



PHỤ THUỘC HÀM VÀ CHUẨN HOÁ QUAN HỆ

NỘI DUNG

- Phụ thuộc hàm (functional dependency - FD)
- Các dạng chuẩn
- Chuẩn hoá quan hệ

PHỤ THUỘC HÀM

(Functional Dependency – FD)

- Tổng quan về phụ thuộc hàm
- Hệ luật dẫn Armstrong
- Bao đóng của tập phụ thuộc hàm
- Bao đóng của tập thuộc tính
- Khóa của lược đồ quan hệ
- Tập phụ thuộc hàm tương đương
- Tập phụ thuộc hàm tối thiểu

TỔNG QUAN

- PHỤ THUỘC HÀM (Functional Dependency)
 - Một công cụ biểu diễn ràng buộc toàn vẹn
 - Ứng dụng làm nguyên tắc để tìm khóa, chuẩn hóa cơ sở dữ liệu,...

TỔNG QUAN

► Định nghĩa

- Cho $Q(A_1, \dots, A_n)$ là một sơ đồ quan hệ với tập thuộc tính $U = \{A_1, \dots, A_n\}$. X và Y là tập con của U . Ta nói $X \rightarrow Y$ (đọc là: X xác định hàm Y) nếu với mỗi cặp bộ u, v của Q sao cho:

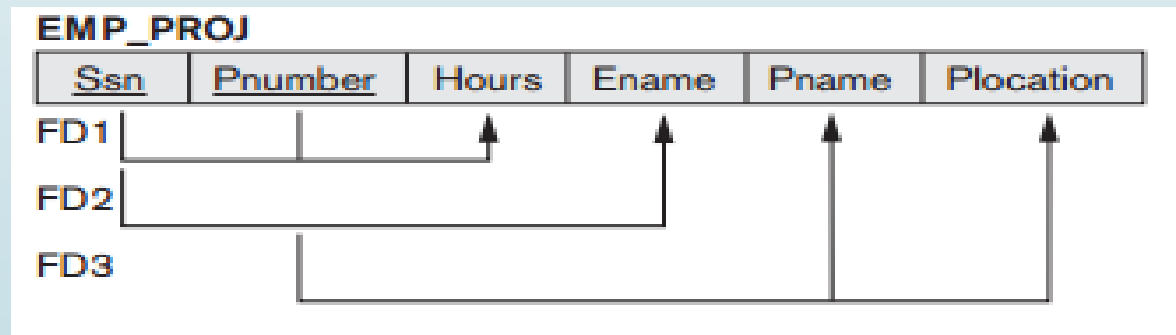
- $u[X] = v[X] \Rightarrow u[Y] = v[Y]$

► Ta cũng có thể đọc :

- Y phụ thuộc hàm vào X trong Q
 - Q thỏa phụ thuộc hàm $X \rightarrow Y$
 - $X \rightarrow Y$ là phụ thuộc hàm trong Q

TỔNG QUAN

- Người ta thường dùng F để chỉ tập các phụ thuộc hàm định nghĩa trên Q . Vì Q hữu hạn nên F cũng hữu hạn, ta có thể đánh số các phụ thuộc hàm của F là f_1, f_2, \dots, f_m .
- Phụ thuộc hàm hiển nhiên: $X \rightarrow X$
 - Quy ước: chỉ cần mô tả các phụ thuộc hàm không hiển nhiên trong tập F , các phụ thuộc hàm hiển nhiên được ngầm hiểu là đã có trong F
- Lược đồ phụ thuộc hàm (A Dependency Diagram)



TỔNG QUAN

► Ví dụ: Cho quan hệ Q (ABCDE) như sau:

A	B	C	D	E
a1	b1	c1	d1	e1
a1	b2	c2	d2	e1
a2	b1	c3	d3	e1
a2	b1	c4	d3	e1
a3	b2	c3	d1	e1

► Phụ thuộc hàm nào sau đây thỏa Q: $A \rightarrow D$, $AB \rightarrow D$, $E \rightarrow A$, $A \rightarrow E$

PHỤ THUỘC HÀM

- Tổng quan về phụ thuộc hàm
- **Hệ luật dẫn Armstrong**
- Bao đóng của tập phụ thuộc hàm
- Bao đóng của tập thuộc tính
- Khóa của lược đồ quan hệ
- Tập phụ thuộc hàm tương đương
- Tập phụ thuộc hàm tối thiểu

HỆ LUẬT DẪN ARMSTRONG

- Để có thể xác định được các phụ thuộc hàm khác từ tập phụ thuộc hàm đã có, ta dùng hệ tiên đề Armstrong (1974)
- Cho $Q(A_1, \dots, A_n)$ là một sơ đồ quan hệ với tập thuộc tính $Q^+ = \{A_1, \dots, A_n\}$. Với $X, Y, Z, W \subseteq Q^+$

1. Luật phản xạ (reflexivity): $X \supseteq Y \Rightarrow X \rightarrow Y$

Quy tắc này đưa ra những phụ thuộc hàm hiển nhiên, đó là những phụ thuộc hàm mà **vế trái bao hàm cả vế phải**. Những phụ thuộc hàm hiển nhiên đều đúng trong mọi quan hệ.

2. Luật tăng trưởng (augmentation): $X \rightarrow Y \Rightarrow XZ \rightarrow YZ$

3. Luật bắc cầu (transitivity): $X \rightarrow Y, Y \rightarrow Z \Rightarrow X \rightarrow Z$

HỆ LUẬT DẪN ARMSTRONG MỞ RỘNG

- 4. Luật hợp (union): $X \rightarrow Y, X \rightarrow Z \Rightarrow X \rightarrow YZ$
- 5. Luật bắc cầu giả (pseudotransitivity): $X \rightarrow Y, WY \rightarrow Z \Rightarrow XW \rightarrow Z$
- 6. Luật phân rã (decomposition): $X \rightarrow Y, Z \subseteq Y \Rightarrow X \rightarrow Z$

PHỤ THUỘC HÀM

- Tổng quan về phụ thuộc hàm (functional dependency)
- Hệ luật dẫn Armstrong
- **Bao đóng của tập phụ thuộc hàm**
- Bao đóng của tập thuộc tính
- Khóa của lược đồ quan hệ
- Tập phụ thuộc hàm tương đương
- Tập phụ thuộc hàm tối thiểu

BAO ĐÓNG CỦA TẬP PHỤ THUỘC HÀM F

- Định nghĩa suy diễn theo logic
- Giả sử:
 - F là tập phụ thuộc hàm trên lược đồ quan hệ $Q(U)$
 - X, Y là các tập con thuộc tính của Q ($X, Y \subseteq U$).
 - Ta nói rằng F suy diễn logic phụ thuộc hàm $X \rightarrow Y$ hay phụ thuộc hàm $X \rightarrow Y$ được suy diễn logic từ F, ký hiệu:

$$F \vdash X \rightarrow Y$$

- Nếu mọi bộ r thỏa các phụ thuộc hàm trong F cũng thỏa phụ thuộc hàm $X \rightarrow Y$.
- Nói cách khác, $X \rightarrow Y$ được suy diễn logic từ F nếu từ F áp dụng luật dẫn Armstrong ta thu được $X \rightarrow Y$.

BAO ĐÓNG CỦA TẬP PHỤ THUỘC HÀM F

- ▶ Ví dụ: Cho tập phụ thuộc hàm $F = \{A \rightarrow B, B \rightarrow CD\}$ ta chứng minh phụ thuộc hàm $AC \rightarrow CD$ được suy diễn logic từ F .
- ▶ Thật vậy, áp dụng luật tăng trưởng ta có:
 - ▶ $A \rightarrow B \Rightarrow AC \rightarrow BC$ (1)
 - ▶ $B \rightarrow CD \Rightarrow BC \rightarrow CD$ (2)
 - ▶ Áp dụng luật bắc cầu cho (1), (2):
 - ▶ $AC \rightarrow BC, BC \rightarrow CD \Rightarrow AC \rightarrow CD$

BAO ĐÓNG CỦA TẬP PHỤ THUỘC HÀM F

- Định nghĩa 1: Cho tập phụ thuộc hàm F trên tập thuộc tính U. Bao đóng (closure) của F, ký hiệu là F^+ , là tập nhỏ nhất các phụ thuộc hàm trên U thoả:

$$F^+ = \{X \rightarrow Y \mid F \vdash X \rightarrow Y\}$$

- Định nghĩa 2: Bao đóng của tập phụ thuộc hàm F là tập hợp tất cả các phụ thuộc hàm có thể suy ra từ F dựa vào các tiên đề Armstrong.

PHỤ THUỘC HÀM

- Tổng quan về phụ thuộc hàm (functional dependency)
- Hệ luật dẫn Armstrong
- Bao đóng của tập phụ thuộc hàm
- **Bao đóng của tập thuộc tính**
- Khóa của lược đồ quan hệ
- Tập phụ thuộc hàm tương đương
- Tập phụ thuộc hàm tối thiểu

BAO ĐÓNG CỦA TẬP THUỘC TÍNH X

► Định nghĩa:

- Cho tập phụ thuộc hàm F trên tập thuộc tính U , Cho $X \subseteq U$.
- Bao đóng của tập thuộc tính X (đối với F), ký hiệu X^+ , là tập sau:

$$X^+ = \{A \mid X \rightarrow A \in F^+\}$$

- Ví dụ: Cho lược đồ quan hệ Q (ABCDEFGH) và tập phụ thuộc hàm $F = \{B \rightarrow A; DA \rightarrow CE; D \rightarrow H; GH \rightarrow C; AC \rightarrow D\}$
- $B^+ = BA$ (do có phụ thuộc hàm $B \rightarrow A$)

BAO ĐÓNG CỦA TẬP THUỘC TÍNH X

Thuật toán tìm bao đóng X

Dữ Liệu Vào $Q, F, X \subseteq Q^+$

Dữ Liệu Ra X^+

Bước 1:

Đặt $X^+ = X$

Bước 2:

Temp = X^+

$\forall f: U \rightarrow V \in F$

if ($U \subseteq X^+$)

$X^+ = X^+ \cup V.$

$F = F - f;$

Bước 3:

if ($X^+ == \text{Temp}$)

“ X^+ chính là kết quả cần tìm “

Dừng thuật toán

else

trở lại Bước 2

BAO ĐÓNG CỦA TẬP THUỘC TÍNH X

- Ví dụ: Cho lược đồ quan hệ $Q(ABCDEFGH)$ và tập phụ thuộc hàm $F = \{$
- $f1: B \rightarrow A;$
 - $f2: DA \rightarrow CE;$
 - $f3: D \rightarrow H;$
 - $f4: GH \rightarrow C;$
 - $f5: AC \rightarrow D \quad \}$

Tìm bao đóng của tập $X = \{AC\}$ dựa trên F

PHỤ THUỘC HÀM

- Tổng quan về phụ thuộc hàm (functional dependency)
- Hệ luật dẫn Armstrong
- Bao đóng của tập phụ thuộc hàm
- Bao đóng của tập thuộc tính
- **Khóa của lược đồ quan hệ**
- Tập phụ thuộc hàm tương đương
- Tập phụ thuộc hàm tối thiểu

KHÓA CỦA LỰC ĐỒ QUAN HỆ

- Định nghĩa
- Thuật toán tìm một khóa của quan hệ Q
- Thuật toán tìm tất cả khóa của quan hệ Q

Định Nghĩa

- $Q(A_1, A_2, \dots, A_n)$ là lược đồ quan hệ.
- Q^+ là tập thuộc tính của Q .
- F là tập phụ thuộc hàm trên Q .
- K là tập con của Q^+ .
- Nói rằng K là một khóa của Q nếu:

1. $K^+ = Q^+$ và
2. Không tồn tại $K' \subset K$ sao cho $K'^+ = Q^+$

- Tập thuộc tính S được gọi là **siêu khóa** nếu $S \supseteq K$
- Thuộc tính A được gọi là thuộc tính khóa nếu $A \in K$ với K là khóa bất kỳ của Q . Ngược lại A được gọi là thuộc tính không khóa.

KHÓA CỦA LỰC ĐỒ QUAN HỆ

- Định nghĩa
- Thuật toán tìm một khóa của quan hệ Q
- Thuật toán tìm tất cả khóa của quan hệ Q

Thuật toán tìm một khóa của quan hệ Q

Bước 1: gán $K = Q^+$

Bước 2: A là một thuộc tính của K, đặt $K' = K - A$.

Nếu $K'^+ = Q^+$ thì gán $K = K'$ thực hiện lại bước 2

- Nếu muốn tìm các khóa khác (nếu có) của quan hệ, ta có thể thay đổi thứ tự loại bỏ các phần tử của K.
- Ví dụ : Cho $Q(A,B,C,D,E,G,H,I)$
 - $F = \{AC \rightarrow B; BI \rightarrow ACD; ABC \rightarrow D; H \rightarrow I; ACE \rightarrow BCG; CG \rightarrow AE\}$
 - Tìm K
- Lần lượt loại các thuộc tính trong K theo thứ tự sau:
 - A, B, D, E, I
 - Ta được một khóa là $\{C,G,H\}$
- Lưu ý: thuật toán này sử dụng trong trường hợp chỉ cần tìm một khóa.

KHÓA CỦA LỰC ĐỒ QUAN HỆ

- Định nghĩa
- Thuật toán tìm một khóa của quan hệ Q
- Thuật toán tìm tất cả khóa của quan hệ Q

Thuật toán tìm tất cả khóa của quan hệ Q

1. Thuật toán cơ bản

- B1: Xác định tất cả các tập con khác rỗng của Q^+ . Kết quả tìm được giả sử là các tập thuộc tính $X_1, X_2, \dots, X_{2^n-1}$
- B2: Tìm bao đóng của các X_i
- B3: Siêu khóa là các X_i có bao đóng đúng bằng Q^+ . Giả sử ta đã có các siêu khóa là $S = \{S_1, S_2, \dots, S_m\}$
- B4: Xây dựng tập chứa tất cả các khóa của Q từ tập S bằng cách xét mọi S_i, S_j con của S ($i \neq j$), nếu $S_i \subset S_j$ thì ta loại S_j ($i, j = 1..m$), kết quả còn lại của S chính là tập tất cả các khóa cần tìm

Thuật toán tìm tất cả khóa của quan hệ Q

- Ví dụ: Tìm tất cả các khóa của lược đồ quan hệ và tập phụ thuộc hàm như sau: $Q(C,S,Z)$; $F = \{f1:CS \rightarrow Z; f2:Z \rightarrow C\}$

X_i	X_i^+	Siêu khóa	Khóa
C	C		
S	S		
CS	CSZ	CS	CS
Z	ZC		
CZ	CZ		
SZ	SZC	SZ	SZ
CSZ	CSZ	CSZ	

Vậy lược đồ quan hệ Q có hai khóa là: $\{C,S\}$ và $\{S,Z\}$

Thuật toán tìm tất cả khóa của quan hệ Q

2. Thuật toán cải tiến: Ta có 1 số qui ước sau

- Tập thuộc tính nguồn (TN): chứa tất cả các thuộc tính có xuất hiện ở vế trái và không xuất hiện ở vế phải của các phụ thuộc hàm và các thuộc tính không xuất hiện ở cả vế trái lẫn vế phải của các phụ thuộc hàm.
- Tập thuộc tính đích (TD): chứa tất cả các thuộc tính có xuất hiện ở vế phải và không xuất hiện ở vế trái của các phụ thuộc hàm.
- Tập thuộc tính trung gian (TG): chứa tất cả các thuộc tính xuất hiện ở cả vế trái lẫn vế phải của các phụ thuộc hàm.

Thuật toán tìm tất cả khóa của quan hệ Q

Bước 1: Tạo tập thuộc tính nguồn TN, tập thuộc tính trung gian TG

Bước 2: If TG = \emptyset then
lược đồ quan hệ chỉ có một khóa K = TN, kết thúc.
Else qua bước 3

Bước 3: Tìm tất cả các tập con Xi của tập trung gian TG

Bước 4: Tìm các siêu khóa Si bằng cách $\forall Xi$
if $(TN \cup Xi)^+ = Q^+$ then $Si = TN \cup Xi$

Bước 5: Tìm khóa bằng cách loại bỏ các siêu khóa không tối thiểu: $\forall Si, Sj \in S$, if $Si \subset Sj$ then Loại Sj ra khỏi Tập siêu khóa S
S còn lại chính là tập khóa cần tìm.

Thuật toán tìm tất cả khóa của quan hệ Q

- Giải lại ví dụ trên: $Q(C,S,Z)$ $F = \{f1:C \rightarrow S; f2:S \rightarrow Z\}$
- Áp dụng thuật toán cải tiến ta có: $TN = \{S\}$; $TG = \{C,Z\}$
- Gọi X_i là các tập con của tập TG

X_i	$TN \cup X_i$	$(TN \cup X_i)^+$	Siêu khóa	Khóa
\emptyset	S	S		
C	CS	$CSZ = Q^+$	CS	CS
Z	ZS	$CSZ = Q^+$	ZS	ZS
CZ	CZS	$CSZ = Q^+$	CZS	

Kết quả quan hệ trên có hai khóa là : $\{S,C\}$ và $\{S,Z\}$

PHỤ THUỘC HÀM

- Tổng quan về phụ thuộc hàm (functional dependency)
- Hệ luật dẫn Armstrong
- Bao đóng của tập phụ thuộc hàm
- Bao đóng của tập thuộc tính
- Khóa của lược đồ quan hệ
- Tập phụ thuộc hàm tương đương
- Tập phụ thuộc hàm tối thiểu

Tập phụ thuộc hàm tương đương

- Hai tập phụ thuộc hàm F và G là tương đương nếu: $F^+ = G^+$
 - Ta nói F phủ G hay G phủ F .
- Thuật toán kiểm tra F và G tương đương:
 - F phủ G : $\forall X \rightarrow Y \in G$, tính X^+ từ F , sau đó kiểm tra xem $Y \in X^+$
 - G phủ F : $\forall X \rightarrow Y \in F$, tính X^+ từ G , sau đó kiểm tra xem $Y \in X^+$

PHỤ THUỘC HÀM

- Tổng quan về phụ thuộc hàm (functional dependency)
- Hệ luật dẫn Armstrong
- Bao đóng của tập phụ thuộc hàm
- Bao đóng của tập thuộc tính
- Khóa của lược đồ quan hệ
- Tập phụ thuộc hàm tương đương
- Tập phụ thuộc hàm tối thiểu

Tập phụ thuộc hàm tối thiểu

- Tập phụ thuộc hàm là tối thiểu nếu nó thoả mãn các điều kiện sau:
 1. Chỉ có một thuộc tính nằm ở phía bên tay phải của tất cả các phụ thuộc hàm trong F .
 2. Không thể bỏ đi bất kỳ một phụ thuộc hàm nào trong F mà vẫn có được một tập phụ thuộc hàm tương đương với F (tức là, không có phụ thuộc hàm dư thừa).
 3. Không thể thay thế bất kỳ phụ thuộc hàm $X \rightarrow A$ nào trong F bằng phụ thuộc hàm $Y \rightarrow A$, với $Y \subset X$ mà vẫn có được một tập phụ thuộc hàm tương đương với F (tức là, không có thuộc tính dư thừa trong phụ thuộc hàm)

Tập phụ thuộc hàm tối thiểu

► Thuật toán: Tìm tập phụ thuộc hàm tối thiểu G của F

1. Đặt $G := F$.

2. Thay thế tất cả các phụ thuộc hàm $X \rightarrow \{A_1, A_2, \dots, A_n\}$ trong G bằng n phụ thuộc hàm: $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_n$.

3. Với mỗi phụ thuộc hàm $X \rightarrow A$ trong G , với mỗi thuộc tính B trong X nếu $((G - \{X \rightarrow A\}) \cup \{(X - \{B\}) \rightarrow A\})$ là tương đương với G , thì thay thế $X \rightarrow A$ bằng $(X - \{B\}) \rightarrow A$ trong G . (Loại bỏ thuộc tính dư thừa trong phụ thuộc hàm)

4. Với mỗi phụ thuộc hàm $X \rightarrow A$ trong G , nếu $(G - \{X \rightarrow A\})$ tương đương với G , thì loại bỏ phụ thuộc hàm $X \rightarrow A$ ra khỏi G (Loại bỏ phụ thuộc hàm dư thừa)

NỘI DUNG

- Phụ thuộc hàm
- **Các dạng chuẩn**
- Chuẩn hoá quan hệ

Các dạng chuẩn

- **Dạng chuẩn 1 (First Normal Form)**
- Dạng chuẩn 2 (Second Normal Form)
- Dạng chuẩn 3 (Third Normal Form)
- Dạng chuẩn Boyce _Codd (Boyce-Codd Normal Form)

Dạng chuẩn 1 (First Normal Form – 2NF)

Một quan hệ ở dạng chuẩn 1 nếu các giá trị của tất cả thuộc tính trong quan hệ là **nguyên tử** (tức là chỉ có 1 giá trị tại một thời điểm)

DEPARTMENT			
DNAME	<u>DNUMBER</u>	DMGRSSN	DLOCATIONS
Research	5	333445555	{Bellaire, Sugarland, Houston}
Administration	4	987654321	{Stafford}
Headquarters	1	888665555	{Houston}

Vi phạm dạng chuẩn 1

Dạng chuẩn 1

DEPARTMENT			
DNAME	<u>DNUMBER</u>	DMGRSSN	<u>DLOCATION</u>
Research	5	333445555	Bellaire
Research	5	333445555	Sugarland
Research	5	333445555	Houston
Administration	4	987654321	Stafford
Headquarters	1	888665555	Houston

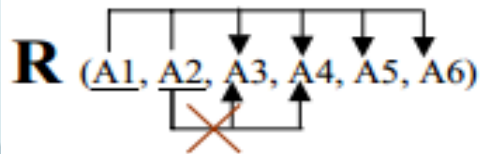
Dư thừa

Dạng chuẩn 2 (Second Normal Form - 2NF)

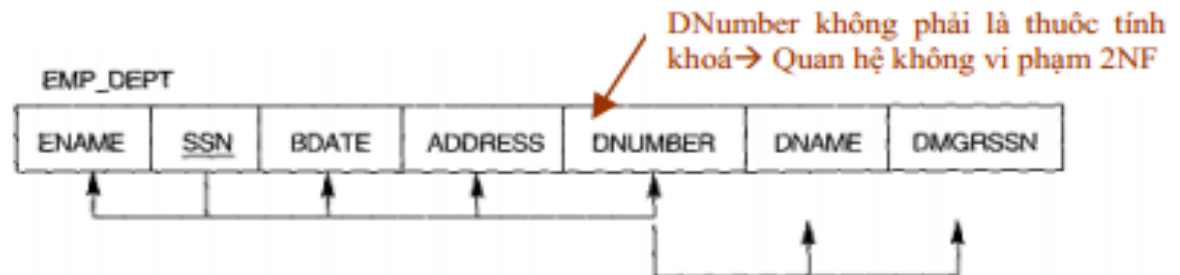
Một quan hệ ở dạng chuẩn 2 nếu:

- Quan hệ đó ở dạng chuẩn 1
- Tất cả các thuộc tính không phải là khóa **phụ thuộc đầy đủ** vào khóa.

Phụ thuộc hàm $Y \rightarrow Z$ là phụ thuộc hàm đầy đủ nếu:
 $\forall A \in Y, (Y - \{A\}) \not\rightarrow Z$



Vi phạm chuẩn 2

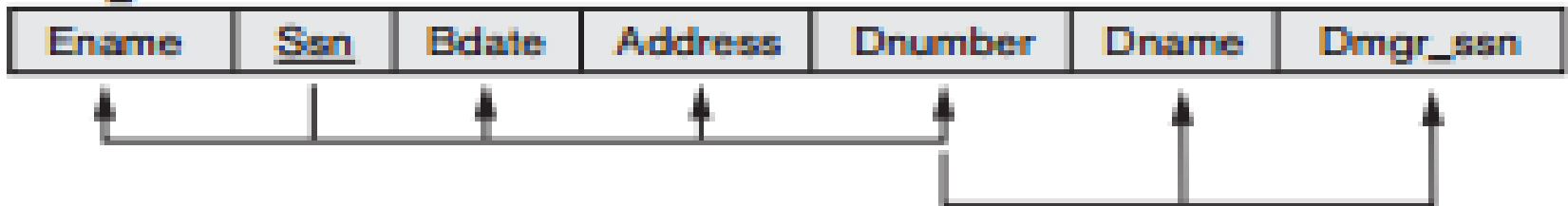


Dạng chuẩn 3 (Third Normal Form – 3NF)

Một quan hệ ở dạng chuẩn 3 nếu:

- Quan hệ ở dạng chuẩn 2
- Và **không có chứa các phụ thuộc hàm phụ thuộc bắc cầu vào khoá.**
 - Phụ thuộc hàm $Y \rightarrow Z$ là phụ thuộc hàm bắc cầu nếu tồn tại hai phụ thuộc hàm: $Y \rightarrow X$ và $X \rightarrow Z$.

EMP_DEPT



Quan hệ EMP_DEPT không phải ở dạng chuẩn 3 vì còn tồn tại phụ thuộc hàm **$Dnumber \rightarrow Dname, Dmgr_Ssn$** là phụ thuộc hàm phụ thuộc bắc cầu vào khoá.

Dạng chuẩn 3 (Third Normal Form – 3NF)

➔ Tách thành 2 quan hệ EMPLOYEE và DEPARTMENT thỏa dạng chuẩn 3

EMPLOYEE

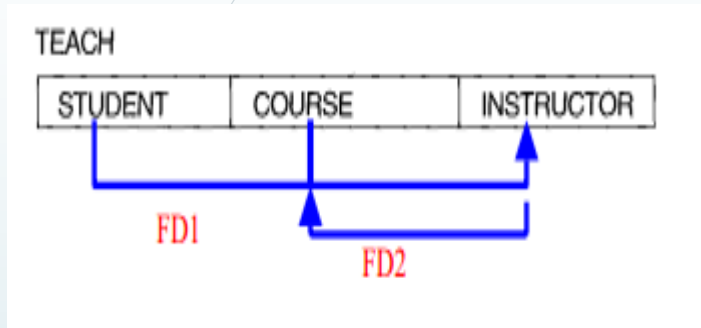
Ename	<u>Ssn</u>	Bdate	Address	Dnumber
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4
Narayan, Ramesh 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>	Dmgr_ssn
Research	5	333445555
Administration	4	987654321
Headquarters	1	888665555

Dạng chuẩn 3 (Third Normal Form – 3NF)

- Quan hệ ở dạng chuẩn 3 vẫn có khả năng dư thừa nếu có thuộc tính không khóa xác định thuộc tính khóa



TEACH

Student	Course	Instructor
Narayan	Database	Mark
Smith	Database	Navathe
Smith	Operating Systems	Ammar
Smith	Theory	Schulman
Wallace	Database	Mark
Wallace	Operating Systems	Ahamad
Wong	Database	Omiecinski
Zelaya	Database	Navathe
Narayan	Operating Systems	Ammar

Dạng chuẩn Boyce _Codd (Boyce-Codd Normal Form - BCNF)

Quan hệ R ở dạng chuẩn BCNF khi tất cả các phụ thuộc hàm $X \rightarrow A$ trong R đều phải có **X là khoá của R**.

Ví dụ:

Quan hệ TEACH được tách thành 2 quan hệ
R1(Student, Instructor) và
R2(Course, Instructor)
đạt dạng chuẩn BCNF

NỘI DUNG

- Phụ thuộc hàm
- Các dạng chuẩn
- **Chuẩn hoá quan hệ**

Chuẩn hoá quan hệ

- **Khái niệm**
- Thuật toán chuẩn hóa lược đồ quan hệ thành BCNF
- Thuật toán chuẩn hóa lược đồ quan hệ thành 3NF

Khái niệm

- Chuẩn hoá là việc phân rã một lược đồ quan hệ thành các lược đồ con ở dạng chuẩn 3 hoặc ở BCNF sao cho vẫn bảo toàn phụ thuộc và không mất thông tin
- Phép phân rã không mất thông tin:
 - Cho R là một lược đồ quan hệ, phép rã $\rho=(R_1, R_2, \dots, R_n)$ và D là tập các phụ thuộc dữ liệu. Phép phân rã ρ không mất thông tin nếu khi thực hiện phép toán kết nối tự nhiên các quan hệ thành phần R_1, R_2, \dots, R_n ta vẫn nhận được kết quả của quan hệ ban đầu

Chuẩn hoá quan hệ

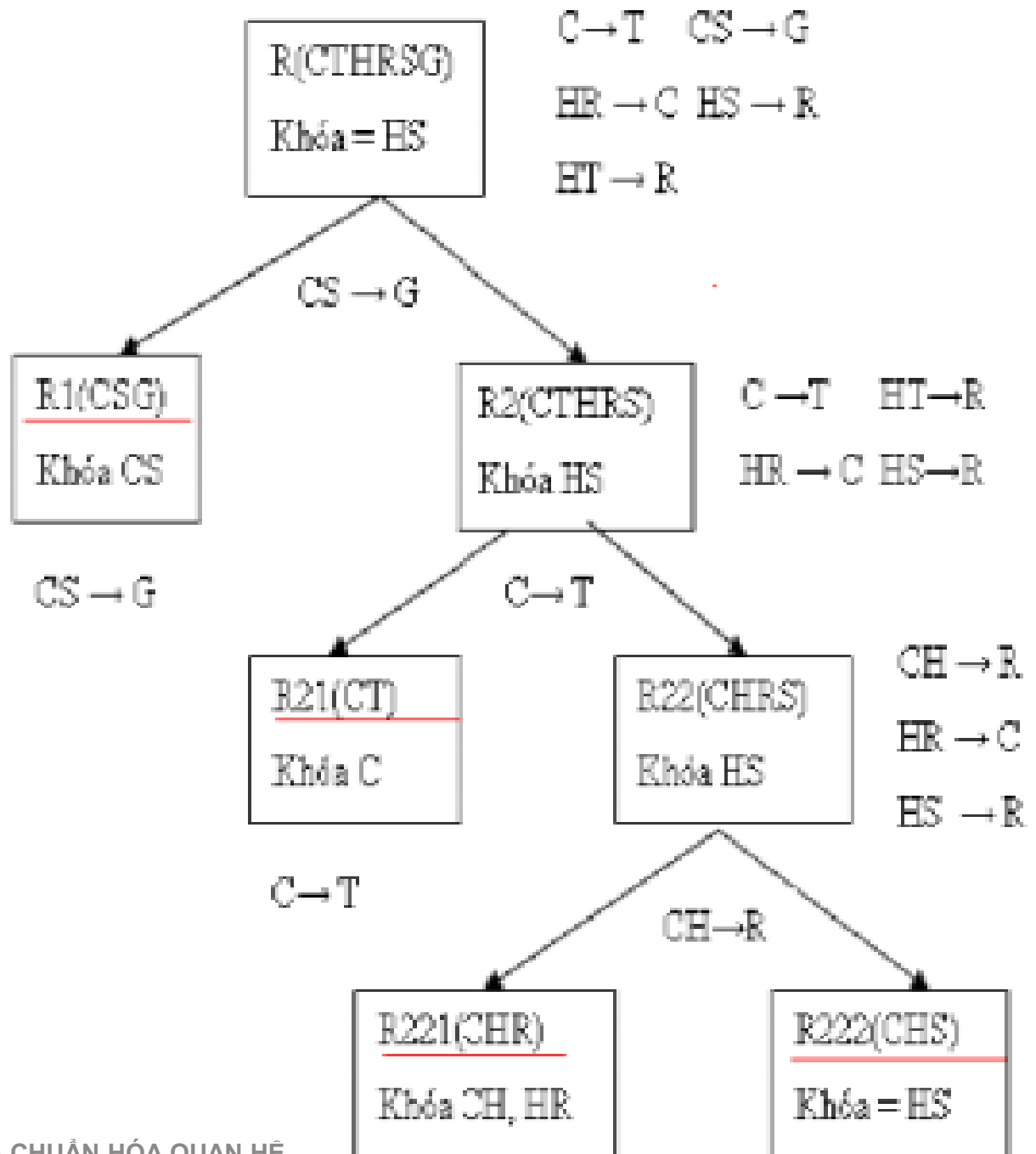
- Khái niệm
- Thuật toán chuẩn hóa lược đồ quan hệ thành BCNF
- Thuật toán chuẩn hóa lược đồ quan hệ thành 3NF

Thuật toán chuẩn hóa lược đồ quan hệ thành BCNF

1. Ban đầu phép tách p chỉ bao gồm R .
 2. Nếu S là một lược đồ thuộc p và S chưa ở dạng BCNF thì chọn phụ thuộc hàm $X \rightarrow A$ thỏa trong S , trong đó X không chứa khóa của S và $A \notin X$. {phụ thuộc hàm vi phạm định nghĩa dạng chuẩn BCNF}
 - Thay thế S trong p bởi S_1 và S_2 như sau $S_1 = XA$, $S_2 = S \setminus A$.
- Quá trình trên tiếp tục cho đến khi tất cả các lược đồ quan hệ đều ở dạng BCNF

Ví dụ: Quan hệ
 $R(CTHRSG)$
 $F = \{C \rightarrow T, HR \rightarrow C, HT \rightarrow R, CS \rightarrow G, HS \rightarrow R\}$
 Khóa của R là HS

$R = \{R1, R21, R221, R222\}$



Chuẩn hoá quan hệ

- Khái niệm
- Thuật toán chuẩn hóa lược đồ quan hệ thành BCNF
- Thuật toán chuẩn hóa lược đồ quan hệ thành 3NF

Thuật toán chuẩn hóa lược đồ quan hệ thành 3NF

- 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 p 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.
- Tại mỗi bước kiểm tra lược đồ R, nếu mỗi thuộc tính không khóa không phụ thuộc bắc cầu vào khóa chính, thì R đã ở dạng 3NF, ngược lại áp dụng bước 3 để tách tiếp.

Thuật toán chuẩn hóa lược đồ quan hệ thành 3NF

- Ví dụ: Cho lược đồ quan hệ $R(C, T, H, R, S, G)$ với tập phụ thuộc hàm tối thiểu $F: C \rightarrow T, HR \rightarrow C, HT \rightarrow R, CS \rightarrow G, HS \rightarrow R$.
 - 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
 - Bước 3:
 - $C \rightarrow T$ vi phạm 3NF (phụ thuộc bắc cầu vào khoá) \rightarrow tách R thành $R_1(C, T)$ và $R_2(C, H, R, S, G)$.
 - $CS \rightarrow G$ vi phạm 3NF (phụ thuộc bộ phận vào khoá) \rightarrow tách R_2 thành $R_{21}(C, S, G)$ và $R_{22}(C, H, R, S)$.
 - $HR \rightarrow C$ vi phạm 3NF \rightarrow tách R_{22} thành $R_{221}(H, R, C)$ và $R_{222}(H, S, R)$

$R \rightarrow R_1, R_{21}, R_{221}, R_{222}$

Tóm tắt

- Các thao tác trên dữ liệu (thêm, sửa, xóa) trên CSDL chưa chuẩn hóa có thể gây ra các bất thường trên CSDL.
- Chuẩn hóa CSDL để đảm bảo dữ liệu không dư thừa và loại bỏ các bất thường trên CSDL
- 3 bước chuẩn hóa thông thường:
 - Chuẩn hóa về dạng 1NF: Loại bỏ dữ liệu dư thừa
 - Chuẩn hóa về dạng 2NF: Loại bỏ các phụ thuộc hàm bộ phận
 - Chuẩn hóa về dạng 3NF: Loại bỏ các phụ thuộc hàm bắc cầu
 - Chuẩn hóa về dạng BCNF: Loại bỏ các phụ thuộc hàm có vế trái không là khóa