CƠ SỞ DỮ LIỆU CHƯƠNG 6: PHỤ THUỘC HÀM VÀ DẠNG CHUẨN

ThS. LÊ NGÔ THỰC VI



MỤC TIÊU



- 1. Hiểu được ý nghĩa phụ thuộc hàm, dạng chuẩn và các thành phần liên quan, các bài toán vận dụng.
- 2. Xác định dạng chuẩn của lược đồ CSDL quan hệ, cải tiến dạng chuẩn lược đồ CSDL.



NỘI DUNG

- 1. Các vấn đề gặp phải khi tổ chức CSDL
- 2. Phụ thuộc hàm (Hệ tiên đề Armstrong, Bao đóng, thuật toán tìm khóa)
- 3. Các dạng chuẩn (1, 2, 3, Boyce Codd)



GIỚI THIỆU

1



CHẤT LƯỢNG CỦA MỘT LƯỢC ĐỒ CSDL QUAN HỆ

- Ngữ nghĩa của các quan hệ
- Thể hiện sự dư thừa thông tin
- Những khó khăn trong vấn đề:
 - Thêm
 - Xóa
 - Cập nhật
- Thể hiện giá trị NULL trong các bộ



CHẤT LƯỢNG CỦA MỘT LƯỢC ĐỒ CSDL QUAN HỆ

Thiết kế một lược đồ quan hệ tốt và xấu?

- Đánh giá (không theo chuẩn) một lược đồ quan hệ dựa trên tiêu chí
 - Ngữ nghĩa của các thuộc tính.
 - Trùng lắp dữ liệu.
 - Giá trị Null trong các bộ.
- Định nghĩa các dạng chuẩn được xem là cách đánh giá theo chuẩn của lược đồ quan hệ



CHẤT LƯỢNG CỦA MỘT LƯỢC ĐỒ CSDL QUAN HỆ

- Đánh giá (không theo chuẩn) một lược đồ quan hệ
 - Ngữ nghĩa của các thuộc tính: Làm thế nào để hiểu được giá trị thuộc tính được lưu trữ trong một bộ?
 - Thiết kế một lược đồ để dễ dàng giải thích nghĩa của nó
 - Giữ cho các thuộc tính trên quan hệ hoặc thực thể không được trùng lắp



CHẤT LƯỢNG CỦA MỘT LƯỢC ĐỒ CSDL QUAN HỆ

- Đánh giá (không theo chuẩn) một lược đồ quan hệ
 - Trùng lắp dữ liệu
 - Khi một phần của dữ liệu có thể suy ra từ một số phần dữ liệu khác thì ta nói có sự trùng lắp dữ liệu

THI	MASV	<u>MAMH</u>	HOTEN	TENMONHOC	DIEM
	sv01	CSDL	Lê Nguyên Khôi	Cơ sở dữ liệu	7
	sv01	HDT	Lê Nguyên Khôi	Hướng đối tượng	2
	sv01	XSTK	Lê Nguyên Khôi	Xác suất thống kê	7
	sv02	CTRR	Hoàng Hải Minh	Cấu trúc rời rạc	9
	sv02	XSTK	Hoàng Hải Minh	Xác suất thống kê	5
rường Đại họ	sv03	CSDL	Nguyễn Thị Hà	Cơ sở dữ liệu	5

Thực hiện bởi Tr



- Trùng lắp dữ liệu: các vấn đề
- a. Sửa dữ liệu: phải sửa luôn các giá trị của các bộ liên quan

THI	MASV	<u>MAMH</u>	HOTEN	TENMONHOC	DIEM
	sv01	CSDL	Lê Nguyên Khôi	Cơ sở dữ liệu	7
	sv01	HDT	Lê Nguyên Khôi	Hướng đối tượng	2
	sv01	XSTK	Lê Nguyên Khôi	Xác suất thống kê	7
	sv02	CTRR	Hoàng Hải Minh	Cấu trúc rời rạc	9
	sv02	XSTK	Hoàng Hải Minh	Xác suất thống kê	5
	sv03	CSDL	Nguyễn Thị Hà	Cơ sở dữ liệu	5



- Trùng lắp dữ liệu: các vấn đề
- b. Thêm dữ liệu: phải thêm chính xác dữ liệu, nếu không sẽ dẫn đến sự mâu thuẫn với các dữ liệu đã nhập.
- c. Xóa dữ liệu: có khả năng sẽ làm mất thông tin

THI	MASV	<u>MAMH</u>	HOTEN	TENMONHOC	DIEM
	sv01	CSDL	Lê Nguyên Khôi	Cơ sở dữ liệu	7
	sv01	HDT	Lê Nguyên Khôi	Hướng đối tượng	2
	sv01	XSTK	Lê Nguyên Khôi	Xác suất thống kê	7
	sv02	CTRR	Hoàng Hải Minh	Cấu trúc rời rạc	9
	sv02	XSTK	Hoàng Hải Minh	Xác suất thống kê	5
	sv03	CSDL	Nguyễn Thị Hà	Cơ sở dữ liệu	5
?	sv04	XSTK	Đặng Thị Yến Vy	XS Thống kê	null



- Trùng lắp dữ liệu: giải quyết => phân rã lược đồ
 Sự phân rã là chia quan hệ gốc thành một vài quan hệ nhỏ hơn.
 Hai đặc tính quan trọng cần phải duy trì trong suốt quá trình phân rã:
 - ✓ Sự bảo toàn thông tin: có thể tìm được bất kỳ thể hiện nào của quan hệ gốc từ các quan hệ được phân rã (tất cả thuộc tính phải được tìm thấy trong các quan hệ được phân rã)
 - ✓ Bảo toàn phụ thuộc: các ràng buộc trên quan hệ gốc phải được đảm bảo còn giữ trên các quan hệ được phân rã.



Trùng lắp dữ liệu: giải quyết => phân rã lược đồ

SINHVIEN	<u>MASV</u>	HOTEN
	sv01	Lê Nguyên Khôi
	sv02	Hoàng Hải Minh
	sv03	Nguyễn Thị Hà

MONHOC	<u>MAMH</u>	TENMONHOC
	CSDL	Cơ sở dữ liệu
	HDT	Hướng đối tượng
	XSTK	Xác suất thống kê

1111		
MASV	<u>MAMH</u>	DIEM
sv01	CSDL	7
sv01	HDT	2
sv01	XSTK	7
sv02	CTRR	9
sv02	XSTK	5
sv03	CSDL	5



CHẤT LƯỢNG CỦA MỘT LƯỢC ĐỒ CSDL QUAN HỆ

- Đánh giá (không theo chuẩn) một lược đồ quan hệ
 - Giá trị Null trong các bộ: xảy ra khi
 - Thuộc tính không/chưa được áp dụng
 - Giá trị tại thuộc tính chưa xác định
 - Chưa nhập giá trị
- Tốn nhiều không gian lưu trữ
- Không chắc chắn cho ra kết quả mong muốn khi sử dụng chúng trong hàm Sum, Count



PHỤ THUỘC HÀM

2



2.1. Phụ thuộc hàm - Functional dependences (FDs)

Phụ thuộc hàm (FDs) là một loại RBTV biểu diễn mối liên hệ ngữ nghĩa giữa các thuộc tính trong cùng một quan hệ.

Dịnh nghĩa:

- X, Y là hai tập thuộc tính trên quan hệ R
- ☐ Ta nói Y phụ thuộc hàm trên X hay Y phụ thuộc X

Ký hiệu: X → Y, nếu và chỉ nếu: mỗi giá trị của X trong R <u>xác định</u> duy nhất một giá trị của Y trong R

☐ X là vế trái của phụ thuộc hàm (yếu tố quyết định cho PTH), Y là vế phải của phụ thuộc hàm.

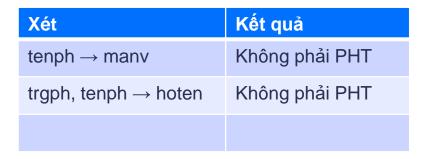


2.1. Phụ thuộc hàm

Ví dụ: cho quan hệ NHANVIEN như sau

manv	hoten	dchi	tenph	trgph
nv01	Nguyễn Minh Anh	Hà Nội	Kế toán	Lê Hải Anh
nv02	Lê Hải Anh	Hà Nội	Kế toán	Lê Hải Anh
nv03	Nguyễn Hà Lê	Đà Nẵng	Kế toán	Lê Hải Anh
nv04	Trần Đăng Hoàng	Đà Nẵng	Dữ liệu	Trần Đăng Hoàng
nv05	Đỗ Mỹ Linh	Cần Thơ	Dữ liệu	Trần Đăng Hoàng

Xét các tính chất sau	Phụ thuộc hàm
Với mỗi manv có duy nhất một hoten.	manv → hoten
Với mỗi manv có duy nhất một tenph.	manv → tenph
Với mỗi tenph có duy nhất một trgph.	tenph → trgph
Với mỗi trgph có duy nhất một tenph.	trgph → tenph
Với mỗi manv có duy nhất một hoten, dchi.	manv → hoten, dchi







2.1. Phụ thuộc hàm

ĐẶC TRƯNG CỦA PTH

- Nhận diện PTH: Việc nhận diện PTH dựa vào ý nghĩa của thuộc tính và mối quan hệ của chúng trong quan hệ.
 - Dựa vào dữ liệu: manv → hoten và hoten → manv
 - Tuy nhiên chỉ có PTH: manv → hoten
- PTH phải được định nghĩa trên lược đồ quan hệ
 - Thỏa với mọi thể hiện của quan hệ

manv	hoten		
nv01	Nguyễn Minh Anh		
nv02	Lê Hải Anh		
nv03	Nguyễn Hà Lê		
nv04	Trần Đăng Hoàng		
nv05	Đỗ Mỹ Linh		



2.2. Suy dẫn từ tập PTH

Gọi F là tập các phụ thuộc hàm cho trước trên 1 quan hệ

Khả năng suy dẫn nhằm khám phá thêm tập PTH là rất cần thiết để thiết kế các lược đồ quan hệ đạt chất lượng tốt.

Định nghĩa luật suy dẫn/diễn:

 $X \to Y$ được suy ra từ F, hay F suy ra $X \to Y$ nếu bất kỳ bộ của quan hệ nếu thỏa F thì cũng thỏa $X \to Y$





Hệ tiên đề ARMSTRONG

- 1. Luật phản hồi/phản xạ:
- 2. Luật cộng/tăng trưởng:
- 3. Luật bắc cầu:
- 4. Luật bắc cầu giả:
- 5. Luật hợp:
- 6. Luật phân rã:

$$\forall \ Y {\subseteq} X, \ X \to Y$$

Nếu
$$X \rightarrow Y$$
 và $Z \subseteq W$ thì $XW \rightarrow YZ$

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

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

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

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



2.2. Suy dẫn từ tập PTH

Ví dụ: Xét lược đồ R thỏa tập PTH F

 $F = \{f1: A \rightarrow B, f2: A \rightarrow C, f3: BC \rightarrow D\}$, chứng minh R có thỏa pth $A \rightarrow D$?

Lời giải:

- 1. A → B (giả thiết)
- 2. $A \rightarrow C$ (giả thiết)
- 3. A \rightarrow BC (từ 1,2: luật hợp)
- 4. BC → D (giả thiết)
- 5. $A \rightarrow D$ (từ 3,4: luật bắc cầu)

Từ F suy dẫn được $A \rightarrow D$, vậy R thỏa pth $A \rightarrow D$



2.2. Suy dẫn từ tập PTH - Bài tập Luật suy diễn

- 1. Cho F = {CD \rightarrow H, B \rightarrow EG, E \rightarrow AD}. Chứng minh: BC \rightarrow H
- 2. Cho F = {AB \rightarrow C, B \rightarrow D, CD \rightarrow E, CE \rightarrow GH} Chứng minh: AB \rightarrow GH
- 3. Cho F = {AB \rightarrow E, AG \rightarrow I, BE \rightarrow I, E \rightarrow G, GI \rightarrow H} Chứng minh: AB \rightarrow GH



2.3. Hệ quả từ tập PTH

Cho F là tập các PTH định nghĩa trên R

Nếu có 1 PTH f khác cũng được thỏa với mọi thể hiện của R thì ta gọi f là
 hệ quả của F

⇒ PTH hệ quả của F là PTH được suy dẫn từ F

Ví dụ: Xét lược đồ R(A, B, C, G, H, I) và PTH F định nghĩa trên R

 $F = \{f1: A \rightarrow B, f2: A \rightarrow C, f3: CG \rightarrow H, f4: CG \rightarrow I, f5: B \rightarrow H\}$

Hỏi: A → H có phải là PTH hệ quả của F không?



2.3. Hệ quả từ tập PTH - Bao đóng của tập PTH

- Cho F là tập các PTH định nghĩa trên R
- Tập hợp các PTH hệ quả từ F được gọi là bao đóng của F

Ký hiệu: F+

 $F \subseteq F^+$

- Nếu F = F + thì F là họ đầy đủ của các phụ thuộc hàm.
- Nếu F quá lớn, tìm F+ sẽ khó khăn và tốn thời gian

Bài toán thực tế:

Cho trước một tập PTH F và PTH f: $X \rightarrow Y$

Xác định f có thuộc bao đóng F+ không? Hay f có phải là PTH hệ quả của F không

Giải quyết: - Tìm bao đóng F+

- Kiểm tra f có nằm trong F+ không



2.3. Hệ quả từ tập PTH - Bao đóng của tập PTH

THUẬT TOÁN TÌM F+

Bước 1:

 $F^+ = F$

Bước 2:

Repeat

- Áp dụng các luật dẫn Armstrong để tìm phụ thuộc hàm hệ quả f từ F⁺
 - Bổ sung f vào F⁺

Until F^+ không thể thay đổi được nữa hay không thể tìm được f khác từ F^+



2.3. Hệ quả từ tập PTH - Bao đóng của tập PTH

Ví dụ:

$$F = \{AE \rightarrow C, CG \rightarrow A, BD \rightarrow G, GA \rightarrow E, H \rightarrow D \}$$

 $CG \rightarrow D$ có thuộc F^+ không?

Lời giải:

Phức tạp nếu dùng hệ dẫn luật Armstrong để xác định bao đóng của F.

- ⇒ Cần phải có cách khác ??
- ⇒ Chuyển sang bài toán thành viên: pth f: X → Y
- Ta chỉ cần tìm bao đóng của tập thuộc tính X dựa trên F
- Kiểm tra Y có thuộc bao đóng của X hay không



Bao đóng của tập thuộc tính X đối với tập phụ thuộc hàm F

Ký hiệu: X^+_F là tập tất cả các thuộc tính Y có thể suy dẫn từ X nhờ F

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

Nhận xét:

- Là tập hợp những VP của các PTH có VT là X nằm trong F
- X ⊆ X⁺_F
- X ⊆ R^{+:} R⁺ là X⁺_F lớn nhất



B1.
$$X_F^+ = X$$

Tìm các PTH trong F có VT là các thuộc tính nằm trong X⁺_F có VP không nằm trong X⁺_F

B2. **Lặp** {

Nếu (có f : U → V thuộc F) \mathbf{va} (U $\subseteq X^{+}_{F}$) và (V $\not\subset X^{+}_{F}$)

Thì $X^+_F = X^+_F \cup V$

} cho đến khi $(X^+_F = R^+)$ hoặc

(không còn thay đổi được nữa)



Ví du 1:

Cho Q = ABCD và F = {A \rightarrow B, A \rightarrow C, CD \rightarrow A}. Tìm A^+_F

Lời giải:

$$A^{+}_{F} = A$$
 $A^{+}_{F} = AB \text{ (Vi } A \rightarrow B)$
 $A^{+}_{F} = ABC \text{ (Vi } A \rightarrow C)$
 $Var{e}_{F} = ABC$



Ví dụ 2:

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 AC_F^+ ?

Lời giải:

 $AC_F^+ = AC$

 $AC_F^+ = ACD \text{ (vi } AC \rightarrow D)$

 $AC_F^+ = ACDE \text{ (vì DA} \rightarrow CE)$

 $AC_F^+ = ACDEH (vi D \rightarrow H)$

Vậy: AC+_F = ACDEH



Bài toán thành viên

Cho tập thuộc tính Q, tập phụ thuộc hàm F trên Q và một phụ thuộc hàm $X \to Y$ trên Q. Xác định $X \to Y \in F^+$ hay không?

Hướng dẫn

Dựa vào tính chất

$$X \rightarrow Y \in F^+ \Leftrightarrow Y \subseteq X^+_F$$

Ta tìm bao đóng X⁺_F
 Nếu Y ⊆ X⁺_F thì X → Y ∈ F⁺
 ngược lại X → Y ∉ F⁺



Ví dụ 2: đầy đủ

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

 $F = \{B \rightarrow A , DA \rightarrow CE, D \rightarrow H, GH \rightarrow C, AC \rightarrow D\}.$

Hỏi: AC → E có thuộc F⁺ không?

Lời giải

Ta có AC+_F = ACDEH (*Lưu ý*: phải trình bày đầy đủ tìm bao đóng)

Vì E \subseteq AC $^+$ _F nên AC \rightarrow E \in F $^+$



2.4. Hệ quả từ tập PTH – Bài tập

Bài tập 1:

Cho Q(ABCDEG) và $F = \{BD \rightarrow C, AEG \rightarrow BC, CG \rightarrow AE, B \rightarrow CG\}$

Hỏi: $B \rightarrow D \in F^+$?

Bài tập 2:

Cho lược đồ Q(ABCDEG) và $F = \{AE \rightarrow C, CG \rightarrow A, BD \rightarrow G, GA \rightarrow E \}$

Hỏi: BDC \rightarrow E \in F $^+$?

Bài tập 3:

Cho Q(ABCDEGH) và $F = \{ B \rightarrow A, D \rightarrow CE, D \rightarrow H, GH \rightarrow C, AC \rightarrow D \}$

Chứng minh: $AC \rightarrow E \in F^+$



Định nghĩa

- Cho lược đồ quan hệ Q(A1, A2, ..., An), Q⁺ là tập thuộc tính của Q, F là tập phụ thuộc hàm trên Q, K là tập con của Q⁺. Khi đó K gọi là một khóa của Q nếu:
 - (i) $K_F^+ = Q_F^+$
 - (ii) Không tồn tại K'⊂K sao cho K'+_F = Q+
- Thuộc tính A được gọi là thuộc tính khóa nếu A∈ K, trong đó K là khóa của Q. Ngược lại thuộc tính A được gọi là thuộc tính không khóa.
- K' được gọi là siêu khóa nếu K' ⊇ K.

Một quan hệ có thể có nhiều khóa



Các phụ thuộc hàm:

- ma_ct → sohd, masp, sl
- sohd, masp → ma_ct, sl

Siêu khóa:

{ma_ct}

{ma_ct, sohd}, {ma_ct, masp}, {ma_ct, sl}, {ma_ct, sohd, masp}, {ma_ct, sohd, sl}, {ma_ct, masp, sl}, {ma_ct, sohd, masp, sl}
{sohd, masp}, {sohd, masp, sl}

Khóa: {ma_ct}, {sohd, masp}

- Thuộc tính khóa: ma_ct, sohd, masp
- Thuộc tính không khóa: sl

CTHD

Ma_ct	Sohd	Masp	Sl
1	hd01	bb01	20
2	hd01	bb02	5
3	hd02	bt01	13
4	hd02	bc04	20
5	hd03	bc04	4



Thuật toán xác định khóa của quan hệ:

- 1. Xây dựng các tổ hợp thuộc tính có thể có từ Q+
- 2. Tìm tập S chứa tất cả các tổ hợp K ⊆ Q⁺ thỏa điều kiện K → Q⁺, mỗi tổ hợp K như vậy là một siêu khóa của Q.
- 3. ∀K∈ S
 Nếu ∃K' | K' ⊂ K thì loại K ra khỏi S

Vấn đề: Số tổ hợp có thể có từ Q⁺ sẽ rất lớn nếu Q⁺ lớn → Cần giới hạn số tổ hợp cần khảo sát



Giới hạn số lượng tổ hợp thuộc tính:

- Tập thuộc tính nguồn, ký hiệu là N
 - A là một thuộc tính nguồn nếu ¬∃f: X→Y ∈ F | A∈Y
 - Nhận xét: mọi thuộc tính nguồn phải xuất hiện trong mọi khóa của Q
- Tập thuộc tính đích:
 - B là một thuộc tính đích nếu ¬∃f: X→Y ∈ F| B∈X
 - Nhận xét: thuộc tính đích không xuất hiện trong bất kỳ khóa nào của Q
- Tập thuộc tính trung gian, ký hiệu là TG: các thuộc tính còn lại để xác định các tổ hợp



2.5. Phụ thuộc hàm – Bài toán tìm Khóa

Bước 1:

- Tìm tập nguồn N.
- Nếu N+_F = Q+ thì chỉ có 1 khoá là N, ngược lại qua bước 2.

Bước 2:

- Tìm tập trung gian TG.
- Tìm tập tất cả các tập con X; của tập TG.

Bước 3: Tìm tập S chứa mọi siêu khóa S_i

- Với mỗi X_i , nếu $(N \cup X_i)^+_F = Q^+$ thì $S_i = (N \cup X_i)^+$
- Nếu: (N ∪ X_i)+_F= Q+ khi đó N ∪ X_i là một khóa.

Khi đó: loại bỏ các trường hợp X_i : $X_i \subset X_j$

VD: $X_i = AB$, $X_j = ABC$. Ta thấy $X_i \subset X_j$, nếu X_i là khóa thì không cần xét trường hợp X_j nữa.



2.5. Phụ thuộc hàm – Bài toán tìm Khóa

- B1: Xác định
 - TN
 - TG
- B2: Lập bảng (Xi tổ hợp từ TG)

N	\mathbf{X}_{i}	$TN \cup X_i$	$(\operatorname{TN} \cup \mathbf{X}_i)_F^+$	K



2.5. Phụ thuộc hàm – Bài toán tìm Khóa

Ví dụ: Cho lược đồ quan hệ Q(A, B, C) và tập phụ thuộc hàm

$$F=\{AB \rightarrow C, C \rightarrow A\}$$

Tìm mọi khóa của Q.

Lời giải:

- Bước 1: $N = \{B\}, N_F^+ = B \neq Q^+$
- Bước 2: TG = {AC}, tập các tập con trung gian là CTG = {A,C, AC}
- Bước 3:

N	X_i	$N \cup X_i$	$(\mathbf{N} \cup \mathbf{X}_i)_F^+$	
В	A	BA	BAC=Q+	Khóa là BA. Loại các phần tử trong CTG chứa A: AC
В	C	BC	BCA=Q+	Khóa là BC

Kết luận: tập khoá S = {BA, BC}



2.5. Phụ thuộc hàm – Bài tập tìm khóa

Bài tập 1: Cho lược đồ quan hệ Q(ABCD) và tập phụ thuộc hàm $F=\{A \rightarrow B, A \rightarrow CD, BC \rightarrow D\}$. Tìm tất cả các khóa của Q

Bài tập 2: Cho lược đồ quan hệ Q(ABCDE) và tập phụ thuộc hàm F={ BC \rightarrow A, A \rightarrow CD, C \rightarrow DE}. Tìm tất cả các khóa của Q.

Bài tập 3: Cho lược đồ quan hệ Q(A, B, C,D) và tập phụ thuộc hàm $F=\{A \rightarrow BCD, CD \rightarrow AB\}$. Tìm tất cả các khóa của Q.

Bài tập 4: cho lược đồ quan hệ Q(A, B, C, D, E, G) và tập phụ thuộc hàm $F=\{AE \rightarrow C, CG \rightarrow A, BD \rightarrow G, GA \rightarrow E\}$. Tìm tất cả các khóa của Q