PHỤ THUỘC HÀM VÀ DẠNG CHUẨN

- Phụ thuộc hàm
- Dạng chuẩn

PHỤ THUỘC HÀM

- Khái niệm phụ thuộc hàm
- Hệ tiên đề phụ thuộc hàm
- Bao đóng của tập phụ thuộc hàm
- Bao đóng của tập thuộc tính
- Phủ và phủ tối tiểu
- Khóa và siêu khóa

KHÁI NIỆM PHỤ THUỘC HÀM

- Giả sử R(U) là một lược đồ quan hệ trên tập thuộc tính U, X và Y là hai tập con (khác rỗng) bất kỳ của U, ta nói X xác định Y, hay Y phụ thuộc hàm vào X, ký hiệu $X \rightarrow Y$, khi và chỉ khi $\forall$ r trên R và $\forall\ t_1, t_2 \in r$, nếu $t_1[X] = t_2[X]$ thì $t_1[Y] = t_2[Y]$

KHÁI NIỆM PHỤ THUỘC HÀM

- Ví dụ
  - Với lược đồ R(U) thì luôn có $U \rightarrow A$, $\forall A \in U$
  - Giả sử X là một khóa của lược đồ $R(\{A_1, A_2, \ldots A_k\})$ thì $X \rightarrow Y$, $\forall\ Y \subseteq \{A_1, A_2, \ldots A_k\}$
  - Trong DUAN(MS_DA, TEN, DIACHI, MS_DV) có phụ thuộc hàm TEN → DIACHI, TEN → DIACHI, MS_DV

HỆ TIÊN ĐỀ PHỤ THUỘC HÀM

- Giả sử F là một tập phụ thuộc hàm trên lược đồ R(U), khi đó
  - Một PTH $X \rightarrow Y$ được gọi là được suy dẫn từ F nếu mọi quan hệ r trên R thỏa tất cả các PTH trong F thì r cũng thỏa $X \rightarrow Y$), ký hiệu $F \vDash X \rightarrow Y$
  - Tập PT hàm S được gọi là được suy dẫn (dẫn xuất) từ F nếu mọi phụ thuộc hàm trong S đều được dẫn xuất từ F

HỆ TIÊN ĐỀ PHỤ THUỘC HÀM

- Hệ tiên đề (luật cơ bản) Armstrong
  - Nếu $Y \subseteq X$ thì $X \rightarrow Y$ (phản xạ)
  - Nếu $Z \subseteq U$ và $X \rightarrow Y$, thì $XZ \rightarrow YZ$ (tăng trưởng)
  - Nếu $X \rightarrow Y$, và $Y \rightarrow Z$ thì $X \rightarrow Z$ (bắc cầu)

HỆ TIÊN ĐỀ PHỤ THUỘC HÀM

- Nếu $X \rightarrow Y$ thì $X \rightarrow Z, \forall Z \subseteq Y$ (luật phân rã)
- Nếu $X \rightarrow Y$, $X \rightarrow V$ thì $X \rightarrow YV$ (luật hợp)
- Nếu $X \rightarrow Y$ và $WY \rightarrow Z$ thì $WX \rightarrow Z$ (luật tựa bắc cầu)

BAO ĐÓNG CỦA TẬP PHỤ THUỘC HÀM

- Ta gọi tập F^+ là bao đóng của F nếu F^+ bao gồm F và tất cả các phụ thuộc hàm được suy dẫn từ F

BAO ĐÓNG CỦA TẬP PHỤ THUỘC HÀM

- Thuật toán tính bao đóng của F
  - Bước 1: $F^+ = F$
  - Bước 2: Áp dụng hệ luật Armstrong để tìm phụ thuộc hàm f được suy dẫn từ F^+ và bổ sung f vào F^+
  - Bước 3: Lặp lại bước 2 cho đến khi không tìm được thêm f nào khác từ F^+

BAO ĐÓNG CỦA TẬP PHỤ THUỘC HÀM

- Ví dụ
  - Cho R(A,B,C) và F={$A \rightarrow B$, $B \rightarrow C$}, tìm F^+
  - Ta có
    - ✓ $F^+ = F$
    - ✓ Ta có f: $A \rightarrow C$ là một phụ thuộc hàm được suy dẫn từ F^+ theo luật bắc cầu, $F^+ = F \cup \{f\} = \{A \rightarrow B, B \rightarrow C, A \rightarrow C\}$, không còn thêm được phụ thuộc hàm nào khác và F^+

BAO ĐÓNG CỦA TẬP THUỘC TÍNH

- Giả sử R(U) là một lược đồ quan hệ, X là một tập thuộc tính trong U và F là một tập phụ thuộc hàm trên R(U). Bao đóng của X theo F, ký hiệu là

$$X_F^+ = \cup\{Y \subseteq U \mid X \rightarrow Y \in F^+\}$$

BAO ĐÓNG CỦA TẬP THUỘC TÍNH

- Thuật toán tính bao đóng của X
  - Bước 1: Khởi tạo $X_F^+ = X$
  - Bước 2: Nếu có phụ thuộc hàm f: $Z \rightarrow V$ trong F, sao cho $Z \subseteq X_F^+$ và $V \not\subset X_F^+$ thì $X_F^+ = X_F^+ \cup V$
  - Bước 3: Lặp lại bước 2 cho đến khi không tìm thêm được phụ thuộc hàm f nào nữa

- Lưu ý: Thuật toán luôn dừng (do tập các thuộc tính hữu hạn), tại điểm dừng, tập X_F^+ chính là bao đóng của tập thuộc tính ban đầu

BAO ĐÓNG CỦA TẬP THUỘC TÍNH

- Ví dụ
  - Cho R(U) với U={A, B, C, D, E, F} với tập phụ thuộc hàm F = {AB → C, BC → AD, D → E, CG → B}. Tìm AB_F^+
  - Theo thuật toán ta có
    - ✓ $AB_F^+ = AB$
    - ✓ $AB_F^+ = AB \cup \{C\} = ABC$ (do $AB \rightarrow C$)
    - ✓ $AB_F^+ = ABC \cup AD = ABCD$ (do $BC \rightarrow AD$)
    - ✓ $AB_F^+ = ABCD \cup \{E\} = ABCDE$ (do $D \rightarrow E$)

BAO ĐÓNG CỦA TẬP THUỘC TÍNH

- Hệ quả: Phụ thuộc hàm f: $X \rightarrow Y$ được suy dẫn từ F, nếu $Y \subseteq X_F^+$
- Ứng dụng
  - Kiểm tra sự tồn tại của phụ thuộc hàm $X \rightarrow Y$
    - ✓ Tính bao đóng của X
    - ✓ Nếu $Y \subseteq X^+$, thì kết luận tồn tại $X \rightarrow Y$
    - ✓ Ngược lại, kết luận không tồn tại $X \rightarrow Y$

BAO ĐÓNG CỦA TẬP THUỘC TÍNH

- Cho lược đồ R(U) với U= {A,B,C,D,E,F} và tập phụ thuộc hàm {AB → C, BC → AD, D → E, CF → B}, kiểm tra sự tồn tại của phụ thuộc hàm AB → D
  - ✓ Dễ thấy $\{A,B\}^+ = \{A,B,C,D,E\} = U$, vì $\{D\} \subseteq \{AB\}^+$, nên tồn tại AB → D
  - ✓ Vì $\{A,B,F\}^+ = \{A,B,C,D,E\}$ nên ABF là một khóa của R(U)

PHỦ VÀ PHỦ TỐI TIỂU

- Cho lược đồ R(U), hai tập phụ thuộc hàm F_1 và F_2 trên R(U) được gọi là tương đương, ký hiệu $F_1 \equiv F_2$, nếu và chỉ nếu $F_1^+ = F_2^+$
- $F_1 \equiv F_2$ khi và chỉ khi F_1 được suy diễn từ F_2 và ngược lại
- Ví dụ: $F_1 = \{A \rightarrow BC, A \rightarrow D, CD \rightarrow E\}$ và $F_2 = \{A \rightarrow BCD, CD \rightarrow E\}$ là tương đương

PHỦ VÀ PHỦ TỐI TIỂU

- Một tập phụ thuộc hàm F được gọi là phủ của tập phụ thuộc hàm E nếu mọi phụ thuộc hàm trong E đều thuộc F^+ (nghĩa là mọi phụ thuộc hàm trong E đều được suy dẫn từ F)
  - Ví dụ: F={A→B, B → C} là một phủ của E = {A→B, B→C, A→C}

PHỦ VÀ PHỦ TỐI TIỂU

- Giả sử E là một tập phụ thuộc hàm, một tập phụ thuộc hàm F được gọi là phủ tối tiểu của E nếu F là phủ của E và thỏa ba điều kiện sau
  - Mọi phụ thuộc hàm trong F chỉ có một thuộc tính ở vế phải
  - Không tồn tại phụ thuộc hàm (dư thừa) $X \rightarrow A$ thuộc F sao cho $F \equiv F - \{X \rightarrow A\}$
  - Không tồn tại phụ thuộc hàm (không đầy đủ) $X \rightarrow A$ thuộc F sao cho có một tập $Z \subset X$ mà $F \equiv (F - \{X \rightarrow A\}) \cup \{Z \rightarrow A\}$

PHỦ VÀ PHỦ TỐI TIỂU

- Ví dụ: Giả sử R(U) là một lược đồ với tập phụ thuộc hàm E={A→B, A→C, B→A, B→C, C→A, C→B, AB→C, AC→B, BC→A}.
- Thì có hai phủ tối tiểu của E
  - F_1 = {A→B, B→C, C→A}
  - F_2 = {A → B, B → A, B → C, C → B}

PHỦ VÀ PHỦ TỐI TIỂU

Thuật toán tìm phủ tối tiểu của E

1. Set F = E.

2. Replace each functional dependency $X \to \{A_1, A_2, ..., A_n\}$ in F by the n functional dependencies $X \to A_1$, $X \to A_2$, ..., $X \to A_n$

3. for each $X \to A$ in F

    for each attribute B that is an element of X

        if $\{ \{F - \{X \to A\} \} \cup \{ (X - \{B\}) \to A\} \}$ is equivalent to F

            then replace $X \to A$ with $(X - \{B\}) \to A$ in F

4. For each remaining functional dependency $X \to A$ in F

    if $\{F - \{X \to A\} \}$ is equivalent to F,

        then remove $X \to A$ from F

PHỦ VÀ PHỦ TỐI TIỂU

Ví dụ

Cho tập E = {B→A, D→A, AB→D} các phụ thuộc hàm trên R(A,B,D), tìm phủ tối tiểu của E

1. Đặt F = E

2. Tất cả các phụ thuộc hàm đã ở dạng chuẩn (vế phải chỉ một thuộc tính)

3. Kiểm tra xem AB→D có thể được thay thế bởi B→D hay A→D không

Từ B →A, ta có BB → AB hay B→AB, vì AB→D suy ra B → D. Vậy có thể thay AB→D bởi B→D nên F ≡ {B→A, D→A, B→D}

4. Trong F có B → D và D → A suy ra B → A, nên loại bỏ B → A . Vậy F ≡ {B→D, D→A} là phủ tối tiểu của E

SIÊU KHÓA

- Cho R(U) là một lược đồ quan hệ, một tập con X của U được gọi là một siêu khóa của R nếu có phụ thuộc hàm $X \rightarrow U$ (hay $\forall$ t_1, t_2 $\in$ r trên R(U), nếu $t_1[X] = t_2[X]$ thì $t_1[U] = t_2[U]$)

SIÊU KHÓA

- Ví dụ
  - Trong DUAN(<u>MS_DA</u>, TEN, DIACHI, MS_DV), thì {<u>MS_DA</u>, TEN}, MS_DA và TEN là các siêu khóa

SIÊU KHÓA

- Tìm khóa (siêu khóa) X của R(U)
  - ✓ Tính bao đóng của X
  - ✓ Nếu $X^+ = U$, thì X là siêu khóa.
  - ✓ Ngược lại, kết luận X không phải là siêu khóa
- Cho lược đồ R(U) với U= {A,B,C,D,E,F} và tập phụ thuộc hàm {AB $\rightarrow$ C, BC $\rightarrow$ AD, D $\rightarrow$ E, CF $\rightarrow$ B}, kiểm tra tập {A,B,F} có là siêu khóa hay không
  - ✓ Vì $\{A,B,F\}^+ = \{A,B,C,D,E\}$ nên ABF là một khóa của R(U)

KHÓA

- Cho R(U) là một lược đồ quan hệ, một tập con X của U được gọi là một khóa của R nếu thõa mãn hai điều kiện
  - X là một siêu khóa
  - Không tồn tại tập con thực sự Z của X sao cho Z $\rightarrow$ U

KHÓA

- Ví dụ
  - Trong DUAN(MS_DA, TEN, DIACHI, MS_DV), thì MS_DA và TEN là các khóa

KHÓA

- Cho R(U) và tập phụ thuộc hàm F, thuật toán tính các khóa của R(U) như sau
  - Bước 1: Xác định tập các tổ hợp các thuộc tính 2^U, $\mathcal{K} = \varnothing$
  - Bước 2: $\forall$ tổ hợp $K \subseteq 2^U$
    - Nếu $K \rightarrow U$ (hay $K_F^+ = U$)
    - $\mathcal{K} = \mathcal{K} \cup \{K\}$ (thêm một siêu khóa K vào $\mathcal{K}$)
  - Bước 3: $\forall K \in \mathcal{K}$
    - Nếu $\exists K' \in \mathcal{K}$ sao cho $K' \subset K$
    - $\mathcal{K} = \mathcal{K} - \{K\}$ (thì loại siêu khóa K không phải là khóa ra khỏi $\mathcal{K}$)

KHÓA

- Cho R(U) và tập phụ thuộc hàm F
  - Một thuộc tính được gọi là thuộc tính đích nếu nó chỉ xuất hiện ở vế phải của các phụ thuộc hàm trong F^+
  - Một thuộc tính được gọi là thuộc tính nguồn nếu nó chỉ xuất hiện ở vế trái của các phụ thuộc hàm trong F^+
  - Nhận xét:
    - Thuộc tính đích không xuất hiện trong bất kỳ khóa nào
    - Mọi thuộc tính nguồn phải xuất hiện trong tất cả các khóa

KHÓA

- Cho R(U) và tập phụ thuộc hàm F, thuật toán tính các khóa của R(U) cải tiến như sau
  - Bước 1: Xác định các tập thuộc tính đích D và nguồn N và các tổ hợp các thuộc tính $2^{U-D\cup N}$, $\mathcal{K} = \varnothing$
  - Bước 2: $\forall$ tổ hợp $K^* \subseteq 2^{U-D\cup N}$
    - Nếu $N\cup K^* \rightarrow U$ (hay $(N \cup K^*)_F^+ = U$)
    - $\mathcal{K} = \mathcal{K} \cup \{N\cup K^*\}$ (thêm một siêu khóa $N\cup K^*$ vào $\mathcal{K}$)
  - Bước 3: $\forall K \in \mathcal{K}$
    - Nếu $\exists K' \in \mathcal{K}$ sao cho $K' \subset K$
    - $\mathcal{K} = \mathcal{K} - \{K\}$ (thì loại siêu khóa K không phải là khóa ra khỏi $\mathcal{K}$)

KHÓA

- Ví dụ: Cho R(A,B,C,D,E,G) với tập phụ thuộc hàm F={AD→B, EG →A, BC →G}, xác định các khóa của R
- Tính toán các khóa như sau:
  - ✓ Tập D=∅ (không có thuộc tính đích)
  - ✓ Tập N={C,D,E}
  - ✓ $2^{U-D \cup N}$ = {G, A, B, GA, GB, AB, GAB}
  - ✓ Theo thuật giải (tính bao đóng của 7 tổ hợp thuộc tính) chỉ có 3 tổ hợp K_1= CDEG, K_2= CDEA, K_3= CDEB là các khóa của R

DẠNG CHUẨN

- Các vấn đề về thiết kế CSDL
- Phân rã lược đồ quan hệ
- Dạng chuẩn thứ nhất
- Dạng chuẩn thứ hai
- Dạng chuẩn thứ ba
- Dạng chuẩn Boyce Codd
- Phân rã về dạng chuẩn BCNF
- Phân rã về dạng chuẩn ba

CÁC VẤN ĐỀ VỀ THIẾT KẾ CSDL

- Dư thừa dữ liệu (dữ liệu lặp đi lặp lại nhiều lần gây lảng phí không gian lưu trữ)
- Xung đột dữ liệu (dữ liệu không nhất quán trong các không gian lưu trữ liên quan)
- Mất mát dữ liệu (dữ liệu bị mất trong quá trình xử lý)
- Một CSDL được thiết kế tốt phải hạn chế dư thừa, xung đột và mất mát dữ liệu trong quá trình xử lý dữ liệu

PHÂN RÃ LƯỢC ĐỒ QUAN HỆ

- Một phân rã của lược đồ R(U), ký hiệu $\Omega = [R_1(U_1), R_2(U_2), \ldots, R_k(U_k)]$, là một phép phân hoạch thỏa mãn hai điều kiện
  - $U_i \subseteq U$
  - $U = U_1 \cup U_2 \cup \ldots \cup U_k$
- Một phân rã $D = [R_1(U_1), R_2(U_2), \ldots, R_k(U_k)]$ của lược đồ R(U) là đảm bảo duy trì tập thuộc tính nếu $R_1 \cup R_2 \cup \ldots \cup R_k = R$

PHÂN RÃ LƯỢC ĐỒ QUAN HỆ

- Cho tập phụ thuộc hàm F trên R, phép chiếu F trên $R_i \subseteq R$, được ký hiệu $\pi_{Ri}(F)$, là tập phụ thuộc hàm $X \rightarrow Y$ trong F^+ sao cho $X \cup Y$ là tập con của R_i.
- Phép chiếu của F trên mỗi lược đồ R_i trong phân rã D của R là một tập phụ thuộc hàm trong F^+ sao cho tất các thuộc tính ở vế trái và phải của phụ thuộc hàm là ở trong R_i.
- Phân rã $D = \{R_1, R_2, ..., R_m\}$ của R được gọi là duy trì phụ thuộc hàm theo F nếu $((\pi_{R1}(F)) \cup ... \cup (\pi_{Rm}(F)))^+ = F^+$.

PHÂN RÃ LƯỢC ĐỒ QUAN HỆ

- Phân rã $D = \{R_1, R_2, ..., R_m\}$ của R theo tập phụ thuộc hàm F được gọi là không mất mát thông tin (lossless) nếu mọi quan hệ r trên R thỏa F thì $(\pi_{R1}(r)*\ldots*\pi_{Rm}(r)) = r$.

PHÂN RÃ LƯỢC ĐỒ QUAN HỆ

- Cho R(U) là một lược đồ quan hệ, F là một tập phụ thuộc hàm trên R(U) thì $\Omega = [R_1(U_1), R_2(U_2)]$ là một phân rã không mất mát thông tin đối với F khi và chỉ khi $U_1 \cap U_2 \rightarrow U_1 - U_2$ hoặc $U_1 \cap U_2 \rightarrow U_2 - U_1$

DẠNG CHUẨN THỨ NHẤT

- Một lược đồ quan hệ được gọi là ở dạng chuẩn thứ nhất (1NF) khi và chỉ khi các thuộc tính của nó chỉ nhận các giá trị nguyên tố

Quan hệ không 1NF

MS	TEN	DIADIEM
1	ABC	TP. Cần Thơ
		TP. Đà Nẵng
		TP. Hà Nội
		TP. Hồ Chí Minh
2	XYZ	TP. Đà Nẵng
		TP. Hà Nội
		TP. Hải Phòng

Quan hệ là 1NF

MS	TEN	DIADIEM
1	ABC	TP. Cần Thơ
1	ABC	TP. Đà Nẵng
1	ABC	TP. Hà Nội
1	ABC	TP. Hồ Chí Minh
2	XYZ	TP. Đà Nẵng
2	XYZ	TP. Hà Nội
2	XYZ	TP. Hải Phòng

DẠNG CHUẨN THỨ HAI

- Một lược đồ quan hệ được gọi là ở dạng chuẩn thứ hai (2NF) khi và chỉ khi nó ở dạng chuẩn thứ nhất và mọi thuộc tính không khóa phụ thuộc hàm đầy đủ vào khóa chính
- Ví dụ: Các quan hệ trên lược đồ THAMGIA_DA(MS_NV, MS_DA, SOGIO, TEN_NV, TEN_DA, DIADIEM) với các phụ thuộc hàm MS_NV, MS_DA → SOGIO, MS_NV → TEN_NV, MS_DA → TEN_DA, DIADIEM không ở dạng chuẩn thứ 2

DẠNG CHUẨN THỨ HAI

- Một lược đồ R được gọi là dạng chuẩn 2 (tổng quát) nếu mọi thuộc tính không khóa của R là không phụ thuộc bộ phận vào bất kỳ khóa nào của R

DẠNG CHUẨN THỨ BA

- Một lược đồ quan hệ được gọi là ở dạng chuẩn thứ ba (3NF) khi và chỉ khi nó ở chuẩn thứ hai và không tồn tại thuộc tính không khóa của R phụ thuộc bắc cầu vào khóa chính
- Ví dụ
  - Các qua hệ trên lược đồ R(S,A,I,P), F ={SI→P, S→A}, không ở 3NF, vì thuộc tính không khóa A phụ thuộc bắc cầu vào khóa chính SI
  - Các qua hệ trên lược đồ R(C, S, Z), F ={CS→Z, Z→C}, là ở 3NF, vì mọi thuộc tính điều là thuộc tính khóa (CS và SZ là khóa)

DẠNG CHUẨN THỨ BA

- Một lược đồ R được gọi là dạng chuẩn 3 (tổng quát) nếu mọi phụ thuộc hàm không tầm thường X→A trong R, thì X là một siêu khóa hoặc A là một thuộc tính khóa của R

DẠNG CHUẨN BOYCE CODD

- Một lược đồ quan hệ R được gọi là ở dạng chuẩn Boyce – Codd (BCNF) khi và chỉ khi nó đạt dạng chuẩn thứ ba và với mọi phụ thuộc hàm $X \rightarrow A$ không tầm thường trên R ($A \notin X$) thì X là một siêu khóa
- Ví dụ: Các quan hệ trên R(C, S, Z), F ={$CS \rightarrow Z$, $Z \rightarrow C$}, là ở 3NF, nhưng không ở BCNF vì trong $Z \rightarrow C$, Z không phải là siêu khóa

PHÂN RÃ VỀ DẠNG CHUẨN BA

- Input: Lược đồ R(U) với tập phụ thuộc hàm F
- Output: Tập các lược đồ R_i dạng 3NF duy trì tập phụ thuộc hàm là kết quả của phép phân rã từ R
  - 1. Tìm một phủ tối tiểu G của F
  - 2. Với mỗi vế trái các phụ thuộc hàm $X \rightarrow A_1$, $X \rightarrow A_2$, ..., $X \rightarrow A_k$ trong G tạo một lược đồ $R_i(X, A_1, A_2, \ldots, A_k\}$
  - 3. Đặt mọi thuộc tính còn lại (chưa xuất hiện trong bất kỳ lược đồ nào) vào trong một lược đồ duy nhất (để đảm bảo duy trì các thuộc tính)

PHÂN RÃ VỀ DẠNG CHUẨN BA

- Ví dụ
  - R(A,B,C,D) với F ={B→C, B→D} được phân rã thành dạng chuẩn ba là R_1(B,C,D) với khóa là B

PHÂN RÃ VỀ DẠNG CHUẨN BA

Ví dụ

- Cho lược đồ U(Emp_ssn, Pno, Esal, Ephone, Dno, Pname, Plocation), trong đó Emp_ssn, Esal, Ephone là mã số, lương và số điện thoại nhân viên, Pno, Pname, Plocation là mã số, tên và vị trí dự án. Dno mã số đơn vị.
- Các phụ thuộc hàm trên U là: f_1: Emp_ssn → {Esal, Ephone, Dno}, f_2: Pno → { Pname, Plocation}, f_3: Emp_ssn, Pno → {Esal, Ephone, Dno, Pname, Plocation}

PHÂN RÃ VỀ DẠNG CHUẨN BA

Bước 1

- Từ f_3 dễ thấy {Emp_ssn, Pno} là một khóa của U, nên F= {Emp_ssn → Esal, Ephone, Dno; Pno→Pname, Plocation; Emp_ssn, Pno→Esal, Ephone, Dno, Pname, Plocation} là tập phụ thuộc hàm trên.
- Áp dụng thuật toán tìm phủ tối tiểu ta có Pno là dư thừa trong Emp_ssn, Pno → Esal, Ephone, Dno và Emp_ssn là dư thừa trong Emp_ssn, Pno→Pname, Plocation. Vì vậy phủ tối tiểu là G= {Emp_ssn → Esal, Ephone, Dno; Pno → Pname, Plocation} (chỉ chứa f_1 và f_2, f_3 dư thừa)

PHÂN RÃ VỀ DẠNG CHUẨN BA

Bước 2

- Dạng chuẩn 3: R_1(Emp_ssn, Esal, Ephone, Dno), R_2 (Pno, Pname, Plocation), các thuộc tính trong U đã ở trong R_1 và R_2
- Lưu ý dạng chuẩn 3 duy trì tập phụ thuộc hàm trên phủ tối tiểu, không duy trì tập phụ thuộc hàm gốc F, khóa {Emp_ssn, Pno} bị mất trong tập các lược đồ dạng chuẩn 3 dẫn đến mất thông tin

PHÂN RÃ VỀ DẠNG CHUẨN BA

- Có thể dư thừa dữ liệu
- Không tổn thất thông tin
- Bảo toàn phụ thuộc hàm

PHÂN RÃ VỀ DẠNG CHUẨN BA

- Input: Lược đồ R(U) với tập phụ thuộc hàm F
- Output: Tập các lược đồ R_i dạng 3NF là kết quả của phép phân rã từ R
  - Bược 1: Tìm một phủ tối thiểu G của F
  - 2. Với mỗi vế trái các phụ thuộc hàm $X \rightarrow A_1$, $X \rightarrow A_2$, ..., $X \rightarrow A_k$ trong G tạo một lược đồ $R_i(X, A_1, A_2, \ldots, A_k\}$
  - 3. Nếu không có lược đồ nào chứa một khóa của R thì tạo thêm một lược đồ chứa tất cả các thuộc tính của khóa
  - Bước 4: Loại bỏ các lược đồ dư thừa trong tập các lược đồ kết quả (một lược đồ Q được gọi là dư thừa nếu Q là kết quả được chiếu của một lược đồ S khác)

PHÂN RÃ VỀ DẠNG CHUẨN BA

Ví dụ

- Áp dụng thuật toán cho ví dụ trước ta có dạng chuẩn 3 duy trì phụ thuộc hàm, không mất thông tin: R_1(Emp_ssn , Esal, Ephone, Dno), R_2 (Pno, Pname, Plocation), R_3(Emp_ssn, Pno)
- Lưu ý: R_3(Emp_ssn, Pno) là lược đồ bổ sung trong bước 3 của thuật toán. Không có lược đồ dư thừa trong các lược đồ kết quả.

PHÂN RÃ VỀ DẠNG CHUẨN BCNF

- Hạn chế dư thừa dữ liệu
- Không tổn thất thông tin
- Không bảo toàn phụ thuộc hàm (vì mọi PTH phải chứa khóa)

PHÂN RÃ VỀ DẠNG CHUẨN BCNF

- Input: Lược đồ R(U) với tập phụ thuộc hàm F
- Output: Tập các lược đồ $R_i(U_i)$ là kết quả của phép phân rã R(U) thành dạng BCNF không mất mát thông tin
  - Bước 1: Kiểm tra R(U) có vi phạm BCNF hay không, nếu không trả về kết quả là R(U)
  - Bước 2: Nếu R(U) vi phạm, chọn một phụ thuộc hàm $X \rightarrow A$ mà X không là siêu khóa, phân rã R(U) thành hai lược đồ $R_1(XA)$ và $R_2(U\text{-}A)$ (lưu ý $R_1 \cap R_2 = X$, $X \rightarrow R_1 - R_2 = A$, trong đó $U_1 = XA$ và $U_2 = U - A$)
  - Bước 3: Tính tập phụ thuộc hàm F_1^+, F_2^+ trên R_1, R_2
  - Bước 4: Lặp lại bước 2 và 3 cho R_1 và R_2 cho đến khi mọi lược đồ đều ở dạng chuẩn BCNF

PHÂN RÃ VỀ DẠNG CHUẨN BCNF

- Ví dụ: Với R(A,B,C,D), F ={A→B, B→C, CD→A}
  - Khóa là CD nên R không là BCNF
  - Từ A→B phân rã R thành 2 lược đồ dạng BCNF là R_1(A,B) với F_1^+ ={A→B} và R_2(A, C, D) với F_2^+ ={CD→A}