

PHÂN TÍCH THIẾT KẾ HỆ THỐNG THÔNG TIN

Information System Analysis and Design

Số tín chỉ: 4(3,1)

Số tiết: 75 tiết (45 LT- 30 TH)

Giảng viên: **TS. Đinh Thị Thu Hương.**

(Mobile: 0903087599 – e-mail: huongdtt@sgu.edu.vn)

Bộ môn: Khoa học máy tính – Khoa CNTT, SGU.

Chương 7: Lý thuyết chuẩn hóa cơ sở dữ liệu

- ☐ Dư thừa dữ liệu
- ☐ Phụ thuộc hàm.
- ☐ Bao đóng.
- ☐ Khóa của lược đồ quan hệ.
- ☐ Phủ tối thiểu.
- ☐ Dạng chuẩn.
- ☐ Chuẩn hóa cơ sở dữ liệu.
- ☐ Bài tập

Dư thừa dữ liệu - (Data redundancy)

- Mục đích của thiết kế CSDL là gom các thuộc tính thành các quan hệ sao cho giảm thiểu dư thừa dữ liệu
- Hậu quả của dư thừa dữ liệu:
 - ☐ Lãng phí không gian đĩa
 - ☐ Các bất thường khi cập nhật
- Ba loại bất thường:
 - ☐ Bất thường khi thêm vào
 - ☐ Bất thường khi xóa bỏ
 - ☐ Bất thường khi sửa đổi

Ví dụ

MaSv	HoTen	MaMH	TenMH	SoTC	Điểm
1111	Mai	CSDL	Cơ Sở Dữ Liệu	4	9
1111	Mai	KTMT	Kiến Trúc Máy Tính	4	8
5556	Long	CSDL	Cơ Sở Dữ Liệu	4	8
5556	Long	KTMT	Kiến Trúc Máy Tính	4	8
9876	Son	CSDL	Cơ Sở Dữ Liệu	4	7

Phụ thuộc hàm (Functional dependencies-FDs)

- Phụ thuộc hàm mô tả mối liên hệ giữa các thuộc tính
- Dựa vào phụ thuộc hàm để thiết kế lại CSDL, loại bỏ các dư thừa dữ liệu
- Có thể biểu diễn RBTV bằng phụ thuộc hàm.
- Ứng dụng của phụ thuộc hàm là giải quyết các bài toán về :

Tìm khóa.

Tìm phủ tối thiểu.

Chuẩn hoá cơ sở dữ liệu.

Phụ thuộc hàm (Functional dependencies-FDs)

- Cho lược đồ quan hệ $R(U)$, r là 1 quan hệ bất kỳ trên R , X và Y là 2 tập thuộc tính con.
- **Định nghĩa:** Phụ thuộc hàm (FD) $f: X \rightarrow Y$ trên lược đồ quan hệ R nếu và chỉ nếu mỗi giá trị X trong r có quan hệ chính xác với 1 giá trị Y trong r . Nghĩa là bất kể khi nào 2 bộ của r có cùng giá trị X thì cũng có cùng giá trị Y .

$$\forall t1, t2 \in r(R): t1[X] = t2[X] \Rightarrow t1[Y] = t2[Y]$$

- ∞ X là vế trái, ký hiệu $\text{left}(f)$ hay còn gọi là determinant
- ∞ Y là vế phải, ký hiệu $\text{right}(f)$ hay còn gọi là dependent

Ví dụ

A	B	C	D	E
1	2	3	4	5
1	4	3	4	5
1	2	4	4	1

Kí hiệu nào là phụ thuộc hàm

I. $AB \rightarrow C$

II. $B \rightarrow D$

III. $DE \rightarrow A$

Phụ thuộc hàm (Functional dependencies-FDs)

- Phụ thuộc hàm là 1 **đặc điểm ngữ nghĩa** của các thuộc tính, được xem là 1 **ràng buộc** giữa các thuộc tính.
- Ví dụ: Một nhân viên chỉ có 1 mức lương nhưng nhiều nhân viên có thể có cùng 1 mức lương

$\text{Emp_ID} \rightarrow \text{Salary}$

$\text{Salary} \nrightarrow \text{Emp_ID}$

- **Phụ thuộc hàm được xác định dựa vào quy tắc nghiệp vụ** được xác định trên lược đồ quan hệ

Phụ thuộc hàm (Functional dependencies-FDs)

- Từ quy tắc bảo toàn thực thể \rightarrow nếu X là 1 candidate key thì tất cả các thuộc tính Y của lược đồ R sẽ phải phụ thuộc hàm vào X
- Ví dụ: trong lược đồ PROFESSOR có ProfId là primary key nên:
ProfId \rightarrow Name, Qualification
- Có 1 số FD trong lược đồ sẽ gây ra dư thừa dữ liệu.

Phụ thuộc hàm (Functional dependencies-FDs)

Ví dụ: FD và dư thừa dữ liệu

- Xét lược đồ PERSON(SSN, Name, Address, Hobby) với quy tắc là 1 người có thể có nhiều sở thích (hobby)
 - SSN, Hobby \rightarrow SSN, Name, Address, Hobby
- Bất thường xảy ra khi một người có nhiều sở thích thay đổi địa chỉ

Phụ thuộc hàm (Functional dependencies-FDs)

- Ví dụ: Cho quan hệ **phancong** sau :

Phancong (Phicong, maybay, ngaykh, giokh)

Tùng	83	9/8	10:15a
Tùng	116	10/8	1:25p
Minh	281	8/8	5:50a
Minh	301	12/8	6:35p
Minh	83	13/8	10:15a
Nghia	83	11/8	10:15a
Nghia	116	12/8	1:25p

Phụ thuộc hàm (Functional dependencies-FDs)

- **Quan hệ Phancong** diễn tả phi công nào lái máy bay gì và máy bay khởi hành khi nào. Quan hệ trên phải tuân theo các điều kiện ràng buộc sau :

- Mỗi máy bay có một giờ khởi hành duy nhất.
- Nếu biết phi công, biết ngày giờ khởi hành thì biết được máy bay do phi công lái.
- Nếu biết máy bay, biết ngày giờ khởi hành thì biết phi công lái chuyến máy bay ấy.

PC	MB	NKH	GKH
Tùng	83	9/8	10:15a
Tùng	116	10/8	1:25p
Minh	281	8/8	5:50a
Minh	301	12/8	6:35p
Minh	83	13/8	10:15a
Nghia	83	11/8	10:15a
Nghia	116	12/8	1:25p

Phụ thuộc hàm (Functional dependencies-FDs)

■ Các ràng buộc này là các ví dụ về phụ thuộc hàm và được phát biểu lại như sau :

- MAYBAY xác định GIOKH.
- {PHICONG, NGAYKH, GIOKH} xác định MAYBAY.
- {MAYBAY, NGAYKH} xác định PHICONG

hay

- GIOKH phụ thuộc hàm vào MAYBAY.
- MAYBAY phụ thuộc hàm vào {PHICONG, NGAYKH, GIOKH} .
- PHICONG phụ thuộc hàm vào {MAYBAY, NGAYKH}.

Và được ký hiệu như sau :

- {MAYBAY} → GIOKH
- {PHICONG, NGAYKH, GIOKH} → MAYBAY
- {MAYBAY, NGAYKH} → PHICONG

PC	MB	NKH	GKH
Tùng	83	9/8	10:15a
Tùng	116	10/8	1:25p
Minh	281	8/8	5:50a
Minh	301	12/8	6:35p
Minh	83	13/8	10:15a
Nghia	83	11/8	10:15a
Nghia	116	12/8	1:25p

Phụ thuộc hàm hiển nhiên/tầm thường

Nếu $X \supseteq Y$ thì $X \rightarrow Y$.

- Với r là quan hệ bất kỳ, F là tập phụ thuộc hàm thỏa trên r , ta luôn có

$F \supseteq \{\text{các phụ thuộc hàm hiển nhiên}\}$

- Ví dụ: Name, Address \rightarrow Name

7.2 Hệ tiên đề Armstrong

❑ Phụ thuộc hàm được suy diễn logic từ F

- Phụ thuộc hàm $X \rightarrow Y$ được suy diễn logic từ F nếu một quan hệ r bất kỳ **thỏa mãn tất cả các phụ thuộc hàm của F** thì **cũng thỏa phụ thuộc hàm $X \rightarrow Y$** .
- Ký hiệu $F \models X \rightarrow Y$.

Ví dụ: Phân công (Phicong, Maybay, NgayKH, GioKH)

(1) : $\{\text{MAYBAY}\} \rightarrow \text{GIOKH}$

(2) : $\{\text{MABAY}, \text{NGAYKH}\} \rightarrow \text{PHICONG}$

\Rightarrow (3) $\{\text{MABAY}, \text{NGAYKH}\} \rightarrow \text{PHICONG}, \text{GIOKH}$

(là phụ thuộc hàm suy diễn từ (1) và (2))

7.2 Hệ tiên đề Armstrong

□ Gọi $R(\Omega)$ là lược đồ quan hệ với $\Omega = \{A_1, A_2, \dots, A_n\}$ là tập thuộc tính và X, Y, Z, W là tập con của Ω . (Kí hiệu: $XY = X \cup Y$)

1. **Luật phản xạ** (reflexive rule):

Nếu $Y \subseteq X$ thì $X \rightarrow Y$

2. **Luật tăng trưởng** (augmentation rule):

Nếu $X \rightarrow Y, Z \subseteq \Omega$ thì $XZ \rightarrow YZ$

3. **Luật bắc cầu** (Transitivity Rule)

Nếu $X \rightarrow Y$ và $Y \rightarrow Z$ thì $X \rightarrow Z$

7.2 Hệ tiên đề Armstrong

– Ba hệ quả của tiên đề Armstrong:

1. Luật hợp (Union Rule)

Nếu $X \rightarrow Y$ và $X \rightarrow Z$ thì $X \rightarrow YZ$

2. Luật bắc cầu giả (Pseudotransitivity Rule)

Nếu $X \rightarrow Y$ và $WY \rightarrow Z$ thì $XW \rightarrow Z$

3. Luật phân rã (Decomposition Rule)

Nếu $X \rightarrow Y$ và $Z \subset Y$ thì $X \rightarrow Z$

Bài tập 1:

Hãy chứng minh:

a/ Nếu $W \rightarrow Y, X \rightarrow Z$ thì $WX \rightarrow Y$

7.2 Hệ luật dẫn Armstrong

□ Bao đóng (closure) của tập phụ thuộc hàm F

Là 1 tập phụ thuộc hàm nhỏ nhất chứa F sao cho không thể áp dụng hệ tiên đề Armstrong trên tập này để tạo ra 1 phụ thuộc hàm khác không có trong tập hợp này.

- Ký hiệu: F^+
- F^+ là **tập tất cả** các phụ thuộc hàm được suy diễn logic từ F .

Nếu $F=F^+$ thì F là họ đầy đủ của các PTH

Ví dụ: Cho r_1 quan hệ trên lược đồ quan hệ $Q(A,B,C,D)$ và tập F được cho như sau:

$$F = \{A \rightarrow B; B \rightarrow C; A \rightarrow D; B \rightarrow D\}$$

khi đó $F^+ = \{A \rightarrow B; B \rightarrow C; A \rightarrow D; B \rightarrow D; A \rightarrow BD; A \rightarrow BCD; A \rightarrow C; A \rightarrow CD; A \rightarrow BC; B \rightarrow CD; \dots\}$

Rõ ràng $F \subseteq F^+$

- **Thuật toán tìm bao đóng F^+**

“Áp dụng hệ tiên đề Armstrong cho đến khi không tìm ra thêm phụ thuộc hàm mới”

❑ Các tính chất của tập F^+

- Tính phản xạ: $F \subseteq F^+$
- Tính đơn điệu: Nếu $F \subseteq G$ thì $F^+ \subseteq G^+$
- Tính lũy đẳng: $(F^+)^+ = F^+$.

7.3 Bao đóng của tập thuộc tính X (closures of attribute sets)

□ Định nghĩa

Bao đóng của tập thuộc tính X dựa trên một tập phụ thuộc hàm F (closure of X under F) là 1 tập thuộc tính Y sao cho:

- $\exists X \rightarrow Y \in F^+$
- $\forall X \rightarrow Z \in F^+: Z \subseteq Y$

Hoặc
$$X_F^+ = \{A | X \rightarrow A \in F^+\}$$

- Ví dụ: Cho quan hệ $Q(A,B,C,D,E,G)$ và
 - $F = \{A \rightarrow C; A \rightarrow EG; B \rightarrow D; G \rightarrow E\};$
 - $X = \{A,B\};$
 - $Y = \{C,G,D\}$

7.3 Bao đóng của tập thuộc tính X (closures of attribute sets) trên tập phụ thuộc hàm

❑ Thuật toán tìm bao đóng:

- Tính liên tiếp tập các tập thuộc tính X^0, X^1, X^2, \dots theo phương pháp sau:
 - *Bước 1:* $X^0 = X$
 - *Bước 2:* lần lượt xét các phụ thuộc hàm của F
 - ❖ Nếu $Y \rightarrow Z$ có $Y \subseteq X_i$ thì $X^{i+1} = X^i \cup Z$
 - ❖ Loại phụ thuộc hàm $Y \rightarrow Z$ khỏi F
 - *Bước 3:* Nếu ở bước 2 không tính được X^{i+1} thì X^i chính là bao đóng của X
 - Ngược lại lặp lại bước 2.

7.3 Bao đóng của tập thuộc tính X (closures of attribute sets) trên tập phụ thuộc hàm

Ví dụ 1: Cho lược đồ quan hệ $R(A,B,C,D,E,G,H)$ và tập phụ thuộc hàm

$F = \{B \rightarrow A; DA \rightarrow CE; D \rightarrow H; GH \rightarrow C; AC \rightarrow D\}$. Tìm bao đóng của $X = \{A, C\}$ trên F ?

7.3 Bao đóng của tập thuộc tính X (closures of attribute sets) trên tập phụ thuộc hàm

□ **Ví dụ 2:** cho lược đồ quan hệ: $Q(A,B,C,D,E,G)$

$$F = \{ \begin{array}{l} f_1: A \rightarrow C; \\ f_2: A \rightarrow EG; \\ f_3: B \rightarrow D; \\ f_4: G \rightarrow E \end{array} \}$$

Tìm bao đóng:

- X^+ với $X = \{A,B\}$;
- Y^+ với $Y = \{C,G,D\}$

Sử dụng bao đóng của tập thuộc tính

- ❑ Kiểm tra siêu khóa (Testing for superkey)
- ❑ Kiểm tra một phụ thuộc hàm $X \rightarrow Y$ có được suy dẫn từ F .
- ❑ Kiểm tra 2 tập phụ thuộc hàm tương đương $F^+ = G^+$

Sử dụng bao đóng của tập thuộc tính

❑ Kiểm tra siêu khóa (Testing for superkey)

- Để kiểm tra **X có phải là siêu khóa**: tính X^+ , nếu X^+ chứa tất cả các thuộc tính của R thì X là siêu khóa.
- **X là khóa dự tuyển** (candidate key) nếu không tập con nào trong số các tập con của nó là khóa.

❑ Kiểm tra một phụ thuộc hàm $X \rightarrow Y$ có được suy dẫn từ F .

- Bước 1: tính X^+
- Bước 2: nếu $Y \subseteq X^+$ thì khẳng định $X \rightarrow Y \in F^+$

Kiểm tra thành viên trong F^+

Ví dụ

- Cho $R = \{A, B, C, D, E, G\}$ và $F = \{AB \rightarrow C, BC \rightarrow D, D \rightarrow EG, BE \rightarrow C\}$, $AB \rightarrow EG$ có nằm trong F^+ ?

Cách 1: Theo tiên đề Armstrong

Cách 2: Theo giải thuật

Kiểm tra thành viên trong F^+

Ví dụ

- Cho $R = (A, B, C, G, H, I)$, $F = \{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\}$

Một số thành viên của F^+

Kiểm tra thành viên trong F^+

Ví dụ kiểm tra phụ thuộc hàm

- Cho $F = \{D \rightarrow B, A \rightarrow C, AD \rightarrow E, C \rightarrow B\}$. Kiểm tra F có bao hàm $A \rightarrow B$??
 - Tìm A^+_F ? $\rightarrow A^+_F = \{ACB\}$
 - Do $B \in A^+_F$ nên F bao hàm $A \rightarrow B$

Phụ thuộc hàm tương đương

(equivalences among sets of dependencies)

- Nếu F và G là 2 tập FD. F suy diễn G (F entails G) nếu F suy diễn được tất cả các FD có trong G .
- F và G là tương đương nhau nếu F suy diễn G và G suy diễn F hay $F^+ = G^+$
 - Ký hiệu $F \equiv G$.
 - Ta nói F phủ G nếu $F^+ \supseteq G^+$

Kiểm tra các tập FD tương đương

- Bước 1: tính X^+
- Bước 2: nếu $Y \subseteq X^+$ thì khẳng định $X \rightarrow Y \in F^+$

Kiểm tra các tập FD tương đương

Ví dụ

■ Hãy khảo sát 2 tập FD sau:

□ $F = \{ AC \rightarrow B, A \rightarrow C, D \rightarrow A \}$

□ $G = \{ A \rightarrow B, A \rightarrow C, D \rightarrow A, D \rightarrow B \}$

F và G có tương đương nhau không???

Kiểm tra các tập FD tương đương

Ví dụ: Cho lược đồ quan hệ $Q(ABCDE)$ hai tập phụ thuộc hàm:
 $F = \{A \rightarrow BC, A \rightarrow D, CD \rightarrow E\}$ và $G = \{A \rightarrow BCE, A \rightarrow ABD, CD \rightarrow E\}$

a) F có tương đương với G không?

b) F có tương đương với $G' = \{A \rightarrow BCDE\}$ không?

Giải:

Kiểm tra các tập FD tương đương

Ví dụ: Cho lược đồ quan hệ $Q(ABCDE)$ hai tập phụ thuộc hàm:

$F = \{A \rightarrow BC, A \rightarrow D, CD \rightarrow E\}$ và $G = \{A \rightarrow BCE, A \rightarrow ABD, CD \rightarrow E\}$

a) F có tương đương với G không?

b) F có tương đương với $G' = \{A \rightarrow BCDE\}$ không?

Giải:

Kiểm tra các tập FD tương đương

- Để chứng minh F và G tương đương ta chứng minh:
 - $F^+ \supseteq G$ Bằng cách: $\forall X \rightarrow Y \in G \Rightarrow X \rightarrow Y \in F^+$
 - $G^+ \supseteq F$ Bằng cách: $\forall X \rightarrow Y \in F \Rightarrow X \rightarrow Y \in G^+$
- Ví dụ: Cho lược đồ quan hệ Q(ABCDE) hai tập phụ thuộc hàm:
 $F = \{A \rightarrow BC, A \rightarrow D, CD \rightarrow E\}$ và $G = \{A \rightarrow BCE, A \rightarrow ABD, CD \rightarrow E\}$
F có tương đương với G không?