

Bài tập

1. Cho lược đồ quan hệ R và tập phụ thuộc hàm

$$F = \{AB \rightarrow E, AG \rightarrow I, BE \rightarrow I, E \rightarrow G, GI \rightarrow H\}$$

- a. Chứng minh $AB \rightarrow GH$ sử dụng hệ tiên đề Armstrong
- b. Chứng minh $AB \rightarrow GH$ sử dụng phương pháp tính bao đóng

Bài tập

2. Tìm phủ tối thiểu của F

a. $F = \{AD \rightarrow BC, D \rightarrow AC, AB \rightarrow C, B \rightarrow D\}$

b. $F = \{A \rightarrow BC, A \rightarrow E, A \rightarrow AB, A \rightarrow D, CD \rightarrow E\}$

Giai bai 1b (1)

Cho $F = \{ AB \rightarrow E, AG \rightarrow I, BE \rightarrow I, E \rightarrow G, GI \rightarrow H \}$.
Chứng minh $AB \rightarrow G^+$.

Vòng 1

$X^+ = \{ A, B \}$

- + Xét $AB \rightarrow E$, $AB \subseteq X^+$, thêm E vào X^+ , $X^+ = \{ A, B, E \}$
 $F = F \setminus \{ AB \rightarrow E \}$.
- + Xét $AG \rightarrow I$, AG không thuộc X^+
- + Xét $BE \rightarrow I$, BE không thuộc X^+
- + Xét $E \rightarrow G$, $E \subseteq X^+$, thêm G vào X^+ , $X^+ = \{ A, B, E, G \}$
 $F = F \setminus \{ E \rightarrow G \}$
- + Xét $GI \rightarrow H$, GI không thuộc X^+

Vậy có thay đổi xảy ra với X^+ do với hơn đầu
 \rightarrow cần một vòng lặp nữa.

Giai bai 1b (1)

Vong 2

$$X^+ = \{A, B, E, G\}, F = \{AG \rightarrow I, BE \rightarrow I, GI \rightarrow H\}$$

* Xét $AG \rightarrow I$, $AG \subseteq X^+$, thêm I vào X^+ , $X^+ = \{A, B, E, G, I\}$
 $F = F \setminus \{AG \rightarrow I\}$

* Xét $BE \rightarrow I$, $BE \subseteq X^+$, thêm I vào X^+ , $X^+ = \{A, B, E, G, I\}$
 $F = F \setminus \{BE \rightarrow I\}$

* Xét $GI \rightarrow H$, $GI \subseteq X^+$, thêm H vào X^+ , ~~$X^+ = \{A, B, E, G, I, H\}$~~
 $X^+ = \{A, B, E, G, I, H\}$, $F = F \setminus \{GI \rightarrow H\}$

Do F đã rỗng, kết thúc vòng lặp

$$\text{Vậy } (AB)^+ = \{A, B, E, G, I, H\}$$

Do đó $AB \rightarrow GH$ (điều phải chứng minh)

KI.ONG

Giai bài 2a (1)

$F = (AD \rightarrow BC, D \rightarrow AC, AB \rightarrow C, B \rightarrow D)$ xác định trên lược đồ quan hệ r .

Tìm một phủ tối thiểu của tập phụ thuộc hàm F .

* Đặt $G = F$

* Tách $G = (AD \rightarrow B, AD \rightarrow C, D \rightarrow A, D \rightarrow C, AB \rightarrow C, B \rightarrow D)$

* **Tối giản trái cho G :**

Xét Member ($G, \{A \rightarrow B\}$) có $A_G^+ = \{A\} \Rightarrow D$ ko dư thừa trong $AD \rightarrow B$

Xét Member ($G, \{D \rightarrow B\}$) có $D_G^+ = (D, A, C) \Rightarrow A$ ko dư thừa trong $AD \rightarrow B$

$\Rightarrow G = (AD \rightarrow B, AD \rightarrow C, D \rightarrow A, D \rightarrow C, AB \rightarrow C, B \rightarrow D)$

Xét Member ($G, \{A \rightarrow C\}$) có $A_G^+ = \{A\} \Rightarrow D$ không dư thừa trong $AD \rightarrow C$

Xét Member ($G, \{D \rightarrow C\}$) có $D_G^+ = (D, A, C, B) \rightarrow A$ dư thừa trong $AD \rightarrow C$

$\Rightarrow G = (AD \rightarrow B, D \rightarrow C, D \rightarrow A, D \rightarrow C, AB \rightarrow C, B \rightarrow D) = \{AD \rightarrow B, D \rightarrow C, D \rightarrow A, AB \rightarrow C, B \rightarrow D\}$

Giai bài 2a (2)

$G = \{AD \rightarrow B, D \rightarrow C, D \rightarrow A, AB \rightarrow C, B \rightarrow D\}$

Xét Member $(G, (A \rightarrow C))$ có $A^+_G = \{A\} \Rightarrow B$ không dư thừa trong $AB \rightarrow C$

Xét Member $(G, (B \rightarrow C))$ có $D^+_G = \{B, D, C, A\} \rightarrow A$ dư thừa trong $AB \rightarrow C$

$\Rightarrow G = \{AD \rightarrow B, D \rightarrow C, D \rightarrow A, D \rightarrow C, B \rightarrow C, B \rightarrow D\}$

*** Loại bỏ các phụ thuộc hàm dư thừa**

$G = \{AD \rightarrow B, D \rightarrow C, D \rightarrow A, B \rightarrow C, B \rightarrow D\};$

Xét Member $(G \setminus (AD \rightarrow B), (AD \rightarrow B))$ có $AD^+_G = \{A, D, C\} \Rightarrow \{AD \rightarrow B\}$ ko dư thừa

Xét Member $(G \setminus (D \rightarrow C), (D \rightarrow C))$ có $D^+_G = \{D, A, B, C\} \Rightarrow \{D \rightarrow C\}$ dư thừa

$\Rightarrow G = \{AD \rightarrow B, D \rightarrow A, B \rightarrow C, B \rightarrow D\}$

Xét Member $(G \setminus (D \rightarrow A), \{D \rightarrow A\})$ có $D^+_G = \{D\} \Rightarrow (D \rightarrow A)$ ko dư thừa

Xét Member $(G \setminus (B \rightarrow C), (B \rightarrow C))$ có $B^+_G = \{B, D, A\} \Rightarrow \{B \rightarrow C\}$ ko dư thừa

Xét Member $(G \setminus (B \rightarrow D), (B \rightarrow D))$ có $B^+_G = \{B, C\} \Rightarrow \{B \rightarrow D\}$ ko dư thừa

Vậy phủ tối thiểu của F là $G = \{AD \rightarrow B, D \rightarrow A, B \rightarrow C, B \rightarrow D\}$

Giai bài 2b (1)

$F = (A \rightarrow BC, A \rightarrow E, A \rightarrow AB, A \rightarrow D, CD \rightarrow E)$ Tìm một phủ tối thiểu của tập phụ thuộc hàm F

*** Đặt $G = F$**

* Tách $G = \{A \rightarrow B, A \rightarrow C, A \rightarrow E, A \rightarrow A, A \rightarrow D, CD \rightarrow E\}$

*** Tối giản trái cho G :**

Xét Member ($G, \{C \rightarrow E\}$) có $C_G^+ = \{C\} \Rightarrow D$ không dư thừa trong $CD \rightarrow E$

Xét Member ($G, \{D \rightarrow E\}$) có $D_G^+ = \{D\} \Rightarrow C$ không dư thừa trong $CD \rightarrow E$

$\Rightarrow G = \{A \rightarrow B, A \rightarrow C, A \rightarrow E, A \rightarrow A, A \rightarrow D, CD \rightarrow E\}$

*** Loại bỏ các phụ thuộc hàm dư thừa**

Xét Member ($G \setminus \{A \rightarrow B\}, \{A \rightarrow B\}$) có $A_G^+ = \{A, C, E, D\} \Rightarrow \{A \rightarrow B\}$ ko dư thừa

Xét Member ($G \setminus \{A \rightarrow C\}, \{A \rightarrow C\}$) có $A_G^+ = \{A, B, E, D\} \Rightarrow \{A \rightarrow C\}$ ko dư thừa

Xét Member ($G \setminus \{A \rightarrow E\}, \{A \rightarrow E\}$) có $A_G^+ = (A, B, C, D, E) \Rightarrow \{A \rightarrow E\}$ dư thừa

$\rightarrow G = \{A \rightarrow B, A \rightarrow C, A \rightarrow A, A \rightarrow D, CD \rightarrow E\}$

Giai bài 2b (2)

Xét Member $(G \setminus \{A \rightarrow A\}, \{A \rightarrow A\})$ có $A_G^+ = \{A, B, C, D, E\} \rightarrow \{A \rightarrow A\}$ dư thừa

$\Rightarrow G = \{A \rightarrow B, A \rightarrow C, A \rightarrow D, CD \rightarrow E\}$

Xét Member $(G \setminus \{A \rightarrow D\}, \{A \rightarrow D\})$ có $A_G^+ = \{A, B, C\} \rightarrow \{A \rightarrow D\}$ ko dư thừa

Xét Member $(G \setminus \{CD \rightarrow E\}, \{CD \rightarrow E\})$ có $CD_G^+ = \{C, D\} \Rightarrow \{CD \rightarrow E\}$ ko dư thừa

Vậy phủ tối thiểu của F là $G = \{A \rightarrow B, A \rightarrow C, A \rightarrow D, CD \rightarrow E\}$

CHUẨN HÓA LỢC ĐỒ QUAN HỆ

Giới thiệu về Chuẩn hóa dữ liệu



- ❖ Chuẩn hóa(Normalization) là **quá trình phân tích và thiết kế lại các lược đồ quan hệ (table)** sao cho dữ liệu được tổ chức hợp lý, **giảm thiểu sự dư thừa**, và **tránh các vấn đề bất thường khi cập nhật dữ liệu**.
- ❖ Các dạng chuẩn được đề xuất:
 - **1NF, 2NF, 3NF** (dựa trên phụ thuộc hàm, do Codd đề xuất)
 - **BCNF** (mạnh hơn 3NF, do Boyce & Codd đề xuất)
 - **4NF, 5NF** (dựa trên phụ thuộc đa trị và phụ thuộc nối)
- ❖ Khi chuẩn hóa, cần đảm bảo:
 - **Không mất dữ liệu khi tách bảng** (nối không mất mát)
 - **Giữ nguyên các phụ thuộc** (bảo toàn phụ thuộc)

CHUẨN HÓA DỰA TRÊN KHÓA CHÍNH

Xác định khóa cho một lược đồ quan hệ

- ❖ Định nghĩa: Nếu R là một lược đồ quan hệ với các thuộc tính A_1, A_2, \dots, A_n và một tập các phụ thuộc hàm F trong đó $X \subseteq \{A_1, A_2, \dots, A_n\}$ thì X là một khóa của R khi:
 - 1. $X \rightarrow A_1 A_2 \dots A_n \in F^+$, và
 - 2. Không có tập con thực sự nào $Y \subseteq X$ mà $Y \rightarrow A_1 A_2 \dots A_n \in F^+$

Về cơ bản, định nghĩa này cho thấy phải tạo ra bao đóng của tất cả các tập con có thể có của R và quyết định xem tập nào suy diễn ra được tất cả các thuộc tính của lược đồ.

Ví dụ: Xác định khóa

Cho $r = (C, T, H, R, S, G)$ với tập phụ thuộc hàm

$F = \{C \rightarrow T, HR \rightarrow C, HT \rightarrow R, CS \rightarrow G, HS \rightarrow R\}$

Bước 1: Tính $(A_i)^+$ với $1 \leq i \leq n$

$C^+ = \{CT\}, T^+ = \{T\}, H^+ = \{H\}$

$R^+ = \{R\}, S^+ = \{S\}, G^+ = \{G\}$

\Rightarrow Không có thuộc tính đơn nào là khóa của r

$$\binom{6}{1} = \frac{6!}{1! \times (6-1)!} = \frac{720}{120} = 6$$

Bước 2: Tính $X^+ = (A_i A_j)^+$ với $1 \leq i \leq n, 1 \leq j \leq n$

$(CT)^+ = \{C, T\},$

$(CH)^+ = \{CHTR\},$

$(CR)^+ = \{CRT\}$

$(CS)^+ = \{CSGT\},$

$(CG)^+ = \{CGT\},$

$(TH)^+ = \{THRC\}$

$(TR)^+ = \{TR\},$

$(TS)^+ = \{TS\},$

$(TG)^+ = \{TG\}$

$(HR)^+ = \{HRCT\},$

$(HS)^+ = \{HSRCTG\},$

$(HG)^+ = \{HG\}$

$(RS)^+ = \{RS\},$

$(RG)^+ = \{RG\},$

$(SG)^+ = \{SG\}$

Tập thuộc tính (HS) là một khóa của r

$$\binom{6}{2} = \frac{6!}{2! \times (6-2)!} = \frac{720}{48} = 15$$

Ví dụ: Xác định khóa

Bước 3: Tính $(A_i A_j A_k)^+$ với $1 \leq i \leq n$, $1 \leq j \leq n$, $1 \leq k \leq n$

$$(CTH)^+ = \{CTHR\},$$

$$(CTS)^+ = \{CTSG\},$$

$$(CHR)^+ = \{CHRT\},$$

$$(CHG)^+ = \{CHGTR\},$$

$$(CRG)^+ = \{CRGT\},$$

$$(THR)^+ = \{THRC\},$$

$$(THG)^+ = \{THGRC\},$$

$$(TRG)^+ = \{TRG\},$$

$$(HRS)^+ = \{HRSCTG\},$$

$$(HSG)^+ = \{HSGRCT\},$$

$$(CTR)^+ = \{CTR\}$$

$$(CTG)^+ = \{CTG\}$$

$$(CHS)^+ = \{CHSTRG\}$$

$$(CRS)^+ = \{CRSTG\}$$

$$(CSG)^+ = \{CSGT\}$$

$$(THS)^+ = \{THSRCG\}$$

$$(TRS)^+ = \{TRS\}$$

$$(TSG)^+ = \{TSG\}$$

$$(HRG)^+ = \{HRGCT\}$$

$$(RSG)^+ = \{RSG\}$$

$$\binom{6}{3} = \frac{6!}{3! \times (6-3)!} = \frac{720}{36} = 20$$

Các siêu khóa được đánh dấu màu đỏ.

Ví dụ: Xác định khóa

Bước 4: Tính $(A_i A_j A_k A_r)^+$ với $1 \leq i \leq n, 1 \leq j \leq n, 1 \leq k \leq n, 1 \leq r \leq n$

$$(CTHR)^+ = \{CTHR\},$$

$$(CTHG)^+ = \{CTHGR\},$$

$$(CHRG)^+ = \{CHRG T\},$$

$$(THRS)^+ = \{THRSCG\},$$

$$(TRSG)^+ = \{TRSG\},$$

$$(CTRS)^+ = \{CTRS\},$$

$$(CSHG)^+ = \{CSHGTR\},$$

$$(CTRG)^+ = \{CTRG\}$$

$$(CTHS)^+ = \{CTHSRG\}$$

$$(CHRS)^+ = \{CHRSTG\}$$

$$(CRSG)^+ = \{CRSGT\}$$

$$(THRG)^+ = \{THRGC\}$$

$$(HRSG)^+ = \{HRSGCT\}$$

$$(CTSG)^+ = \{CTSG\}$$

$$(THSG)^+ = \{THSGRC\}$$

$$\binom{6}{4} = \frac{6!}{4! \times (6-4)!} = \frac{720}{48} = 15$$

Các siêu khóa được đánh dấu màu đỏ.

Ví dụ: Xác định khóa

Bước 5: Tính $(A_i A_j A_k A_r A_s)^+$ với $1 \leq i \leq n$, $1 \leq j \leq n$, $1 \leq k \leq n$,
 $1 \leq r \leq n$, $1 \leq s \leq n$

$$(CTHRS)^+ = \{CTHSRG\}$$

$$(CTHRG)^+ = \{CTHGR\}$$

$$(CTHSG)^+ = \{CTHSGR\}$$

$$(CHRSR)^+ = \{CHRSRT\}$$

$$(CTRSG)^+ = \{CTRSG\}$$

$$(THRSG)^+ = \{THRSGC\}$$

Các siêu khóa được đánh dấu màu đỏ

$$\binom{6}{5} = \frac{6!}{5! \times (6-5)!} = \frac{720}{120} = 6$$

Ví dụ: Xác định khóa

Bước 6: Tính $(A_i A_j A_k A_r A_s A_t)^+$ với $1 \leq i \leq n, 1 \leq j \leq n, 1 \leq k \leq n,$
 $1 \leq r \leq n, 1 \leq s \leq n, 1 \leq t \leq n$

$$(CTHRSG)^+ = \{CTHSRG\}$$

Siêu khóa được đánh dấu màu đỏ.

$$\binom{6}{6} = \frac{6!}{6! \times (6-6)!} = \frac{720}{720} = 1$$

Với 6 thuộc tính, số trường hợp phải xét là:

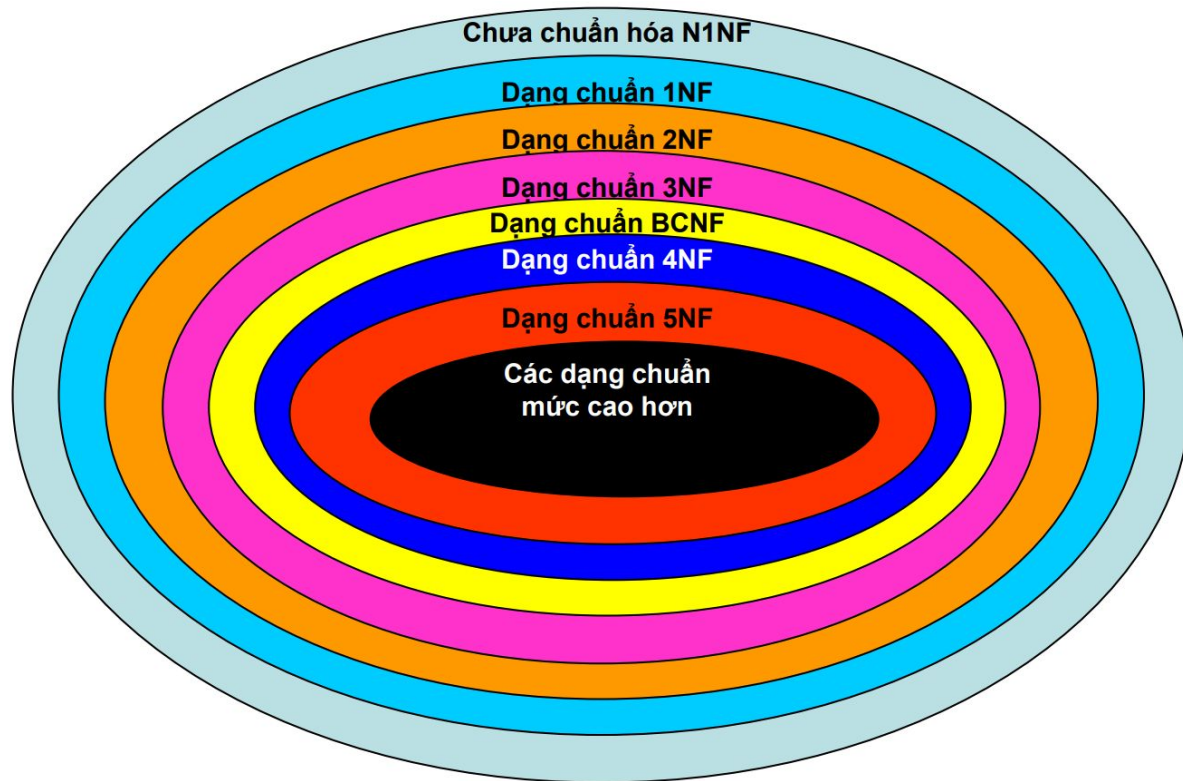
$$\binom{6}{1} + \binom{6}{2} + \binom{6}{3} + \binom{6}{4} + \binom{6}{5} + \binom{6}{6} = 6 + 15 + 20 + 15 + 1 = 63$$

Bài tập: Tìm tất cả các khóa của $R = (A, B, C, D)$ với tập phụ thuộc hàm $F = \{A \rightarrow B, B \rightarrow C\}$

Chuẩn hóa dựa trên khóa chính

- ❖ **Chuẩn hóa** là một kỹ thuật chính thức dùng cho việc phân tích các quan hệ dựa trên khóa chính (hoặc khóa dự bị), các thuộc tính và các phụ thuộc hàm.
- ❖ Kỹ thuật chuẩn hóa liên quan đến tập các luật được sử dụng để kiểm tra các quan hệ sao cho CSDL có thể đạt chuẩn hóa tới một mức nào đó.
- ❖ Khi quan hệ vi phạm một luật, cần phải **tách** nó ra thành một số các quan hệ khác nhỏ hơn.
- ❖ Việc chuẩn hóa được thực hiện bao gồm một số các bước, mỗi bước liên quan đến một dạng chuẩn cụ thể với các tính chất rõ ràng

Quan hệ giữa các dạng chuẩn



Các yêu cầu về Chuẩn hóa



- ❖ Đối với mô hình quan hệ, một vấn đề rất quan trọng và thiết yếu là phải nhận ra được một quan hệ vừa tạo ra đã ở dạng chuẩn 1 (1NF) hay chưa. Tất cả các dạng chuẩn ở mức cao hơn sau đó là tùy theo từng trường hợp, có thể có hoặc không.
- ❖ Tuy nhiên, để tránh trường hợp dị thường thông tin khi cập nhật dữ liệu, người thiết kế CSDL thường được khuyến nghị là phải đưa toàn bộ các quan hệ trong CSDL về ít nhất là dạng chuẩn 3 (3NF).
- ❖ Việc chuẩn hóa sẽ được thực hiện từ dạng chưa chuẩn hóa, đưa về dạng chuẩn 1, sau đó đưa về dạng chuẩn 2, dạng chuẩn 3, ... (đến các dạng chuẩn ở các mức cao hơn).

Dạng chưa chuẩn hóa (non-first normal form - N1NF)

- ❖ Các quan hệ ở dạng chưa chuẩn hóa đồng nghĩa với việc chúng chưa ở dạng chuẩn 1.
- ❖ **Các quan hệ chưa ở dạng chuẩn 1** chứa một hoặc một số thuộc tính không nguyên tố, các thuộc tính lặp, và các thuộc tính dẫn xuất.
- ❖ **Thuộc tính chứa giá trị nguyên tố**: là những thuộc tính chứa giá trị đơn và không thể phân rã được nữa.

Ví dụ

Chuyển đổi các thuộc tính ghép về các thuộc tính đơn.

Không nguyên tố

Họ và tên	
Họ	Tên
Nguyễn văn	Nam
Vũ Văn	Bình

nguyên tố

Họ	Tên
Nguyễn văn	Nam
Vũ Văn	Bình

Họ và tên
Nguyễn văn Nam
Vũ Văn Bình

nguyên tố

Ví dụ

Chuyển đổi các thuộc tính đa trị thành các thuộc tính đơn trị.

<u>SN</u>	Tên	Tuổi	Môn học
0012	Nam	20	C++, Java
0123	Bình	19	Java
0032	Linh	18	CSDL
0133	Trang	20	Toán, C++

<u>SN</u>	Tên	Tuổi
0012	Nam	20
0123	Bình	19
0032	Linh	18
0133	Trang	20

<u>SN</u>	Môn học
0012	C++, Java
0123	Java
0032	CSDL
0133	Toán, C++

Dạng Chuẩn 1NF



1. Mọi giá trị thuộc tính của quan hệ đều ở dạng **nguyên tố**.
2. Không có thuộc tính **đa trị**.
3. Không có thuộc tính **dẫn xuất**.

Dạng Chuẩn 1NF

- ❖ Giả sử K là khóa của lược đồ quan hệ R và A là thuộc tính đa trị hoặc là nhóm thuộc tính lặp
- ❖ Để chuẩn hóa R về 1NF:
 - Loại A ra khỏi R
 - Tạo một lược đồ quan hệ mới $S(K, A)$ với khóa chính của S là:
 - $\{K, A\}$ nếu A là thuộc tính đa trị
 - $\{K, K_A\}$ nếu A là nhóm thuộc tính lặp, là khóa bộ phận của A

Ví dụ Chuẩn 1NF

1. Lược đồ **DONVI**(MaDV, TenDV, MaNQL, Diadiem), với Diadiem là thuộc tính đa trị:
 - Loại Diadiem ra khỏi DONVI: DONVI1(MaDV, TenDV, MaNQL)
 - Tạo lược đồ quan hệ mới DIADIEMDV(MaDV, Diadiem)
2. Lược đồ **NHANVIEN_LAMVIEC**(MaDA, TenDA, MaNV, Sogio), trong đó (MaNV, Sogio) là nhóm thuộc tính lặp
 - Loại (MaNV, Sogio) ra khỏi NHANVIEN_LAMVIEC:
NHANVIEN_LAMVIEC1(MaDA, TenDA)
 - Tạo lược đồ quan hệ mới NHANVIEN_SOGIO(MaDA, MaNV, Sogio)

Dạng Chuẩn 2NF

- ❖ **Dạng chuẩn 2 dựa trên khái niệm phụ thuộc hàm đầy đủ.**
- ❖ **Một phụ thuộc hàm $X \rightarrow Y$ là một phụ thuộc hàm đầy đủ** nếu bỏ đi bất kỳ một thuộc tính $A \in X$ thì phụ thuộc hàm đó không còn đúng nữa
- ❖ Nếu tồn tại một thuộc tính B sao cho $(X - B) \rightarrow Y$, chúng ta nói rằng Y phụ thuộc hàm bộ phận vào X

Dạng Chuẩn 2NF

- ❖ Một lược đồ quan hệ R là ở **Dạng chuẩn 2 (2NF - second normal form)** nếu:
 - R là 1NF, và
 - Mỗi thuộc tính không khóa A trong R phụ thuộc hàm đầy đủ vào khóa chính của R

Dạng Chuẩn 2NF

- ❖ Giả sử R là lược đồ quan hệ ở dạng chuẩn 1NF và không là 2NF, có khóa K
- ❖ Để chuẩn hóa R về 2NF, đối với mỗi nhóm thuộc tính **không khóa** X phụ thuộc hàm bộ phận vào K:
 - Loại X ra khỏi R
 - Gọi khóa bộ phận của K xác định hàm X là $K_x : K_x \rightarrow X$
 - Tạo lược đồ quan hệ mới $S(\underline{K_x}, X)$ với K_x là khóa chính

Ví dụ Dạng Chuẩn 2NF

- ❖ Lược đồ quan hệ ở dạng chuẩn 1NF
NHANVIEN_DUAN(MaNV, MaDA, Sogio, Hoten, TenDA, Diadiem) với các phụ thuộc hàm:
 - {MaNV, MaDA} → {Sogio, Hoten, TenDA, Diadiem}
 - MaNV → Hoten (phụ thuộc hàm bộ phận)
 - MaDA → {TenDA, Diadiem} (phụ thuộc hàm bộ phận)
- ❖ **NHANVIEN_DUAN** không là 2NF do Hoten, TenDA và Diadiem phụ thuộc hàm bộ phận vào khóa chính {MaNV, MaDA}
- ❖ Chuẩn hóa về 2NF:
 - Loại Hoten, TenDA, Diadiem ra khỏi NHANVIEN_DUAN:
NHANVIEN_DUAN1(MaNV, MaDA, Sogio)
 - Tạo lược đồ quan hệ mới 1: N1(MaNV, Hoten)
 - Tạo lược đồ quan hệ mới 2: N2(MaDA, TenDA, Diadiem)

Dạng Chuẩn 3NF

- ❖ Dạng chuẩn 3 dựa trên khái niệm phụ thuộc hàm bậc cầu.
- ❖ Một phụ thuộc hàm $X \rightarrow Y$ trong một lược đồ quan hệ R là một **phụ thuộc hàm bậc cầu** nếu tồn tại một tập hợp thuộc tính Z của R không phải là khóa và không phải là tập con của khóa sao cho $X \rightarrow Z$ và $Z \rightarrow Y$ đều đúng
- ❖ 1 lược đồ quan hệ R là ở **dạng chuẩn 3 (3NF - third normal form)** nếu:
 - R là 2NF, và
 - Không có thuộc tính không khóa nào của R phụ thuộc hàm bậc cầu vào khóa chính

Dạng Chuẩn 3NF

- ❖ Giả sử R là lược đồ quan hệ có khóa chính K; R ở dạng chuẩn 2NF và không ở dạng chuẩn 3NF
- ❖ Để chuẩn hóa R về 3NF, với mỗi thuộc tính X phụ thuộc bắc cầu vào K:
 - Loại bỏ X ra khỏi R
 - Gọi Y là thuộc tính bắc cầu, ta có: $K \rightarrow Y$ và $Y \rightarrow X$
 - Tạo lược đồ quan hệ mới S(Y, X) với Y là khóa chính

Ví dụ Dạng Chuẩn 3NF

- ❖ **LAMVIEC**(MaNV, Hoten, Ngaysinh, MaDV, TenDV, MaDA, TenDA) với các thuộc tính đơn, ở dạng chuẩn 2NF vì khóa chính chỉ có một thuộc tính (→ không có phụ thuộc hàm bộ phận), có các phụ thuộc hàm:
 - $\text{MaNV} \rightarrow \{\text{Hoten}, \text{Ngaysinh}, \text{MaDV}, \text{TenDV}, \text{MaDA}, \text{TenDA}\}$
 - $\text{MaDV} \rightarrow \text{TenDV}$
 - $\text{MaDA} \rightarrow \text{TenDA}$
- ❖ LAMVIEC **không ở dạng chuẩn 3NF** vì các phụ thuộc hàm **bắc cầu**
 - $\text{MaNV} \rightarrow \text{MaDV} \rightarrow \text{TenDV}$ và $\text{MaNV} \rightarrow \text{MaDA} \rightarrow \text{TenDA}$
- ❖ Chuẩn hóa LAMVIEC về 3NF:
 - Loại $\{\text{TenDV}, \text{TenDA}\}$ phụ thuộc bắc cầu vào khóa chính:
 $\text{LAMVIEC}_1(\text{MaNV}, \text{Hoten}, \text{Ngaysinh}, \text{MaDV}, \text{MaDA})$
 - Tạo lược đồ quan hệ mới N1(MaDV, TenDV)
 - Tạo lược đồ quan hệ mới N2(MaDA, TenDA)

Ví dụ Dạng Chuẩn 3NF

- ❖ Cho $R = (A, B, C, D)$, khóa $K = \{A, B\}$ tập phụ thuộc hàm $F = \{AB \rightarrow C, C \rightarrow D\}$
- ❖ Kiểm tra dạng Chuẩn 3NF
 - R không là dạng chuẩn 3NF vì D là một thuộc tính không khóa mà lại phụ thuộc bắc cầu vào khóa AB.

Thực hiện tách về dạng chuẩn 3NF

- $R_1 = \{A, B, C\}$ với khóa $K_1 = \{A, B\}$, tập phụ thuộc hàm là $\{AB \rightarrow C\}$
- $R_2 = \{C, D\}$ với khóa $K_2 = \{C\}$, tập phụ thuộc hàm là $\{C \rightarrow D\}$

Ví dụ Dạng Chuẩn 3NF

- ❖ Cho $R = (A, B, C, D)$, và tập phụ thuộc hàm $F = \{AB \rightarrow C, C \rightarrow B, A \rightarrow D\}$
- ❖ Khóa $K = \{A, B\}$
 - R không là dạng chuẩn 2NF vì $\{A \rightarrow D\}$ là phụ thuộc một phần

Thực hiện tách về dạng chuẩn 2NF

- $R_1 = \{A, B, C\}$ với khóa $K_1 = \{A, B\}$, tập phụ thuộc hàm là $\{AB \rightarrow C, C \rightarrow B\}$
- $R_2 = \{A, D\}$ với khóa $K_2 = \{A\}$, tập phụ thuộc hàm là $\{A \rightarrow D\}$

Kiểm tra chuẩn 3NF \Rightarrow OK

Dạng Chuẩn **Boyce-Codd**

- ❖ Một lược đồ quan hệ R là **ở dạng chuẩn Boyce-Codd (BCNF - Boyce-Codd normal form)** nếu
 - ❖ R là 3NF, và
 - ❖ Không có **thuộc tính khóa** phụ thuộc hàm vào **thuộc tính không khóa** trong R

Dạng Chuẩn Boyce-Codd (BCNF)

- ❖ Giả sử R là lược đồ quan hệ có khóa K, R là 3NF và không là BCNF
- ❖ Để chuẩn hóa R về BCNF, đối với mỗi thuộc tính không khóa A xác định hàm thuộc tính khóa $A \rightarrow KS$, với $KS \subset K$:
 - Loại KS ra khỏi R và bổ sung A vào khóa chính của R. Khóa chính mới của R là: $K - KS \cup \{A\}$
 - Tạo lược đồ mới $S(\underline{A}, KS)$ với A là khóa chính

Ví dụ Dạng Chuẩn Boyce-Codd (BCNF)

- ❖ Lược đồ $R(\underline{A1}, \underline{A2}, A3, A4, A5, A6)$ có các phụ thuộc hàm:
 - $\{A1, A2\} \rightarrow \{A3, A4, A5, A6\}$: $\{A1, A2\}$ là khóa chính
 - $A4 \rightarrow A2$
 - $A6 \rightarrow A1$
- ❖ R là 2NF (không có phụ thuộc hàm bộ phận) và 3NF (không có phụ thuộc hàm bắc cầu từ khóa chính)
- ❖ R **không là BCNF** vì có **thuộc tính khóa** phụ thuộc hàm **thuộc tính không khóa**
- ❖ Chuẩn hóa về BCNF:
 - Loại **A2** ra khỏi R và bổ sung **A4** vào khóa chính: $R(\underline{A1}, \underline{A4}, A3, A5, A6)$
 - Tạo quan hệ mới $S1(\underline{A4}, A2)$ với **A4** là khóa chính
 - Loại **A1** ra khỏi R và bổ sung **A6** vào khóa chính: $R(\underline{A6}, \underline{A4}, A3, A5)$
 - Tạo quan hệ mới $S2(\underline{A6}, A1)$ với **A6** là khóa chính

Tổng kết về Chuẩn hóa



- ❖ Chuẩn hóa về 1NF: loại bỏ dữ liệu dư thừa
- ❖ Chuẩn hóa về 2NF: loại bỏ các phụ thuộc hàm bộ phận
- ❖ Chuẩn hóa về 3NF: loại bỏ các phụ thuộc hàm bắc cầu

Ví dụ

❖ Cho 1 tập phụ thuộc hàm F, có khóa là {MãNV, MãTruSo}

{MãNV, MãTruSo, MãPhongHop, TênPhongHop, TênNV, TênTruSo, ThoiGianSuDung}

- $\text{MãPhongHop} \rightarrow \text{MãNV}$
- $\text{MãPhongHop} \rightarrow \text{TênPhongHop}$
- $\text{MãNV} \rightarrow \text{TênNV}$
- $\text{MãTruSo} \rightarrow \text{TênTruSo}$
- $\{\text{MãNV}, \text{MãTruSo}\} \rightarrow \{\text{ThoiGianSuDung}, \text{MãPhongHop}, \text{TênPhongHop}\}$

Hãy chuẩn hóa về dạng BCNF

Ví dụ

F {MãNV, MãTruSo, MãPhongHop, TênPhongHop, TênNV, TênTruSo, ThoiGianSuDung}

Kiểm tra 2NF \Rightarrow Chưa thỏa mãn 2 NF do:

- TênNV phụ thuộc một phần vào khóa chính
- TênTruSo phụ thuộc một phần vào khóa chính

$R1 = \{\underline{\text{MãNV}}, \underline{\text{MãTruSo}}, \text{ThoiGianSuDung}, \text{MãPhongHop}, \text{TenPhongHop}\},$

$F1 = \{\text{MãNV}, \text{MãTruSo} \rightarrow \{\text{MãPhongHop}, \text{TênPhongHop}, \text{ThoiGianSuDung}\},$
 $\{\text{MãPhongHop} \rightarrow \text{TênPhongHop}\}$

$R2 = \{\underline{\text{MãNV}}, \text{TênNV}\}, K2 = \{\text{MãNV}\}, F2 = \{\text{MãNV} \rightarrow \text{TênNV}\}$

$R3 = \{\text{MãTruSo}, \text{TênTruSo}\}, K3 = \{\text{MãNV}\}, F3 = \{\text{MãTruSo} \rightarrow \text{TênTruSo}\}$

Ví dụ

F {MãNV, MãTruSo, MãPhongHop, TênPhongHop, TênNV, TênTruSo, ThoiGianSuDung}

Kiểm tra 3NF \Rightarrow Chưa thỏa mãn 3NF do:

- TênPhongHop phụ thuộc hàm bắc cầu khóa chính thông qua MãPhongHop

R1 = {MãNV, MãTruSo, MãPhongHop},

F1 = {MãNV, MãTruSo} \rightarrow {MãPhongHop, ThoiGianSuDung}

R2 = {MãNV, TênNV}, K2 = {MãNV}, F2 = {MãNV \rightarrow TênNV}

R3 = {MãTruSo, TênTruSo}, K3 = {MãNV}, F3 = {MãTruSo \rightarrow TênTruSo}

R4 = {MãPhongHop, TênPhongHop} F4 = {MãPhongHop \rightarrow TênPhongHop}

Ví dụ

F {MãNV, MãTruSo, MãPhongHop, TênPhongHop, TênNV, TênTruSo, ThoiGianSuDung}

Kiểm tra BCNF \Rightarrow Chưa thỏa mãn BCNF do:

- MãNV là thuộc tính khóa, nhưng phụ thuộc vào thuộc tính không khóa là MãPhongHop

$R2 = \{\underline{\text{MãNV}}, \text{TênNV}\}, F2 = \{\text{MãNV} \rightarrow \text{TênNV}\}$

$R3 = \{\underline{\text{MãTruSo}}, \text{TênTruSo}\}, F3 = \{\text{MãTruSo} \rightarrow \text{TênTruSo}\}$

$R4 = \{\underline{\text{MãPhongHop}}, \text{TênPhongHop}\}, F4 = \{\text{MãPhongHop} \rightarrow \text{TênPhongHop}\}$

$R5 = \{\text{MãPhongHop}, \text{MãTruSo}, \text{ThoiGianSuDung}\}$

$F5 = \text{MãPhongHop}, \text{MãTruSo}, \text{ThoiGianSuDung}$

$R6 = \{\underline{\text{MãPhongHop}}, \text{MãNV}\}, F6 = \{\text{MãPhongHop} \rightarrow \text{MãNV}\}$

Ví dụ

F {MãNV, MãTruSo, MãPhongHop, TênPhongHop, TênNV, TênTruSo, ThoiGianSuDung}

Gộp R4, R6 \Rightarrow R46

R2 = {MãNV, TênNV}, F2 = {MãNV \rightarrow TênNV}

R3 = {MãTruSo, TênTruSo}, F3 = {MãTruSo \rightarrow TênTruSo}

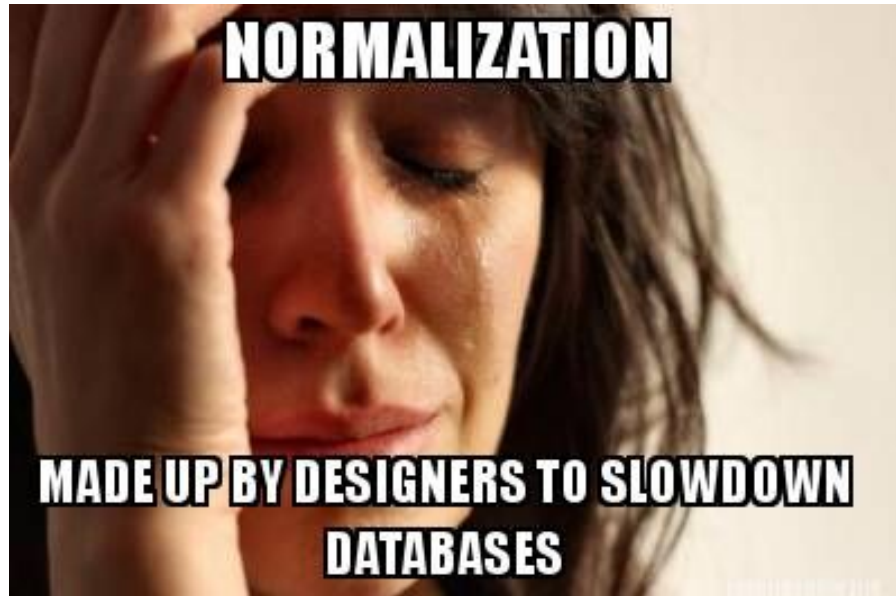
R5 = {MãPhongHop, MãTruSo, ThoiGianSuDung}

F5 = MãPhongHop, MãTruSo, ThoiGianSuDung

R46 = {MãPhongHop, MãNV, TênPhongHop}

F46 = {MãPhongHop \rightarrow MãNV, MãPhongHop \rightarrow TênPhongHop}

Chuẩn hóa có thể giảm hiệu năng



Khử chuẩn hóa (Denormalization)



- ❖ Denormalization sẽ phải đánh đổi một phần về *Toàn vẹn dữ liệu*
- ❖ Denormalization không hẳn là 1 CSDL không được chuẩn hóa
- ❖ Denormalization có thể được thực hiện **sau khi** đã **chuẩn hóa** CSDL
- ❖ Một số ưu điểm của Denormalization:
 - **Tăng tốc độ truy vấn** khi giảm quá trình JOIN, thường áp dụng cho các truy vấn READ (đọc) lặp lại thường xuyên và phức tạp
 - **Thống kê nhanh** bằng cách lưu lại các giá trị như tổng (sum), trung bình (avg) thẳng vào database để giảm thời gian tính toán lại các thống kê này

Khử chuẩn hóa (Denormalization)

Customers Table:

CustomerID	Name	Address
101	Sam	New York
102	Jim	Canada

Normalized Orders Table:

OrderID	CustomerID	Product	Quantity	Price_in_Dollars
1	101	Laptop	1	1100
2	101	Mouse	2	25
3	102	Keyboard	1	50

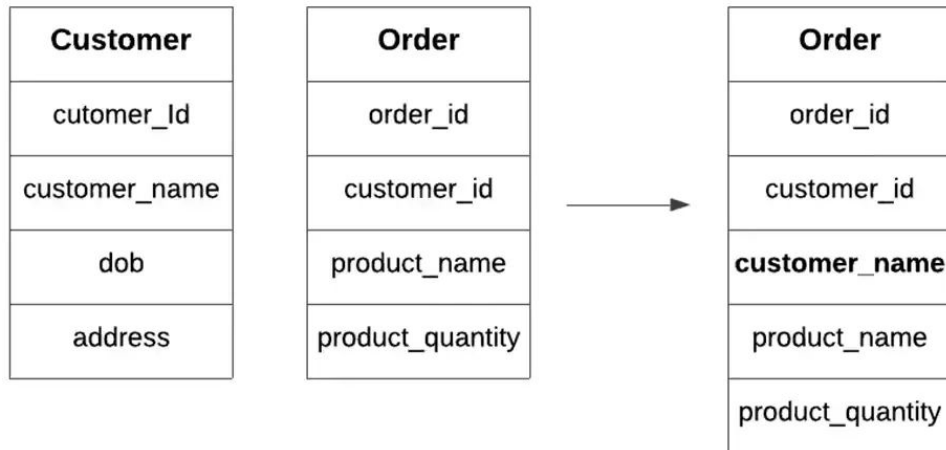
Denormalized Orders Table:

OrderID	CustomerID	Name	Product	Quantity	Price_in_Dollars
1	101	Sam	Laptop	1	1100
2	101	Sam	Mouse	2	25
3	102	Jim	Keyboard	1	50

6 phương pháp Denormalization

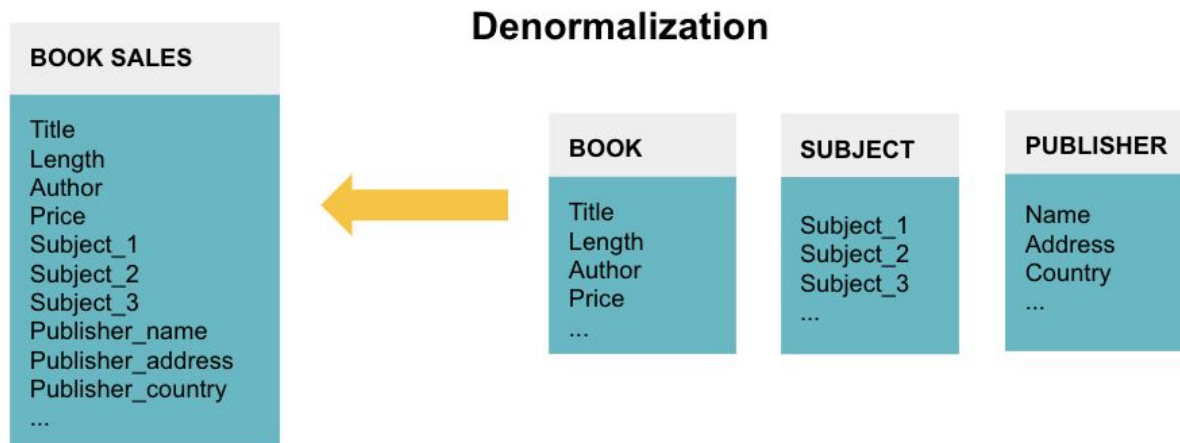
Denormalization 01: Pre-Joining Tables

- ❖ Sử dụng khi việc JOIN mất nhiều công sức tính toán
- ❖ Thực hiện lắp lại các cột cụ thể từ *các bảng đã được chuẩn hóa*
- ❖ Ví dụ, thông tin tên của Customer sẽ được copy sang bảng Order



Denormalization 02: Mirrored Tables

- ❖ Thực hiện tạo ra 1 bảng copy một phần hoặc toàn bộ từ nhiều bảng khác
- ❖ Thay vì lưu 3 bảng book, subject, publisher, thì giờ chỉ lưu chung vào 1 bảng



Denormalization 03: Tách bảng

- ❖ **Tách theo chiều ngang:** Tách dựa trên logic về khía cạnh nào đó như vùng miền, khoa, thời gian, ...

Student_Id	Name	Department	Age	Address
1000	George	Computer Science	21	New York
1001	Alex	Civil	20	Canada
1002	James	Computer Science	22	Maple
1003	Samuel	Engineering	23	Los Angeles

Computer Science Students Table:

Student_Id	Name	Department	Age	Address
1000	George	Computer Science	21	New York
1002	James	Computer Science	22	Maple

Civil Students Table:

Student_Id	Name	Department	Age	Address
1001	Alex	Civil	20	Canada

Mechanical Students Table:

Student_Id	Name	Department	Age	Address
1003	Samuel	Mechanical	23	Los Angeles

Denormalization 03: Tách bảng

- ❖ **Tách theo chiều dọc:** Sử dụng khi có một số trường cần sử dụng thường xuyên hơn so với một số trường khác

Patient_Id	Name	Address	Medical_History
2000	Jacob	New York	Diabetes, Hypertension
2001	Zerah	Canada	Asthma
2002	Savio	Maple	No known conditions

Patient_Details Table:

Patient_Id	Name	Address
2000	Jacob	New York
2001	Zerah	Canada
2002	Savio	Maple

truy cập
thường
xuyên

Patient_Medical_History Table:

Patient_Id	Medical_History
2000	Diabetes, Hypertension
2001	Asthma
2002	No known conditions

Denormalization 04: Thêm thuộc tính dẫn xuất

- ❖ **Tính trước thuộc tính dẫn xuất và bổ sung vào bảng:** Tránh phải tính lại giá trị này trong mỗi truy vấn trong tương lai

Student Table:

Student_Id	Name	Age	Department	Total_Marks
1001	Allen	21	Computer Science	175
1002	Smith	22	Mechanical	160
1003	Bailey	20	Civil	179

Student_Grades Table:

Student_Id	Assignment_Name	Marks
1001	Assignment_1	85
1001	Assignment_2	90
1002	Assignment_1	78
1002	Assignment_2	82
1003	Assignment_1	88
1003	Assignment_2	91

Denormalization 05: Tạo ra Materialized View

- ❖ **Materialized View** tính toán trước các kết quả của truy vấn thường xuyên sử dụng, sau đó lưu vào 1 bảng **thật** trong database

```
CREATE MATERIALIZED VIEW Product_Sales_View AS
SELECT p.Product_Name, SUM(o.Quantity *
p.Price_in_Dollars) AS Total_Sales
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
GROUP BY p.Product_Name;
```

Denormalization 06: Lưu dư thừa khóa ngoại



A (cha)

└─ **B (con của A)**

└─ **C (con của B, và lưu luôn khóa ngoại A)**

- ❖ Một School có nhiều Class.
- ❖ Một Class có nhiều Student.
- ❖ Nhưng trong bảng Student, lưu luôn:
 - class_id (khóa ngoại về Class)
 - school_id (khóa ngoại thừa về School)