

Chương 7

Phụ thuộc hàm và Chuẩn hóa cơ sở dữ liệu

Chuẩn hóa lược đồ CSDL

- Chuẩn hóa là gì?
 - Các dạng chuẩn là gì?
 - Các dạng chuẩn
 - Dạng 1 (1 Normal Form - 1NF).
 - Dạng 2 (2 Normal Form - 2NF).
 - Dạng 3 (3 Normal Form - 3NF).
 - Dạng Boyce - Codd (Boyce - Codd Normal Form - BCNF).
-

Dạng chuẩn 1 (1)

- Lược đồ quan hệ R được gọi là thuộc dạng chuẩn 1 nếu và chỉ nếu mọi thuộc tính của R là thuộc tính đơn.

PHONGBAN

TenPB	<u>MaPB</u>	TrPhg	CacTruso
Nghien cuu	5	333445555	Tan Binh, Thu Duc
Hanh chinh	4	987654321	Go Vap

← Không thuộc
dạng chuẩn 1

PHONGBAN

TenPB	<u>MaPB</u>	TrPhg	<u>Truso</u>
Nghien cuu	5	333445555	Tan Binh
Nghien cuu	5	333445555	Thu Duc
Hanh chinh	4	987654321	Go Vap

← Thuộc dạng chuẩn 1

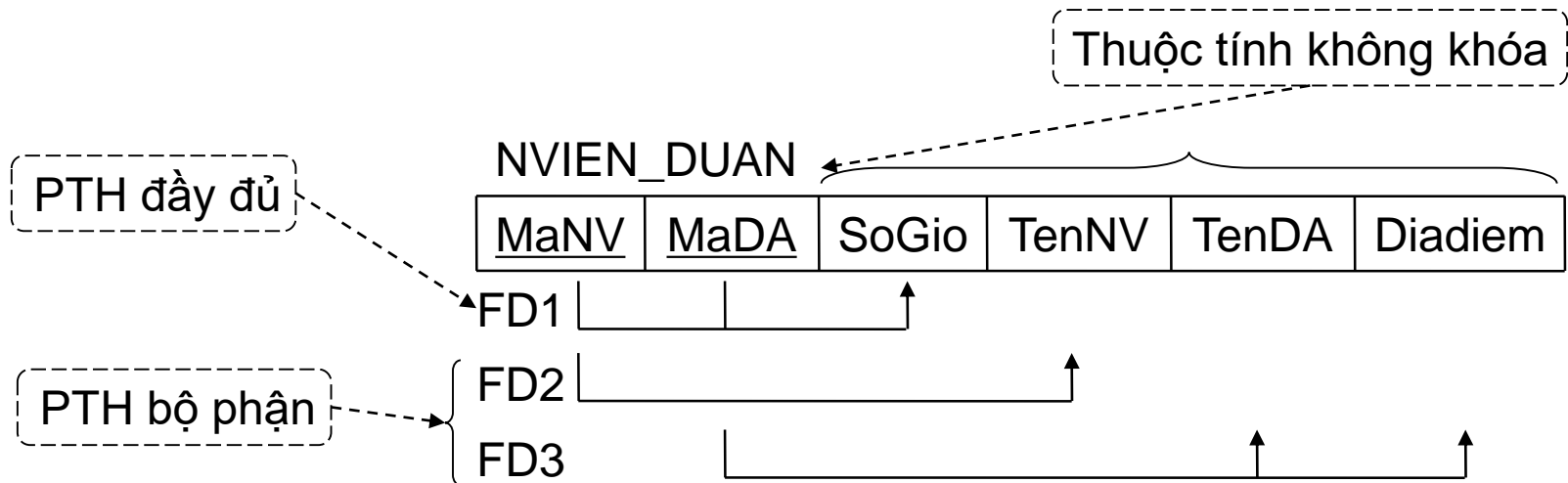
Dạng chuẩn 1 (2)

- Nhận xét

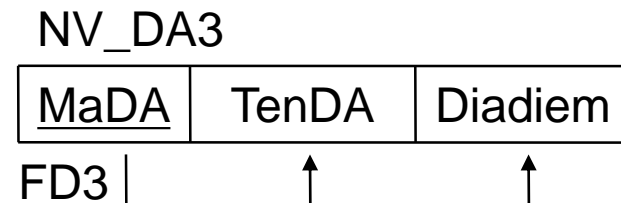
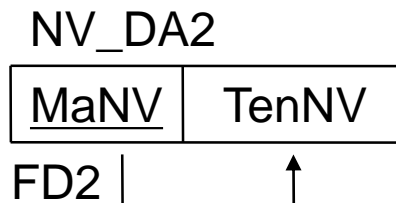
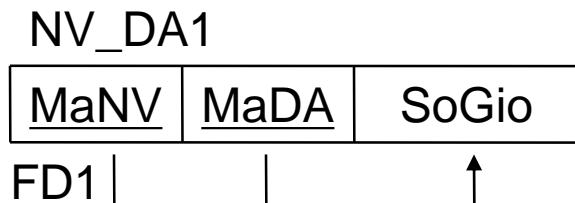
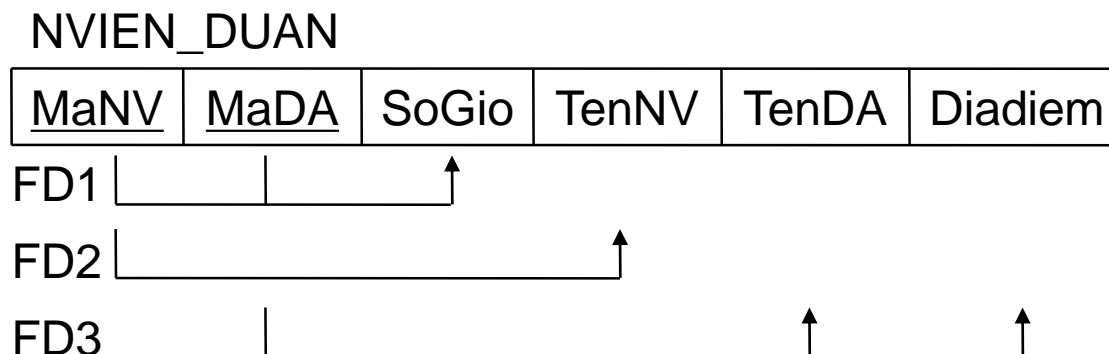
- Mọi lược đồ quan hệ đều thuộc dạng chuẩn 1.
- Dạng chuẩn 1 có thể dẫn đến sự trùng lặp dữ liệu. Do đó gây ra các dị thường về cập nhật dữ liệu.

Dạng chuẩn 2 theo khóa chính (1)

- Lược đồ quan hệ R được gọi là thuộc dạng chuẩn 2 nếu mọi thuộc tính không khóa của R phụ thuộc đầy đủ vào khóa chính của R.
- $R(U)$, $K \subseteq U$ là khóa của R
 - $A \in U$ là thuộc tính không khóa nếu $A \notin K$.
 - $X \rightarrow Y$ là PTH đầy đủ nếu $\forall A \in X$ thì $(X - \{A\}) \rightarrow Y$ không đúng trên R.
Ngược lại $X \rightarrow Y$ là PTH bộ phận.
- Ví dụ

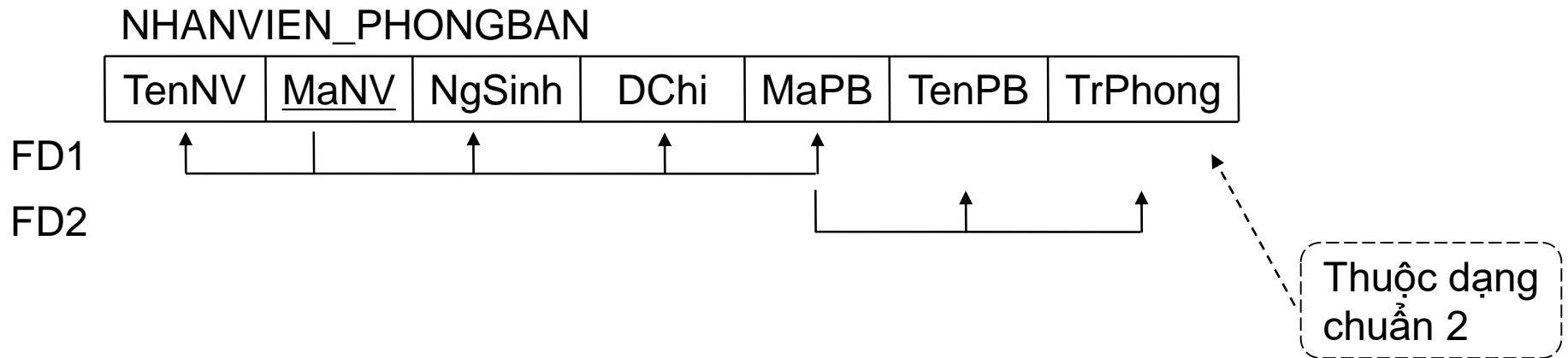


Dạng chuẩn 2 theo khóa chính (2)



3 lược đồ NV_DA1, NV_DA2, NV_DA3 thuộc dạng chuẩn 2

Dạng chuẩn 2 theo khóa chính (3)

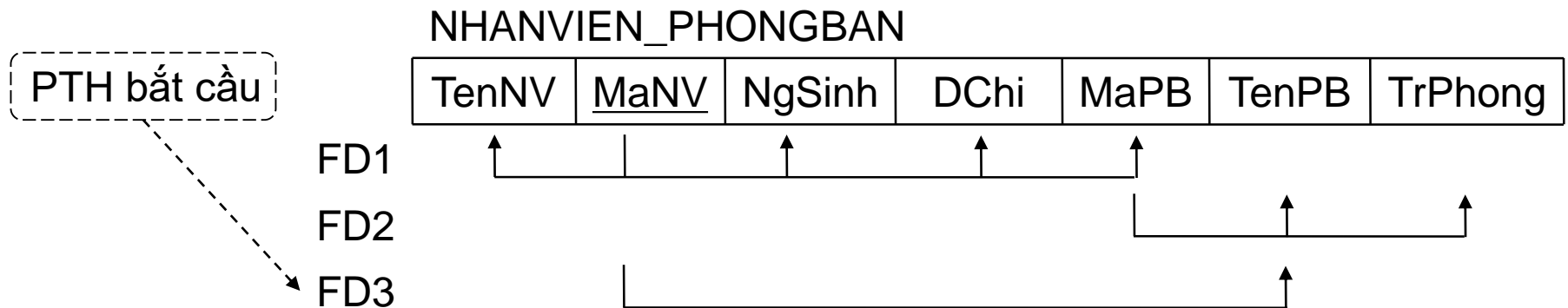


■ Nhận xét

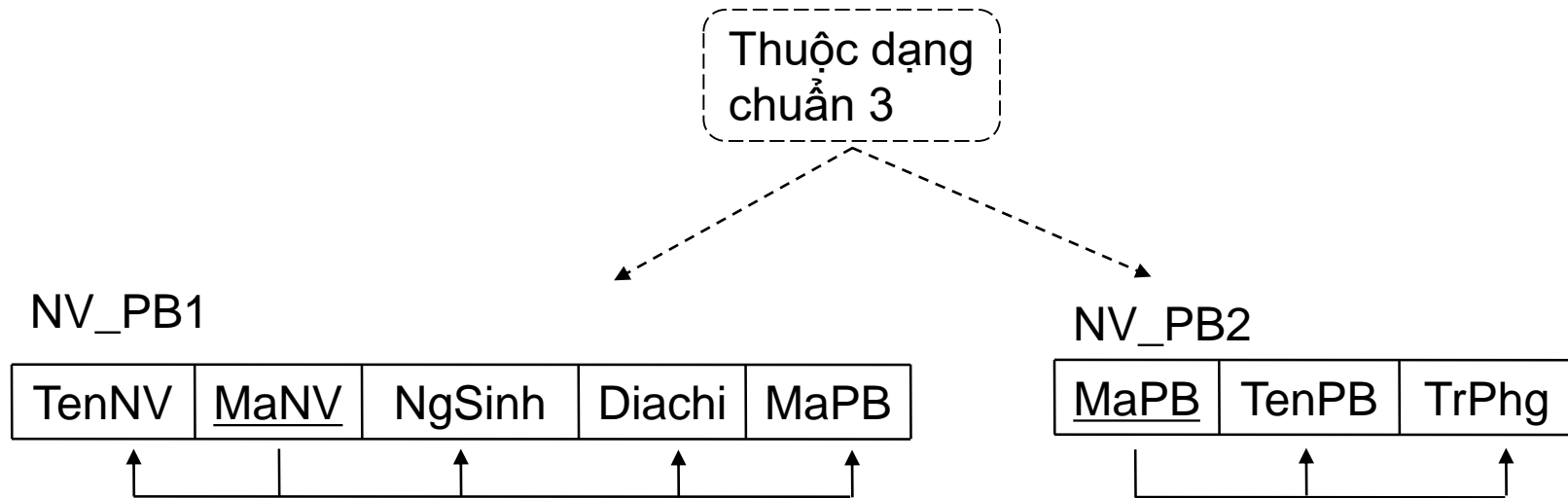
- Mọi lược đồ quan hệ thuộc dạng chuẩn 2 cũng thuộc dạng chuẩn 1.
- Nếu R chỉ có một khóa K và $\text{card}(K) = 1$ thì R thuộc dạng chuẩn 2.
- Còn xuất hiện sự trùng lặp dữ liệu. Do đó gây ra các dị thường về cập nhật dữ liệu.

Dạng chuẩn 3 theo khóa chính (1)

- Lược đồ quan hệ R được gọi là thuộc dạng chuẩn 3 nếu
 - R thuộc dạng chuẩn 2.
 - Mọi thuộc tính không khóa của R không phụ thuộc bắt cầu vào khóa chính của R.
- Cho R(U)
 - $X \rightarrow Y$ là PTH bắt cầu nếu $\exists Z \subseteq U$, Z không là khóa và cũng không là tập con của khóa của R mà $X \rightarrow Z$ và $Z \rightarrow Y$ đúng trên R.
- Ví dụ



Dạng chuẩn 3 theo khóa chính (2)

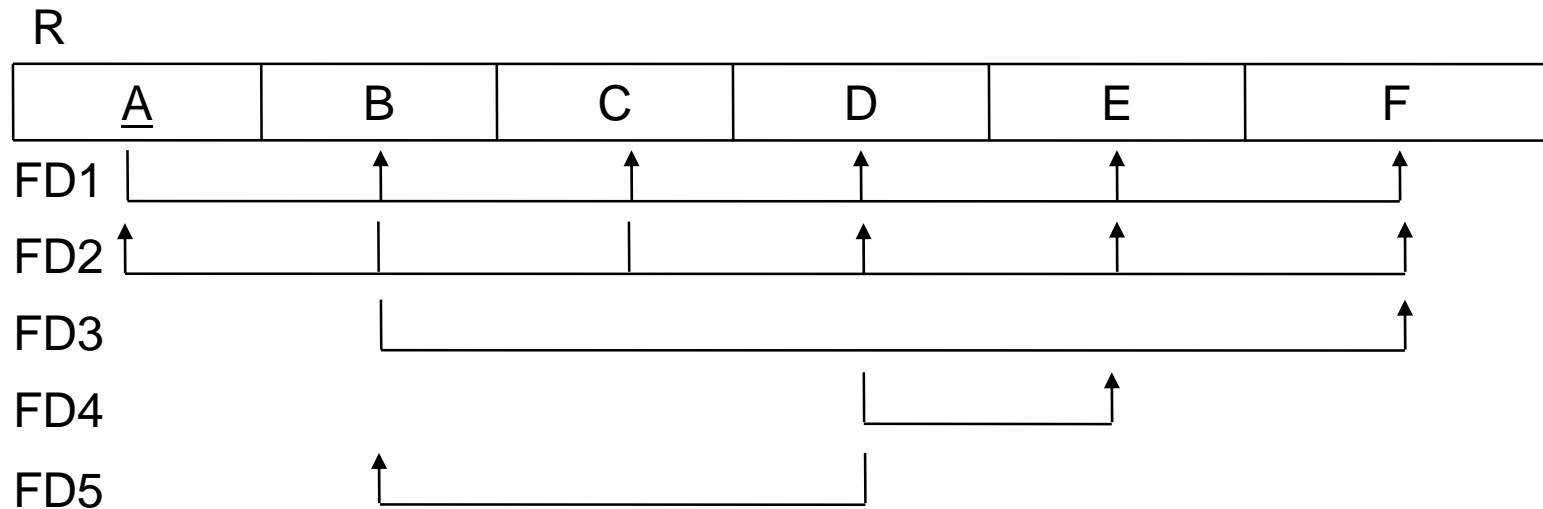


■ Nhận xét

- Mọi lược đồ quan hệ thuộc dạng chuẩn 3 cũng thuộc dạng chuẩn 2.
 - PTH bắt cầu là nguyên nhân dẫn đến trùng lặp dữ liệu.
 - Dạng chuẩn 3 là dạng chuẩn tối thiểu trong thiết kế CSDL.
-

Dạng chuẩn 2 tổng quát

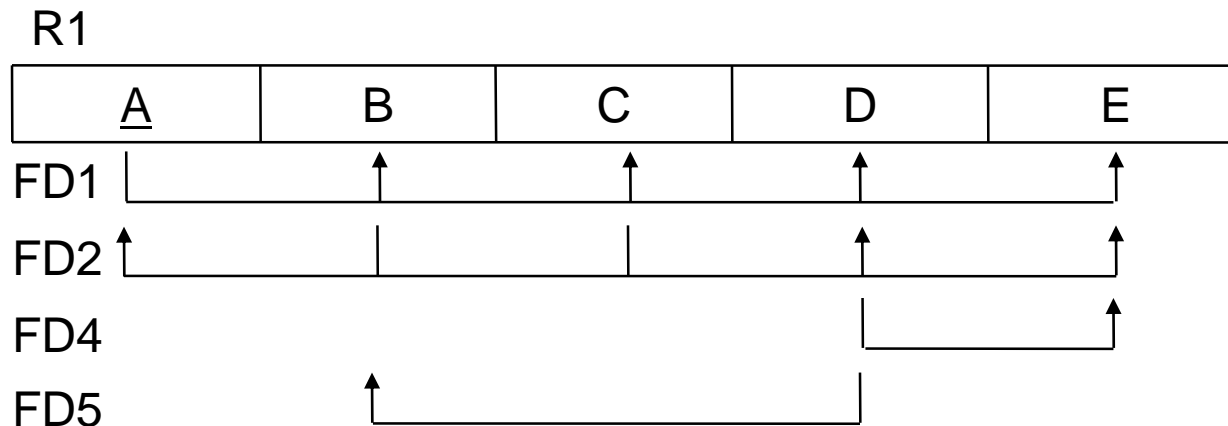
- Lược đồ quan hệ R được gọi là thuộc dạng chuẩn 2 nếu mọi thuộc tính không khóa của R phụ thuộc đầy đủ vào các khóa của R.
- Cho R(ABCDEF) có 2 khóa là A và BC.



Lược đồ R không thuộc dạng chuẩn 2

Dạng chuẩn 3 tổng quát

- Lược đồ quan hệ R được gọi là thuộc dạng chuẩn 3 nếu PTH không hiển nhiên $X \rightarrow A$ đúng trên R thì
 - X là siêu khóa của R, hoặc
 - A là thuộc tính khóa của R.
- R1(ABCDE) có 2 khóa là A và BC.

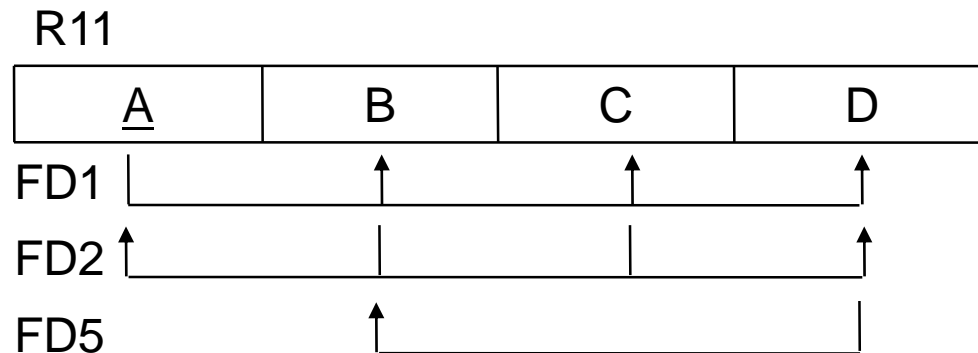


Lược đồ bên
thuộc dạng
chuẩn 2,
nhưng không
thuộc dạng
chuẩn 3

- Nhận xét
 - Định nghĩa tổng quát cho phép kiểm tra dạng chuẩn 3 mà không cần kiểm tra dạng chuẩn 2.

Dạng chuẩn Boyce - Codd (1)

- Lược đồ quan hệ R được gọi là thuộc dạng chuẩn BC nếu PTH không hiển nhiên $X \rightarrow Y$ đúng trên R thì X là siêu khóa của R.
- R11(ABCD)



Lược đồ R11 thuộc dạng chuẩn 3, nhưng không thuộc dạng chuẩn BC

Dạng chuẩn Boyce - Codd (2)

R11

<u>A</u>	B	C	D
1	a	a	1
2	a	b	1
3	b	a	2
4	b	b	2

Trùng lặp dữ liệu



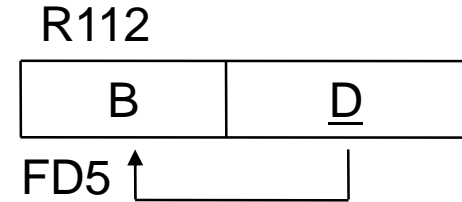
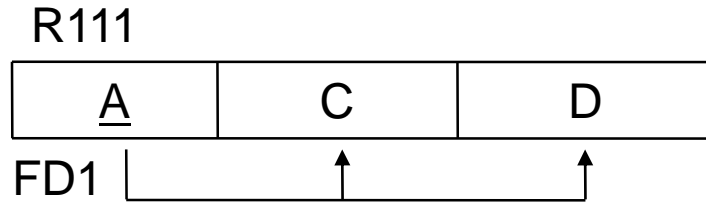
R111

<u>A</u>	C	D
1	a	1
2	b	1
3	a	2
4	b	2

R112

<u>D</u>	B
1	a
2	b

Dạng chuẩn Boyce - Codd (3)



2 lược đồ trên thuộc dạng chuẩn BC

■ Nhận xét

- Mọi lược đồ quan hệ thuộc dạng chuẩn BC cũng thuộc dạng chuẩn 3.
 - Dạng chuẩn BC đơn giản và chặt chẽ hơn dạng chuẩn 3.
 - Mục tiêu của quá trình chuẩn hóa là đưa các lược đồ quan hệ về dạng chuẩn 3 hoặc BC.
-