

Chương 7

Phụ thuộc hàm và Chuẩn hóa cơ sở dữ liệu

Nội dung trình bày

- Nguyên tắc thiết kế các lược đồ quan hệ.
 - Phụ thuộc hàm.
 - Các dạng chuẩn.
 - Một số thuật toán chuẩn hóa.
-

Phân rã lược đồ quan hệ

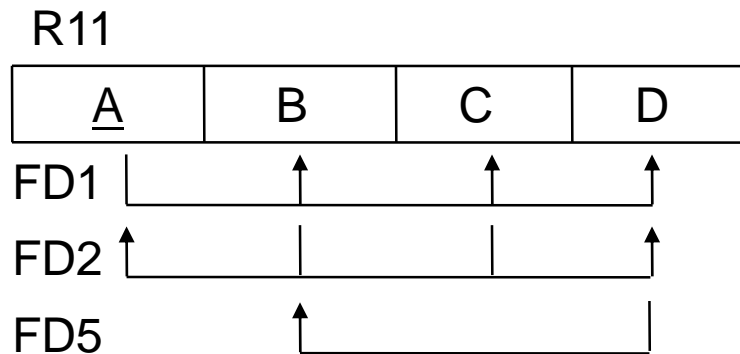
- Lược đồ quan hệ chung $R(A_1, \dots, A_n)$
 - Tập hợp tất cả các thuộc tính của các thực thể.
 - Xác định tập PTH F trên R .
 - Phân rã
 - Sử dụng các thuật toán chuẩn hóa để tách R thành tập các lược đồ $D = \{R_1, \dots, R_m\}$.
 - Yêu cầu
 - Bảo toàn thuộc tính.
 - Các lược đồ R_i phải ở dạng chuẩn 3 hoặc Boyce-Codd.
-

Phân rã bảo toàn PTH (1)

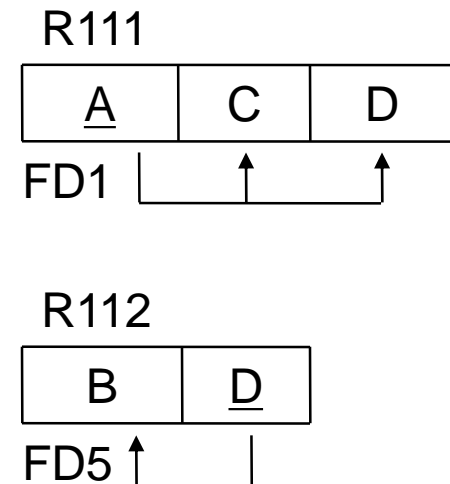
■ Tính chất bảo toàn PTH

- Xét lược đồ R và tập PTH F . Giả sử R được phân rã thành $D = \{R_1, \dots, R_m\}$.
 - Đặt $\pi_{R_i}(F) = \{X \rightarrow Y \in F^+ : X \cup Y \subset R_i\}$.
 - D được gọi là phân rã bảo toàn phụ thuộc hàm đối với F nếu $(\pi_{R_1}(F) \cup \dots \cup \pi_{R_m}(F))^+ = F^+$.

■ Ví dụ



----->



Phân rã bảo toàn PTH (2)

R11	<u>A</u>	B	C	D
	1	α	β	2
	2	β	γ	3
	3	α	δ	2

R111	<u>A</u>	C	D
	1	β	2
	2	γ	3
	3	δ	2
	4	β	4

R112	<u>D</u>	B
	2	α
	3	β
	4	α

<u>A</u>	B	C	D
1	α	β	2
...
4	α	β	4

Thêm bộ (4, β , 4) vào R111
và (4, α) vào R112
thì trạng thái csdl sẽ không
thỏa PTH FD2

Phân rã bảo toàn PTH (3)

▪ Định lý 7.1

- Tồn tại một phân rã bảo toàn PTH $D = \{R_1, \dots, R_m\}$ của lược đồ R đối với tập PTH F sao cho các R_i ở dạng chuẩn 3.

▪ Thuật toán 7.4

- Nhập: $R(U)$, $U = \{A_1, \dots, A_n\}$ và tập PTH F .
 - Xuất: $D = \{R_1, \dots, R_m\}$, R_i ở dạng chuẩn 3.
 - *B1*:
 - Tìm phủ tối thiểu G của F .
 - *B2*:
 - Với mỗi $X \rightarrow A_i \in G$, xây dựng lược đồ $R_i(U_i)$, $U_i = X \cup \{A_i\}$. Khóa chính của R_i là X .
-

Phân rã bảo toàn PTH (4)

- *B3*:
 - Giả sử xong B2 ta có các lược đồ R_1, \dots, R_m . Nếu $U_1 \cup \dots \cup U_m \neq U$ thì xây dựng thêm lược đồ $R_{m+1}(U_{m+1})$, $U_{m+1} = U - (U_1 \cup \dots \cup U_m)$. Khóa chính của R_{m+1} là U_{m+1} .
- *B4*:
 - Xuất các lược đồ R_i .

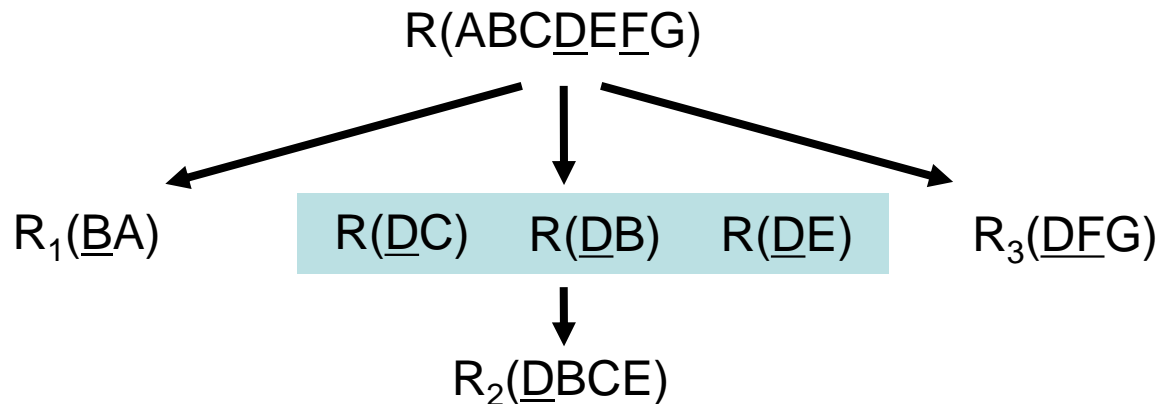
Ví dụ phân rã bảo toàn PTH (1)

- Cho
 - $R(ABCDEFGG)$
 - $F = \{B \rightarrow A, D \rightarrow C, D \rightarrow EB, DF \rightarrow G\}$
- Tách về dạng chuẩn 3, bảo toàn PTH

- $B1$:

- Phủ tối thiểu $G = \{B \rightarrow A, D \rightarrow C, D \rightarrow B, D \rightarrow E, DF \rightarrow G\}$.

- $B2$:



- $B3$:

- Xuất $D = \{R_1, R_2, R_3\}$.

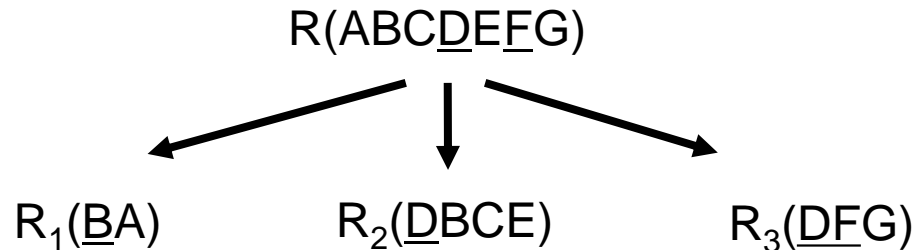
Ví dụ phân rã bảo toàn PTH (2)

- Cho
 - $R(ABCDEFGHI)$
 - $F = \{B \rightarrow A, D \rightarrow C, D \rightarrow EB, DF \rightarrow G\}$
- Tách về dạng chuẩn 3, bảo toàn PTH

- $B1$:

- Phủ tối thiểu $G = \{B \rightarrow A, D \rightarrow C, D \rightarrow B, D \rightarrow E, DF \rightarrow G\}$.

- $B2$:



- $B3$:

- Vì $U_1 \cup U_2 \cup U_3 = \{ABCDEF\}$ nên đặt $R_4(HI)$.

- $B4$:

- $D = \{R_1, R_2, R_3, R_4\}$.
-

Phân rã không mất thông tin (1)

▪ Tính chất không mất thông tin

- Xét lược đồ R và tập PTH F . Giả sử R được phân rã thành $D = \{R_1, \dots, R_m\}$.
 - D được gọi là phân rã không mất thông tin đối với F nếu với mọi trạng thái $r \in R$ thì $(\pi_{R_1}(r) * \dots * \pi_{R_m}(r)) = r$.

▪ Định lý 7.2

- Phân rã $D = \{R_1(U_1), R_2(U_2)\}$ của $R(U)$ không mất thông tin đối với tập PTH F nếu và chỉ nếu:
 - $(U_1 \cap U_2) \rightarrow (U_1 - U_2) \in F^+$, hoặc
 - $(U_1 \cap U_2) \rightarrow (U_2 - U_1) \in F^+$.

▪ Định lý 7.3

- Nếu phân rã $D = \{R_1, \dots, R_m\}$ của R không mất thông tin đối với F và phân rã $D_i = \{Q_1, \dots, Q_k\}$ của R_i không mất thông tin đối với $\pi_{R_i}(F)$ thì $D' = \{R_1, \dots, R_{i-1}, Q_1, \dots, Q_k, R_{i+1}, \dots, R_m\}$ của R cũng không mất thông tin.
-

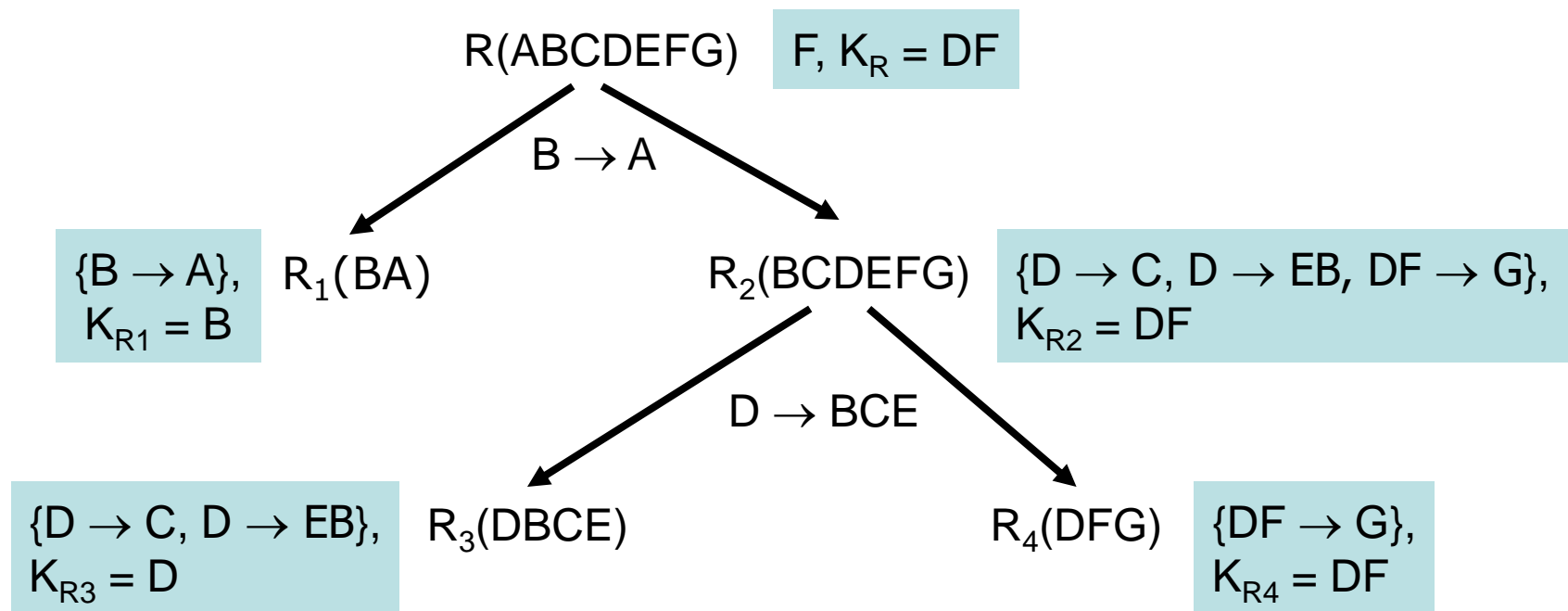
Phân rã không mất thông tin (2)

▪ Thuật toán 7.5

- Nhập: $R(U)$, $U = \{A_1, \dots, A_n\}$ và tập PTH F .
 - Xuất: $D = \{R_1, \dots, R_m\}$, R_i ở dạng chuẩn Boyce-Codd.
 - $B1$:
 - $D = \{R\}$;
 - $B2$:
 - Nếu có lược đồ $Q(U_Q) \in D$ không ở dạng chuẩn BC thì
 - + Tìm $X \rightarrow Y \in \pi_Q(F)$ làm Q vi phạm điều kiện BC.
 - + $D = (D - \{Q\}) \cup Q_1(U_{Q1}) \cup Q_2(U_{Q2})$ với $U_{Q1} = U_Q - Y$ và $U_{Q2} = X \cup Y$.
 - + Quay lại $B2$.
 - Ngược lại, chuyển sang $B3$.
 - $B3$:
 - Xuất D .
-

Ví dụ phân rã không mất thông tin (1)

- Cho:
 - $R(ABCDEFG)$
 - $F = \{B \rightarrow A, D \rightarrow C, D \rightarrow EB, DF \rightarrow G\}$
- Tách về dạng chuẩn BC, không mất thông tin.



Ví dụ phân rã không mất thông tin (2)

