# CHƯƠNG IV

# CÁC LOẠI PHỤ THUỘC DỮ LIỆU

# Dư thừa dữ liệu

- Cho quan hệ :
- DULIEU(MSSV,hoten,diachi,MM,tenmon,diem)

- Cho các quan hệ:
- SINHVIEN(MSSV,hoten,diachi)
- MONHOC(MM,tenmon)
- HOC(MSSV,MM,diem)

# Dư thừa dữ liệu

- Xét quan hệ CÁ_NHÂN (id, hoten, diachi, sothich) với các thể hiện:

| id | hoten | diachi | sothich |
|---|---|---|---|
| 10110100 | John Doe | 123 Lý Tự Trọng | Bơi lội |
| 10110100 | John Doe | 123 Ly Tự Trọng | Bida |
| 21345622 | Huỳnh Huy | 2 Võ Thị Sáu | Cầu Lông |
| 21345622 | Huỳnh Huy | 2 Võ Thị Sáu | Bóng chuyền |
| 55555555 | Lê Văn Tám | 411 30/4 | Leo núi |

# Dư thừa dữ liệu

CÁ_NHÂN (id, hoten, diachi)
THICH (id,sothich)

- Xét quan hệ CÁ_NHÂN (id, hoten, diachi, sothich) với các thể hiện:

| id | hoten | diachi | sothich |
|---|---|---|---|
| 10110100 | John Doe | 123 Lý Tự Trọng | Bơi lội |
| 10110100 | John Doe | 123 Ly Tự Trọng | Bida |
| 21345622 | Huỳnh Huy | 2 Võ Thị Sáu | Cầu Lông |
| 21345622 | Huỳnh Huy | 2 Võ Thị Sáu | Bóng chuyền |
| 55555555 | Lê Văn Tám | 411 30/4 | Leo núi |

- Xét 4 bộ đầu tiên:
  - Nhiều thông tin lặp lại (id, hoten, diachi) => lưu trữ **dư thừa** cho cùng thông tin => Đây không là vấn đề chính
  - Vấn đề chính là giữ cho các bản sao **dư thừa** luôn nhất quán trong CSDL và điều này phải được thực hiện một cách hiệu quả.
- => Dư thừa có thể dẫn đến bất thường dữ liệu

# Dị thường dữ liệu

- Dị thường dữ liệu là
  - Dữ liệu không đồng nhất

  Ví dụ, cùng MSSV → có 2 họ tên tên khác nhau

- Sự mâu thuẫn như vậy có thể phát sinh khi
  - Có một bộ đặc biệt được lưu trữ tại nhiều địa điểm (các bản sao);
  - Nhưng không phải tất cả các bản sao đều được cập nhật.

# Các tiêu chí đánh giá thiết kế LĐQH

- Đảm bảo rằng ngữ nghĩa của các thuộc tính là rõ ràng trong lược đồ
- Giảm thông tin dư thừa trong các bộ. Dư thừa dữ liệu gây:
  - Dị thường dữ liệu khi thêm hoặc sửa
  - Mất thông tin khi xoá
- Giảm giá trị NULL trong các bộ. Các giá trị NULL làm:
  - Lãng phí không gian lưu trữ
  - Khó thực hiện việc chọn, các hàm kết tập và nối kết
- Không chấp nhận khả năng tạo ra các bộ giả (spurious tuples)
  - Sinh ra do kết nối các quan hệ không dựa trên khoá chính và khoá ngoài.

# Giới thiệu PTH

- Khái niệm quan trọng nhất trong lý thuyết thiết kế lược đồ quan hệ là phụ thuộc hàm (PTH)
- PTH là công cụ hình thức để phân tích các lược đồ quan hệ :
  - Cho phép phát hiện và
  - Mô tả một số các vấn đề thừa và dị thường dữ liệu
- Một PTH là một ràng buộc giữa hai tập thuộc tính từ một cơ sở dữ liệu.
- PTH được sử dụng để xác định các dạng chuẩn (Normal Form).

# Định nghĩa

- Cho lược đồ quan hệ R(U) với:
  - U={ A1, A2, ..., An},
  - $X \neq \emptyset$, $Y \neq \emptyset$, $X \subseteq U$, $Y \subseteq U$

**Định nghĩa**

*X xác định Y hay Y phụ thuộc hàm vào X , nếu và chỉ nếu với mỗi giá trị của X xác định duy nhất một giá trị của Y, hay:*

$$\forall\, t1, t2 \in R,\ t1[X]=t2[X] \text{ thì } t1[Y]=t2[Y]$$

- Ký hiệu **X** $\rightarrow$ **Y**
- X là vế trái và Y là vế phải của PTH

# Ví dụ

| R(U) | A | B |
|---|---|---|
| | 1 | 4 |
| | 1 | 5 |
| | 3 | 7 |

$A \rightarrow B$ ?

$B \rightarrow A$ ?

MSSV → hoten

Hoten → MSSV

# Các tính chất phụ thuộc hàm

- $X \supseteq Y$, thì $X \rightarrow Y$
- $X \rightarrow Y$ và $Y \rightarrow Z$, thì $X \rightarrow Z$
- $X \rightarrow Y$ thì $XZ \rightarrow YZ$
- $X \rightarrow X$
- $X \rightarrow Y$, $W$ là bộ phận của $Y$, $XZ \rightarrow Y \setminus W$
- $X \rightarrow Y$, $Z \rightarrow W$ thì $XZ \rightarrow YW$
- $X \rightarrow Y$, $XZ \rightarrow Y$
- $X \rightarrow Y$ và $X \rightarrow Z$ thì $X \rightarrow YZ$
- $X \rightarrow YZ$ thì $X \rightarrow Y$ và $X \rightarrow Z$
- $X \rightarrow YZ$ và $Z \rightarrow AW$ thì $X \rightarrow YAW$

# Ví dụ

- Cho quan hệ R với tập phụ thuộc hàm F :

F= { $A \rightarrow B$
$BC \rightarrow D$
$D \rightarrow E$
$AC \rightarrow D$
$AC \rightarrow E$ }

| R | A | B | C | D | E |
|---|---|---|---|---|---|
| | a1 | b1 | c1 | d3 | e2 |
| | a1 | b1 | c3 | d4 | e3 |
| | a2 | b2 | c4 | d2 | e1 |
| | a3 | b1 | c1 | d3 | e2 |
| | a2 | b2 | c4 | d2 | e1 |

$AC \rightarrow BE$ ?
$AB \rightarrow E$ ?

# Luật suy diễn - Hệ tiên đề Armstrong

- Cho lược đồ quan hệ R(U), U={ A1, A2, …, An}, $X \neq \emptyset$ , $Y \neq \emptyset$, X, Y, Z, $W \subseteq U$
- Hệ tiên đề Armstrong gồm các luật sau:và $X \rightarrow Z$

| | | |
|---|---|---|
| – | *Phản xạ:* | Nếu $Y \subseteq X$ Thì $X \rightarrow Y$ |
| – | *Tăng trưởng:* | Nếu $X \rightarrow Y$ Thì $XZ \rightarrow YZ$ |
| – | *Bắc cầu:* | Nếu $X \rightarrow Y$ và $Y \rightarrow Z$ Thì $X \rightarrow Z$ |

- 3 luật trên có thể suy diễn ra các luật sau

| | | |
|---|---|---|
| – | *Hợp:* | Nếu $X \rightarrow Y$ và $X \rightarrow Z$, Thì $X \rightarrow YZ$ |
| – | *Giả bắc cầu:* | Nếu $X \rightarrow Y$ và $YZ \rightarrow W$, Thì $XZ \rightarrow W$ |
| – | *Phân rã:* | Nếu $X \rightarrow YZ$ , Thì $X \rightarrow Y$ và $X \rightarrow Z$ |

# Sử dụng hệ tiên đề Armstrong

- Sử dụng hệ tiên đề Armstrong để suy diễn một phụ thuộc hàm mới từ một tập các phụ thuộc hàm cho trước
- Ví dụ: Cho quan hệ R với tập PTH F như sau:

F= { A → C
B → H
C,G → H,I
A,B → I }

- Chứng minh rằng AG → I được suy diễn từ F
  - Ta có: A → C, C,G → H,I => AG → H, I (tựa bắc cầu)
  => AG → I (phân rã)

=> Vậy AG → I được suy diễn từ F

# Bao đóng (Closure)

- Bao đóng của tập các PTH
- Bao đóng của tập các thuộc tính

# Bao đóng của tập các PTH

- Cho lược đồ quan hệ R(U), U={ A1, A2, …, An}
- F là tập PTH trên R

Bao đóng của F, ký hiệu $F^+$ bao gồm:
- F
- Và các PTH được suy diễn từ F

- F gọi là đầy đủ nếu F = F+
- Trên thực tế, việc tính F+ khó thực hiện vì có thể dẫn đến sự bùng nổ tổ hợp

=> Thay vào đó, ta sẽ xét xem một PTH dạng X → Y có thuộc F+ hay không, nghĩa là X → Y được suy diễn từ F không ?

# Bao đóng của tập các PTH

- Ví dụ:
- Cho F= { $AB \rightarrow C$
  $BC \rightarrow D$
  $D \rightarrow E, G$ }
  - Chứng minh rằng $AB \rightarrow E \in F^+$

  - Ta có: $\left.\begin{array}{l} AB \rightarrow C \\ BC \rightarrow D \end{array}\right\}$ => $AB \rightarrow D$ (tựa bắc cầu)
    và ta có $D \rightarrow E,G$ => $D \rightarrow E$ (phân rã)
    => $AB \rightarrow E$ (bắc cầu)

  => Vậy $AB \rightarrow E \in F^+$ hay $AB \rightarrow E$ được suy diễn từ F

# Bao đóng của tập thuộc tính

- Cho lược đồ quan hệ R(U), U = {A1, A2, …, An}
- X là tập các thuộc tính trên U, F là tập các PTH trên R

- Bao đóng của tập thuộc tính X đối với F, ký hiệu $X^+$ bao gồm tập các thuộc tính PTH vào X, nghĩa là:
  - $X^+ = \{ A \in U \mid X \rightarrow A \in F^+ \}$

- Nhận xét:

– Làm thế nào biết được một PTH X → Y có được suy diễn từ F không ?

$$X \rightarrow Y \in F^+ \iff X^+ \supseteq Y$$

– Nếu $X^+ = U$ thì X là siêu khóa của R

# Thuật toán tìm $X^+$

- Dữ liệu vào: lược đồ quan hệ R với tập thuộc tính U, tập PTH F và $X \subseteq U$
- Dữ liệu ra : $X^+$
- Giải thuật:

– Bước 1: $X^+ = X$

– Bước 2: Nếu tồn tại $A \rightarrow B \in F$ và $A \subseteq X^+$ thì :

$$X^+ = X^+ \cup B$$

– Lặp lại bước 2 cho đến khi không thể thêm thuộc tính cho $X^+$ hoặc tất cả các PTH đã được xét.

– Bước 3 : Kết quả là $X^+$

# Thuật toán tìm $X^+$

Ví dụ: cho tập PTH

$$F = \{\ A \rightarrow C,\quad B \rightarrow H,\quad C,G \rightarrow H,I,\quad A,B \rightarrow I\ \}$$

- Tìm $(A)^+$ $(AB)^+$
- PTH $A \rightarrow H$ có suy diễn được từ F không ?
  Hay $A \rightarrow H$ có thuộc $F^+$ không ?
- PTH $AB \rightarrow CH$ có suy diễn được từ F không ?
  Hay $AB \rightarrow CH$ có thuộc $F^+$ không ?

- Tìm $(A)^+$
  - Bước 1: khởi tao

    $A^+ = A$
  - Bước 2: lặp
    - $A^+ = A + C = AC$ $(A \rightarrow C)$
    - $A^+ = AC + \varnothing$
  - B3: $A^+ = AC$

- Ta có $A^+ = AC$ không chứa H

  $\Rightarrow A \rightarrow H \notin F^+$

# Thuật toán tìm $X^+$

Ví dụ: cho tập PTH

$F = \{ A \rightarrow C$
$B \rightarrow H$
$C,G \rightarrow H,I$
$A,B \rightarrow I \}$

- Tìm $(A)^+$ $(AB)^+$
- PTH $A \rightarrow H$ có suy diễn được từ F không ? Hay $A \rightarrow H$ có thuộc $F^+$ không ?
- PTH $AB \rightarrow CH$ có suy diễn được từ F không ? Hay $AB \rightarrow CH$ có thuộc $F^+$ không ?

- Tìm $(AB)^+$
  - Bước 1

    $(AB)^+ = AB$
  - Bước 2 lặp
    - $(AB)^+ = AB + C + H + I$

      $= ABCHI$ ($A \rightarrow C$, $B \rightarrow H$, $AB \rightarrow I$ )
    - $AB^+ = ABCHI + \varnothing$
  - B3: $AB^+ = ABCHI$
- Tương tự ta có: $(AB)^+ = ABCHI$

Ta có: $(AB)^+ = ABCHI \supseteq CH \Rightarrow AB \rightarrow CH \in F^+$

# PHỤ THUỘC HÀM THỪA

- Định nghĩa: Cho s = (Ω,f)

  PTH X→ Y ϵ f là thừa khi và chỉ khi

  $$(f \setminus X \rightarrow Y)^{+} = f^{+}$$

- Thuật toán: Tìm L → R có thừa không?

1. Đặt ∑ = f
2. Gán ∑ = ∑ \ L → R, Nếu ∑ = Φ → kết thúc và L → R không thừa. Ngược lại,
3. Đặt T = **L**
4. Nếu tồn tại X → Y mà X thuộc T, gán T = TY. Nếu R thuộc T → L → R là thừa. Ngược lại, xét tiếp các PTH còn lại cho đến khi hết các PTH.

# PTH thừa

Ví dụ: F={ $A \rightarrow B$, $B \rightarrow C$, $D \rightarrow BE$, $A \rightarrow C$ }

PTH thừa ???

=> $A \rightarrow C$ vì PTH này được suy diễn từ hai PTH $A \rightarrow B$ và $B \rightarrow C$

# THUỘC TÍNH THỪA

Để phát hiện thuộc tính thừa ở vế trái trong PTH $L_1L_2\ldots L_i\ldots L_n \rightarrow R$, ta thực hiện như sau

Với mọi $L_i$ ở vế trái, nếu $(L_1L_2\ldots.L_n \backslash L_i)^+$ chứa R thì $L_i$ là thuộc tính thừa

# Phủ tối tiểu

Tập PTH tối tiểu: F được gọi là tối tiểu nếu F thỏa các điều kiện sau:

– Mọi PTH trong F chỉ có một thuộc tính ở vế phải

– Không tồn tại PTH thừa

– Không tồn tại PTH mà vế trái của nó có thuộc tính thừa

*=> Phủ tối tiểu của tập PTH F là tập F' tương đương với F*

*=> Mọi tập PTH đều có ít nhất một tập PTH tối tiểu*

# PTH rút gọn tự nhiên

- Tập PTH F được gọi là rút gọn tự nhiên nếu
  - Vế phải và vế trái của mọi PTH không có thuộc tính chung
    - Nếu có thuộc tính chung thì bỏ thuộc tính chung đó ở vế phải
  - Hai PTH khác nhau có vế trái giống nhau
    - Nếu có hai PTH cùng vế trái dạng $L_1 \rightarrow R_1$ và $L_1 \rightarrow R_2$ thì gom hai PTH đó lại thành $L_1 \rightarrow R_1R_2$
- Ví dụ: Tìm tập PTH rút gọn tự nhiên của F

F= {AD → CD

B → H

C,G → H

C,G → I

A,B → I }

=> $F_{TN}$= {AD → C

B → H

C,G → H,I

A,B → I }

# Bài toán tìm phủ tối tiểu

- Các bước tìm tập PTH tối tiểu của F:
  – Tách các PTH sao cho VP có 1 thuộc tính (dùng phân rã)
  – Loại bỏ các PTH thừa
  – Loại bỏ các thuộc tính thừa ở VT
  – Tìm tập PTH rút gọn tự nhiên của F

# Ví dụ Phủ tối tiểu

- Cho F={ A → B, B → C, D → BE, A → C}, tìm F tối tiểu ?

1) Tách các PTH sao cho VP có 1 thuộc tính
   - F={ A → B,
     B → C,
     D → B,
     D → E,
     A → C}

2) Loại bỏ PHT thừa
   - Ta có: A → B và B → C => A → C (thừa => loại PTH này)

3) Không có thuộc tính thừa ở VT

4) F rút gọn tự nhiên
   - F={ A → B,
     B → C,
     D → BE }

# Ví dụ

- Tìm tập PTH tối tiểu của:

  $F=\{A \rightarrow BC,$

  $B \rightarrow CE,$

  $A \rightarrow E,$

  $AC \rightarrow H,$

  $D \rightarrow B\ \}$

# Ví dụ

- Tìm tập PTH tối tiểu của:
  $F=\{A \to BC, B \to CE, A \to E, AC \to H, D \to B\}$

- B1: $F=\{A \to B,$
  $A \to C,$
  $B \to C,$
  $B \to E,$
  $A \to E,$
  $AC \to H,$
  $D \to B\}$

B2

$F=\{A \to B,$
$B \to C,$
$B \to E,$
$AC \to H,$
$D \to B\}$

Ta có: $A \to B, B \to C \Rightarrow A \to C$ (thừa)
$A \to B, B \to E \Rightarrow A \to E$ (thừa)

# Ví dụ

- $F=\{A\rightarrow BC,\ B \rightarrow CE,\ A \rightarrow E,\ AC \rightarrow H,\ D \rightarrow B\}$
- B3: $F=\{A\rightarrow B,$
  $B \rightarrow C,$
  $B \rightarrow E,$
  $AC \rightarrow H,$
  $D \rightarrow B\}$

$F=\{A\rightarrow B,\ B \rightarrow C,\ B \rightarrow E,\ A \rightarrow H,\ D \rightarrow B\}$

B4

$F=\{A\rightarrow BH,\ B \rightarrow CE,\ D \rightarrow B\}$ là tập tối tiểu

Ta có $AC \rightarrow H$ có vế trái có 2 thuộc tính

Do $A\rightarrow B$ và $B \rightarrow C \Rightarrow A\rightarrow C$

=> thuộc tính C trong $AC \rightarrow H$ thừa

# Bài tập

- **Bài 1**: Cho quan hệ sau KHACHHANG(id, ten, tpho) như sau
  - Các phát biểu sau đúng hay sai:

1) ***Id → ten***
2) ***id → tpho***
3) ***ten → id***
4) ***ten → tpho***
5) ***tpho → id***
6) ***id → tpho, ten***

| | | |
|---|---|---|
| 001 | Albert | Bruxelles |
| 002 | Francois | Liege |
| 003 | Brabo | Anvers |
| 004 | Albert | Anvers |
| 005 | Leon | Liege |
| 006 | Philippe | Bruxelles |
| 007 | Brabo | Anvers |

# Bài tập

- **Bài 2**: Cho R(A, B, C, D, E) và F= {AB → C, B → D, CD → E}
  - Tính $(BC)^+$
  - Chứng minh rằng $F^+$ chứa AB → E
  - Chứng minh rằng AB là khóa

# Bài tập

- **Bài 2**: Cho R(A, B, C, D, E) và F= {AB $\rightarrow$ C, B $\rightarrow$ D, CD $\rightarrow$ E}
  - Tính $(BC)^+$ = BC
    - $(BC)^+$ = BC +D=BCD (B $\rightarrow$ D)
    - $(BC)^+$ =BCD + E = BCDE (CD $\rightarrow$ E)

.$(BC)^+$ =BCDE + $\varnothing$ = BCDE

  - Chứng minh rằng $F^+$ chứa AB $\rightarrow$ E
    - (AB)+= ABCDE $\supseteq$ E => AB $\rightarrow$ E $\in$ $F^+$
  - Chứng minh rằng AB là khóa

Ta có (AB)+= ABCDE =U => AB là siêu khoá

- Mặt khác: A+=A => A không là SK

B+=BD => B không là SK

AB là Sk nhỏ nh

=> AB là khoá

# Bài tập

**Bài 3**: Cho quan hệ

•**KhamBenh**(idBN, tenBN, idBsi, tenBsi, ngaykham, loaibenh)

•Một bác sĩ có thể khám nhiều bệnh liên quan nhiều bệnh nhân

- Xác định tất cả các PTH của quan hệ KhamBenh biết rằng mỗi bệnh nhân có thể khám nhiều lần trong ngày nhưng không quá 1 lần với cùng bác sĩ.

# Bài tập

**Bài 4**:

Cho F = {AB → C, A → D, D → E, AC → B}

Hai PTH AB → E và D → C có được suy diễn từ F hay không?

# Cám ơn