các thuộc tính khóa (nếu cần) để có một lược đồ quan hệ.
- Các thuộc tính và các phụ thuộc hàm còn lại nhóm vào trong một quan hệ.
- Lặp lại quá trình với lược đồ chưa đạt 2NF cho đến khi tất cả các lược đồ quan hệ con (tách ra từ lược đồ quan hệ ban đầu) đều là 2NF.

Các dạng chuẩn – Dạng chuẩn 2 (2NF)

- Chuẩn hóa lược đồ về 2NF (tt)

❖ Ví dụ: Cho LĐQH sau:

CungCap(maNCC, sanPham, tenNCC, diaChi, gia)

F = {maNCC → tenNCC, diaChi
maNCC, sanPham → gia}

Phụ thuộc hàm vi phạm

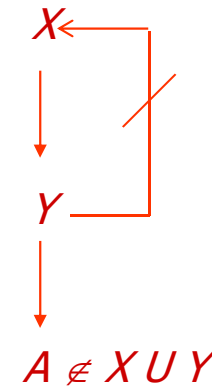
❖ Phân rã lược đồ CungCap thành 2 lược đồ:

- NhaCungCap(maNCC, tenNCC, diaChi)
 - FD: {maNCC → tenNCC, diaChi}
 - Khóa: maNCC
- CungCap(maNCC, sanPham, gia)
 - FD: {maNCC, sanPham → gia}
 - Khóa: maNCC, sanPham

Các dạng chuẩn – Dạng chuẩn 3 (3NF)

❖ Phụ thuộc hàm bậc cầu

- Cho lược đồ quan hệ $R(U, F)$ A là một thuộc tính của U . A được gọi là phụ thuộc hàm bậc cầu vào X trên R nếu tồn tại một tập con Y của U sao cho:
- $X \rightarrow Y$, $Y \rightarrow A$ nhưng $Y \not\rightarrow X$ với $A \notin X \cup Y$



❖ Định nghĩa dạng chuẩn 3 (3NF):

- LĐQH R được gọi là ở dạng chuẩn 3 nếu:
 - R ở dạng chuẩn 2
 - Mọi thuộc tính không khóa của R đều **không phụ thuộc bậc cầu vào khóa chính**

Các dạng chuẩn – Dạng chuẩn 3 (3NF) (tt)

❖ Ví dụ:

- $R = (\underline{\text{Cust_ID}}, \text{Name}, \text{Salesperson}, \text{Region})$
 $F = \{ \text{Cust_ID} \rightarrow \{\text{Name}, \text{Salesperson}, \text{Region}\}, \text{Salesperson} \rightarrow \text{Region} \}$
- Khóa $K = \text{Cust_ID}$
- Phụ thuộc hàm bậc cầu:
 $\text{Cust_ID} \rightarrow \text{Salesperson} \rightarrow \text{Region}$
Do đó lược đồ quan hệ không đạt dạng chuẩn 3.

Các dạng chuẩn – Dạng chuẩn 3 (3NF) (tt)

❖ Định nghĩa dạng chuẩn 3 (3NF):

- LĐQH R được gọi là ở dạng chuẩn 3 nếu:
 - R ở dạng chuẩn 2
 - Mọi thuộc tính không khóa của R đều **không phụ thuộc bắc cầu** vào khóa chính

❖ ĐN khác:

- R là 3NF nếu với mọi PTH $X \rightarrow A$ trong F, $A \notin X$ thì
 - X là siêu khóa hoặc
 - A là thuộc tính khóa (nguyên tố).

Các dạng chuẩn – Dạng chuẩn 3 (3NF)

- Kiểm tra LĐQH đạt 3NF

❖ **Cách làm:**

- Xác định khóa của lược đồ quan hệ
 - Kiểm tra LĐQH đã là 2NF chưa, nếu đã là 2NF, duyệt lần lượt từng phụ thuộc hàm trong F, nếu tồn tại một phụ thuộc hàm $X \rightarrow A$ mà:
 - X không là khóa và
 - A là thuộc tính không khóa
- Kết luận vi phạm dạng chuẩn 3
- Nếu không có phụ thuộc hàm nào vi phạm dạng chuẩn 3 thì R là đạt dạng chuẩn 3.

Các dạng chuẩn – Dạng chuẩn 3 (3NF)

- Kiểm tra LĐQH đạt 3NF (tt)

❖ Ví dụ:

- $R = (C, S, Z), \quad F = \{CS \rightarrow Z, Z \rightarrow C\}$
- Khóa dự tuyển: CS và SZ
- Tất cả các thuộc tính đều là thuộc tính khóa
→ R là 3NF

❖ Ví dụ:

- $R = (A, S, I, P), \quad F = \{SI \rightarrow P, S \rightarrow A\},$
- Khóa: SI
- $S \rightarrow A$ nhưng S là tập con của khóa và A là thuộc tính không khóa → R không là 2NF → R không là 3NF

Các dạng chuẩn – Dạng chuẩn 3 (3NF) (tt)

❖ **Nhận xét:**

- Từ định nghĩa ta có nhận xét rằng, nếu R là 3NF thì nó là 2NF, trong dạng chuẩn 2NF ta chỉ cấm các thuộc tính không khóa phụ thuộc vào tập con thực sự của khóa (là tập có bao đóng khác U),
- Trong 3NF, ta cấm các thuộc tính không khóa phụ thuộc vào mọi tập có bao đóng khác U (trong đó có tập con thực sự của khóa).

Các dạng chuẩn – Dạng chuẩn 3 (3NF)

- Cách chuẩn hóa lược đồ về 3NF

❖ **Cách làm:**

- Xác định PTH vi phạm dạng chuẩn 3 (3NF)
- Tách các nhóm thuộc tính có PTH bậc cầu vào khóa thành một QH. Khóa của QH mới này chính là thuộc tính mà chúng có phụ thuộc hàm.
- Các thuộc tính và các PTH còn lại nhóm vào trong một quan hệ.
- Lặp lại quá trình với lược đồ chưa đạt 3NF cho đến khi tất cả các LĐQH con (tách ra từ lược đồ quan hệ ban đầu) đều là 3NF.

Các dạng chuẩn – Dạng chuẩn 3 (3NF)

- Cách chuẩn hóa lược đồ về 3NF (tt)

(a) Quan hệ SALES với các dữ liệu mẫu

Cust_ID	Name	Salesperson	Region
8023	Anderson	Smith	South
9167	Bancroft	Hicks	West
7924	Hobbs	Smith	South
6837	Tucker	Hernandez	East
8596	Eckersley	Hicks	West
7018	Arnold	Faulb	North

(b) Quan hệ SALES với phụ thuộc hàm bậc cao

$CustID \rightarrow Name$

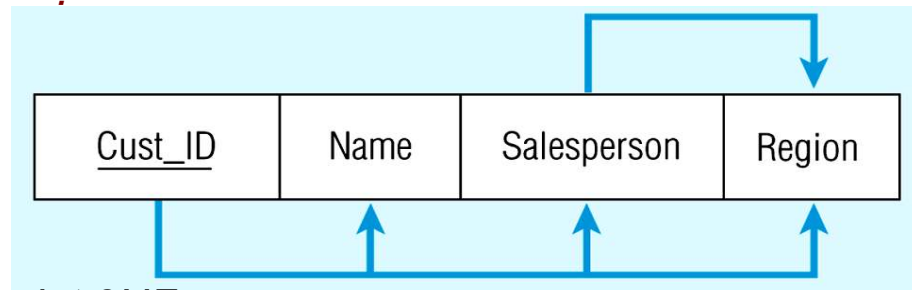
$CustID \rightarrow Salesperson$

$CustID \rightarrow Region$

$Salesperson \rightarrow Region$

$CustID \rightarrow Salesperson \rightarrow Region$

PTH bậc cao \rightarrow không đạt 3NF



Các dạng chuẩn – Dạng chuẩn 3 (3NF)

(c) Phân rã quan hệ SALES

Không còn phụ thuộc hàm bậc cầu nữa ...

Cả hai quan hệ bây giờ là ở dạng chuẩn 3

SALES1			SPERSON	
Cust_ID	Name	Salesperson	Salesperson	Region
8023	Anderson	Smith	Smith	South
9167	Bancroft	Hicks	Hicks	West
7924	Hobbs	Smith	Hernandez	East
6837	Tucker	Hernandez	Faulb	North
8596	Eckersley	Hicks		
7018	Arnold	Faulb		

(d) Quan hệ ở dạng chuẩn 3

CustID → Name

CustID → Salesperson



Các dạng chuẩn – Dạng chuẩn 3 (3NF)

- Cách chuẩn hóa lược đồ về 3NF (tt)

❖ Ví dụ: Xét lại lược đồ csdl KeHoach

■ Thi (ngThi, gioThi, phThi, maMH, gvDay)

- FD: { maMH → gvDay
ngThi, gioThi, phThi → maMH }
- Khóa: {ngThi, gioThi, phThi }



■ CoiThi (giamThi, ngThi, gioThi, phThi)

- FD: {giamThi → ngThi, gioThi, phThi}
- Khóa: giamThi



Các dạng chuẩn – Tóm lại

Dạng chuẩn	Thuộc tính không khóa phụ thuộc vào tập con của khóa	Thuộc tính không khóa phụ thuộc vào tập khác tập con của khóa	Thuộc tính khóa phụ thuộc tập khác	Thuộc tính khóa và thuộc tính không khóa phụ thuộc vào khóa
2NF	X	V	V	V
3NF	X	X	V	V

Các dạng chuẩn – Dạng chuẩn Boyce codd

❖ **ĐN:**

- R là dạng chuẩn Boyce-Codd nếu với mọi $X \rightarrow A$, $A \notin X$ thì X là siêu khóa.

❖ **Kiểm tra lược đồ là BCNF**

- Duyệt qua lần lượt từng phụ thuộc hàm $X \rightarrow A$ trong F
- Nếu tồn tại một phụ thuộc hàm mà X không phải là siêu khóa thì kết luận lược đồ không là BCNF
- Ngược lại, nếu mọi fd $X \rightarrow A$ trong F đều có vế trái là siêu khóa thì kết luận lược đồ là BCNF

Các dạng chuẩn – Dạng chuẩn Boyce codd

❖ Ví dụ:

- $R = (C, S, Z), \quad F = \{CS \rightarrow Z, Z \rightarrow C\}$

Khóa dự tuyển: CS và SZ \Rightarrow Dạng chuẩn 3

Có $Z \rightarrow C$, nhưng Z không là siêu khóa

\Rightarrow Không phải BCNF

❖ **Định lý:** Nếu lược đồ quan hệ R với tập phụ thuộc hàm F đạt BCNF thì nó cũng đạt 3NF.

- Trong 3NF chỉ cấm các thuộc tính không khóa phụ thuộc vào tập có bao đóng khác U, còn trong BCNF, ta cấm tất cả các thuộc tính phụ thuộc vào tập có bao đóng khác U.

Các dạng chuẩn – Dạng chuẩn Boyce codd

❖ **Mệnh đề:**

- Mọi lược đồ quan hệ gồm hai thuộc tính thì ở dạng BCNF
- Lược đồ $R(U, F)$ với tập F chỉ gồm một phụ thuộc hàm thì ở dạng BCNF
- Nếu $R(U, F)$ không đạt BCNF thì tồn tại hai thuộc tính A, B sao cho:

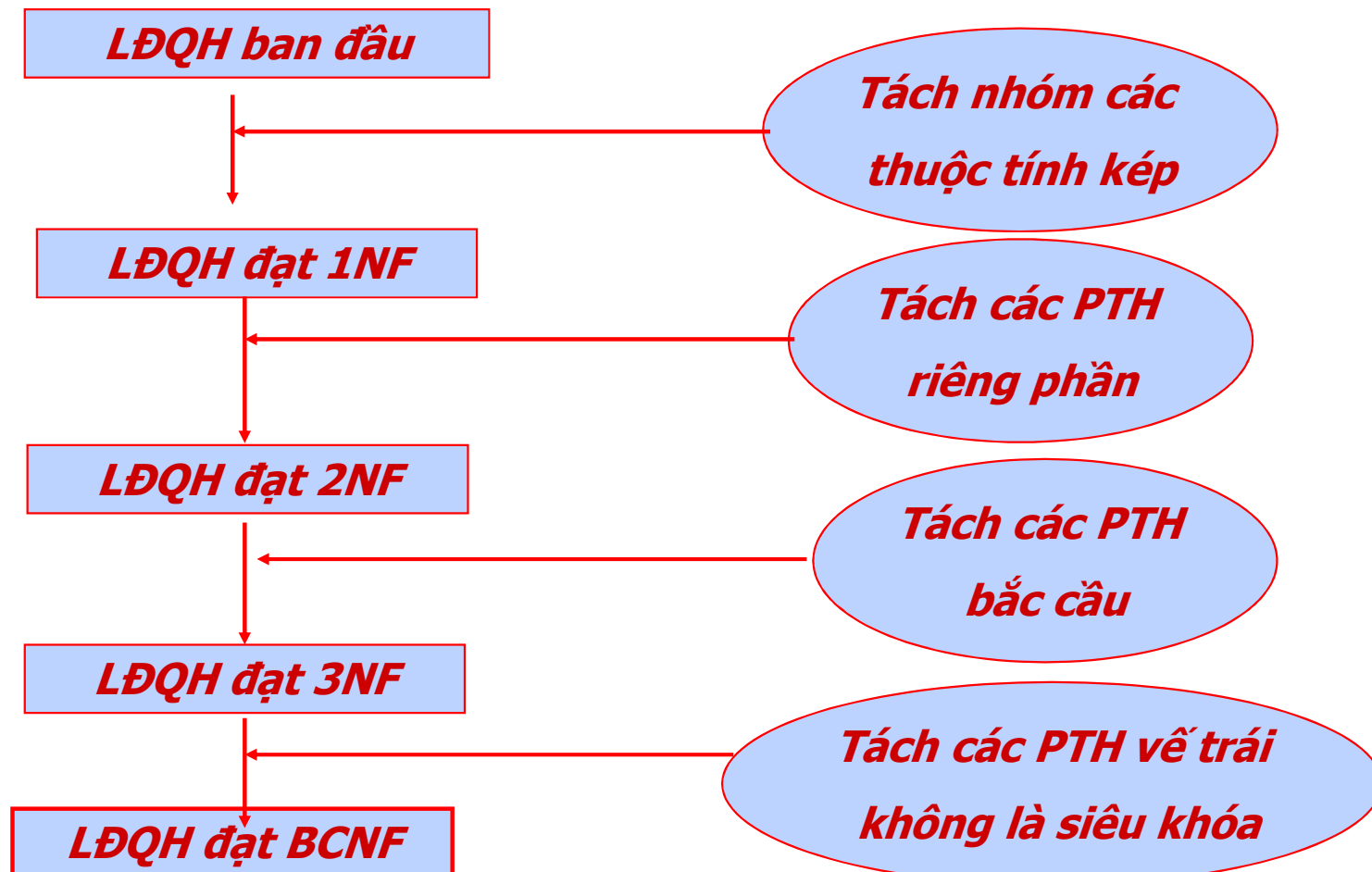
$$R \setminus (A \cup B) \rightarrow A$$

Các dạng chuẩn – Tóm lại

Dạng chuẩn	Thuộc tính không khóa phụ thuộc vào tập con của khóa	Thuộc tính không khóa phụ thuộc vào tập khác tập con của khóa	Thuộc tính khóa phụ thuộc tập khác	Thuộc tính khóa và thuộc tính không khóa phụ thuộc vào khóa
2NF	X	V	V	V
3NF	X	X	V	V
BCNF	X	X	X	V

The Key, The Whole Key, Nothing but the Key, so help me Codd!!!

Các bước chuẩn hóa



PHÂN RÃ LỰC ĐỒ QUAN HỆ

Phân rã các lược đồ quan hệ

❖ **Phân rã (decomposition):**

- một lược đồ quan hệ $R(U)$ là thay R bằng tập các lược đồ con $R_1(U_1), R_2(U_2), \dots, R_k(U_k)$ của R sao cho $U_1 \cup U_2 \cup \dots \cup U_k = U$.

❖ **Phân rã bảo toàn nội dung:**

- Phép kết tự nhiên các phân rã cho trở lại R

❖ **Phân rã bảo toàn phụ thuộc hàm:**

- Gọi $F_i = \{X \rightarrow Y \text{ đúng trong } F, XY \subseteq U_i\}$ là tập phụ thuộc hàm của R_i
- $\cup F_i$ tương đương F

Phân rã các lược đồ quan hệ
- Kiểm tra phân rã bảo toàn thông tin

- ❖ **I: $R(U, F)$, một phân rã của R , $p = (R_1, R_2, \dots, R_n)$**
- ❖ **O: Cho câu trả lời phân rã có mất thông tin hay không?**
- ❖ **Phương pháp:**
 - (1). Lập bảng ban đầu là ma trận m hàng (ứng với m lược đồ con R_i) và n cột (ứng với n thuộc tính). Phần tử (i, j) của ma trận được xác định theo công thức:
 - $(i, j) = a_j$ nếu $A_j \in R_i$
 - $(i, j) = b_{ij}$ nếu $A_j \notin R_i$ trong đó, $a_j, b_{ij} \in \text{Dom}(A_j)$

Phân rã các lược đồ quan hệ

- Kiểm tra phân rã bảo toàn thông tin (tt)

■ (2). Biến đổi bảng

- Với mỗi FD nếu phát hiện trên bảng có hai hàng giống nhau trên X và khác nhau trên Y \rightarrow làm cho hai hàng đó cũng giống nhau trên Y. (ưu tiên làm bằng ký hiệu là a_j).
- Tiếp tục áp dụng các phụ thuộc hàm cho bảng (kể cả việc lặp lại các phụ thuộc hàm đã áp dụng) 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.

Phân rã các lược đồ quan hệ

- Kiểm tra phân rã bảo toàn thông tin (tt)

- (3) Xem bảng kết quả:
 - Nếu trong bảng cuối cùng có chứa hàng gồm toàn ký hiệu a (tức là hàng (a_1, a_2, \dots, a_n)), ta kết luận $p = (R_1, R_2, \dots, R_m)$ là phân tách không mất thông tin (tức cho câu trả lời đúng)
 - Trường hợp ngược lại, p là phân tách mất thông tin.

❖ **Nhận xét:** thuật toán dừng khi xuất hiện hàng toàn a_j

❖ **Ví dụ:**

Cho $R(U, F)$ $U = BOISQD$

$F = \{S \rightarrow D \quad I \rightarrow B \quad IS \rightarrow Q \quad B \rightarrow O\}$

Phân rã $p = (SD, IB, ISQ, BO)$ có kết nối mất thông tin không?

Bảng ban đầu

	B	O	I	S	Q	D
SD	b11	b12	b13	a4	b15	a6
IB	a1	b22	a3	b24	b25	b26
ISQ	b31	b32	a3	a4	a5	b36
BO	a1	a2	b43	b44	b45	b46

S→D

	B	O	I	S	Q	D
SD	b11	b12	b13	a4	b15	a6
IB	a1	b22	a3	b24	b25	b26
ISQ	b31	b32	a3	a4	a5	a6
BO	a1	a2	b43	b44	b45	b46

I→B

	B	O	I	S	Q	D
SD	b11	b12	b13	a4	b15	a6
IB	a1	b22	a3	b24	b25	b26
ISQ	a1	b32	a3	a4	a5	a6
BO	a1	a2	b43	b44	b45	b46

B→O

	B	O	I	S	Q	D
SD	b11	b12	b13	a4	b15	a6
IB	a1	a2	a3	b24	b25	b26
ISQ	a1	a2	a3	a4	a5	a6
BO	a1	a2	b43	b44	b45	b46

***Dòng toàn a ⇒
phép tách
không mất
thông tin***

Phân rã các lược đồ quan hệ

- Kiểm tra phân rã bảo toàn thông tin (tt)

❖ **Định lý: (Delobel)**

- Cho $R(U, F)$ là một lược đồ quan hệ, khi đó $p = (R1, R2)$ là phân tách có kết nối không tồn thất khi và chỉ khi:

$$(R1 \cap R2) \rightarrow (R1 - R2) \text{ hay } (R1 \cap R2) \rightarrow (R2 - R1)$$

❖ **Nói cách khác:**

- Với $R(U, F)$ là một lược đồ quan hệ, giả sử $X \rightarrow Y \in F^+$ khi đó $p = (XY, X(U \setminus Y))$ là phân tách có kết nối không tồn thất với F .

❖ **Dạng phát biểu này của định lý sẽ là cơ sở cho việc kiểm tra phép tách BCNF nêu ra ở phần sau.**

Phân rã các lược đồ quan hệ

- Kiểm tra phân rã bảo toàn phụ thuộc hàm

- ❖ **I:** một phân rã $p = (R_1, \dots, R_k)$ và tập F
- ❖ **O:** khẳng định xem p có bảo toàn F hay không?
- ❖ **Phương pháp:**
 - Đặt $G = \bigcap_{i=1}^k R_i(F)$.
 - Mục đích của ta là: kiểm tra xem G có tương đương với F hay không.
 - Để xem G có tương đương với F hay không, ta phải xét mỗi phụ thuộc $X \rightarrow Y$ trong F và xác định xem X^+ , được tính ứng với G , có chứa Y hay không. Thủ thuật để tính X^+ mà không cần có G là xét lặp đi lặp lại kết quả tính bao đóng X^+ ứng với các hình chiếu của F trên các R_i .

Phân rã các lược đồ quan hệ

- Kiểm tra phân rã bảo toàn phụ thuộc hàm (tt)

❖ Ví dụ: Cho LĐQH $R(ABCDE)$ $F = \{E \rightarrow AC, A \rightarrow B, CD \rightarrow A, AB \rightarrow C, E \rightarrow B\}$

Cho phân rã sau của R , $p = (ACE, ABCD)$

Hỏi phân rã p có bảo toàn phụ thuộc hàm không?

❖ Cách làm:

▪ Đặt $R_1(ACE)$, khi đó $F_1 = \{E \rightarrow AC\}$

$R_2(ABCD)$, khi đó $F_2 = \{A \rightarrow B, CD \rightarrow A, AB \rightarrow C\}$

Đặt $G = F_1 \cup F_2 = \{E \rightarrow AC, A \rightarrow B, CD \rightarrow A, AB \rightarrow C\}$

Dễ thấy $G \subset F$ do đó F phủ G (1)

Ngoài ra, tất cả các PTH trong F đều thuộc G trừ PTH $E \rightarrow B$. Mà ta có $(E_G)^+ \supset B$, do đó $G \vdash E \rightarrow B$, do đó G phủ F (2)

(1) và (2) $\rightarrow F$ và G tương đương

\rightarrow Phân rã bảo toàn phụ thuộc hàm.

Phân rã các lược đồ quan hệ (tt)

❖ **Chú ý:** Hai tính chất kết nối không /tồn thất và bảo toàn tập F là độc lập với nhau.

- **Ví dụ:** $R(CSZ)$, $F = \{CS \rightarrow Z, Z \rightarrow C\}$, khi đó phân tách $p = (SZ, CZ)$ có kết nối không tồn thất vì $(SZ \cap CZ) \rightarrow (CZ - SZ)$ (nghĩa là $Z \rightarrow C$)

Tuy nhiên, chiếu của F trên SZ chỉ cho những phụ thuộc tầm thường không bảo toàn tập F.

- Cũng vậy, có những phân rã bảo toàn F nhưng không có tính chất nối không mất thông tin. Ví dụ với $F = \{A \rightarrow B, C \rightarrow D\}$, $U = ABCD$ và $p = (AB, CD)$

Phân rã các lược đồ quan hệ

❖ Giải thuật:

Input: Nhập: R và F là phủ tối thiểu

Output:

1. Tạo ra R_1 với tập các thuộc tính không có trong vế trái hay vế phải của bất kỳ phụ thuộc hàm nào.

■ 2. Nếu có phụ thuộc hàm nào trong F chứa toàn bộ các thuộc tính của R thì tạo ra R.

■ 3. Ngược lại:

• Tạo ra lược đồ XA cho mỗi $X \rightarrow A$

• Tuy nhiên nếu $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_n$ thì thay thế tập thuộc tính $XA_1A_2\dots A_n$ cho XA_i ($i = 1 \dots n$)

❖ Đánh giá:

■ Phân rã này bảo toàn thông tin và phụ thuộc hàm

Phân rã các lược đồ quan hệ

❖ Lược đồ với **F** là phủ tối thiểu

- $R = CTHRSG$, $F = \{C \rightarrow T, HR \rightarrow C, HT \rightarrow R, CS \rightarrow G, HS \rightarrow R\}$

❖ Kết xuất:

- CT $F_1 = \{C \rightarrow T\}$
- CHR $F_2 = \{HR \rightarrow C\}$
- THR $F_3 = \{HT \rightarrow R\}$
- CSG $F_4 = \{CS \rightarrow G\}$
- HRS $F_5 = \{HS \rightarrow R\}$

Phân rã các LĐQH- thành BCNF

- Vào: LĐQH R và tập pth F.
- Ra: Phép tách của R không mất mát thông tin sao cho mỗi LĐQH trong phép tách đều ở BCNF đối với phép chiếu của F trên lược đồ đó.

Các bước của thuật toán:

- Ban đầu phép tách ρ chỉ bao gồm R.
- Các bước tiếp: Nếu S là một lược đồ thuộc ρ và S chưa ở BCNF thì chọn PTH $X \rightarrow A$ thỏa trong S, trong đó X không chứa khóa của S và $A \notin X$. //PTH vi phạm định nghĩa BCNF.

Phân rã các LĐQH- thành BCNF (tt)

- Thay thế S trong ρ bởi S1 và S2 với
$$S1 = XA, S2 = S \setminus A.$$
- Theo định lý Rissanen, phép tách S thành S1 và S2 là phép tách không mất thông tin đối với tập PTH trên S vì
$$S1 \cap S2 = X, X \rightarrow S1 \setminus S2 = A \text{ (PTH } X \rightarrow A)$$
- Quá trình tiếp tục cho đến khi tất cả các lược đồ đều ở BCNF

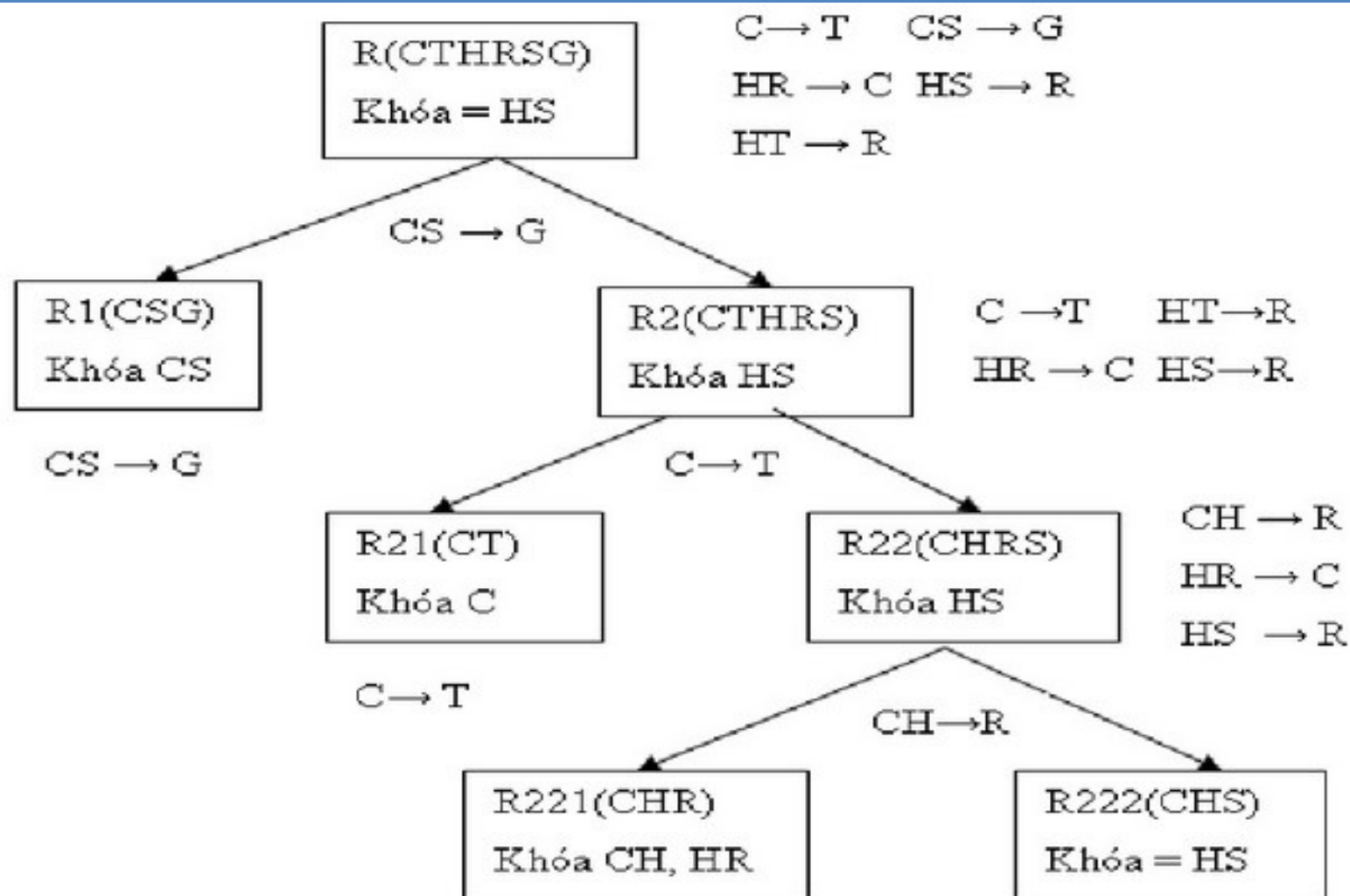
Ví dụ tách LĐQH thành BCNF

- **Ví dụ:** Cho lược đồ R(CTHRSG)
- Trong đó C: course ;T: teacher; H:hour; R: room; S:student; G:group).
- Ta có tập pth F:
- $C \rightarrow T$: Mỗi course có một thầy (teacher) duy nhất.
- $HR \rightarrow C$: Tại một thời điểm (hour) ở tại phòng học (room) chỉ có một course duy nhất.
- $HT \rightarrow R$: Tại một thời điểm (hour) và một teacher chỉ ở một room duy nhất
- $CS \rightarrow G$: Một sinh viên học một course thì chỉ ở một lớp duy nhất.
- $HS \rightarrow R$: Một sinh viên, ở một thời điểm nhất định chỉ ở trong một phòng duy nhất.
- **Khóa của R là HS.**
- Tách lược đồ R thành các lược đồ con ở dạng BCNF.

Nhận xét

- ❖ Lược đồ này có duy nhất một khóa là SH.
- ❖ Lược đồ này chưa ở dạng BCNF.
- ❖ Ta thấy lược đồ $\alpha=(U,F)$ có PTH $CS \rightarrow G$ vi phạm điều kiện BCNF nên ta tách lược đồ thành các lược đồ $U_1=CGS$, $U_2=CTHRS$
- ❖ Ta thấy lược đồ $\alpha_2=(U_2, F_2)$ có PTH $C \rightarrow T$ vi phạm BCNF nên ta tách thành các lược đồ $U_3=CT$, $U_4=CHRS$
- ❖ Ta thấy lược đồ $\alpha_4=(U_4, F_4)$ có PTH $CH \rightarrow R$ vi phạm BCNF nên ta tách thành các lược đồ $U_5=CHR$, $U_6=CHS$

Ví dụ tách LĐQH thành 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 $S1 = (XY, F1)$
- $S2 = ((S - Y) \cup X, F2)$ với $F1, F2$ là tất cả các FD của F'

End

Return Decomposition

Ví dụ

- ❖ 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

- ❖ 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 $F \rightarrow AH \in R2$ và $A \rightarrow DE \in R1$
 - $BH \rightarrow E$ có thể suy diễn được dựa vào $(BH)^+$ từ $R1, R2$
- ➔ Phân rã $R1, R2$ bảo toàn phụ thuộc hàm

Ví dụ

❖ 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

Tổng kết về chuẩn hóa CSDL

- ❖ **Mục tiêu của các dạng chuẩn trên quan hệ là hạn chế những dư thừa trong dữ liệu lưu trữ.**
- ❖ **Những dư thừa dữ liệu là do phát sinh từ mối liên quan giữa các mục dữ liệu thể hiện qua các PTH → dẫn đến dị thường khi thêm, xóa, cập nhật dữ liệu.**
- ❖ **Trong 2NF và 3NF vẫn còn dư thừa dữ liệu**
 - 2NF: loại được sự phụ thuộc riêng phần vào khóa đối với thuộc tính không nguyên tố, vẫn còn tồn tại phụ thuộc bắc cầu vào khóa.
 - 3NF: Loại được sự phụ thuộc bắc cầu vào khóa đối với thuộc tính không nguyên tố.

Tổng kết về chuẩn hóa CSDL (tt)

- ❖ **BCNF:** Loại bỏ mọi dư thừa dữ liệu mà PTH có thể tạo ra.
- ❖ Để đạt dạng chuẩn cao, ít dư thừa dữ liệu một LĐQH thường được phân rã thành nhiều LĐQH con.
- ❖ Một LĐQH luôn có thể được phân rã thành các LĐQH con thỏa 3NF vừa bảo toàn nội dung vừa bảo toàn PTH.
- ❖ Một LĐQH luôn có thể được phân rã thành các LĐQH con đạt BCNF bảo toàn nội dung nhưng có thể không bảo toàn được tập PTH.

Bài tập

❖ 1. Cho biết dạng chuẩn của các LĐQH sau:

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

❖ 2. Cho LĐQH $Q(CDEGHK)$ 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
- d) Phân rã Q thành một LĐQH đạt BCNF hoặc 3NF, chỉ ra tập PTH và khóa cho mỗi lược đồ con.

❖ Bài tập:

3. Cho $Q(ABCD)$ và $F = \{ A \rightarrow C; D \rightarrow C; BD \rightarrow A \}$

■ Xét phân rã $Q1(AB); Q2(ACD); Q3(BCD)$

a) Tìm các khóa của Q

b) Tìm các phụ thuộc hàm được chiếu trên từng quan hệ, suy ra các khóa của các quan hệ con.

c) Lược đồ này ở dạng chuẩn mấy?

d) CM lược đồ phân rã trên không bảo toàn thông tin.

4. Cho một quan hệ phổ quát liên quan đến ứng dụng quản lý bán hàng của một công ty: $Q(\text{MAKH}, \text{TENKH}, \text{MAHD}, \text{NGAYHD}, \text{MAHG}, \text{TENHG}, \text{SOLG}, \text{DG})$

- Ý nghĩa: MAKH: Mã khách hàng; TEHKH: Tên khách hàng; MAHD: Mã số hóa đơn, NGÀYHD: Ngày lập hóa đơn; MAHG: Mã số hàng hóa, TENHG: Tên hàng hóa; SOLG : Số lượng hàng của một mặt hàng trên 1 dòng chi tiết hóa đơn.; DG: Đơn giá bán của một mặt hàng.
- Với tập phụ thuộc hàm : $F_Q = \{ \text{MAKH} \rightarrow \text{TENKH}; \text{MAHG} \rightarrow \text{TENHG}; \text{TENHG} \rightarrow \text{MAHG}; \text{MAHD} \rightarrow \text{MAKH}, \text{NGAYHD}; \text{MAHD}, \text{MAHG} \rightarrow \text{SOLG}, \text{DG} \}$

❖ Yêu Cầu:

- Hãy đánh giá dạng chuẩn của quan hệ Q.
- Xác định một lược đồ CSDL khác đạt dạng chuẩn cao nhất.

5. Xét LĐQH $Q(\text{MsKH}, \text{TP}, \text{CTyVC}, \text{MsHH}, \text{SL})$

- MsKH: Mã số Khách hàng. TP: Thành phố của nhà cung cấp. CtyVC: công ty vận chuyển hàng. MsHH: mã hàng hóa. SL: số lượng.
- $F = \{ f1: \text{MsKH} \rightarrow \text{TP}, \text{CTyVC}; f2: \text{MsKH}, \text{MsHH} \rightarrow \text{SL}; f3: \text{TP} \rightarrow \text{CTyVC} \}$
 - Xác định khóa của quan hệ
 - Xác định dạng chuẩn của quan hệ
 - Có thể tách quan hệ trên thành các quan hệ nào để lđcsdl đạt dạng chuẩn cao nhất.

6. Cho $Q(\text{Chuyên viên}(\text{CV}), \text{Văn Phòng}(\text{VP}), \text{Cổ Đông}(\text{CĐ}), \text{Số Lượng Cổ Phần}(\text{SL}), \text{Cổ Phần}(\text{CP}), \text{Lãi}(\text{CP}(\text{L}))$).

- $F_Q = \{f_1: \text{CP} \rightarrow \text{L}; f_2: \text{CĐ} \rightarrow \text{CV}; f_3: \text{CĐ}, \text{CP} \rightarrow \text{SL}; f_4: \text{CV} \rightarrow \text{VP}\}$
- Xác định khóa của quan hệ.
- Xác định phân rã đạt dạng chuẩn BCK, bảo toàn thông tin và PTH.
- Xét phân rã: $C1 = \{ \langle Q1(\text{CĐ}, \text{CP}, \text{SL}, \text{L}), F1=\{f_1; f_3\} \rangle,$
 – $\langle Q2(\text{CĐ}, \text{CV}, \text{VP}), F2=\{f_2; f_4\} \rangle$
 – Chỉ ra những trùng lặp thông tin của phân rã trên
- Chỉ ra những bất tiện trong quá trình khai thác CSDL trên.
 - Xét phân rã: $C2 = \{ \langle Q1(\text{CP}, \text{CĐ}, \text{VP}), F1= \emptyset \rangle, \langle Q2(\text{CP}, \text{CĐ}, \text{SL}), F2= \{f_3\} \rangle,$
 – $\langle Q3(\text{CĐ}, \text{CV}), F3=\{f_2\} \rangle, \langle Q4(\text{CP}, \text{L}), F4= \{f_1\} \rangle$
 - C2 có bảo toàn thông tin và bảo toàn phụ thuộc hàm không ?

Bài Tập:

7. Cho LĐCSDL có các phụ thuộc hàm $F = \{ f1: A \rightarrow BC; f2: C \rightarrow B \}$ và 2 quan hệ sau: $Q1(A \ B \ C)$, $Q2(B \ C)$.

- a) Xác định tập phụ thuộc hàm trên từng quan hệ
- b) Xác định dạng chuẩn cao nhất của LĐCSDL.

8. Cho $Q(ABCD)$ có $F = \{ A \rightarrow B; B \rightarrow C; A \rightarrow D; D \rightarrow C \}$

Gọi $C = \{ Q_1(AB); Q_2(AC); Q_3(BD) \}$

a) C có bảo toàn thông tin hay không?

b) C có bảo toàn phụ thuộc hàm hay không?

9. Gọi $F = (AB \rightarrow C; A \rightarrow D; BD \rightarrow C)$

a) Hãy đưa ra một phân rã của $Q(ABCD)$ đạt DC3 và bảo toàn phụ thuộc

c) Trình bày những PTH chiếu trên các quan hệ con của phân rã

d) Kết quả của câu (a) có bảo toàn thông tin hay không? Nếu không, có thể sửa lại như thế nào để phân rã bảo toàn thông tin và vẫn bảo toàn PTH.

10. Cho $Q(SDIBQO)$ với $F_Q = \{ S \rightarrow D; I \rightarrow B; IS \rightarrow Q; B \rightarrow 0 \}$
Hãy tìm phân rã đạt DC3, bảo toàn PTH, bảo toàn thông tin

11. CHUYỀN_MÔN(MÃGV, MÔN) đạt DC1 không? Nếu không hãy nêu cách khắc phục.

MAGV	MÔN
GV01	PASC, CTDL
GV02	CSDL, PT

Bài tập

12. ĐẶT_HÀNG (SốĐH, MãHH, NgàyĐH, MãKH, SL)

$$F = \{ \text{SốĐH} \rightarrow \text{NgàyĐH}, \text{MãKH}; \\ \text{SốĐH}, \text{MãHH} \rightarrow \text{SL} \}$$

Đạt dạng chuẩn 2 không? Nếu không hãy giải thích lý do và nêu cách khắc phục?

13. GIẢNG_DẠY (MãLớp, MãSốGV, TênGV, Địa chỉ)

$$F = \{ \text{MãLớp} \rightarrow \text{MãSốGV}; \\ \text{MãSốGV} \rightarrow \text{TênGV}, \text{Địa chỉ} \}$$

Đạt dạng chuẩn 3 không? Nếu không hãy giải thích lý do và nêu cách khắc phục?



Bài tập

14. Cho $R(CIDBKHLMG)$

$F = \{C \rightarrow IDBK, D \rightarrow B, K \rightarrow H\}$

Tìm 1 phân rã bảo đảm tính bảo toàn thông tin và đạt dạng chuẩn BCNF.

15. Bài tập Sách Phương pháp giải bài tập cơ sở dữ liệu quan hệ, Nguyễn Đức Thuần, Trương Ngọc Châu.

