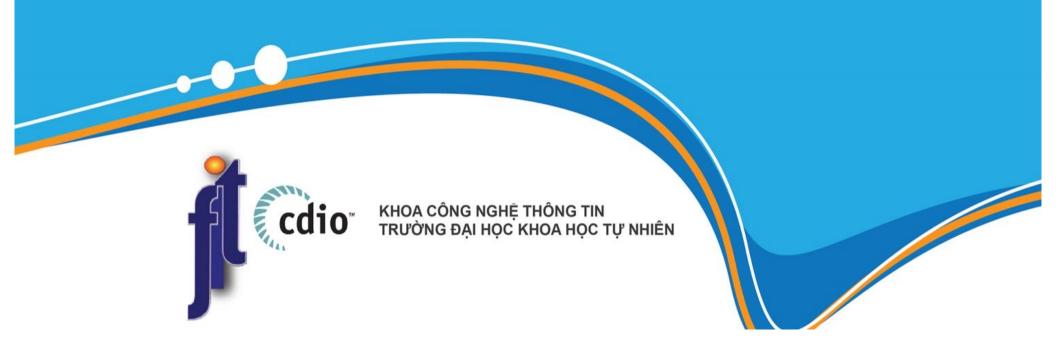
# CHƯƠNG 8 Phụ thuộc hàm và dạng chuẩn (Functional Dependencies & Normal Form)



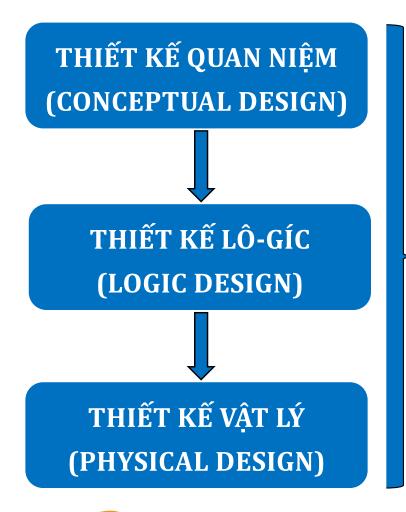


## Nội dung chi tiết

- Các bước thiết kế CSDL quan hệ
- Sự trùng lắp thông tin
- ☐ Phụ thuộc hàm
- Luật dẫn Armstrong
- Bài toán xác định khoá
- Các dạng chuẩn



#### Các bước thiết kế CSDL quan hệ



- Dáp ứng các yêu cầu thông tin cho các người dùng và các ứng dụng cụ thể.
- Cung cấp kiến trúc lưu trữ dữ liệu đúng đắn, tự nhiên và dễ hiểu.
- Hỗ trợ một số yêu cầu về hiệu năng: thời gian hồi đáp, thời gian xử lý, không gian lưu trữ, ...



#### Các bước thiết kế CSDL quan hệ

- Một lược đồ CSDL **thiết kế tốt** được thể hiện:
  - Ngữ nghĩa của quan hệ, thuộc tính được thể hiện rõ ràng
  - Các thuộc tính trên quan hệ không được trùng lắp
  - Giảm trùng lắp giữa các bộ trên một quan hệ (dư thừa thông tin)
  - Giảm **giá trị rỗng** (Null) trên các quan hệ.
- Lược đồ **thiết kế trùng lắp** thông tin dẫn đến:
  - Tốn không gian lưu trữ
  - Dị thường khi thực hiện các phép cập nhật dữ liệu
- ☐ Kiểm tra lược đồ tốt hay xấu?
  - Bằng kinh nghiệm
  - Bằng tiêu chuẩn dạng chuẩn



#### Các bước thiết kế CSDL quan hệ

- Các bước cơ bản của thiết kế csdl ở mức quan niệm:
  - Thiết kế mô hình DL ở mức quan niệm (ER).
  - Chuyển mô hình DL mức quan niệm sang mức logic (chuyển mô hình ER sang mô hình DL quan hệ).
  - Với mỗi quan hệ, xác định tập PTH nhận diện.
  - Đánh giá chất lượng của lược đồ
  - Nâng chuẩn của 1 lược đồ quan hệ bằng cách loại:
    - PTH không đầy đủ
    - PTH bắc cầu vào khóa chính của lược đồ
  - Đánh giá chất lượng của lược đồ qua mỗi lần nâng chuẩn bằng tiêu chí dạng chuẩn



#### Nội dung chi tiết

- Các bước thiết kế CSDL quan hệ
- ☐ Sự dư thừa thông tin
- ☐ Phụ thuộc hàm
- Luật dẫn Armstrong
- Bài toán xác định khoá
- Các dạng chuẩn



#### Sự dư thừa thông tin

Du thừa!

Khi một phần của dữ liệu có thể suy ra từ một số thành phần dữ liệu khác thì ta nói có sự trùng lắp dữ liệu.

				1			)
MAGV	TENGV	NGSINH	DCHI		МАВМ	TENBM	TRGBM
009	Tiên	11/02/1960	119 Cống Quỳnh, Tp HCM		5	НТТТ	005
005	Tùng	20/08/1962	222 Nguyễn Văn Cừ, Tp HCM		5	нттт	005
007	Hằng	11/3/1954	332 Nguyễn Thái Học, Tp HCM		4	MMT	008
001	Như	01/02/1967	291 Hồ Văn Huê, Tp HCM		4	MMT	008
004	Hùng	04/03/1967	95 Bà Rịa, Vũng Tàu		5	НТТТ	005
003	Tâm	04/05/1957	34 Mai Thị Lự, Tp HCM		5	нттт	005
008	Quang	01/09/1967	80 Lê Hồng Phong, Tp HCM		4	MMT	008
006	Vinh	01/01/1065	4E Trucha Vucana Hà Nài		1	CNDM	006

<u>Sự dư thừa thông tin:</u> Thông tin về tên bộ môn và người trưởng bộ môn lặp lại nhiều lần



#### Sự dư thừa thông tin

- Các dị thường khi thực hiện thao tác cập nhật:
  - Thêm: Thêm 1 giáo viên thì phải thêm thông tin bộ môn hoặc phải để giá trị cho các thuộc tính liên quan đến bộ môn là Null. Không thể thêm 1 bộ môn mới mà chưa có giáo viên nào thuộc về bộ môn đó.
  - Sửa: khi sửa thông tin bộ môn, thì phải sửa đến tất cả các bộ liên quan nếu không sẽ dẫn tới trình trạng không nhất quán.
  - Xóa: có thể bị mất thông tin. Xem xét khi xóa giáo viên có mã 006?

MAGV	TENGV	NGSINH	DCHI	MABM	TENBM	TRGBM		
009	Tiên	11/02/1960	119 Cống Quỳnh, Tp HCM	5	НТТТ	005		
005	Tùng	20/08/1962	222 Nguyễn Văn Cừ, Tp HCM	5	НТТТ	005		
007	Hằng	11/3/1954	332 Nguyễn Thái Học, Tp HCM	4	MMT	008		
001	Như	01/02/1967	291 Hồ Văn Huê, Tp HCM	4	MMT	008		
004	Hùng	04/03/1967	95 Bà Rịa, Vũng Tàu	5	НТТТ	005		
003	Tâm	04/05/1957	34 Mai Thị Lự, Tp HCM	5	НТТТ	005		
008	Quang	01/09/1967	80 Lê Hồng Phong, Tp HCM	4	MMT	008		
006	Vinh	01/01/1965	45 Trưng Vương, Hà Nội	1	CNPM	006		



#### Sự dư thừa thông tin

#### Tốn không gian lưu trữ

	88								
MAGV	TENGV	NGSINH	DCHI	MABM	TENBM	TRGBM			
009	Tiên	11/02/1960	119 Cống Quỳnh, Tp HCM	5	нттт	005			
005	Tùng	20/08/1962	222 Nguyễn Văn Cừ, Tp HCM	5	нттт	005			
007	Hằng	11/3/1954	332 Nguyễn Thái Học, Tp HCM	4	MMT	008			
001	Như	01/02/1967	291 Hồ Văn Huê, Tp HCM	4	MMT	008			
004	Hùng	04/03/1967	95 Bà Rịa, Vũng Tàu	5	НТТТ	005			
003	Tâm	04/05/1957	34 Mai Thị Lự, Tp HCM	5	нттт	005			
008	Quang	01/09/1967	80 Lê Hồng Phong, Tp HCM	4	MMT	008			
006	Vinh	01/01/1965	45 Trưng Vương, Hà Nội	1	CNPM	006			

Lược đồ 1

## Tính kích thước lưu trữ của mỗi lược đồ khi thêm 10 giáo viên cho bộ môn HTTT? (Giả sử kích thước mỗi thuộc tính là 20 bytes)

MAGV	TENGV	NGSINH	DCHI	МАВМ
009	Tiên	11/02/1960	119 Cống Quỳnh, Tp HCM	5
005	Tùng	20/08/1962	222 Nguyễn Văn Cừ, Tp HCM	5
007	Hằng	11/3/1954	332 Nguyễn Thái Học, Tp HCM	4
001	Như	01/02/1967	291 Hồ Văn Huê, Tp HCM	4
004	Hùng	04/03/1967	95 Bà Rịa, Vũng Tàu	5
003	Tâm	04/05/1957	34 Mai Thị Lự, Tp HCM	5
008	Quang	01/09/1967	80 Lê Hồng Phong, Tp HCM	4
006	Vinh	01/01/1965	45 Trưng Vương, Hà Nội	1

MABM	TENBM	TRGBM
5	нттт	005
5	нттт	005
4	MMT	008
4	MMT	008
5	нттт	005
5	нттт	005
4	MMT	008
1	CNPM	006

Lược đồ 2



#### Nội dung chi tiết

- Các bước thiết kế CSDL quan hệ
- Sự trùng lắp thông tin và các dị thường
- ☐ Phụ thuộc hàm
- Bài toán xác định khoá
- Các dạng chuẩn



- Cho R (A1, A2, ..., An), r(R), ký hiệu  $R+ = \{A1, A2, ..., An\}$ .
- $\square$  Phụ thuộc hàm giữa 2 tập thuộc tính X, Y  $\subseteq$  R+.
  - Ký hiệu : X → Y, X là vế trái, Y là vế phải.
  - $\square$  Ý nghĩa:  $\forall r \in R$ , t1,  $t2 \in r$ , nếu t1[X] = t2[X] thì t1[Y] = t2[Y].
  - Ví du:

MAGV → TENGV,NGSINH,DCHI MABM MABM → {TENBM, TRGBM}

MAGV	TENGV	NGSINH	DCHI	МАВМ	TENBM	TRGBM
009	Tiên	11/02/1960	119 Cống Quỳnh, Tp HCM	5	нттт	005
005	Tùng	20/08/1962	222 Nguyễn Văn Cừ, Tp HCM	5	нттт	005
007	Hằng	11/3/1954	332 Nguyễn Thái Học, Tp HCM	4	MMT	008
001	Như	01/02/1967	291 Hồ Văn Huê, Tp HCM	4	MMT	008
004	Hùng	04/03/1967	95 Bà Rịa, Vũng Tàu	5	нттт	005
003	Tâm	04/05/1957	34 Mai Thị Lự, Tp HCM	5	нттт	005
800	Quang	01/09/1967	80 Lê Hồng Phong, Tp HCM	4	MMT	008
006	Vinh	01/01/1965	45 Trưng Vương, Hà Nội	1	CNPM	006



#### Phụ thuộc hàm (PTH)

- ☐ Nhận xét
  - PTH thể hiện các ràng buộc từ thế giới thực.
  - ☐ Khi K là khóa của R thì ta có phụ thuộc hàm: K→ R+
  - PTH dùng để đánh giá một kết quả thiết kế CSDL.



- Nhận diện PTH: Việc nhận diện PTH dựa vào ý nghĩa của thuộc tính và mối quan hệ của chúng trong quan hệ.
- □ Ví dụ:

MAGV	TENGV	NGSINH	DCHI	MABM	TENBM	TRGBM
009	Tiên	11/02/1960	119 Cống Quỳnh, Tp HCM	5	НТТТ	005
005	Tùng	20/08/1962	222 Nguyễn Văn Cừ, Tp HCM	5	НТТТ	005
007	Hằng	11/3/1954	332 Nguyễn Thái Học, Tp HCM	4	MMT	008
001	Như	01/02/1967	291 Hồ Văn Huê, Tp HCM	4	MMT	008
004	Hùng	04/03/1967	95 Bà Rịa, Vũng Tàu	5	НТТТ	005
003	Tâm	04/05/1957	34 Mai Thị Lự, Tp HCM	5	НТТТ	005
008	Quang	01/09/1967	80 Lê Hồng Phong, Tp HCM	4	MMT	008
006	Vinh	01/01/1965	45 Trưng Vương, Hà Nội	1	CNPM	006

- Dựa vào dữ liệu trên :
  - $\square$  MAGV  $\rightarrow$  TENGV
  - $\square$  TENGV  $\rightarrow$  MAGV
- ☐ Tuy nhiên, chỉ có pth:
  - $\square$  MAGV  $\rightarrow$  TENGV



- Luật dẫn Armstrong
  - Luật phản hồi

$$\forall Y \subseteq X, X \rightarrow Y$$

Luật cộng

(FD2)

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

Thì X, W  $\rightarrow$  Y, Z

Luật bắc cầu

(FD3)

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

Thì  $X \rightarrow Z$ 



#### Luật dẫn khác:

Luật phân rã

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

Thì X, W 
$$\rightarrow$$
 Z

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

Thì 
$$X \rightarrow Y$$
, Z

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

Thì 
$$X \rightarrow Z$$



# Suy dẫn từ tập PTH

- Gọi **F** là tập PTH trên 1 quan hệ R
  - → Suy luận "quan hệ phải thỏa một tập PTH khác nào đó"
- Vai trò: Khám phá thêm tập PTH hỗ trợ cho đánh giá và thiết kế các lược đồ quan hệ đạt chất lượng tốt
- f là một PTH được <u>suy dẫn</u> từ F, ký hiệu F ⊢ f, nếu:
  - Tồn tại một chuỗi phụ thuộc hàm  $f_1$ ,  $f_2$ ,... $f_n$ , với:
    - $\blacksquare$   $f_n = f$
    - $f_i \in F$  hoặc được suy từ những phụ thuộc hàm  $f_i$ , j=1..i-1 nhờ vào <u>luật dẫn</u>
- $\square$  F' là tập PTH suy dẫn từ F qua tập luật dẫn R (F  $\subseteq$  F' )



#### Ví dụ

☐ Xét lược đồ R(A,B,C) thỏa tập PTH

$$F = \{ f1: A \rightarrow B \\ f2: B \rightarrow C \}$$

☐ Ta có thể suy diễn R còn thỏa PTH

f3: 
$$A \rightarrow C$$



PTH đầy đủ

Xét pth  $X \rightarrow Y$  được định nghĩa trên R

Nếu ⊣∃X' ⊂ X sao cho X' → Y

Thì Y phụ thuộc đầy đủ vào X

- Ví dụ:
  - Cho R(A, B, C, D, E, I)
  - Tập PTH:  $F = \{A \rightarrow BCD, BCD \rightarrow E, CD \rightarrow EI\}$
  - − Pth BCD→E là phụ thuộc hàm đầy đủ không?



#### Một số khái niệm

□ **PTH đầy đủ** (Full functional dependency)

Xét  $X \to Y$ : Nếu  $\neg \exists X' \subset X$  sao cho  $X' \to Y$ Thì Y phụ thuộc hàm đầy đủ vào X(Hoặc không phụ thuộc hàm vào tập con nào của X)

PTH riêng phần (Partial functional dependency)

Xét  $X \to Y$ : Nếu  $\exists X' \subset X$  sao cho  $X' \to Y$ Thì Y phụ thuộc riêng phần vào X(Hoặc *không phụ thuộc hàm đầy đủ vào X*)



#### Một số khái niệm

VD1: SinhVien (mãsv, họtên, mãlớp, tênlớp, mônhọc, điểm)

Phụ thuộc hàm đầy đủ (KHÔNG phụ thuộc một phần X)

- f1: Mãsv → Họtên, mãlớp
- f2: Mãlớp → Tênlớp
- f5: Mãsv, Mônhọc → Điểm

Phụ thuộc hàm riêng phần (phụ thuộc một phần X)

- f7: Mãsv, Tênsv, Mônhọc → Điểm
- f8: Mãsv, Mãlóp → Tênlóp

**VD2**: R(A, B, C, D, E, I),  $F = \{A \rightarrow BCD, BCD \rightarrow E, CD \rightarrow EI\}$ BCD $\rightarrow E$  là phụ thuộc hàm đầy đủ không?



#### Một số khái niệm

#### PTH bắc cầu

- transitive dependency
- $X \rightarrow A$  được gọi là phụ thuộc bắc cầu nếu tồn tại Y **thoả đồng thời 4 điều kiện** sau:
  - $(1) X \rightarrow Y,$
  - $(2) Y \rightarrow A,$
  - (3) Y **→** X,
  - (4) A ∉ XY.

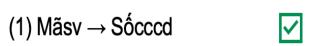
SinhVien (mãsv, họtên, mãlớp, tênlớp, mônhọc, điểm, sốcccd, ngày cấp sốcccd)

Xác định phụ thuộc hàm bắc cầu?

#### f1: Mãsv → tênlớp

- (1) Mãsv → Mãlớp
- (2) Mãlớp → Tênlớp
- (3) Mãlớp <del>X</del> Mãs∨
- (4) Tênlớp ∉ {Mãsv, Mãlớp} 🔽

→ Phụ thuộc bắc cầu



f2: Mãsv → ngàycấp\_sốcccd

- (2) Sốcccd → Ngàycấp\_sốcccd ✓
- (3) Sốcccd → Mãsv
- → KHÔNG phụ thuộc bắc cầu



#### Nội dung chi tiết

- Các bước thiết kế CSDL quan hệ
- Sự trùng lắp thông tin
- ☐ Phụ thuộc hàm
- Bài toán xác định khoá
- Các dạng chuẩn



#### Tìm khoá với tập tố hợp giới hạn

- B1 Gọi left(f) là vế trái, right(f) là vế phải của phụ thuộc hàm f.
- B2 Tính  $\mathcal{N} := U \bigcup_{\forall f \in F} right(f)$
- B3 Nếu  $\mathcal{N}^+_{\mathscr{F}} = U$  thì R chỉ có một khóa là  $\mathscr{N}$ .
- B4 Nếu N⁺<sub>F</sub>⊂ U thì
  - B4.1Tính  $\mathscr{D} := \bigcup_{\forall f \in \mathscr{F}} right(f) \bigcup_{\forall f \in F} left(f)$
  - **B4.2Tinh**  $\mathcal{L} := U \mathcal{N}^+_{\mathscr{F}} \mathcal{D}$
  - B5  $\mathcal{K} := \emptyset$ ;

Với mỗi tập con  $\mathcal{L}_i \subseteq \mathcal{L}$ , nếu  $\{\mathcal{NL}_i\}^+_{\mathscr{F}} = U : \mathcal{K} := \mathcal{K} \cup \{\mathcal{NL}_i\}$ 

Trong khi mà tồn tại  $\mathcal{K}_i$ ,  $\mathcal{K}_j$  thuộc  $\mathcal{K}$  sao cho  $\mathcal{K}_i \subset \mathcal{K}_j$  thì

$$\mathcal{K} := \mathcal{K} - \{\mathcal{K}_i\}$$

R có tập khóa là K.

# Ví dụ

$$R(A,B,C,D,E,F)$$

$$U = \{ABCDEF\}$$

```
B1 Gọi left(f) là vế trái, right(f) là vế phải của phụ thuộc hàm f.
B2 - Tính \mathcal{N} := U - \bigcup_{\forall f \in F} right(f)
B3 - Nếu \mathcal{N}^+_{\mathscr{F}} = U thì R chỉ có một khóa là \mathscr{N}.
B4 - Nếu \mathcal{N}^+_{\mathscr{F}} \subset U thì
B4.1Tính \mathscr{D} := \bigcup_{\forall f \in \mathscr{F}} right(f) - \bigcup_{\forall f \in F} left(f)
B4.2Tính \mathscr{L} := U - \mathcal{N}^+_{\mathscr{F}} \mathscr{D}
B5 \mathscr{K} := \mathscr{O};
Với mỗi tập con \mathscr{L}_i \subseteq \mathscr{L}, nếu \{\mathcal{N}\mathscr{L}_i\}^+_{\mathscr{F}} = U: \mathscr{K} := \mathscr{K} \cup \{\mathcal{N}\mathscr{L}_i\}
Trong khi mà tồn tại \mathscr{K}_i, \mathscr{K}_j thuộc \mathscr{K} sao cho \mathscr{K}_i \subset \mathscr{K}_j thì \mathscr{K} := \mathscr{K} - \{\mathscr{K}_j\}
R có tập khóa là \mathscr{K}.
```

R có duy nhất 1 khoá là AD



 Tìm khóa của lược đồ quan hệ :

```
R(A,B,C,D,E,F)
U = \{ABCDEF\}
```

Có tập phụ thuộc hàm sau:

```
\label{eq:f1:A} \begin{array}{c} \text{F= \{ } \text{f1: A} \rightarrow \text{D ,} \\ \text{f2: C} \rightarrow \text{AF ,} \\ \text{f3: AB} \rightarrow \text{EC} \\ \text{\}} \end{array}
```

```
Ví dụ
```

```
B1: left(f) = { ABC }, right(f) = { ACDEF }
```

B2: 
$$N = U - right(f) = \{ABCDEF\} - \{ACDEF\} = \{B\}$$

Tìm khóa của lược đồ B3: N<sup>+</sup><sub>F</sub> = { tập thuộc tính xác định/ được dẫn ra từ N } quan hệ:
 = { B } ⊂ U

$$R(A,B,C,D,E,F)$$

$$U = \{ABCDEF\}$$

Có tập phụ thuộc hàm sau:

```
F = \{ \begin{array}{l} \text{f1: A} \rightarrow \text{D} \; , \\ \text{f2: C} \rightarrow \text{AF} \; , \\ \text{f3: AB} \rightarrow \text{EC} \\ \end{array} \}
```



B2: 
$$N = U - right(f) = \{ABCDEF\} - \{ACDEF\} = \{B\}$$

quan hệ:

$$R(A,B,C,D,E,F)$$

$$U = \{ABCDEF\}$$

 Tìm khóa của lược đổ B3: N+ = { tập thuộc tính xác định/ được dân ra từ N }  $= \{B\} \subset U$ 

```
B4.1: D = right(f) - left(f) --tìm tập chỉ bên vế phải
        = { ACDEF } - { ABC } = { DEF }
```

Có tập phụ thuộc hàm

```
F=\{ f1: A \rightarrow D,
       f2: C \rightarrow AF,
       f3: AB → EC
```

B4.2: L = U - D.N+ --tìm tập có khả năng làm khoá = { ABCDEF } - { DEFB } = { AC }



B2: 
$$N = U - right(f) = \{ABCDEF\} - \{ACDEF\} = \{B\}$$

quan hệ:

$$R(A,B,C,D,E,F)$$

$$U = \{ABCDEF\}$$

 Tìm khóa của lược đồ B3: N+ = { tập thuộc tính xác định/ được dẫn ra từ N }  $= \{B\} \subset U$ 

Có tập phụ thuộc hàm

F= 
$$\{ f1: A \rightarrow D, f2: C \rightarrow AF, f3: AB \rightarrow EC \}$$

B4.2: 
$$L = U - D.N_F^+$$
 --tim tập có khả năng làm khoá = { ABCDEF } - { DEFB } = { AC }

B5: Tìm Li từ tổ hợp các thuộc tính trong L Số lượng Li =  $n = 2^n - 1$ , Li =  $\{A, C, AC\}$ 

# Codio Ví du

B2: 
$$N = U - right(f) = \{ ABCDEF \} - \{ ACDEF \} = \{ B \}$$

quan hệ:

$$R(A,B,C,D,E,F)$$

$$U = \{ABCDEF\}$$

Có tập phụ thuộc hàm

$$F = \{ \begin{array}{l} f1 \colon A \to D \;, \\ f2 \colon C \; \to AF \;, \\ f3 \colon AB \to EC \\ \} \end{array}$$

B4.2: L = U - D.N+ --tìm tập có khả năng làm khoá = { ABCDEF } - { DEFB } = { AC }

B5: Tìm Li từ tổ hợp các thuộc tính trong L Số lượng Li =  $n = 2^n - 1$ , Li =  $\{A, C, AC\}$ Với  $Li = \{A\}$ { NLi }+<sub>F</sub> = { BADECF } = U → K1 = { BA } là khoá → Loại tất cả tập liên quan A ra khỏi Li



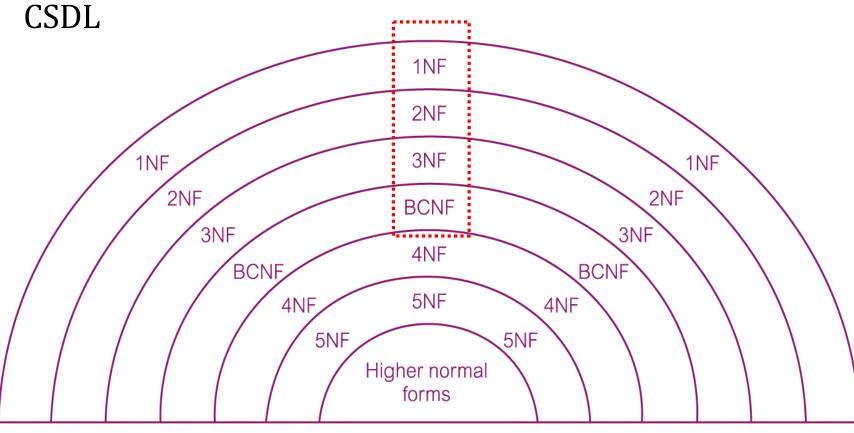
#### Nội dung chi tiết

- Các bước thiết kế CSDL quan hệ
- Sự trùng lắp thông tin
- ☐ Phụ thuộc hàm
- ☐ Khoá
- Các dạng chuẩn



#### Dạng chuẩn

Mục đích: đánh giá mức độ trùng lắp dữ liệu của lược đồ





# Dạng chuẩn 1 (DC1)

- Một quan hệ ở dạng chuẩn 1 không có các trường lặp và các trường kép, còn được gọi là cấu trúc phẳng (tất cả các giá trị tại các thuộc tính phải là giá trị nguyên tố).
- Nhận xét:
  - Mọi lược đồ quan hệ đều thỏa DC 1.
  - DC1 còn tình trạng trùng lắp dữ liệu cao, nên gây ra các bất thường về cập nhật dữ liệu.

Ví dụ: cho quan hệ THUENHA như sau

_		•		<u>.                                    </u>					
	MANT	TENNT	MANHA	DCHI_NHA	NGAYTHUE_BT	NGAYTHUE_KT	GIATHUE	MACHUNHA	TENCHUNHA
	CR76	L.V.Hùng	PG4	125 Lê Văn Sỹ, Q3	01/08/2012	01/09/2013	50tr	CO40	N.T Lan
			PG16	432 CMT8, QTB	01/09/2007	01/10/2011	150tr	CO72	B.T.Thanh
	CR56	H.V.Gia	PG4	125 Lê Văn Sỹ, Q3	01/08/2007	01/07/2012	50tr	CO40	N.T.Lan
†	Phi	dạng chuẩn	PG16	432 CMT8, QTB	01//01/2012	01/01/2013	150tr	CO72	B.T.Thanh
£			P036	124 Tô Ký, Q12	01/01/2010	01/01/2014	200tr	CO20	N.T.Phuong



# 👊 Biến đổi phi dạng chuẩn sang DC1

#### Phương pháp

- □ Phương pháp 1: Điền đầy đủ các dữ liệu vào các chỗ trống bằng dữ liệu trùng lắp  $\rightarrow$  dẫn đến nhiều dữ liệu bị trùng lắp trên quan hệ.
- Phương pháp 2: Thay thế các giá trị không nguyên tố bằng cách xác định tập thuộc tính làm khóa chính và sau đó tách thành một quan hệ mới → tạo ra hai hay nhiều quan hệ mới, và sẽ làm giảm bớt sự trùng lắp thông tin.



# Riến đổi phi dạng chuẩn sang DC1

#### Phương pháp 1:

MANT	TENNT	MANHA	DCHI_NHA	NGAYTHUE_BT	NGAYTHUE_KT	GIATHUE	MACHUNHA	TENCHUNHA
CR76	L.V.Hùng	PG4	125 Lê Văn Sỹ, Q3	01/08/2012	01/09/2013	50tr	CO40	N.T Lan
		PG16	432 CMT8, QTB	01/09/2007	01/10/2011	150tr	CO72	B.T.Thanh
CR56	H.V.Gia	PG4	125 Lê Văn Sỹ, Q3	01/08/2007	01/07/2012	50tr	CO40	N.T.Lan
		PG16	432 CMT8, QTB	01//01/2012	01/01/2013	150tr	CO72	B.T.Thanh
		PG36	124 Tô Ký, Q12	01/01/2010	01/01/2014	200tr	CO20	N.T.Phuong

MANT	TENNT	MANHA	DCHI_NHA	NGAYTHUE_BT	NGAYTHUE_KT	GIATHUE	MACHUNHA	TENCHUNHA
, CR76	L.V.Hùng	PG4	125 Lê Văn Sỹ, Q3	01/08/2012	01/09/2013	50tr	CO40	N.T Lan
CR76	L.V.Hùng	PG16 \	432 CMT8, QTB	01/09/2007	01/10/2011	150tr	CO72	B.T.Thanh
CR56	H.V.Gia	PG4	125 Lê Văn Sỹ, Q3	01/08/2007	01/07/2012	50tr	CO40	N.T.Lan
CR56	H.V.Gia	PG16 /	432 CMT8, QTB	01//01/2012	01/01/2013	150tr	CO72	B.T.Thanh
CR56	H.V.Gia	PG36	124 Tô Ký, Q12	01/01/2010	01/01/2014	200tr	CO20	N.T.Phuong



# 🕬 Biến đổi phi dạng chuẩn sang DC1

Phương pháp 2: xác định khóa và tách bảng

#### **NGUOI\_THUE (MANT, TENNT)**

MANT	TENNT
CR76	L.V.Hùng
CR56	H.V.Gia

#### NHA\_CHO\_THUE (MANT, MANHA, DCHI\_NHA, NGAYTHUE\_BT, NGAYTHUE\_KT, GIATHUE, MACHUNHA, TENCHUNHA)

MANT	MANHA	DCHI_NHA	NGAYTHUE_BT	NGAYTHUE_KT	GIATHUE	MACHUNHA	TENCHUNHA
CR76	PG4	125 Lê Văn Sỹ, Q3	01/08/2012	01/09/2013	50tr	CO40	N.T Lan
CR76	PG16	432 CMT8, QTB	01/09/2007	01/10/2011	150tr	CO72	B.T.Thanh
CR56	PG4	125 Lê Văn Sỹ, Q3	01/08/2007	01/07/2012	50tr	CO40	N.T.Lan
CR56	PG16	432 CMT8, QTB	01//01/2012	01/01/2013	150tr	CO72	B.T.Thanh
CR56	PG36	124 Tô Ký, Q12	01/01/2010	01/01/2014	200tr	CO20	N.T.Phuong



# Dạng chuẩn 2 (DC2)

- Một số khái niệm bổ sung
  - Thuộc tính khóa (không khóa): A là một thuộc tính khóa (không khóa) của Q nếu:
    - A ∈ Q+
    - A có (không có) tham gia vào một khóa của Q
  - Ví dụ 4.5: Q1(MNOPX); Q2(PY)
    - M,N,O là các thuộc tính khóa của Q1, P là thuộc tính khóa của Q2 nhưng là thuộc tính không khóa của Q1
  - Thuộc tính phụ thuộc đầy đủ (không phụ thuộc đầy đủ) vào một tập thuộc tính:
    - A là một thuộc tính phụ thuộc đầy đủ (không phụ thuộc đầy đủ) vào X nếu  $X \rightarrow A$  là một phụ thuộc hàm đầy đủ (không đầy đủ)



## Dạng chuẩn 2 (DC2)

- Một lược đồ đạt dạng chuẩn 2 nếu như lược đồ đó đã đạt DC1 và các thuộc tính không khóa phụ thuộc đầy đủ vào thuộc tính khóa.
- Nhận xét:
  - Mọi lược đồ quan hệ đạt DC2 cũng đạt DC1.
  - Nếu R chỉ có 1 khóa và khóa chỉ có 1 thuộc tính thì R đạt DC2.
  - DC2 còn xuất hiện sự trùng lặp dữ liệu nên vẫn còn xuất hiện các dị thường về cập nhật dữ liệu.
- ☐ Ví dụ: NGUOI\_THUE (MANT, TENNT)

MANT	TENNT
CR76	L.V.Hùng
CR56	H.V.Gia

Xét quan hệ NGUOI\_THUE: PK = {MANT}, F = {MANT → TENNT} NGUOI\_THUE có đạt DC2 ?

► NGUOI\_THUE: thỏa DC2



## Dạng chuẩn 2 (DC2)

## ☐ Ví dụ:

### NHA\_CHO\_THUE (MANT, MANHA, DCHI\_NHA, NGAYTHUE\_BT, NGAYTHUE\_KT, GIATHUE, MACHUNHA, TENCHUNHA)

MANT	MANHA	DCHI_NHA	NGAYTHUE_BT	NGAYTHUE_KT	GIATHUE	MACHUNHA	TENCHUNHA
CR76	PG4	125 Lê Văn Sỹ, Q3	01/08/2012	01/09/2013	50tr	CO40	N.T Lan
CR76	PG16	432 CMT8, QTB	01/09/2007	01/10/2011	150tr	CO72	B.T.Thanh
CR56	PG4	125 Lê Văn Sỹ, Q3	01/08/2007	01/07/2012	50tr	CO40	N.T.Lan
CR56	PG16	432 CMT8, QTB	01//01/2012	01/01/2013	150tr	CO72	B.T.Thanh
CR56	PG36	124 Tô Ký, Q12	01/01/2010	01/01/2014	200tr	CO20	N.T.Phuong

Xét quan hệ NHA\_CHO\_THUE:
PK = {MANT, MANHA}
F = { MANHA → DCHI\_NHA, GIATHUE, MACHUNHA, TENCHUNHA
Vậy NHA\_CHO\_THUE có thỏa DC2?

► NHA\_CHO\_THUE: không thỏa DC2 → chỉ đạt DC1



## Biến đổi DC1 sang DC2

- □ B1: Xác định khóa chính trên quan hệ bị DC1: NHA\_CHO\_THUE với PK = {MANT, MANHA}
- B2: Xác định các pth gây ra thuộc tính không khóa không phụ thuộc đầy đủ vào khóa
- MANHA → DCHI\_NHA, GIA\_THUE, MACHUNHA, TENCHUNHA
- B3: Nếu có tồn tại pth không đầy đủ trên khóa chính thì xóa chúng ở quan hệ cũ và đưa chúng vào một quan hệ mới.
  - ■Xóa pth MANHA → DCHI\_NHA, GIA\_THUE, MACHUNHA, TENCHUNHA bằng cách đưa chúng vào quan hệ mới NHA\_THUE (MANHA, DCHI\_NHA, GIA\_THUE, MACHUNHA, TENCHUNHA).
  - DĐổi tên quan hệ cũ thành TT\_THUE\_NHA(<u>MANT,MANHA</u>,
    NGAYTHUE\_BT, NGAYTHUE\_KT)



## Biến đổi DC1 sang DC2

### DC<sub>1</sub>

### NHA\_CHO\_THUE (MANT, MANHA, DCHI\_NHA, NGAYTHUE\_BT, NGAYTHUE\_KT, GIATHUE, MACHUNHA, TENCHUNHA)

MANT	MANHA	DCHI_NHA	NGAYTHUE_BT	NGAYTHUE_KT	GIATHUE	MACHUNHA	TENCHUNHA
CR76	PG4	125 Lê Văn Sỹ, Q3	01/08/2012	01/09/2013	50tr	CO40	N.T Lan
CR76	PG16	432 CMT8, QTB	01/09/2007	01/10/2011	150tr	CO72	B.T.Thanh
CR56	PG4	125 Lê Văn Sỹ, Q3	01/08/2007	01/07/2012	50tr	CO40	N.T.Lan
CR56	PG16	432 CMT8, QTB	01//01/2012	01/01/2013	150tr	CO72	B.T.Thanh
CR56	PG36	124 Tô Ký, Q12	01/01/2010	01/01/2014	200tr	CO20	N.T.Phuong

### NHA\_THUE (MANHA, DCHI\_NHA, GIATHUE, MACHUNHA, TENCHUNHA)

MANHA	DCHI_NHA	GIATHUE	MACHUNHA	TENCHUNHA
PG4	125 Lê Văn Sỹ, Q3	50tr	CO40	N.T Lan
PG16	432 CMT8, QTB	150tr	CO72	B.T.Thanh
PG36	124 Tô Ký, Q12	200tr	CO20	N.T.Phuong

### TT\_THUE\_NHA (MANT, MANHA, NGAYTHUE\_BT, NGAYTHUE\_K

MANT	MANHA	NGAYTHUE_BT	NGAYTHUE_KT
CR76	PG4	01/08/2012	01/09/2013
CR76	PG16	01/09/2007	01/10/2011
CR56	PG4	01/08/2007	01/07/2012
CR56	PG16	01//01/2012	01/01/2013
CR56	PG36	01/01/2010	01/01/2014 <sub>4.0</sub>

DC<sub>2</sub>



- Khái niệm bổ sung: Thuộc tính phụ thuộc bắc cầu vào một tập thuộc tính :
  - A phụ thuộc bắc cầu vào X nếu có 4 điều kiện sau:
  - $X \to Y \in F+ \qquad (i)$
  - $Y \to A \in F+$  (ii)

  - $\blacksquare \quad A \notin (X \cup Y) \qquad (iv)$



- Khái niệm bổ sung: Thuộc tính phụ thuộc bắc cầu vào một tập thuộc tính
- - $\square$  P có phụ thuộc bắc cầu vào NO (NO  $\rightarrow$  P)?
    - NO  $\rightarrow$  M  $\Rightarrow$  NO  $\rightarrow$ MN : thỏa (i)
    - MN → P:thỏa (ii)
    - MN  $\rightarrow$  O  $\Rightarrow$  MN  $\rightarrow$ NO : không thỏa (iii)
  - $\square$  R có phụ thuộc bắc cầu vào NO (NO  $\rightarrow$  R)?
    - NO  $\rightarrow$  MN và MN $\rightarrow$ P  $\Rightarrow$  NO  $\rightarrow$ P (i)
    - $P \rightarrow R (ii)$
    - $P \rightarrow NO \notin F+ (iii)$
    - $\mathbb{R} \notin NOP (iv)$

P không phụ thuộc bắc cầu vào NO

R phụ thuộc bắc cầu vào NO



Một lược đồ R ở dạng chuẩn 3 khi nó đạt dạng chuẩn 2 và tất cả các thuộc tính không khóa không phụ thuộc bắt cầu vào khóa

☐ Ví dụ:

NHA\_THUE (MANHA, DCHI\_NHA, GIATHUE, MACHUNHA, TENCHUNHA)

MANHA	DCHI_NHA	GIATHUE	MACHUNHA	TENCHUNHA
PG4	125 Lê Văn Sỹ, Q3	50tr	CO40	N.T Lan
PG16	432 CMT8, QTB	150tr	CO72	B.T.Thanh
PG36	124 Tô Ký, Q12	200tr	CO20	N.T.Phuong

PK = {MANHA}

F= { f1: MANHA → DCHI\_NHA,GIA\_THUE, MACHUNHA

f2: MACHUNHA → TENCHUNHA }

NHA THUE có đạt DC3?

### Ta có:

MANHA → MACHUNHA

MACHUNHA → TENCHUNHA

⇒ MANHA → TENCHUNHA (do bắt cầu)

► NHA\_THUE: không thỏa DC3, do TENCHUNHA phụ thuộc bắt cầu vào thuộc tính khóa



## Biến đổi DC2 lên DC3

- Xác định thuộc tính khóa trong quan hệ đạt dạng chuẩn 2 NHA\_THUE: PK = {MANHA}
- Xác định phụ thuộc hàm trong quan hệ gây ra thuộc tính không khóa bắt cầu vào khóa

## **MACHUNHA** → **TENCHUNHA**

- Xóa pth gây ra tính chất bắt cầu bằng cách đưa nó vào quan hệ mới.
  - Xóa pth MACHUNHA → TENCHUNHA từ quan hệ NHA\_THUE và đưa vào quan hệ mới CHU\_NHA(MACHUNHA,TEN CHUNHA)

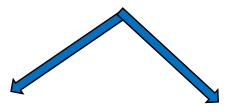


## Biến đổi DC2 lên DC3

### NHA\_THUE (MANHA, DCHI\_NHA, GIATHUE, MACHUNHA, TENCHUNHA)

### DC2

MANHA	DCHI_NHA	GIATHUE	MACHUNHA	TENCHUNHA
PG4	125 Lê Văn Sỹ, Q3	50tr	CO40	N.T Lan
PG16	432 CMT8, QTB	150tr	CO72	B.T.Thanh
PG36	124 Tô Ký, Q12	200tr	CO20	N.T.Phuong



### NHA\_THUE (MANHA, DCHI\_NHA, GIATHUE, MACHUNHA)

MANHA	MANHA DCHI_NHA		MACHUNHA
PG4	125 Lê Văn Sỹ, Q3	50tr	CO40
PG16	432 CMT8, QTB	150tr	CO72
PG36	124 Tô Ký, Q12	200tr	CO20

### CHU\_NHA(MACHUNHA, TENCHUNHA)

MACHUNHA	TENCHUNHA
CO40	N.T Lan
CO72	B.T.Thanh
CO20	N.T.Phuong

DC3

DC3



- ☐ Nhận xét:
  - Mọi lược đồ quan hệ đạt DC3 cũng đạt DC2.
  - Phụ thuộc hàm bắc cầu là nguyên nhân dẫn đến trùng lắp dữ liệu và làm cho lược đồ không đạt DC3.
  - DC3 không đạt được là do sự xuất hiện của các thuộc tính không khóa trong lược đồ suy ra được lẫn nhau.
  - DC3 là dạng chuẩn tối thiểu trong thiết kế CSDL



# Dạng chuẩn BCK (Boyce Codd Kent)

Một lược đồ chỉ đạt dạng chuẩn BCK nếu khi mỗi xác định (pth) có vế trái đều là siêu khóa/khóa

### **NGUOI\_THUE (MANT, TENNT)**

			<u> </u>	1
	MANT	TENNT		
			DOM	
	CR76	L.V.Hùng	BCK	
	CR56	H.V.Gia		
			S DC BCK	
]	MANT	→ TEN		_
				7

### TT\_THUE\_NHA (MANT, MANHA, NGAYTHUE\_BT, NGAYTHUE\_KT)

MANT	MANHA	NGAYTHUE_BT	NGAYTHUE_KT	
CR76	PG4	01/08/2012	01/09/2013	
CR76	PG16	01/09/2007	01/10/2011	BCK
CR56	PG4	01/08/2007	01/07/2012	
CR56	PG16	01//01/2012	01/01/2013	
CR56	PG36	01/01/2010	01/01/2014	

MANT, MANHA → NGAYTHUE\_BT, NGAYTHUE\_KT

### **BCK**

### NHA\_THUE (MANHA, DCHI\_NHA, GIATHUE, MACHUNHA)

MANHA	DCHI_NHA	GIATHUE	MACHUNHA
PG4	125 Lê Văn Sỹ, Q3	50tr	CO40
PG16	432 CMT8, QTB	150tr	CO72
PG36	124 Tô Ký, Q12	200tr	CO20

MANHA → DCHI\_NHA,GIA\_THUE, MACHUNHA, TENCHUNHA

### CHU\_NHA(MACHUNHA, TENCHUNHA)

MACHUNHA	TENCHUNHA	
CO40 CO72	N.T Lan B.T.Thanh	вск
CO20	N.T.Phuong	

MACHUNHA → TENCHUNHA



# Dạng chuẩn BCK (Boyce Codd Kent)

## ☐ Ví dụ:

### PHONGVAN(MAUV, NGAYPV, GIOPV, MANV, MAPHG)

MAUV	NGAYPV	GIOPV	MANV	MAPHG
CR76	13/05/2005	10:30	SG5	G101
CR56	13/05/2005	12:00	SG5	G101
CR74	13/05/2005	12:00	SG37	G102
CR56	1/07/2005	10:30	SG5	G201

```
F = { f1: MAUV, NGAYPV → GIOPV, MANV, MAPHG
f2: MANV, NGAYPV, GIOPV → MAUV
f3: MAPHG, NGAYPV, GIOPV → MAUV, MANV
f4: MANV, NGAYPV → MAPHG
}
```

```
Khóa chính = {MAUV, NGAYPV}
Khóa ứng viên = { (MANV, NGAYPV, GIOPV); (MAPHG, NGAYPV, GIOPV)}
⇒ PHONGVAN: Không đạt DC BCK, do PTH f4 có vế trái không là khóa/khóa ứng viên.
```



## Biến đổi DC3 lên DC BCK

- Xác định phụ thuộc hàm A → B của Q, trong đó A ≠ B và A không là siêu khóa/khóa.
- Phân rã quan hệ gốc Q thành hai quan hệ Q1 = {A,B}, Q2 = {tập các thuộc tính còn lại của Q } {B}
- Lặp lại qui trình trên cho Q2 đến khi không thể tiếp tục.
- Quan hệ Q1 và các Qi phân rã được từ Q2 là quan hệ cuối cùng đạt chuẩn BCK.



## Biến đổi DC3 lên DC BCK

☐ Ví dụ:

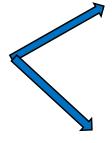
## PHONGVAN(MAUV, NGAYPV, GIOPV, MANV, MAPHG)

MAUV	NGAYPV	GIOPV	MANV	MAPHG
CR76	13/05/2005	10:30	SG5	G101
CR56	13/05/2005	12:00	SG5	G101
CR74	13/05/2005	12:00	SG37	G102
CR56	1/07/2005	10:30	SG5	G201



## NV\_PHONG (MANV,NGAYPV, MAPHG)

MANV	NGAYPV	MAPHG		
SG5	13/05/2005	G101		
SG5	13/05/2005	G101		
SG37	13/05/2005	G102		
SG5	1/07/2005	G201		



### PHONGVAN(MAUV, NGAYPV, GIOPV, MANV)

MAUV	NGAYPV	GIOPV	MANV
CR76	13/05/2005	10:30	SG5
CR56	13/05/2005	12:00	SG5
CR74	13/05/2005	12:00	SG37
CR56	1/07/2005	10:30	SG5



# Dạng chuẩn BCK

- ☐ Nhận xét:
  - Mọi lược đồ quan hệ đạt DC BCK cũng thuộc DC3.
  - Mục tiêu của quá trình thiết kế CSDL là đưa các lược đồ quan hệ về DC3 hoặc BCK.



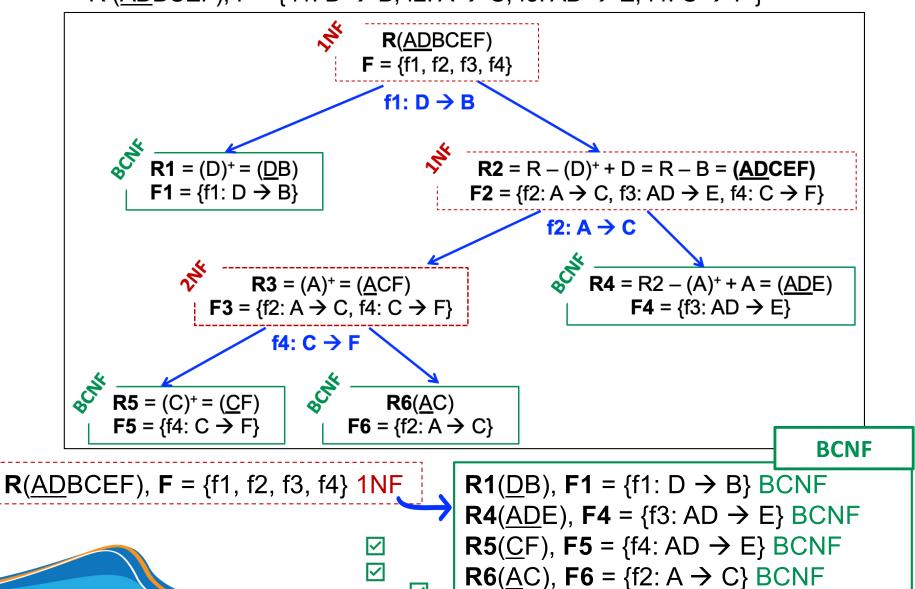
# Tóm tắt dạng chuẩn

DC	Tiêu chí	Vi phạm	Nguyên tắc chuẩn hoá
DC1	Không đa trị, không kết hợp.	Tìm thấy ít nhất một	Đa trị: Tách ra thành nhiều
(1NF)		thuộc tính bị đa trị	dòng, <u>HOẶC</u> , tách ra bảng
		hoặc kết hợp.	khác, kéo theo khoá của quan
			hệ cũ để kết bảng.
			Kết hợp: Tách thành phần
			con thành nhiều cột thuộc tính
			đơn khác nhau.
DC2	Tất cả thuộc tính không khoá phụ	Có ít nhất một X → Y	Tách tất cả các trường hợp
(2NF)	<b>thuộc đầy đủ vào khoá</b> (thuộc	với X là phần con của	X→Y vi phạm ra bảng khác
	tính không khoá không phụ thuộc	khoá và Y là không	(giữ lại X ở bảng cũ để kết
	một phần vào khoá)	khoá.	bảng).
DC3	Tất cả không khoá <b>KHÔNG PHỤ</b>	Tìm thấy ít nhất một	Tách tất cả các trường hợp
(3NF)	THUỘC bắc cầu vào bất kỳ khoá	X → Y trong đó: X	X→Y vi phạm ra (giữ lại X ở
	nào (các tập thuộc tính <mark>không</mark>	không chứa khoá <u>và</u>	bảng cũ để kết bảng).
	khoá không suy ra lẫn nhau)	Y là không phải thuộc	
		tính khoá.	
DCBC	Mọi phụ thuộc hàm vế trái phải là	Tìm thấy ít nhất một X	Tách tất cả các trường hợp
(BCN	siêu khoá (chứa bất kỳ khoá ứng	→ Y trong đó: X	X→Y vi phạm ra (giữ lại X ở
<b>F</b> )	viên nào) (các phần con của khoá	không chứa khoá.	bảng cũ để kết bảng).
	không suy ra lẫn nhau)		



## Phân rã

R (ADBCEF),  $F = \{ f1: D \rightarrow B, f2: A \rightarrow C, f3: AD \rightarrow E, f4: C \rightarrow F \}$ 



 $\overline{\mathbf{A}}$ 



## Luyện tập #1

R(ABCDEG),  $F = \{AB \rightarrow C; CD \rightarrow E; AG \rightarrow B; B \rightarrow D; A \rightarrow D\}$ 

- a) Tìm các khoá có thể có cho quan hệ R.
- b) Tìm ra các điểm trùng lắp thông tin trên R.
- c) Xác định dạng chuẩn của R.
- d) Phân rã R thành các lược đồ CSDL mới đạt ít nhất DC3.



## Luyện tập #2

R(ABCDEG),  $F = \{f_1: AD \rightarrow B; f_2: EG \rightarrow A; f_3: BC \rightarrow G\}$ 

- a) Tìm các khoá có thể có cho quan hệ R.
- b) Tìm ra các điểm trùng lắp thông tin trên R.
- c) Xác định dạng chuẩn của R.
- d) Phân rã R thành các lược đồ CSDL mới đạt ít nhất DC3.



