Trường Đại học Công nghệ Thông tin

Khoa Hệ thống Thông tin

Chương 6

PHỤ THUỘC HÀM VÀ DẠNG CHUẨN

Cao Thị Nhạn

NỘI DUNG

1. Phụ thuộc hàm
   1.1 Hệ luật dẫn Amstrong
   1.2 Bao đóng của tập thuộc tính
   1.3 Phủ tối thiểu
   1.4 Khóa
   1.5 Thuật toán tìm khóa
2. Dạng chuẩn
   2.1 Dạng chuẩn 1
   2.2 Dạng chuẩn 2
   2.3 Dạng chuẩn 3
   2.4 Dạng chuẩn Boyce Codd
   2.5 Dạng chuẩn của lược đồ quan hệ

Phụ thuộc hàm

- Với DEAN(MaDA, TenDA, DdiemDA, Phong)
  - MaDA xác định TenDA (hay TenDA phụ thuộc vào MaDA)
  - MaDA xác định DDiemDA (hay DDiemDA phụ thuộc vào MaDA)
  - MaDA xác định Phong (hay Phong phụ thuộc vào MaDA)
- Ký hiệu:
  - MaDA → TenDA
  - MaDA → DdiemDA
  - MaDA → Phong

Nhận xét các PTH trên

Phụ thuộc hàm

- Với:
  - $Q(A_1, A_2, \ldots, A_n)$ là quan hệ;
  - $Q^+ = \{A_1, A_2, \ldots, A_n\}$;
  - X, Y là hai tập con của Q^+;
  - t_1, t_2 là hai bộ bất kỳ của Q.
- Khi đó: $X \rightarrow Y \Leftrightarrow (t_1.X = t_2.X \Rightarrow t_1.Y = t_2.Y)$
- X xác định Y hay Y phụ thuộc hàm vào X
- X được gọi là vế trái phụ thuộc hàm, Y được gọi là vế phải phụ thuộc hàm.

Phụ thuộc hàm

- Xét CTIETHD (SoHD, MaHang, SoLuong, DonGia, ThanhTien)
  - {SoHD, MaHang} → SoLuong
  - {SoHD, MaHang} → DonGia
  - {SoHD, MaHang} → ThanhTien
  - {SoLuong, DonGia} → ThanhTien

Nhận xét các PTH trên

- Tồn tại nhiều phụ thuộc hàm trên một quan hệ
- Tập các phụ thuộc hàm được ký hiệu F

Hệ luật dẫn Amstrong

- Gọi F là tập các phụ thuộc hàm, R là quan hệ
- $X \rightarrow Y$ được suy ra từ F nếu bất kỳ bộ của quan hệ R thỏa F thì cũng thỏa $X \rightarrow Y$
- Ký hiệu: $F \models X \rightarrow Y$

Hệ luật dẫn Amstrong

- Với X, Y, Z, W $\subseteq$ U. Phụ thuộc hàm có các tính chất sau:
  - F1) Tính phản xạ: Nếu Y $\subseteq$ X thì X $\rightarrow$ Y
  - F2) Tính tăng trưởng: Nếu X $\rightarrow$ Y thì XZ $\rightarrow$ YZ
  - F3) Tính bắc cầu: Nếu {X $\rightarrow$ Y, Y $\rightarrow$ Z} thì X $\rightarrow$ Z
- Một số tính chất được suy ra từ hệ luật dẫn Amstrong
  - F4) Tính kết hợp: Nếu {X $\rightarrow$ Y, X $\rightarrow$ Z} thì X $\rightarrow$ YZ
  - F5) Tính phân rã: Nếu {X $\rightarrow$ YZ, X $\rightarrow$ Y} thì X $\rightarrow$ Z
  - F6) Tính tựa bắc cầu: Nếu {X $\rightarrow$ Y, YZ $\rightarrow$ W} thì XZ $\rightarrow$ W

Hệ luật dẫn Amstrong
Ví dụ

- Cho quan hệ R(A, B, C, D, E, G, H) và tập phụ thuộc hàm F={AB→C, B→D, CD→E, CE→GH, G→A}
- Chứng tỏ AB→E được suy diễn từ F.

1. AB→C
2. AB→AB (tính phản xạ)
3. AB→B (tính phân rã)
4. B→D
5. AB→D (tính bắc cầu 3+4)
6. AB→CD (tính hợp 1+5)
7. CD→E
8. AB→E (tính bắc cầu 6+7)

Bao đóng tập thuộc tính

- Bao đóng tập phụ thuộc hàm

  Bao đóng của tập phụ thuộc hàm F, ký hiệu F^+ là tập tất cả các phụ thuộc hàm được suy ra từ F.

- Bao đóng tập thuộc tính

  Bao đóng của tập thuộc tính X đối với tập phụ thuộc hàm F, ký hiệu là X^+_F là tập tất cả các thuộc tính A có thể suy dẫn từ X nhờ tập bao đóng của các phụ thuộc hàm F^+

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

Bao đóng tập thuộc tính

- **Input:** (Q,F), $X \subseteq Q^+$
- **Output:** X^+_F

Nhận xét

Bước 1: Tính dãy $X^{(0)}$, $X^{(1)}$,..., $X^{(i)}$:

- $X^{(0)} = X$
- $X^{(i+1)} = X^{(i)} \cup Z$, $\exists(Y \rightarrow Z) \in F(Y \subseteq X^{(i)})$,

  loại $(Y \rightarrow Z)$ ra khỏi F

- Dừng khi $X^{(i+1)} = X^{(i)}$ hoặc khi $X^{(i)}=Q^+$

Bước 2: Kết luận $X^+_F = X^{(i)}$

Bao đóng tập thuộc tính

Cho LDQH R(A,B,C,D,E,G,H) và tập PTH

F={f1: B $\rightarrow$ A, f2: DA $\rightarrow$ CE, f3: D $\rightarrow$ H, f4: GH $\rightarrow$ C, f5: AC $\rightarrow$ D}

Tìm AC^+_F

Bước 1: X^0 = AC

Bước 2: Từ f1 đến f4 không thoả, f5 thoả nên X^1= AC ∪ D = ACD

Lặp lại bước 2: f1 không thoả,

f2 thỏa nên X^2=ACD ∪ CE = ACDE

f3 thỏa nên X^2=ACDE ∪ H =ACDEH

f4 không thỏa, f5 đã thỏa

Lặp lại bước 2: f2, f3 và f5 đã thỏa, f1 và f4 không thỏa. Nên X^3=X^2=ACDEH

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→Y trên Q. Câu hỏi đặt ra rằng X→Y ∈ F^+ hay không?
- Giải quyết: $X \rightarrow Y \in F^{+} \Leftrightarrow Y \subseteq X_F^{+}$

Nhận xét

Bài toán thành viên
Ví dụ

- Cho lược đồ quan hệ R(A,B,C,D,E,G,H) và tập phụ thuộc hàm F={f1: B→A , f2: DA→CE, f3: D→H, f4: GH→C, f5: AC→D}
- Cho biết AC→E ∈ F^+?

$$AC_F^+ = ACDEH$$

Vì E ∈ AC^+_F nên AC→E ∈ F^+

Phủ tối thiểu

- **Hai tập PHT tương đương**: Hai tập phụ thuộc hàm F và G tương đương nếu $F^+ = G^+$. Ký hiệu $G \equiv F$
- **PTH có thuộc tính vế trái dư thừa**
  - Cho F là tập PTH trên lược đồ quan hệ Q; $Z \rightarrow Y \in F$ là PTH có thuộc tính vế trái dư thừa nếu tồn tại $A \in Z$ mà

$$F = F - (Z \rightarrow Y) \cup ((Z - A) \rightarrow Y)$$

  - Ngược lại $Z \rightarrow Y$ là PTH có thuộc tính vế trái không dư thừa hay Y phụ thuộc đầy đủ vào Z.
  - Ví dụ: R(A, B, C, D) và F={ A→B , BC→D, C→D}. Khi đó BC→D là phụ thuộc hàm có thuộc tính vế trái dư thừa

Phủ tối thiểu

- **Thuật toán tìm các PTH đầy đủ tương ứng**
  - Với mỗi PTH $X \rightarrow Y$, $X=A_1A_2...A_n$, $n>=2$, đặt $Z=X$
  - Với mỗi A_i, thực hiện:
    - $Tam = Z \setminus A_i$
    - Nếu $Tam \rightarrow Y \in (F-\{X \rightarrow Y\})^+$ thì $Z=Tam$
- Ví dụ: R(A,B,C,D), $F=\{A \rightarrow B, BC \rightarrow D, C \rightarrow D\}$
  - Xét $BC \rightarrow D$: có $C^+_{F-\{BC \rightarrow D\}} = CD$, vì $D \subseteq C^+_F$ nên B là thuộc tính thừa → loại bỏ B
  - Vậy $F=\{A \rightarrow B, C \rightarrow D\}$

Phủ tối thiểu

- **Tập PTH có vế phải một thuộc tính:** Mỗi tập PTH F đều tương đương với một tập PTH G mà vế phải của các PTH thuộc G chỉ gồm một thuộc tính
- **Tập phụ thuộc hàm không dư thừa:** F là tập PTH không dư thừa nếu không tồn tại $F' \subset F$ sao cho $F' \equiv F$.
- **Thuật toán loại những PTH dư thừa**

  Với mỗi phụ thuộc hàm $X \rightarrow Y \in F$, nếu $X \rightarrow Y$ là thành viên của $F - \{X \rightarrow Y\}$ thì loại $X \rightarrow Y$ khỏi F.

Phủ tối thiểu

- **Phủ tối thiểu (PTT):** F được gọi là phủ tối thiểu của tập phụ thuộc hàm (hay tập PTH tối thiểu) nếu thỏa:
  1. Tập PTH có vế trái không dư thừa
  2. Tập PTH có vế phải 1 thuộc tính
  3. Tập PTH không dư thừa

Nhận xét

- **Thuật toán tìm PTT:**
  1. Loại các thuộc tính vế trái dư thừa của mọi PTH
  2. Phân rã các PTH có vế phải nhiều thuộc tính thành các PTH có vế phải một thuộc tính
  3. Loại các PTH dư thừa

Phủ tối thiểu

- Bài tập tìm PTT

1. Cho R(A, B, C, D) và F={AB→CD, B→C, C→D}. Tìm PTT của F
2. Cho R(A, B, C, D, E, G, H) và F={B→A, A→BC, AB→G, GH→E, BCG→A}. Tìm PTT của F

Khóa

- Định nghĩa

Cho lược đồ quan hệ Q(A1, A2, ..., An), Q^+ là tập thuộc tính của quan hệ 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) $K^+_F = Q^+$

(ii) Không tồn tại $K' \subset K$ *sao cho* $K'^+_F = Q^+$

Thuộc tính A được gọi là thuộc tính khóa nếu $A \in K$. Ngược lại thuộc tính A được gọi là thuộc tính không khóa.

K' được gọi là siêu khóa nếu $K \subseteq K'$.

Khóa

- Thuật toán tìm 1 khóa: sử dụng đồ thị có hướng
- Bước 1: Xây dựng đồ thị
  - Mỗi nút là tên một thuộc tính
  - Cung nối 2 nút thể hiện phụ thuộc hàm
  - Nút gốc: chỉ có hướng mũi tên ra (chỉ nằm ở vế trái PTH)
  - Nút lá: chỉ có các hướng mũi tên vào (chỉ nằm ở vế phải)

Nhận xét

Khóa

- Thuật toán tìm 1 khóa: sử dụng đồ thị có hướng
- Bước 2:

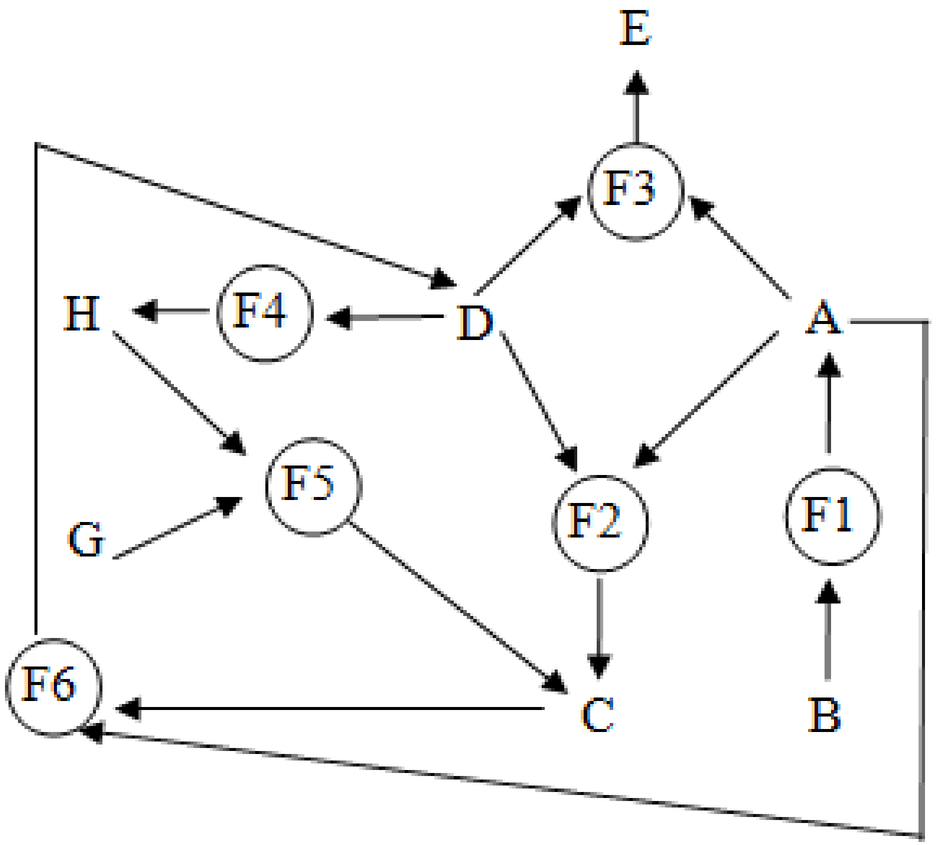
Nhận xét

  - Xuất phát từ tập nút gốc X, tìm bao đóng X^+_F
  - Nếu $X^+_F = Q^+$ thì X là khóa, ngược lại bổ sung một thuộc tính không thuộc nút lá vào X rồi thực hiện tìm bao đóng của X.
  - Dừng khi tìm được một khóa của R

Khóa

- Cho lược đồ quan hệ R(A, B, C, D, E, G, H) và tập PTH F={B→A, DA→CE, D→H, GH→C, AC→D}
- Tìm một khóa của R
- Các nút gốc: BG
- Các nút lá: E
- Các nút trung gian:

ACDH



Khóa

- BG^+_F=BGA, vì $BG^+_F \neq Q^+$ nên BG không là khóa.
- Bổ sung D
- BGD^+_F=BGDACEH, vì $BGD^+_F = Q^+$ nên BGD là khóa.
- Nhận xét:
  - Ưu điểm
  - Hạn chế

Khóa

- Thuật toán tìm mọi khóa
- Tập thuộc tính nguồn, ký hiệu N, chứa những thuộc tính chỉ xuất hiện ở vế trái của các PTH
- Tập thuộc tính trung gian, ký hiệu là TG, chứa những thuộc tính vừa xuất hiện ở vế trái, vừa xuất hiện ở vế phải của các PTH

Khóa

- Thuật toán tìm mọi khóa

1. Tính N^+_F, nếu $N^+_F = Q^+$, chỉ có một khóa duy nhất, ngược lại qua bước 2

Nhận xét

2. Tìm TG, tính mọi tập con Xi của tập TG
3. Với mỗi Xi, nếu $(N \cup Xi)^+_F = Q^+$ thì $Si = (N \cup Xi)$, loại các tập Xj: $Xi \subset Xj$
4. Kết luận tập các khóa K={Si}

Khóa

- Hãy tìm mọi khóa cho các ví dụ sau

1. Cho R(A, B, C), F={AB→C, C→A}.
2. R(A, B, C, D, E, G, H), F={B→A, DA→CE, D→H, GH→C, AC→D}

NỘI DUNG

1. Phụ thuộc hàm
   - 1.1 Hệ luật dẫn Amstrong
   - 1.2 Bao đóng của tập thuộc tính
   - 1.3 Phủ tối thiểu
   - 1.4 Khóa
   - 1.5 Thuật toán tìm khóa
2. Dạng chuẩn
   - 2.1 Dạng chuẩn 1
   - 2.2 Dạng chuẩn 2
   - 2.3 Dạng chuẩn 3
   - 2.4 Dạng chuẩn Boyce Codd
   - 2.5 Dạng chuẩn của lược đồ quan hệ

Dạng chuẩn 1

- *Lược đồ Q ở dạng chuẩn 1 nếu mọi thuộc tính đều mang giá trị nguyên tố.*
- Giá trị nguyên tố là giá trị không phân nhỏ được nữa.
- Các thuộc tính đa trị (multi-valued), thuộc tính đa hợp(composite) không là nguyên tố.

Dạng chuẩn 1

- HOADON(<u>MaHD</u>, MaKH, NgayHD, CtietMua, SoTien)

MaHD	MaKH	NgayHD	CtietMua (Tên hàng, Số lượng, ĐVT)	SoTien
HD01	KH01	15-10-15	Bánh Orion, 1, gói	25.000
			Kẹo mút Chupa Chups, 2, cây	2.000
HD02	KH01	18-10-15	Gạo, 2, kg	30.000
HD03	KH02	24-10-15	Đường, 1, kg	15.000
			Bánh AFC, 2, gói	24.000

- CTietMua không là nguyên tố nên lược đồ quan hệ HOADON không đạt DC1

Dạng chuẩn 2

- *Lược đồ Q ở dạng chuẩn (DC) 2 nếu thỏa:*
  1. *Q đạt dạng chuẩn 1, và*
  2. *Mọi thuộc tính không khóa của Q đều phụ thuộc đầy đủ vào khóa*
- Kiểm tra DC 2
  1. Tìm mọi khóa của Q
  2. Với mỗi khóa K, tìm bao đóng của tập tất cả các tập con thực sự S_i của K
  3. Nếu tồn tại bao đóng S_i^+ chứa thuộc tính không khóa thì Q không đạt DC 2, ngược lại Q đạt DC 2.

Dạng chuẩn 2

- Ví dụ 1.
  - Cho Q1 (A, B, C, D), F={A→B, B→DC}
  - Lược đồ chỉ có một khóa là A, nên mọi thuộc tính đều phụ thuộc đầy đủ vào khóa. Do vậy Q1 đạt DC 2.
- Ví dụ 2.
  - Cho Q2 (A, B, C, D), F={AB→D, C→D}
  - Lược đồ có khóa là ABC, ngoài ra còn có C⊂ABC và C→D, với D là thuộc tính không khóa (nghĩa là thuộc tính D không phụ thuộc đầy đủ vào khóa). Do vậy Q2 không đạt DC 2.

Dạng chuẩn 2

- Ví dụ 3.
  - Cho SINHVIEN (<u>MSSV, MaMH</u>, TenSV, DiaChi, Diem)
  - F={MSSV, MaMH → Diem, MSSV→TenSV, DiaChi}
  - Quan hệ SINHVIEN không đạt dạng chuẩn 2 vì:
    1. Khóa: {MSSV, MaMH}
    2. Thuộc tính TenSV, DiaChi là thuộc tính không khoá chỉ phụ thuộc vào MSSV: như vậy không phụ thuộc đầy đủ vào khoá.

Dạng chuẩn 3

- *Lược đồ Q ở dạng chuẩn (DC) 3 nếu mọi phụ thuộc hàm $X \rightarrow A \in F$, với $A \notin X$ đều có:*

    1. *X là siêu khóa, hoặc*
    2. *A là thuộc tính khóa*

Dạng chuẩn 3

- Kiểm tra DC 3
  1. Tìm mọi khóa của Q
  2. Phân rã vế phải của mọi phụ thuộc hàm trong F để tập F trở thành tập phụ thuộc hàm có vế phải một thuộc tính
  3. Nếu mọi phụ thuộc hàm $X \rightarrow A \in F$, mà $A \notin X$ đều thỏa:
     - X là siêu khóa (vế trái chứa một khóa), hoặc
     - A là thuộc tính khóa (vế phải là tập con của khóa)

  thì Q đạt DC 3, ngược lại Q không đạt DC 3.

Dạng chuẩn 3

- Ví dụ 1: Q (A, B, C, D), F={AB→D, C→D}

1. Q có một khóa là ABC
2. Mọi phụ thuộc hàm trong F đều đã có vế phải một thuộc tính.
3. Với AB→D
   - Vế trái (AB) không phải là siêu khóa.
   - Vế phải (D) không là thuộc tính khóa
   - Vậy Q không đạt DC 3

Dạng chuẩn Boyce Codd (BC)

- *Lược đồ Q ở dạng chuẩn BC nếu mọi phụ thuộc hàm $X \rightarrow A \in F$, với $A \notin X$ đều có X là siêu khóa*

Dạng chuẩn BC

- Kiểm tra DC BC

1. Tìm mọi khóa của Q
2. Phân rã vế phải của mọi phụ thuộc hàm trong F để tập F trở thành tập phụ thuộc hàm có vế phải một thuộc tính
3. Nếu mọi phụ thuộc hàm $X \rightarrow A \in F$, mà $A \notin X$ đều thỏa:
   - X là siêu khóa (vế trái chứa một khóa)

   thì Q đạt DC BC, ngược lại Q không đạt DC BC.

Dạng chuẩn BC

- Ví dụ: Q (A, B, C, D, E, I), F={ACD→EBI, CE→AD}

1. Q có hai khóa là {ACD, CE}
2. Phân rã vế phải của các phụ thuộc hàm trong F, ta có

   F={ACD→E, ACD→B, ACD→I, CE→A, CE→D}
3. Mọi phụ thuộc hàm trong F đều có vế trái là một siêu khóa
4. Vậy Q đạt dạng chuẩn BC

Kiểm tra dạng chuẩn

- DC của lược đồ quan hệ là dạng chuẩn cao nhất của lược đồ quan hệ
- **Kiểm tra dạng chuẩn của lược đồ quan hệ Q**

1. Tìm mọi khóa của Q
2. Kiểm tra DC BC, nếu đúng thì kết luận Q đạt DC BC, ngược lại qua bước 3.
3. Kiểm tra DC 3, nếu đúng thì kết luận Q đạt DC 3, ngược lại qua bước 4.
4. Kiểm tra DC 2, nếu đúng thì kết luận Q đạt DC 2, ngược lại kết luận Q đạt DC 1.

- Dạng chuẩn của một lược đồ CSDL là dạng chuẩn thấp nhất trong các dạng chuẩn của các lược đồ quan hệ con.