

---

# Mô hình dữ liệu quan hệ - Lý thuyết thiết kế CSDL

---

Phần 2:

Chuẩn hóa và các Dạng chuẩn

---

# Mục đích

Giúp nắm được các khái niệm và vấn đề:

- Các dạng chuẩn (Normal Forms): 1NF, 2NF, 3NF, BCNF
- Bao đóng (Closure)
- Giải thuật tìm tất cả các khóa
- Tại sao và làm thế nào để chuẩn hóa

---

# Các nội dung chính

1. Các dạng chuẩn
2. Bao đóng
3. Thuật toán tìm toàn bộ các khóa
4. Tập phụ thuộc hàm tối thiểu
5. Các phương pháp chuẩn hóa

# 1. Chuẩn hóa và các dạng chuẩn

- Định nghĩa các dạng chuẩn
  - Dạng chuẩn 1 (1NF - First NF)
  - Dạng chuẩn 2 (2NF - Second NF)
  - Dạng chuẩn 3 (3NF - Third NF)
  - Dạng chuẩn Boyce-Codd (BCNF - Boyce-Codd NF)
- Các phương pháp chuẩn hóa
  - Phép tách (Decomposition)
  - Phép ghép (Composition)

# 1. Dạng chuẩn 1 (1NF)

## ■ Định nghĩa:

- **Giá trị nguyên tố (Atomic Value):** Là giá trị mà không thể bị chia nhỏ hơn được nữa

**Một thuộc tính là nguyên tố** nếu miền giá trị của nó là nguyên tố. Thuộc tính nguyên tố cũng còn được gọi là **thuộc tính đơn**.

- **Dạng chuẩn 1:** một LĐQH R là ở dạng chuẩn 1 nếu như mọi thuộc tính của nó đều nhận giá trị nguyên tố.

- **Lưu ý:** sau này mặc định ta coi các LĐQH đều đã ở dạng chuẩn 1.

# 1. Dạng chuẩn 2 (2NF)

- **Định nghĩa:** một LĐQH R là ở dạng chuẩn 2 nếu nó thỏa mãn 2 điều kiện:
  - R đã ở dạng chuẩn 1
  - Mọi thuộc tính không khóa của R đều phụ thuộc hàm đầy đủ vào khóa của R
- Vd1: cho QH R(A,B,C,D) với tập các PTH F:  
 $AB \rightarrow C; B \rightarrow D;$   
Hỏi R có ở dạng chuẩn 2 không?

# 1. Dạng chuẩn 2

- Vd2: cho QH  $R(A,B,C,D)$  với tập các PTH F:  
 $A \rightarrow B; B \rightarrow C; C \rightarrow D;$   
Hỏi  $R$  có ở dạng chuẩn 2 không?
- Vd3: QH  $Student(ID, Name, Class, Dept, Subject, Mark)$  với các PTH F:  
 $ID \rightarrow Name, Class;$   
 $Class \rightarrow Dept;$   
 $ID, Subject \rightarrow Mark$   
Hỏi  $Student$  có ở dạng chuẩn 2 không?

# 1. Dạng chuẩn 3 (3NF)

- **Định nghĩa:** có 2 cách tương đương để đ/n dạng chuẩn 3:
  - **Cách thứ nhất:** 1 QH R là ở dạng chuẩn 3 nếu thỏa mãn:
    1. R đã ở dạng chuẩn 2
    2. Mọi thuộc tính không khóa của R đều phụ thuộc hàm trực tiếp vào khóa

# 1. Dạng chuẩn 3

## ■ Định nghĩa:

- **Cách thứ hai:** 1 QH R là ở dạng chuẩn 3 nếu thỏa mãn:
  1. R đã ở dạng chuẩn 1
  2. Với mọi PTH có dạng  $X \rightarrow A$  thuộc R (với X là một tập các thuộc tính, còn A là một thuộc tính) thì Hoặc X là một siêu khóa, hoặc A là thuộc tính khóa.

# 1. Dạng chuẩn 3

- Vd2: Xét QH  $R(A,B,C,D)$  với tập các PTH  $F$ :

$$A \rightarrow B; B \rightarrow C; C \rightarrow D;$$

Hỏi  $R$  có ở dạng chuẩn 2, chuẩn 3 không?

- Vd4: Xét QH  $R(A,B,C,D)$  với tập các PTH  $F$ :

$$AB \rightarrow C; CD \rightarrow A;$$

Hỏi  $R$  có ở dạng chuẩn 2, chuẩn 3 không?

# 1. Dạng chuẩn Boyce-Codd (BCNF)

- **Định nghĩa:** một LĐ QH R là ở dạng chuẩn BC nếu nó thỏa mãn:
  1. R đã ở dạng chuẩn 1,
  2. Với mọi PTH có dạng  $X \rightarrow A$  in R (với X là một tập thuộc tính, và A là một thuộc tính) thì X là siêu khóa.

# 1. Dạng chuẩn Boyce-Codd

- Vd4: Xét QH  $R(A,B,C,D)$  với tập các PTH F:

$$AB \rightarrow C; CD \rightarrow A;$$

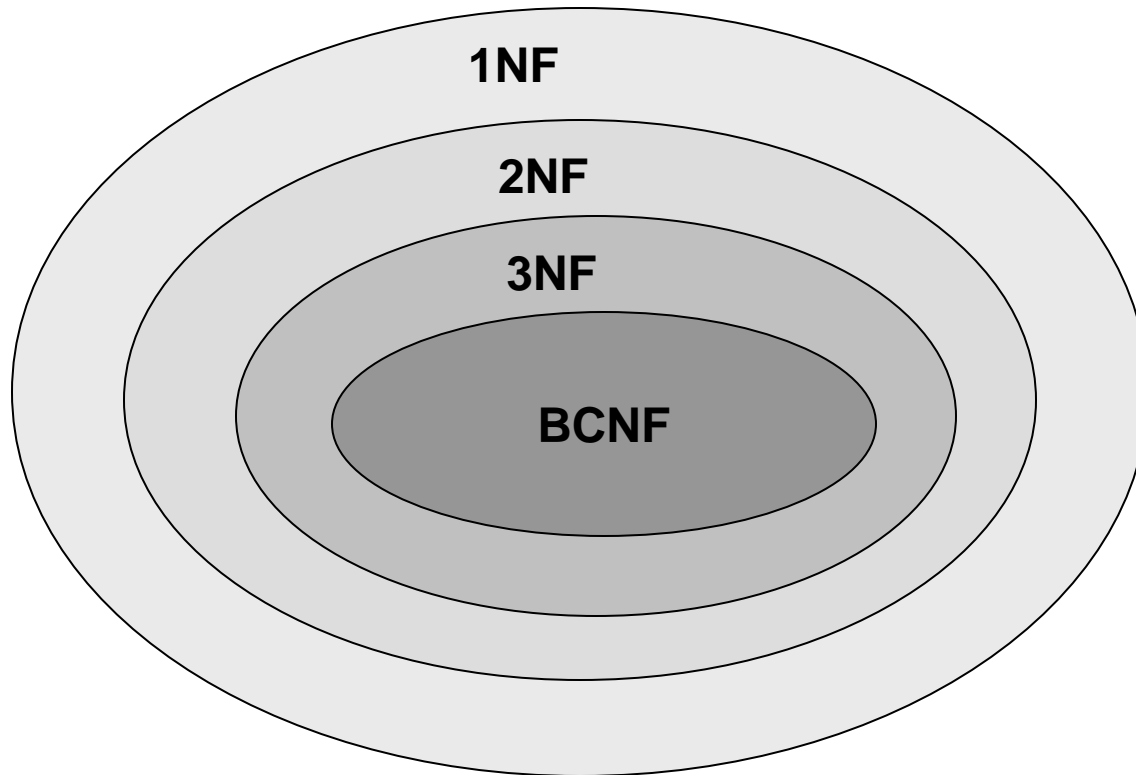
Hỏi R có ở dạng chuẩn 3, chuẩn BC không?

- Vd5: xét QH  $R(A,B,C,D)$  với các PTH F:

$$AB \rightarrow C; AB \rightarrow D;$$

Hỏi R có ở dạng chuẩn BC không?

# So sánh giữa các dạng chuẩn



## 2. Bao đóng (Closure)

- **Định nghĩa:** cho QH  $R$  với tập PTH  $F$ , và tập thuộc tính  $X \subseteq R$ . Ta gọi bao đóng của  $X$  (ký hiệu  $X^+$ ) được định nghĩa như sau:

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

Vd: cho  $R(A,B,C,D)$ , với các PTH  $F$ :

$$A \rightarrow B; BC \rightarrow AD;$$

$$X=(A) \quad \text{thì} \quad X^+=(A,B)$$

$$X=(AC) \quad \text{thì} \quad X^+=(A,B,C,D)$$

## 2. Bao đóng – Thuật toán tìm bao đóng

- **Đầu vào:**  $R(A_1, A_2, \dots, A_n)$  với tập PTH  $F$ ;  $X \subseteq R$
- **Đầu ra:**  $X^+$
- **Giải thuật:**
  - Bước 1:  $X^+ := X$ ;
  - Bước 2: While  $(\exists Y \rightarrow A \in F \text{ so that } (Y \subseteq X^+ \text{ and } A \notin X^+))$   
 $X^+ := X^+ \cup A$
  - Bước 3: return  $X^+$

---

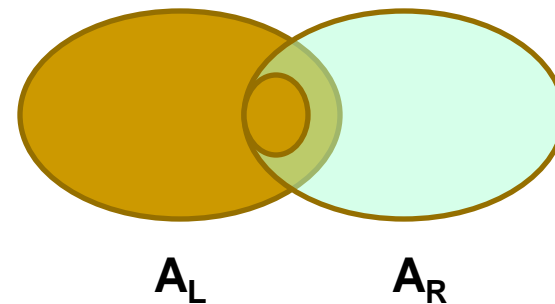
## 2. Bao đóng - ứng dụng

- Giúp tìm khóa trong quan hệ, bởi vì  $K^+ = R$

### 3. Giải thuật tìm tất cả các khóa của lược đồ quan hệ

- **Một tính chất của khóa:** cho LĐQH R với tập các PTH F. Ta gọi  $A_L$ ,  $A_R$  tương ứng là các hợp (unions) của các thuộc tính thuộc về trái và về phải của các PTH trong F. Chúng ta có thể giả sử rằng  $R = A_L \cup A_R$ . Nếu K là một khóa của R thì nó luôn thỏa mãn:

$$A_L \setminus A_R \subseteq K \subseteq A_L$$



### 3. Giải thuật tìm tất cả các khóa của lược đồ quan hệ

- Vd:  $R(A,B,C,D,E)$ , với F:  
 $A \rightarrow B; BC \rightarrow AD; CD \rightarrow E;$

Ta có:  $A_L = (ABCD); A_R = (ABDE)$

$$A_L \setminus A_R = C;$$

Ta tìm được 2 khóa của R là:

$$K1 = AC;$$

$$K2 = BC;$$

# 3. Giải thuật tìm tất cả các khóa của lược đồ quan hệ

- **Đầu vào:** QH  
 $R(A_1, A_2, \dots, A_n)$  với tập các PTH F;
- **Đầu ra:** Tất cả các khóa của R
- **Giải thuật:**
  - Bước 1: Tính  $A_L$  và  $A_R$  của R.
  - Bước 2:  $K := A_L \setminus A_R$ ;  $X := A_L \cap A_R$   
Nếu  $(K^+ = R)$  thì K là khóa duy nhất;  
Return K;  
Else
  - Bước 3: While  $(K^+ \neq R)$  {  
    **Lấy ra một thuộc tính  $A_i$  từ X;**  
     $K := K \cup A_i$ ;  
}
  - Ta sẽ có thêm một khóa K ở đây;
  - Lặp lại Bước 3 cho đến khi đã lấy hết các trường hợp

# Bài tập

- Tìm tất cả các khóa của LĐQH:

R (A,B,C,D,E,F) với các PTH F:

$AB \rightarrow C$ ;

$BC \rightarrow D$ ;

$BD \rightarrow AE$ ;

$AC \rightarrow F$ ;

- Lời giải:

$A_L = ABCD$ ;  $A_R = CDAEF$

$\rightarrow A_L \setminus A_R = B$ ;  $A_L \cap A_R = ACD$

$B^+ = B \neq R$

$(BA)^+ = BACDEF = R \rightarrow K_1 = AB$

$(BC)^+ = BCDAEF = R \rightarrow K_2 = BC$

$(BD)^+ = BDAECF = R \rightarrow K_3 = BD$

**Kết luận:** Lược đồ R có 3 khóa  $K_1, K_2$  và  $K_3$  như trên

# 4. Tập PTH cơ sở tối thiểu (Minimal Basis)

- **Tập PTH cơ sở (Basis):** cho QH  $R$  với tập các PTH  $F$ . Một tập PTH bất kỳ mà tương đương với  $F$  thì được gọi là một cơ sở của  $F$ .
- **Tập cơ sở tối thiểu (Minimal basis /canonical set):** là một tập PTH thỏa mãn 3 điều kiện:
  - Vé phải của tất cả các PTH chỉ có một thuộc tính.
  - Nếu loại bỏ đi bất kỳ PTH nào, thì các PTH còn lại không đủ là cơ sở của  $F$  (không có PTH dư thừa).
  - Với mọi PTH, nếu loại bỏ đi bất kỳ thuộc tính nào khỏi vé trái của PTH đó, thì các PTH còn lại không đủ là cơ sở của  $F$  (không chứa PTH bộ phận)

# Ví dụ

- Cho LĐQH  $R(A,B,C,D,E,F)$  với tập F:  
 $AB \rightarrow CD; C \rightarrow EF;$   
 $BC \rightarrow DE; E \rightarrow F;$   
Hãy tìm cơ sở tối thiểu của F?

- **Giải:** Sử dụng tính chất tách ta có:  
1.  $AB \rightarrow C;$     5.  $BC \rightarrow D;$   
2.  $AB \rightarrow D;$     6.  $BC \rightarrow E;$   
3.  $C \rightarrow E;$       7.  $E \rightarrow F;$   
4.  $C \rightarrow F;$

Do 3., nên PTH 6. bị loại ra.  
Do 3. và 7., nên 4. bị loại ra.  
Do 1. và 5., nên 2. bị loại ra.

Cuối cùng, ta có tập cơ sở tối thiểu của F là:

1.  $AB \rightarrow C;$
2.  $BC \rightarrow D;$
3.  $C \rightarrow E;$
4.  $E \rightarrow F;$

# 4. Tập PTH cơ sở tối thiểu

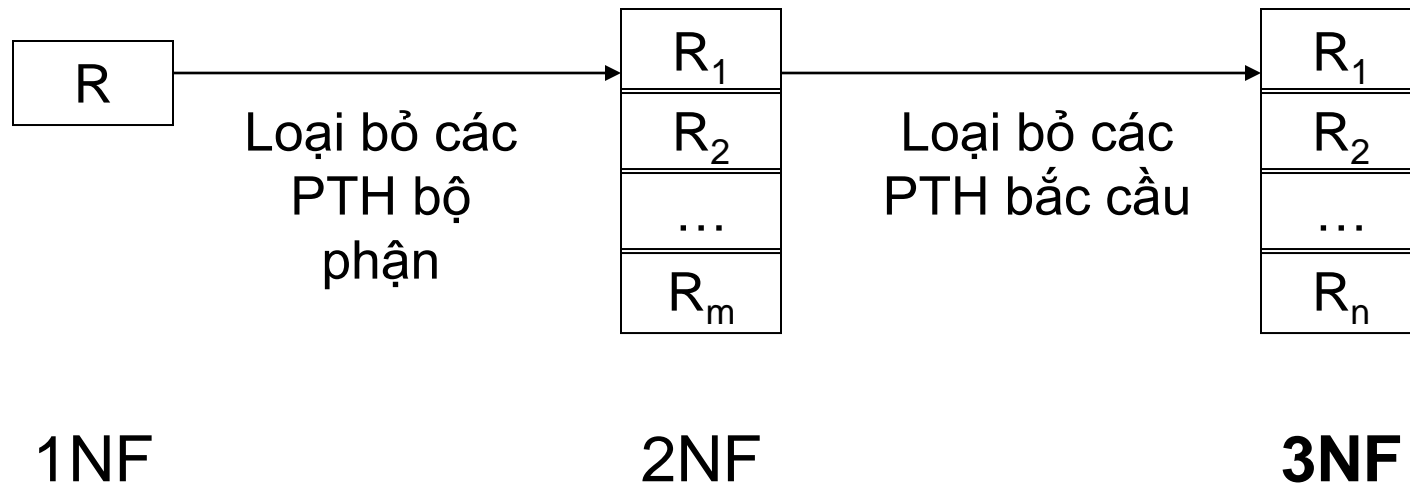
- **Giải thuật tìm tập cơ sở tối thiểu:**
  - **Đầu vào :** QH R với tập PTH F.
  - **Đầu ra:** tập PTH cơ sở tối thiểu của F
  - **Giải thuật:**
    - **Bước 0:** Loại bỏ các PTH tầm thường ( $X \rightarrow Y$  mà  $Y \subseteq X$ ).
    - **Bước 1:** Với mỗi PTH có vế phải nhiều hơn 1 thuộc tính, thì chuyển nó thành các PTH có vế phải có 1 thuộc tính theo tính chất tách.
    - **Bước 2:** Loại bỏ các PTH dư thừa từ Bước 1.
    - **Bước 3:** Loại bỏ các PTH bộ phận từ Bước 2.

# 5. Các phương pháp chuẩn hóa (Normalization Methods)

Có 2 cách tiếp cận trái ngược nhau:

- ❑ **Tách** (Decomposition): tách một lược đồ lớn thành các lược đồ con nhỏ hơn ở dạng chuẩn mong muốn (thường là dạng chuẩn 3).
- ❑ **Ghép** (Composition): Từ các thuộc tính đơn lẻ, ta xây dựng các lược đồ quan hệ ở dạng chuẩn 3.

# 5. Chuẩn hóa- Phương pháp tách



# 5. Chuẩn hóa- Phương pháp tách

## ■ Giải thuật:

- **Đầu vào:** QH R với tập các PTH F mà chưa ở dạng chuẩn 3.
- **Đầu ra:** một phép tách R sao cho bảo toàn F, nối không mất thông tin, và mọi LĐ con đều ở dạng chuẩn 3.
- **Giải thuật:**
  - **Bước 1:** Với mỗi PTH bộ phận hay bắc cầu của một thuộc tính không khóa A vào khóa K của R, ta có thể tìm được một tập các thuộc tính  $X \neq K$  sao cho  $X \rightarrow A$  là PTH đầy đủ hay trực tiếp và không tồn tại  $Y \neq X$  mà  $Y \rightarrow A$ .

Từ  $X \rightarrow A$ , ta tách thành 2 QH R1 và R2 sao cho:

$$R1=XA; \quad R2=R \setminus A;$$

- **Bước 2:** Lặp lại bước 1 cho mỗi lược đồ con mà chưa ở dạng chuẩn 3.

**Lưu ý 1:** điều kiện “không tồn tại  $Y \neq X$  sao cho  $Y \rightarrow A$ ” nhằm đảm bảo rằng không có PTH nào bị mất trong quá trình tách.

**Lưu ý 2:** Các LĐ mà có chung khóa có thể lại được ghép lại thành một lược đồ với khóa chung đó, và hợp các thuộc tính của các lược đồ.

# Ví dụ

- QH Student(ID, name, class, dept, subject, mark) với tập các PTH F:
  - ID → name;
  - ID → class;
  - class → dept;
  - ID, subject → mark;

QH này đã ở dạng chuẩn 3 chưa?

Nếu chưa, hãy tách nó thành các lược đồ con ở dạng chuẩn 3?

# Ví dụ (tiếp)

## ■ Giải:

- Trước tiên tìm khóa của QH  $K=(ID,subject)$ .
- Do đó các thuộc tính không khóa gồm (*name, class, dept, mark*)
- Do *name* PTH bộ phận vào  $K$ , ta tìm thấy  $ID \rightarrow name$  là PTH đầy đủ. Từ đây, ta chia QH thành 2 QH:  $S1(\underline{ID}, name)$  và  $S2(\underline{ID}, class, dept, \underline{subject}, mark)$ .  $S1$  đã ở dạng chuẩn 3, nhưng  $S2$  thì chưa.
- Tương tự trên,  $S2$  có thể được chia thành các QH sau:  
 $S3(\underline{ID}, class); S4(\underline{class}, dept); S5(\underline{ID}, \underline{subject}, mark);$
- Sau khi đổi lại tên các QH, rồi gộp các QH có chung khóa, cuối cùng ta có các QH sau:  
 **$S1(\underline{ID}, name, class); S2(\underline{class}, dept); S3(\underline{ID}, \underline{subject}, mark);$**

# 5. Chuẩn hóa – Phương pháp ghép

## ■ Giải thuật:

- **Đầu vào:** một QH  $R$  chưa ở dạng chuẩn 3, với tập các PTH  $F$ .
- **Đầu ra:** phép tách  $R$  mà bảo toàn các PTH, nối không mất, và các lược đồ con đều ở dạng chuẩn 3.
- **Giải thuật:**
  1. Tìm tập cơ sở tối thiểu của  $F$ . Sau đó đánh thứ tự các PTH từ 1 đến  $n$ .
  2. Với mỗi PTH thứ  $i$   $X \rightarrow A$ ,  $i=1..n$ , ta tạo một QH với lược đồ  $R_i = (XA)$ .
  1. Nếu chưa có lược đồ nào trong số được tạo ra ở bước 2 chứa siêu khóa của  $R$ , thì bổ sung thêm một lược đồ là khóa của  $R$ .

# Ví dụ

- QH Student(ID, name, class, dept, subject, mark) với tập các PTH F:
  - ID → name;
  - ID → class;
  - class → dept;
  - ID, subject → mark;

QH này đã ở dạng chuẩn 3 chưa?

Nếu chưa, hãy tách nó thành các lược đồ con ở dạng chuẩn 3?

# Ví dụ (tiếp)

## ■ Giải:

- Tìm cơ sở tối thiểu:  $F$  đã là cơ sở tối thiểu.
- Từ 4 PTH của  $F$ , ghép lại ta sẽ có 4 quan hệ:

**$S1(\underline{ID}, name); S2(\underline{ID}, class);$**

**$S3(\underline{class}, dept); S4(\underline{ID}, \underline{subject}, mark);$**

- Do  $S1$  và  $S2$  có chung khóa  $ID$ , nên có thể ghép chúng lại thành một quan hệ mới:  $S1(\underline{ID}, name, class)$ . Cuối cùng, ta có 3 quan hệ:

**$S1(\underline{ID}, name, class);$**

**$S2(\underline{class}, dept);$**

**$S3(\underline{ID}, \underline{subject}, mark);$**

# Ví dụ (tiếp)

S1: Student

<u>ID</u>	<u>Name</u>	<u>Class</u>
E1-001	Nguyen Van A	E1
E1-002	Tran Thi B	E1
E2-001	Nguyen Hong C	E2
IT1-001	Tran Thi B	IT1
IT1-002	Le Van D	IT1

S2: Class

<u>Class</u>	<u>Department</u>
E1	Electronics
E2	Electronics
IT1	IT

S3: Examination

<u>ID</u>	<u>Subject</u>	<u>Mark</u>
E1-001	Electronic Circuit	8
E1-001	Digital Technique	7
E1-002	Digital Technique	9
E1-002	Electronic Circuit	8
E2-001	Digital Technique	6
IT1-001	Electronic Circuit	10
IT1-002	Digital Technique	8

---

# Thank you!