

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA HỆ THỐNG THÔNG TIN**

Tài liệu bài giảng:

MÔN CƠ SỞ DỮ LIỆU

Chương 6:

PHỤ THUỘC HÀM VÀ DẠNG CHUẨN

ThS. Dương Phi Long – Email: longdp@uit.edu.vn

NỘI DUNG BÀI HỌC

01



Các vấn đề gặp phải khi tổ chức CSDL

02



Phụ thuộc hàm

03



Dạng chuẩn





Các vấn đề gặp phải khi tổ chức CSDL

1

Các vấn đề gặp phải khi tổ chức CSDL

VD1: SINHVIEN_DIEMTHI (MaSV, MaMH, HoTen, TenMH, Diem)

SINHVIEN_DIEMTHI				
MaSV	MaMH	HoTen	TenMH	Diem
SV01	CSDL	Nguyễn Tuyết An	Cơ sở dữ liệu	10
SV01	NMLT	Nguyễn Tuyết An	Nhập môn lập trình	9.5
SV01	HDT	Nguyễn Tuyết An	Hướng đối tượng	8.5
SV02	CSDL	Trần Ngọc Minh	Cơ sở dữ liệu	8
SV02	CTRR	Trần Ngọc Minh	Cấu trúc rời rạc	5
SV03	NMLT	Phạm Tiến Dũng	Nhập môn lập trình	7
SV03	CTRR	Phạm Tiến Dũng	Cấu trúc rời rạc	7.5



Dư thừa dữ liệu

Gây ra những bất thường, mâu thuẫn dữ liệu

Các vấn đề gặp phải khi tổ chức CSDL

VD1: SINHVIEN_DIEMTHI (MaSV, MaMH, HoTen, TenMH, Diem)

SINHVIEN_DIEMTHI				
MaSV	MaMH	HoTen	TenMH	Diem
SV01	CSDL	Nguyễn Tuyết An	Cơ sở dữ liệu	10
SV01	NMLT	Nguyễn Tuyết Anh	Nhập môn lập trình	9.5
SV01	HDT	Nguyễn Tuyết An	Hướng đối tượng	8.5
SV02	CSDL	Trần Ngọc Minh	Cơ sở dữ liệu	8
SV02	CTRR	Trần Ngọc Minh	Cấu trúc rời rạc	5
SV03	NMLT	Phạm Tiến Dũng	Nhập môn lập trình	7
SV03	CTRR	Phạm Tiến Dũng	Cấu trúc rời rạc	7.5

Bất thường, mâu thuẫn dữ liệu khi thực hiện thao tác **sửa** dữ liệu



Các vấn đề gặp phải khi tổ chức CSDL

VD1: SINHVIEN_DIEMTHI (MaSV, MaMH, HoTen, TenMH, Diem)

SINHVIEN_DIEMTHI				
MaSV	MaMH	HoTen	TenMH	Diem
SV01	CSDL	Nguyễn Tuyết An	Cơ sở dữ liệu	10
SV01	NMLT	Nguyễn Tuyết An	Nhập môn lập trình	9.5
SV01	HDT	Nguyễn Tuyết An	Hướng đối tượng	8.5
SV02	CSDL	Trần Ngọc Minh	Cơ sở dữ liệu	8
SV02	CTRR	Trần Ngọc Minh	Cấu trúc rời rạc	5
SV03	NMLT	Phạm Tiến Dũng	Nhập môn lập trình	7
SV03	CTRR	Phạm Tiến Dũng	Cấu trúc rời rạc	7.5
SV04	Null	Phan Minh Đức	Null	Null

Thao tác **không**
thực hiện được

Bất thường, mâu thuẫn dữ liệu khi thực hiện thao tác **thêm** dữ liệu

Các vấn đề gặp phải khi tổ chức CSDL

VD1: SINHVIEN_DIEMTHI (MaSV, MaMH, HoTen, TenMH, Diem)

SINHVIEN_DIEMTHI				
MaSV	MaMH	HoTen	TenMH	Diem
SV01	CSDL	Nguyễn Tuyết An	Cơ sở dữ liệu	10
SV01	NMLT	Nguyễn Tuyết An	Nhập môn lập trình	9.5
SV01	HDT	Nguyễn Tuyết An	Hướng đối tượng	8.5
SV02	CSDL	Trần Ngọc Minh	Cơ sở dữ liệu	8
SV02	CTRR	Trần Ngọc Minh	Cấu trúc rời rạc	5
SV03	NMLT	Phạm Tiến Dũng	Nhập môn lập trình	7
SV03	CTRR	Phạm Tiến Dũng	Cấu trúc rời rạc	7.5

Bất thường, mâu thuẫn dữ liệu khi thực hiện thao tác **xóa** dữ liệu

Các vấn đề gấp phải khi tổ chức CSDL

VD1: **SINHVIEN** (MaSV, HoTen)

MONHOC (MaMH, TenMH)

DIEMTHI (MaSV, MaMH, Diem)

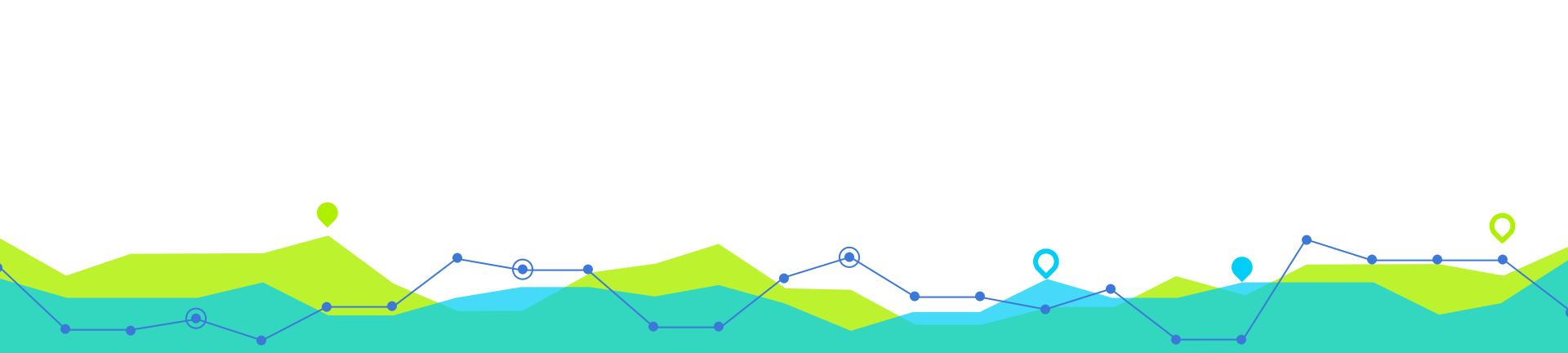
SINHVIEN	
<u>MaSV</u>	<u>HoTen</u>
SV01	Nguyễn Tuyết An
SV02	Trần Ngọc Minh
SV03	Phạm Tiến Dũng

MONHOC	
<u>MaMH</u>	<u>TenMH</u>
CSDL	Cơ sở dữ liệu
NMLT	Nhập môn lập trình
HDT	Hướng đối tượng
CTRR	Cấu trúc rời rạc

DIEMTHI		
<u>MaSV</u>	<u>MaMH</u>	<u>Diem</u>
SV01	CSDL	10
SV01	NMLT	9.5
SV01	HDT	8.5
SV02	CSDL	8
SV02	CTRR	5
SV03	NMLT	7
SV03	CTRR	7.5

2

Phụ thuộc hàm

- 
1. Các khái niệm cơ bản
 2. Hệ luật dẫn Armstrong
 3. Bao đóng
 4. Phủ tối thiểu
 5. Khóa

1. Các khái niệm cơ bản

- Phụ thuộc hàm (PTH) trên quan hệ R biểu diễn *mối liên hệ giữa các tập thuộc tính trong R*
- Ký hiệu: $X \rightarrow Y$
- Định nghĩa:
 - Quan hệ R (A_1, A_2, \dots, A_n), $R^+ = \{A_1, A_2, \dots, A_n\}$
 - $X, Y \subset R^+$
 - $t1, t2$: 2 bộ bất kỳ trên R

Ta nói: $X \rightarrow Y \Leftrightarrow (t1.X = t2.X \Rightarrow t1.Y = t2.Y)$



1. Các khái niệm cơ bản

- Nghĩa là: với 1 giá trị của X thì có một giá trị duy nhất của Y
 - X xác định Y
 - Y phụ thuộc (hàm) vào X
- X là vế trái của PTH, Y là vế phải của PTH
- Có nhiều PTH trên 1 quan hệ, Tập PTH được ký hiệu là F

VD2:

- $MaNV \rightarrow TenNV$
- $MaNV, MaDA \rightarrow ThoiGian$

1. Các khái niệm cơ bản

- **VD3: CTHD** (*SoHD, MaSP, SL, DonGia, ThanhTien*)

Xác định các PTH trong CTHD

CTHD				
SoHD	MaSP	SL	DonGia	ThanhTien
HD01	SP01	5	2.000	10.000
HD01	SP03	2	10.000	20.000
HD02	SP01	5	2.000	10.000
HD02	SP04	2	3.000	6.000
HD03	SP02	4	10.000	40.000
HD03	SP03	4	12.500	50.000
HD03	SP04	8	2.500	20.000

- ✗ 1) $\text{SoHD} \rightarrow \text{MaSP}$
- ✗ 2) $\text{SoHD} \rightarrow \text{SL}$
- ✗ 3) $\text{MaSP} \rightarrow \text{DonGia}$
- ✓ 4) $\text{SoHD}, \text{MaSP} \rightarrow \text{SL}$
- ✓ 5) $\text{SoHD}, \text{MaSP} \rightarrow \text{DonGia}$
- ✗ 6) $\text{SL} \rightarrow \text{ThanhTien}$
- ✗ 7) $\text{DonGia} \rightarrow \text{ThanhTien}$
- ✓ 8) $\text{SL}, \text{DonGia} \rightarrow \text{ThanhTien}$
- ✗ 9) $\text{SoHD}, \text{MaSP} \rightarrow \text{SL}, \text{DonGia}, \text{ThanhTien}$



2. Hệ luật dẫn Armstrong

- Quan hệ $R(A, B, C)$ và F là tập PTH
- $X \rightarrow Y$ được suy ra từ F nếu bất kỳ bộ của quan hệ R thỏa F thì cũng thỏa $X \rightarrow Y$
- Ký hiệu: $F \models X \rightarrow Y$

2. Hệ luât dẫn Armstrong

- Với $X, Y, Z, W \subseteq R^+$. Phụ thuộc hàm có các tính chất sau:
 - F1. **Tính phản xạ** (reflexivity):

Nếu $Y \subseteq X$ thì $X \rightarrow Y$ (PTH hiển nhiên)

VD4: MaSV, TenSV \rightarrow TenSV

2. Hệ luật dẫn Armstrong

- Với $X, Y, Z, W \subseteq R^+$. Phụ thuộc hàm có các tính chất sau:
 - F2. **Tính tăng trưởng** (augmentation):

Nếu $X \rightarrow Y$, thì $XZ \rightarrow YZ$

VD5: $MaSV \rightarrow TenSV$

$\Rightarrow MaSV, NgaySinh \rightarrow TenSV, NgaySinh$



2. Hệ luật dẫn Armstrong

- Với $X, Y, Z, W \subseteq R^+$. Phụ thuộc hàm có các tính chất sau:
 - F3. **Tính bắc cầu** (transitivity):

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

VD6: $MaSV \rightarrow MaLop$

$MaLop \rightarrow TenLop$



$\Rightarrow MaSV \rightarrow TenLop$



2. Hệ luật dẫn Armstrong

- Một số tính chất bổ sung từ Hệ luật dẫn Armstrong:
 - F4. **Tính kết hợp** (*union*):

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

VD7: $MaSV \rightarrow TenSV$
 $MaSV \rightarrow GioiTinh$

$\Rightarrow MaSV \rightarrow TenSV, GioiTinh$



2. Hệ luật dẫn Armstrong

- Một số tính chất bổ sung từ Hệ luật dẫn Armstrong:
 - F5. **Tính phân rã** (decomposition):

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

VD8: $MaSV \rightarrow TenSV, GioiTinh$

$\Rightarrow MaSV \rightarrow TenSV$ và $MaSV \rightarrow GioiTinh$

2. Hệ luật dẫn Armstrong

- Một số tính chất bổ sung từ Hệ luật dẫn Armstrong:
 - F6. **Tính tựa bắc cầu** (psuedotransitivity):

Nếu $X \rightarrow Y$ và $YZ \rightarrow W$, thì $XZ \rightarrow W$

VD9: $MaSV \rightarrow MaLop$

$MaLop, MaMon \rightarrow MaGV$



$\Rightarrow MaSV, MaMon \rightarrow MaGV$

2. Hệ luât dẫn Armstrong

- **VD10:** Cho $R(A, B, C, D)$ và $F = \{A \rightarrow B, A \rightarrow C, BC \rightarrow D\}$.

Chứng minh $A \rightarrow D$ được suy diễn từ F

1. $A \rightarrow B$ (giả thiết)
2. $A \rightarrow C$ (giả thiết)
3. $A \rightarrow BC$ (tính kết hợp 1 và 2)
4. $BC \rightarrow D$ (giả thiết)
5. $A \rightarrow D$ (tính bắc cầu 3 và 4)

2. Hệ luât dẫn Armstrong

- **VD11:** Cho $R(A, B, C, D, E)$ và $F = \{AB \rightarrow D, C \rightarrow A, B \rightarrow E\}$.

Chứng minh: $BC \rightarrow DE$ được suy diễn từ F

1. $C \rightarrow A$ (giả thiết)
2. $AB \rightarrow D$ (giả thiết)
3. $BC \rightarrow D$ (tựa bắc cầu 1 và 2)
4. $B \rightarrow E$ (giả thiết)
5. $BC \rightarrow EC$ (tăng trưởng 4)
6. $BC \rightarrow E$ (phân rã 5)
7. $BC \rightarrow DE$ (kết hợp 3 và 6)

2. Hệ luât dẫn Armstrong

- **VD12:** Cho $R(A, B, C, D, E, G, H)$

$$F = \{AB \rightarrow C, B \rightarrow D, CD \rightarrow E, CE \rightarrow GH, G \rightarrow A\}.$$

Chứng minh: $AB \rightarrow E$ được suy diễn từ F

1. $AB \rightarrow C$ (giả thiết)
2. $AB \rightarrow B$ (phản xạ)
3. $B \rightarrow D$ (giả thiết)
4. $AB \rightarrow D$ (bắc cầu 2 và 3)
5. $AB \rightarrow CD$ (kết hợp 1 và 4)
6. $CD \rightarrow E$ (giả thiết)
7. $AB \rightarrow E$ (bắc cầu 5 và 6)

3. Bao đóng

- **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 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^+

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

3. Bao đóng

Tìm bao đóng của tập thuộc tính X đối với tập phụ thuộc hàm F:

- **Input:** (R, F) , $X \subseteq R^+$
- **Output:** X_F^+
 - **Bước 1:** Tính dãy $X^{(0)}, X^{(1)}, \dots, X^{(i)}$:
 - $X^{(0)} = X$
 - $X^{(i+1)} = X^{(i)} \cup Z, \exists (Y \rightarrow Z) \in F (Y \subseteq X^{(i)})$, loại $(Y \rightarrow Z)$ ra khỏi F
 - Dừng khi $X^{(i+1)} = X^{(i)}$ hoặc khi $X^{(i)} = R^+$
 - **Bước 2:** Kết luận $X_F^+ = X^{(i)}$



3. Bao đóng

VD13: Cho $R(A,B,C,D,E,G,H)$ 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:

- $X^0 = AC$
 - Từ $f1$ đến $f4$ không thoả,
 - $f5$ thoả nên $X^1 = AC \cup D = ACD$



3. Bao đóng

VD13: Cho $R(A,B,C,D,E,G,H)$ 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 (tt):

- Lặp lại bước 1, $X^1 = ACD$
 - $f1$ không thỏa
 - $f2$ thỏa nên $X^1 = ACD \cup E = ACDE$
 - $f3$ thỏa nên $X^1 = ACDE \cup H = ACDEH$
 - $f4$ không thỏa, $f5$ đã thỏa

3. Bao đóng

VD13: Cho $R(A,B,C,D,E,G,H)$ 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 (tt):

- Lặp lại bước 1, $X^2 = ACDEH$
 - $f2, f3, f5$ đã thỏa
 - $f1$ không thỏa, $f4$ không thỏa nên $X^3 = X^2 = ACDEH$

Bước 2: Vậy $AC_F^+ = ACDEH$



3. Bao đóng

- Bài toán thành viên

- Cho quan hệ R , tập phụ thuộc hàm F trên R và một phụ thuộc hàm $X \rightarrow Y$ trên R .
- Câu hỏi đặt ra: $X \rightarrow Y \in F^+$ hay không?

=> Giải quyết: $X \rightarrow Y \in F^+ \Leftrightarrow Y \subseteq X^+$

3. Bao đóng

VD14: Cho $R(A,B,C,D,E,G,H)$ và tập PTH

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

Cho biết $AC \rightarrow E \in F^+$?

Từ kết quả ở VD13, ta có $AC_F^+ = ACDEH$

Vì $E \in AC_F^+$, nên $AC \rightarrow E \in F^+$

4. Phủ tối thiểu

- Các khái niệm:

- Hai tập PTH tương đương
- PTH có thuộc tính về trái dư thừa
- PTH có về phải một thuộc tính
- PTH không dư thừa
- Phủ tối thiểu

- Các thuật toán:

- Tìm PTH đầy đủ (PTH có thuộc tính về trái không dư thừa)
- Tìm PTH loại những PTH dư thừa
- Tìm Phủ tối thiểu



4. Phủ tối thiểu

- **Hai tập PHT tương đương:**
 - Hai tập PTH F và G tương đương nếu $F^+ = G^+$
 - Ký hiệu $G \equiv F$



4. Phủ tối thiểu

- PTH có thuộc tính về trái dư thừa

- Cho F là tập PTH trên R , $X \rightarrow Y \in F$ là PTH có thuộc tính về trái dư thừa nếu:

$$\exists A \in m \text{ mà } F = F - (X \rightarrow Y) \cup ((X - A) \rightarrow Y)$$

- Ngược lại $Z \rightarrow Y$ là PTH có thuộc tính về trái không dư thừa hay Y phụ thuộc đầy đủ vào Z .

VD15: $R(A, B, C, D)$ và $F=\{A \rightarrow B, BC \rightarrow D, C \rightarrow D\}$

Khi đó $BC \rightarrow D$ là PTH có thuộc tính về trái dư thừa

4. Phủ tối thiểu

- **Thuật toán tìm PTH đầy đủ (PTH có vế trái không dư thừa)**
 - Với mỗi PTH $X \rightarrow Y$, $X = A_1A_2 \dots A_n$ và $n \geq 2$, đặt $Z = X$
 - Với mỗi A_i , thực hiện:
 - $Tam = Z \setminus A_i$
 - Nếu $Tam \rightarrow Y \in (F - \{X \rightarrow Y\})^+$ thì $Z = Tam$



4. Phủ tối thiểu

- **Thuật toán tìm PTH đầy đủ (PTH có vế trái không dư thừa)**

VD16: $R(A, B, C, D)$ và $F = \{A \rightarrow B, BC \rightarrow D, C \rightarrow D\}$

- Xét $BC \rightarrow D$:
 - có $C_{F-\{BC \rightarrow D\}}^+ = CD$
 - Vì $D \subseteq C_F^+$ nên B là thuộc tính dư thừa \rightarrow loại bỏ B
- Vậy $F = \{A \rightarrow B, C \rightarrow D\}$

4. 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.



4. Phủ tối thiểu

- PTH không dư thừa

- *F là tập PTH không dư thừa nếu không tồn tại $F' \subset F$ sao cho $F' \equiv F$.*
- *Ngược lại F được gọi là tập PTH dư thừa.*

- Thuật toán loại những PTH dư thừa

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



4. Phủ tối thiểu

- Phủ tối thiểu (PTT) của một tập phụ thuộc hàm

F được gọi là PTT của tập PTH (hay tập PTH tối thiểu) nếu thỏa:

- (i) F có thuộc tính vẽ trái không dư thừa
- (ii) F có vẽ phải một thuộc tính
- (iii) F không dư thừa



4. Phủ tối thiểu

- **Thuật toán tìm Phủ tối thiểu (PTT) của một tập phụ thuộc hàm**
 - **Bước 1:** Phân rã các PTH có vế phải nhiều thuộc tính thành các PTH có vế phải một thuộc tính
 - **Bước 2:** Loại bỏ các thuộc tính vế trái dư thừa của các PTH
 - **Bước 3:** Loại bỏ các PTH dư thừa



4. Phủ tối thiểu

- **VD17:** Cho $R(A,B,C,D)$ và $F = \{AB \rightarrow CD, B \rightarrow C, C \rightarrow D\}$
Tìm phủ tối thiểu?

Bước 1: Phân rã các PTH sao cho vế phải có một thuộc tính

Ta có $F = \{f1: AB \rightarrow C, f2: AB \rightarrow D, f3: B \rightarrow C, f4: C \rightarrow D\}$

Bước 2: Loại bỏ các thuộc tính vế trái dư thừa của các PTH

- $f3: B \rightarrow C, f4: C \rightarrow D$: Không xét vì vế trái chỉ có một thuộc tính.



4. Phủ tối thiểu

- **VD17:** Cho $R(A,B,C,D)$ và $F = \{AB \rightarrow CD, B \rightarrow C, C \rightarrow D\}$
Tìm phủ tối thiểu?

Bước 2: Loại bỏ các thuộc tính về trái dư thừa của các PTH (if)

- Xét $f_1: AB \rightarrow C$:
 - Nếu bỏ A thì $B_{F-\{f_1\}}^+ = BCD$ chứa C nên bỏ A .
 - Nếu bỏ B thì $A_{F-\{f_1\}}^+ = A$ không chứa C , không bỏ B .
- ⇒ Viết lại $f_1: B \rightarrow C$



4. Phủ tối thiểu

- **VD17:** Cho $R(A,B,C,D)$ và $F = \{AB \rightarrow CD, B \rightarrow C, C \rightarrow D\}$
Tìm phủ tối thiểu?

Bước 2: Loại bỏ các thuộc tính về trái dư thừa của các PTH (if)

- Xét $f_2: AB \rightarrow D$:
 - Nếu bỏ A thì $B_{F-\{f_2\}}^+ = BCD$ chứa D , nên bỏ A .
 - Nếu bỏ B thì $A_{F-\{f_2\}}^+ = A$ không chứa D , không bỏ B .
- ⇒ Viết lại $f_2: B \rightarrow D$

⇒ Ta có $\textcolor{red}{F = \{B \rightarrow C, C \rightarrow D\}}$

4. Phủ tối thiểu

- **VD17:** Cho $R(A,B,C,D)$ và $F = \{AB \rightarrow CD, B \rightarrow C, C \rightarrow D\}$
Tìm phủ tối thiểu?

Bước 3: Loại khỏi F các PTH dư thừa

- $f1: B \rightarrow C$: ta có $B_{F-\{f1\}}^+ = B$, $f1$ không dư thừa.
- $f2: C \rightarrow D$: ta có $C_{F-\{f2\}}^+ = D$, $f2$ không dư thừa.

Vậy Phủ tối thiểu là $F = \{B \rightarrow C, C \rightarrow D\}$



4. Phủ tối thiểu

- **VD18:** Cho $R(A, B, C, D, E, G, H)$

và $F = \{B \rightarrow A, A \rightarrow BC, AB \rightarrow G, GH \rightarrow E, BCG \rightarrow A\}$.

Tìm phủ tối thiểu?



5. Khóa

- **Định nghĩa:** Cho $R (A_1, A_2, \dots, A_n)$, $R^+ = \{A_1, A_2, \dots, A_n\}$, F là tập phụ thuộc hàm trên R , K là tập con của R^+ . Khi đó K gọi là một khóa của R nếu:
 - (i) $K_F^+ = R^+$
 - (ii) Không tồn tại $K' \subset K$ sao cho $K'^+_F = R^+$
- Thuộc tính A được gọi là thuộc tính khóa nếu $A \in K$, trong đó K là khóa của R . Ngược lại thuộc tính A được gọi là thuộc tính không khóa.
- K'' được gọi là siêu khóa nếu $K \subseteq K''$.

5. Khóa

b1: Tập nguồn, tập đích, tập trung gian, tập treo

- Thuật toán tìm tất cả Khóa

• **Bước 1:**

- Xác định *tập thuộc tính nguồn* (ký hiệu N), **chứa những thuộc tính chỉ xuất hiện ở vế trái của các PTH.**

- *Tính N_F^+ :*
❖ Nếu $N_F^+ = R^+$: Kết luận Khóa là N
❖ Ngược lại: Tiếp tục đến bước 2

tập đích là những thuộc tính chỉ xuất hiện ở vế phải của phụ thuộc hàm

tập treo là tập các thuộc tính không xuất hiện ở vế nào



5. Khóa

- Thuật toán tìm tất cả Khóa (tt)

- **Bước 2:**

- Xác định *tập thuộc tính trung gian* (ký hiệu TG), chứa những thuộc tính xuất hiện ở cả vế trái và vế phải của các PTH
- Xác định các tập con X_i có thể có của tập TG

- **Bước 3:** $\forall X_i \subseteq TG$, nếu $(N \cup X_i)_F^+ = R^+$

Thì $S_i = N \cup X_i$, loại bỏ các tập X_j : $X_i \subset X_j$

- **Bước 4:** Kết luận tập các khóa $K = \{S_i\}$

5. Khóa

- **VD19:** Cho $R(A, B, C, D, E, G, H)$

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

Tìm Khóa?

Bước 1: Tập nguồn $N = \{B, G\}$

Ta có $BG_F^+ = BGA \neq R^+ \Rightarrow BG$ không là Khóa

Bước 2: Tập trung gian $TG = \{A, C, D, H\}$

⇒ Các tập con khác rỗng của TG là $\{A\}, \{C\}, \{D\}, \{H\}, \{AC\}, \{AD\}, \{AH\}, \{CD\}, \{CH\}, \{DH\}, \{ACD\}, \{ACH\}, \{ADH\}, \{CDH\}, \{ACDH\}$

5. Khóa

- **VD19:** Cho $R(A, B, C, D, E, G, H)$

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

Bước 3:

$N \cup X_i$	$(N \cup X_i)_F^+$	$(N \cup X_i)_F^+ = R^+$	Kết luận
BGA	BGA	Sai	
BGC	$BGCADEH$	Đúng	BGC là 1 khóa, loại xét các tập con chứa C : $AC, CD, CH, ACD, ACH, CDH, ACDH$
BGD	$BGDACEH$	Đúng	BGD là 1 khóa, loại xét các tập con chứa D : AD, DH, ADH
BGH	$BGHACDE$	Đúng	BGH là 1 khóa, loại xét các tập con chứa H : AH

⇒ **Khóa của R là BGC, BGD, BGH**

2

Dạng chuẩn



1. Dạng chuẩn 1
2. Dạng chuẩn 2
3. Dạng chuẩn 3
4. Dạng chuẩn Boyce Codd
5. Dạng chuẩn của lược đồ
quan hệ, lược đồ CSDL

Dạng chuẩn

- Được sử dụng để **chuẩn hóa quan hệ**, đáp ứng các mục tiêu thiết kế:
 - Giảm tối đa trùng lắp thông tin
 - Kiểm tra RBTV dễ dàng
- Đánh giá chất lượng thiết kế của lược đồ CSDL
 - E.F.Codd đưa ra 3 dạng chuẩn (Normal Form)
 - R.F.Boyce và E.F.Codd cải tiến dạng chuẩn gọi dạng chuẩn Boyce-Codd (BC)
- Các dạng chuẩn được định nghĩa **dựa trên khái niệm PTH**



1. Dạng chuẩn 1

- Lược đồ R đạt dạng chuẩn 1 (DC1) nếu *tất cả* 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.
- **VD20:** DiaChi: Số 175 Đường 3/2 Phường 10 Quận 5 không là nguyên tố.



1. Dạng chuẩn 1

- **VD21:** HOADON(MaHD, MaKH, NgayHD, CTietMua, SoTien)

MaHD	MaKH	NgayHD	CTietMua			SoTien
			Tên hàng	Số lượng	ĐVT	
HD01	KH01	15-10-05	Bánh Orion	1	Gói	25.000
			Kẹo mút	2	Cây	2.000
HD02	KH01	18-10-05	Gạo	2	Kg	30.000
HD03	KH02	24-10-05	Đường	1	Kg	15.000
			Bánh AFC	2	Gói	24.000

⇒ CTietMua không là nguyên tố, nên LDQH không thỏa DC1

1. Dạng chuẩn 1

- **VD22:** *THAMGIA (MaNV, HoTen, NgSinh, MaDA, TenDA, ThoiGian)*

THAMGIA					
<u>MaNV</u>	<u>HoTen</u>	<u>NgSinh</u>	<u>MaDA</u>	<u>TenDA</u>	<u>ThoiGian</u>
NV01	Nguyễn Tuyết An	01/10/1978	DA01	Dự án 01	20
			DA02	Dự án 02	15
			DA04	Dự án 04	10
NV02	Trần Ngọc Minh	25/01/1987	DA03	Dự án 03	30
NV03	Phạm Tiến Dũng	12/12/1882	DA02	Dự án 02	20
			DA04	Dự án 04	20

⇒ *THAMGIA không đạt DC1*

1. Dạng chuẩn 1

- **VD23:** *THAMGIA (MaNV, HoTen, NgSinh, MaDA, TenDA, ThoiGian)*

THAMGIA					
<u>MaNV</u>	<u>HoTen</u>	<u>NgSinh</u>	<u>MaDA</u>	<u>TenDA</u>	<u>ThoiGian</u>
NV01	Nguyễn Tuyết An	01/10/1978	DA01	Dự án 01	20
NV01	Nguyễn Tuyết An	01/10/1978	DA02	Dự án 02	15
NV01	Nguyễn Tuyết An	01/10/1978	DA04	Dự án 04	10
NV02	Trần Ngọc Minh	25/01/1987	DA03	Dự án 03	30
NV03	Phạm Tiến Dũng	12/12/1882	DA02	Dự án 02	20
NV03	Phạm Tiến Dũng	12/12/1882	DA04	Dự án 04	20

⇒ THAMGIA đạt DC1 nhưng còn trùng lặp thông tin

2. Dạng chuẩn 2

- Lược đồ R đạt dạng chuẩn 2 (DC2) nếu thỏa:
 - R đạt DC1 và
 - Các thuộc tính không khóa đều **phụ thuộc đầy đủ** vào khóa
- **Kiểm tra DC2:**
 - Tìm tất cả khóa của R
 - Với mỗi khóa K , tìm S_i^+ với S_i là tất cả các tập con thực sự của K
 - Nếu tồn tại S_i^+ chứa thuộc tính không khóa thì R không đạt DC2, ngược lại Q đạt DC 2.

2. Dạng chuẩn 2

- **VD24:** Cho $R1 (A, B, C, D)$

Tập PTH $F = \{A \rightarrow B, B \rightarrow DC\}$

Kiểm tra $R1$ có đạt DC2 không?

- Tìm khóa của $R1$

Tập nguồn $N = A$

Ta có: $A_F^+ = ABCD$

Lược đồ có khóa là A



2. Dạng chuẩn 2

- **VD24:** Cho R1 (A, B, C, D)

Tập PTH $F = \{A \rightarrow B, B \rightarrow DC\}$

Kiểm tra R1 có đạt DC2 không?

- Lược đồ chỉ có một khóa là A và khóa có duy nhất 1 thuộc tính, nên mọi thuộc tính không khóa đều phụ thuộc đầy đủ vào khóa.
- Do vậy R1 đạt DC 2.

2. Dạng chuẩn 2

- **VD25:** Cho R2 (A, B, C, D)

Tập PTH $F = \{AB \rightarrow D, C \rightarrow D\}$

Kiểm tra R2 có đạt DC2 không?

- Tìm khóa của R2

Tập nguồn $N = ABC$

Ta có: $ABC_F^+ = ABCD$

Lược đồ có khóa là ABC



2. Dạng chuẩn 2

- **VD25:** Cho R2 (A, B, C, D)

Tập PTH $F = \{AB \rightarrow D, C \rightarrow D\}$

Kiểm tra R2 có đạt DC2 không?

- Ta thấy $C \subset ABC$ và $C \rightarrow D$, D là thuộc tính không khóa, nên D không phụ thuộc đầy đủ vào khóa.
- Do vậy R2 không đạt DC2.

2. Dạng chuẩn 2

- **VD26:** Cho SINHVIEN (MSSV, MaMH, TenSV, DiaChi, Diem)

Tập PTH $F = \{MSSV, MaMH \rightarrow Diem, MSSV \rightarrow TenSV, DiaChi\}$

Kiểm tra SINHVIEN có đạt DC2 không?

- Tìm khóa của SINHVIEN

Tập nguồn $N = \{MSSV, MaMH\}$

Ta có: $\{MSSV, MaMH\}_F^+ = \{MSSV, MaMH, TenSV, DiaChi, Diem\}$

Lược đồ có khóa là $K = \{MSSV, MaMH\}$



2. Dạng chuẩn 2

- **VD26:** Cho *SINHVIEN* (*MSSV*, *MaMH*, *TenSV*, *DiaChi*, *Diem*)

Tập PTH $F = \{MSSV, MaMH \rightarrow Diem, MSSV \rightarrow TenSV, DiaChi\}$

Kiểm tra *SINHVIEN* có đạt DC2 không?

- Ta thấy $MSSV \subset K$ và $MSSV \rightarrow TenSV, DiaChi$. $TenSV, DiaChi$ là thuộc tính không khóa, nên không phụ thuộc đầy đủ vào khóa.
 - Do vậy *SINHVIEN* không đạt DC2.
- ⇒ Tách thành 2 lược đồ, đạt DC2
- *DANGKY*(*MSSV*, *MaMH*, *Diem*); $F1 = \{MSSV, MaMH \rightarrow Diem\}$
 - *SINHVIEN* (*MSSV*, *TenSV*, *DiaChi*); $F2 = \{MSSV \rightarrow TenSV, DiaChi\}$

2. Dạng chuẩn 2

- **VD27:** Cho NHANVIEN (*MaNV, TenNV, NgSinh, SDT, MaPB, TenPB, TrgPB*)

Tập PTH $F = \{MaNV \rightarrow TenNV, NgSinh, SDT, MaPB; MaPB \rightarrow TenPB, TrgPB\}$

Kiểm tra NHANVIEN có đạt DC2 không?

- Tìm khóa của NHANVIEN

Tập nguồn $N = \{MaNV\}$

Ta có: $MaNV_F^+ = \{MaNV, TenNV, NgSinh, SDT, MaPB, TenPB, TrgPB\}$

Lược đồ có khóa là MaNV



2. Dạng chuẩn 2

- **VD27:** Cho NHANVIEN (*MaNV, TenNV, NgSinh, SDT, MaPB, TenPB, TrgPB*)

Tập PTH $F = \{MaNV \rightarrow TenNV, NgSinh, SDT, MaPB; MaPB \rightarrow TenPB, TrgPB\}$

Kiểm tra NHANVIEN có đạt DC2 không?

- Lược đồ có khóa là *MaNV*, duy nhất 1 thuộc tính, nên các thuộc không khóa đều phụ thuộc đầy đủ vào khóa.
- Do vậy NHANVIEN đạt DC2 nhưng vẫn còn trùng lặp dữ liệu

2. Dạng chuẩn 2

- Nhận xét:

- *Lược đồ đạt DC2 thì cũng đạt DC1*
- *Nếu lược đồ chỉ có 1 khóa và khóa gồm 1 thuộc tính => Đạt DC2*
- *DC2 còn trùng lặp dữ liệu*



2. Dạng chuẩn 3

- *Định nghĩa 1: Lược đồ R đạt dạng chuẩn 3 (DC3) nếu thỏa:*
 - R đạt DC2 và
 - Các thuộc tính không khóa không phụ thuộc bắc cầu vào khóa
- **Phụ thuộc bắc cầu:** Thuộc tính $A \in R^+$ được gọi là phụ thuộc bắc cầu vào tập thuộc tính X nếu $\exists Y \in R^+$:
 - 1) $X \rightarrow Y \in F^+$ và $Y \rightarrow A \in F^+$
 - 2) $Y \rightarrow X \notin F^+$
 - 3) $A \notin (X \cup Y)$



2. Dạng chuẩn 3

- *Định nghĩa 2: Lược đồ R đạt DC3 nếu tất cả các phụ thuộc hàm $X \rightarrow Y \in F$, với $Y \notin X$ đều có:*
 - *X là siêu khóa, hoặc*
 - *Y là thuộc tính khóa*



3. Dạng chuẩn 3

- Kiểm tra DC3:

- Tìm tất cả khóa của R
 - Phân rã vế phải của các PTH trong F thành các PTH có vế phải là thuộc tính
 - Nếu mọi phụ thuộc hàm $X \rightarrow Y \in F$, và $Y \notin X$ đều thỏa:
 - X là siêu khóa (vế trái chứa một khóa), hoặc
 - Y là thuộc tính khóa (vế phải là tập con của khóa)
- ⇒ *Thì R đạt DC 3. Ngược lại R không đạt DC 3.*

3. Dạng chuẩn 3

- **VD28:** Cho $R1 (A, B, C, D)$

Tập PTH $F = \{AB \rightarrow D, C \rightarrow D\}$

Kiểm tra $R1$ có đạt DC3 không?

- Tìm khóa của $R1$

Tập nguồn $N = \{A, B, C\}$

Ta có: $ABC_F^+ = ABCD$

Lược đồ có khóa là ABC



3. Dạng chuẩn 3

- **VD28:** Cho R1 (A, B, C, D)

Tập PTH $F = \{AB \rightarrow D, C \rightarrow D\}$

Kiểm tra R1 có đạt DC3 không?

- Mọi phụ thuộc hàm trong F đều đã có vế phải thuộc tính.
 - Xét $AB \rightarrow D$, có
 - Vế trái (AB) không phải là siêu khóa, và
 - Vế phải (D) không là thuộc tính khóa
- ⇒ Do vậy R1 không đạt DC3.

3. Dạng chuẩn 3

- **VD29:** Cho NHANVIEN (*MaNV, TenNV, NgSinh, SDT, MaPB, TenPB, TrgPB*)

Tập PTH $F = \{MaNV \rightarrow TenNV, NgSinh, SDT, MaPB; MaPB \rightarrow TenPB, TrgPB\}$

Kiểm tra NHANVIEN có đạt DC3 không?

- Tìm khóa của NHANVIEN

Tập nguồn $N = \{MaNV\}$

Ta có: $MaNV_F^+ = \{MaNV, TenNV, NgSinh, SDT, MaPB, TenPB, TrgPB\}$

Lược đồ có khóa là MaNV

3. Dạng chuẩn 3

- **VD29:** Cho NHANVIEN (*MaNV, TenNV, NgSinh, SDT, MaPB, TenPB, TrgPB*)

Tập PTH $F = \{MaNV \rightarrow TenNV, NgSinh, SDT, MaPB; MaPB \rightarrow TenPB, TrgPB\}$

Kiểm tra NHANVIEN có đạt DC3 không?

- Phân rã vế phải của các PTH trong F , ta có:

$F = MaNV \rightarrow TenNV; MaNV \rightarrow NgSinh; MaNV \rightarrow SDT; MaNV \rightarrow MaPB; MaPB \rightarrow TenPB; MaPB \rightarrow TrgPB\}$

- Xét $MaPB \rightarrow TenPB$ có vế trái không là siêu khóa và vế phải không là thuộc tính khóa

⇒ NHANVIEN Không đạt DC3. Thuộc tính không khóa *TenPB* phụ thuộc bắc cầu vào khóa chính *MaNV*, gây trùng lặp dữ liệu.

3. Dạng chuẩn 3

- Nhận xét:

- *Lược đồ đạt DC3 thì cũng đạt DC2*
- *Phụ thuộc bắc cầu gây nên trùng lắp dữ liệu*
- *Khi thiết kế CSDL, yêu cầu tối thiểu đạt DC3*



4. Dạng chuẩn Boyce Codd

- Lược đồ R đạt dạng chuẩn Boyce Codd (DC BC) nếu tất cả các phụ thuộc hàm $X \rightarrow Y \in F$, với $Y \notin X$ đều có X là siêu khóa.



4. Dạng chuẩn Boyce Codd

- Kiểm tra DCBC:

- Tìm tất cả khóa của R
 - Phân rã vế phải của các PTH trong F thành các PTH có vế phải một thuộc tính
 - Nếu mọi phụ thuộc hàm $X \rightarrow Y \in F$, và $Y \notin X$ đều thỏa: X là siêu khóa (vế trái chứa một khóa)
- ⇒ *Thì R đạt DC BC. Ngược lại R không đạt DC BC.*

4. Dạng chuẩn Boyce Codd

- **VD30:** Cho $R (A, B, C, D, E, I)$

Tập PTH $F = \{ACD \rightarrow EBI, CE \rightarrow AD\}$

Kiểm tra R có đạt DC BC không?

- Tìm khóa của R :

R có 2 khóa là $\{ACD, CE\}$

- Phân rã vế phải của các PTH trong F , ta có:

$F = \{ACD \rightarrow E, ACD \rightarrow B, ACD \rightarrow I, CE \rightarrow A, CE \rightarrow D\}$

- Tất cả PTH trong F đều có vế trái là một siêu khóa

⇒ Do vậy R đạt DC BC.

4. Dạng chuẩn Boyce Codd

- **VD31:** Cho $R (A, B, C, D)$

Tập PTH $F = \{A \rightarrow BCD, BC \rightarrow AD, D \rightarrow B\}$

Kiểm tra R có đạt DC BC không?

- Tìm khóa của R :

R có 3 khóa là $\{A, BC, CD\}$

- Phân rã vế phải của các PTH trong F , ta có:

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

- Xét $D \rightarrow B$ có vế trái không chứa 1 khóa

⇒ Do vậy R không đạt DC BC.

4. Dạng chuẩn Boyce Codd

- **Nhận xét:**

- *Lược đồ đạt DC BC thì cũng đạt DC3*
- *Vẫn có thể trùng lặp thông tin*

5. Dạng chuẩn của Lược đồ quan hệ, Lược đồ CSDL

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

5. Dạng chuẩn của Lược đồ quan hệ, Lược đồ CSDL

- Kiểm tra dạng chuẩn của lược đồ quan hệ R

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

Tổng kết chương



Phụ thuộc hàm

1. Các khái niệm cơ bản
2. Hệ luật dẫn Armstrong
3. Bao đóng
4. Phủ tối thiểu
5. Khóa



Dạng chuẩn

1. Dạng chuẩn 1
2. Dạng chuẩn 2
3. Dạng chuẩn 3
4. Dạng chuẩn Boyce Codd
5. Dạng chuẩn của lược đồ quan hệ, lược đồ CSDL



THANKS!

Any questions?

