

Khoa Toán, Đại học Khoa học Huế

Cơ sở dữ liệu

Database

TS. Nguyễn Hoàng Sơn

Nội dung môn học

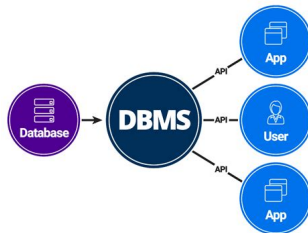
- ① Tổng quan
- ② Các mô hình dữ liệu
- ③ Phụ thuộc hàm
- ④ Khoá lược đồ quan hệ
- ⑤ Chuẩn hoá lược đồ quan hệ
- ⑥ Tài liệu tham khảo

Nội dung môn học

- ➊ Tổng quan
- ➋ Các mô hình dữ liệu
- ➌ Phụ thuộc hàm
- ➍ Khoá lược đồ quan hệ
- ➎ Chuẩn hoá lược đồ quan hệ
- ➏ Tài liệu tham khảo

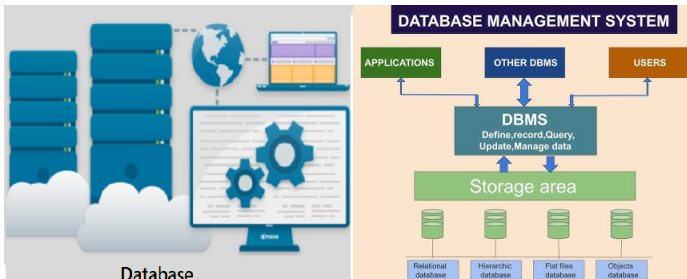
Chương 1. Tổng quan

- 1 Hệ cơ sở dữ liệu
- 2 Kiến trúc ba mức trừu tượng của hệ CSDL
- 3 Lược đồ và thể hiện của CSDL
- 4 Sự độc lập của DL
- 5 Mô hình DL
- 6 Hệ quản trị CSDL
- 7 Quy trình thiết kế CSDL
- 8 Vai trò con người trong hệ CSDL



Hệ cơ sở dữ liệu

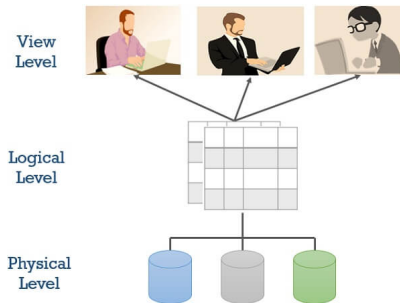
- ① Cơ sở dữ liệu (Database)
- ② Hệ quản trị CSDL (DB Management System - DBMS)
- ③ Hệ CSDL (DB System)



Database System

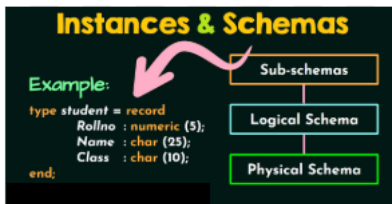
Kiến trúc ba mức trừu tượng

- ANSI-PARC:
 - ① Mức vật lý
 - ② Mức logic
 - ③ Mức khung nhìn



Lược đồ và thể hiện: Định nghĩa

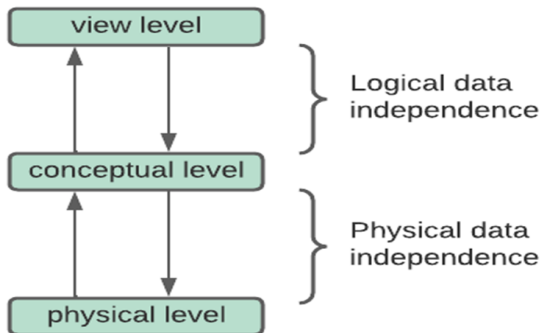
- **Lược đồ CSDL**: 1 mô tả logic về cấu trúc của CSDL (bảng, cột, kiểu DL, ràng buộc, mối quan hệ, chỉ mục, khung nhìn).
 - ① Lược đồ vật lý
 - ② Lược đồ logic
 - ③ Lược đồ khung nhìn
- **Thể hiện CSDL**: trạng thái hiện tại của CSDL.



Sự độc lập của DL

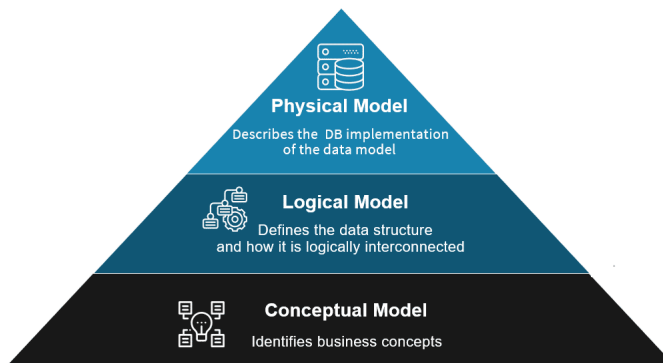
- Cho phép sự thay đổi ở 1 mức trừu tượng thì không ảnh hưởng đến các mức khác:

- ① Độc lập DL vật lý
- ② Độc lập DL logic



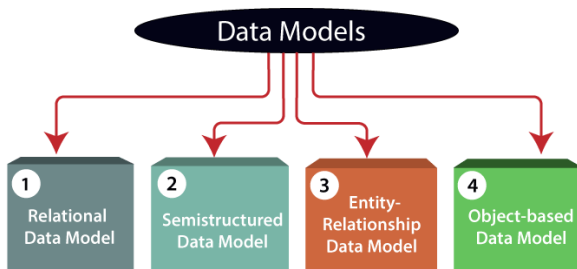
Mô hình dữ liệu: Khái niệm- (1)

- Tổ chức, biểu diễn, quản lý DL trong HT CSDL:
 - ① Tập các ký hiệu để mô tả DL
 - ② Tập các phép toán thao tác trên DL đó.
- Phân loại các **MHDL**:



Mô hình dữ liệu: Khái niệm- (2)

- Một số MHDL phổ biến:

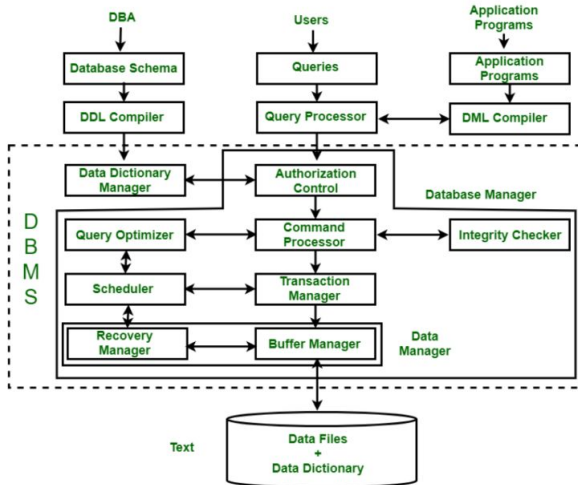


Hệ quản trị CSDL: Giới thiệu

- Hệ thống phần mềm dùng để quản lý, tổ chức CSDL.
- **Các chức năng chính:**
 - ① Quản lý DL, bảo mật DL, duy trì tính toàn vẹn DL
 - ② Hỗ trợ NN truy vấn, truy cập đồng thời và phục hồi DL
- **Một số loại DBMS:**
 - ① Quan hệ: *MySQL, Microsoft SQL Server*
 - ② Hướng đối tượng: *ObjectDB, DB4o*
 - ③ NoSQL: *MongoDB, Redis, Neo4j, Cassandra*
 - ④ Phân tán: *Google Spanner, Apache Cassandra, Amazon DynamoDB*

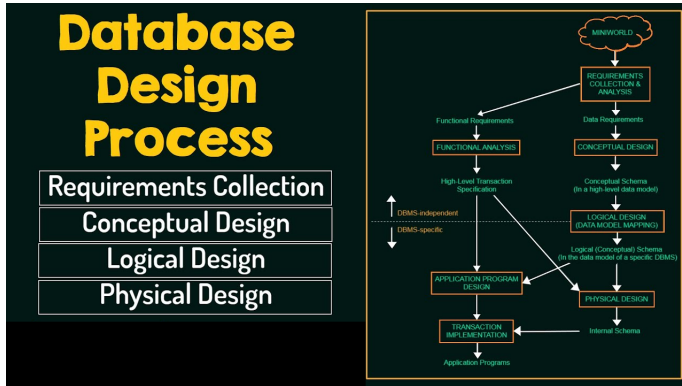
Hệ quản trị CSDL: Kiến trúc

- Kiến trúc của DBMS:**



Architecture of DBMS

Quy trình thiết kế CSDL



Quy trình thiết kế CSDL: Giới thiệu

- ① **Thu thập yêu cầu:** (mục tiêu, hoạt động chính)
- ② **Thiết kế khái niệm**
- ③ **Thiết kế logic**
- ④ **Thiết kế vật lý**
 - **Một số bước khác:**
- ① **Thực hiện CSDL**
 - *Triển khai CSDL theo thiết kế*
- ② **Kiểm thử và tối ưu hóa**
 - *Đảm bảo CSDL hoạt động đúng, đạt hiệu suất tốt*
- ③ **Vận hành và bảo trì**
 - *Đảm bảo CSDL hoạt động ổn định, phù hợp với các thay đổi*

Quy trình thiết kế CSDL: Ví dụ- (1)

Ví dụ 1.2. (Thiết kế CSDL cho 1 HT quản lý SV)

1. Thu thập yêu cầu

- Cần lưu trữ thông tin SV (Mã SV, Tên, Ngày sinh)
- SV đăng ký nhiều môn học, mỗi môn có 1 mã duy nhất
- Cần lưu điểm của SV cho từng môn học

3. Thiết kế logic

Bảng Student:

StudentID (PK)	Name	DOB
1	Alice	2002-05-10
2	Bob	2001-09-15

Bảng Course:

CourseID (PK)	CourseName
C1	Toán
C2	Lý

Bảng Enrollment:

StudentID (FK)	CourseID (FK)	Grade
1	C1	8.5
1	C2	7.0

2. Thiết kế khái niệm

- Thực thể SV (Student) có thuộc tính: Mã SV, Tên, Ngày sinh
- Thực thể Môn học (Course) có thuộc tính: Mã môn, Tên môn
- Quan hệ Đăng ký (Enrollment) giữa SV và Môn học có thuộc tính: Điểm số

4. Thiết kế vật lý

- Chọn KDL:

- + StudentID: INT
- + Name: VARCHAR(50)
- + DOB: DATE

- Tạo chỉ mục trên StudentID và CourseID để tối ưu truy vấn

5. Tạo bảng bằng SQL

...

6. Kiểm thử

- Chạy thử truy vấn lấy danh sách SV đăng ký môn Toán
- Kiểm tra tốc độ truy vấn khi có DL lớn

7. Vận hành

- Thiết lập sao lưu DL mỗi ngày
- Giám sát hiệu suất hệ thống

Quy trình thiết kế CSDL: Ví dụ- (2)

Ví dụ 1.2. (tiếp theo)

- Tạo bảng bằng SQL:

```
CREATE TABLE Student (  
    StudentID INT PRIMARY KEY,  
    Name VARCHAR(50),  
    DOB DATE  
);  
  
CREATE TABLE Course (  
    CourseID VARCHAR(10) PRIMARY KEY,  
    CourseName VARCHAR(100)  
);  
  
CREATE TABLE Enrollment (  
    StudentID INT,  
    CourseID VARCHAR(10),  
    Grade FLOAT,  
    PRIMARY KEY (StudentID, CourseID),  
    FOREIGN KEY (StudentID) REFERENCES Student(StudentID),  
    FOREIGN KEY (CourseID) REFERENCES Course(CourseID)  
);
```

Vai trò con người trong hệ CSDL

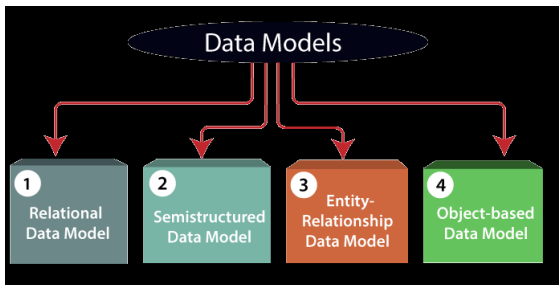
- 1 **Người quản trị CSDL (Administrator)**
- 2 **Người thiết kế CSDL (Designer)**
- 3 **Người lập trình UD (Programmer)**
- 4 **Người sử dụng cuối (End)**

Nội dung môn học

- ① Tổng quan
- ② Các mô hình dữ liệu
- ③ Phụ thuộc hàm
- ④ Khoá lược đồ quan hệ
- ⑤ Chuẩn hoá lược đồ quan hệ
- ⑥ Tài liệu tham khảo

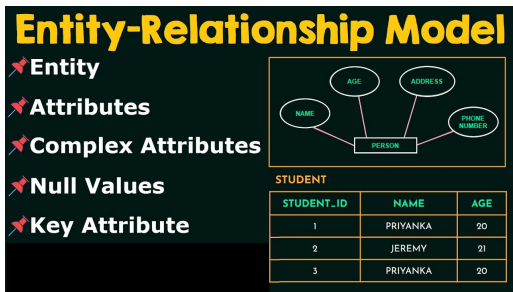
Chương 2. Các mô hình dữ liệu

- 1 Mô hình dữ liệu thực thể - mối quan hệ
(Entity-Relationship data model, ER)
- 2 Mô hình dữ liệu quan hệ (Relational data model)



Mô hình dữ liệu ER

- ① Các thành phần chính của MHDL ER
- ② Môi quan hệ *isa*
- ③ Sơ đồ MHDL ER



Các thành phần chính của MHDL ER - (1)

- **Thực thể:** 1 ĐT cụ thể hoặc trừu tượng trong thế giới thực (*sinh viên, môn học*).
 - ① **Thực thể mạnh:** độc lập (*hoá đơn*)
 - ② **Thực thể yếu:** phụ thuộc (*hoá đơn chi tiết*)
- **Thuộc tính:** các đặc điểm, thông tin dùng để mô tả TTH.

Thuộc tính	Đặc điểm	Ví dụ
Đơn giản	Không thể chia nhỏ	<i>Mã SV</i>
Phức hợp	Có thể tách thành nhiều thành phần con	<i>Họ tên, Ngày sinh</i>
Đơn trị	Chỉ có 1 giá trị duy nhất tại 1 thời điểm	<i>CCCD</i>
Đa trị	Có thể có nhiều giá trị	<i>Số ĐT</i>
Suy diễn	Có thể tính toán từ thuộc tính khác	<i>Tuổi</i>

Các thành phần chính của MHDL ER - (2)

- **Khoá**: tập thuộc tính giúp **phân biệt duy nhất** từng TTH
 - $\leadsto$ Mỗi TTH có thể có nhiều khoá

Các loại khoá	Đặc điểm	Ví dụ
Khoá dự tuyển (Candidate, CK)	khoá có thể chứa <i>NULL</i>	<i>CCCD</i> <i>MSV</i>
Khoá chính (Primary, PK)	khoá dự tuyển được chọn, không được chứa <i>NULL</i>	<i>MSV</i>
Khoá ngoại (Foreign, FK)	khoá dự tuyển tham chiếu đến khóa chính của TTH khác	<i>Did</i>
Khoá tổng hợp (Composite, CoK)	khóa chính được tạo từ các khóa chính khác	$\{MaHĐ, MaSP\}$

Các thành phần chính của MHDL ER - (3)

- **Mối quan hệ:** sự liên kết giữa các TTH.

Quan hệ	Đặc điểm	Ví dụ
1-1 (One-to-One)	Mỗi TTH A chỉ liên kết với đúng 1 TTH B, và ngược lại	<i>Quản lý</i>
1-n (One-to-Many)	1 TTH A có thể liên kết với nhiều TTH B, nhưng mỗi TTH B chỉ thuộc về 1 TTH A	<i>Làm việc trong</i>
n-m (Many-to-Many)	1 TTH A có thể liên kết với nhiều TTH B, và ngược lại	<i>Đăng ký môn học</i>

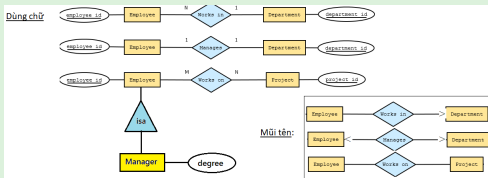
Mối quan hệ *isa*

- Biểu diễn **sự kế thừa** giữa các thực thể/phân cấp.
 - “Xe đạp”, “Xe máy” **isa** “Phương tiện giao thông”
 - ~→
 - Thực thể **con**: “Xe đạp”, “Xe máy”
 - Thực thể **cha**: “Phương tiện giao thông”
- **Đặc điểm của *isa***:
 - ① Kế thừa
 - ② Phân cấp
- ~→
 - ① Ưu điểm: Tái sử dụng, tính tổ chức, mở rộng dễ dàng
 - ② Hạn chế: Một số DBMS không hỗ trợ *isa*, số lượng lớp trong phân cấp nhiều ⇒ khó quản lý.

Sơ đồ MHDL ER - (1)

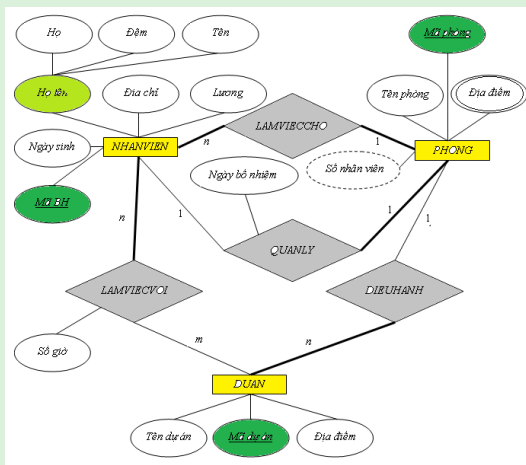
- Biểu diễn **đồ họa cấu trúc logic** của CSDL.
 - ① **Hình chữ nhật**: thực thể
 - ② **Hình elip**: thuộc tính
 - ① Thuộc tính của khóa chính: gạch chân
 - ② Thuộc tính đa trị: thuộc tính được nối với thực thể bằng đường kẻ đôi
 - ③ Thuộc tính suy diễn: hình elip với nét đứt
 - ④ Thuộc tính phức hợp: nhiều hình elip liên kết với nó
 - ③ **Hình thoi**: mối quan hệ
 - ④ **Tam giác**: mối quan hệ *isa*

Ví dụ 2.1.



Sơ đồ MHDL ER - (2)

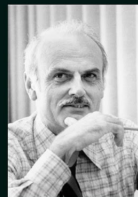
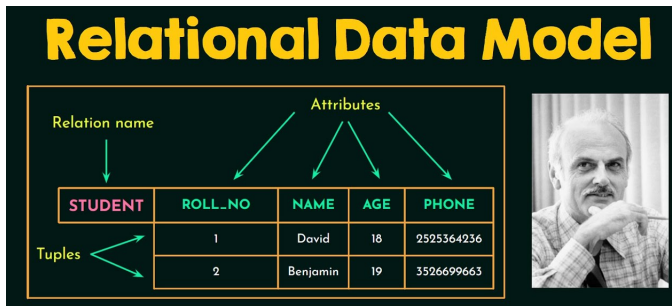
Ví dụ 2.2. (Sơ đồ MHDL ER)



Một hệ thống quản lý nhân sự với các thực thể chính

Mô hình dữ liệu quan hệ

- ① Khái niệm quan hệ
- ② Khoá quan hệ
- ③ Đại số quan hệ
- ④ Chuyển đổi mô hình



Edgar F. Codd (1969)

Khái niệm quan hệ: Định nghĩa- (1)

- Giả sử:
 - Tập các thuộc tính: $U = \{A_1, A_2, \dots, A_n\} \neq \emptyset$
 - Tương ứng: mỗi A_i có *miền giá trị* $dom(A_i)$.

- **Quan hệ** trên U :

$$R = \{(x_1, x_2, \dots, x_n) : x_i \in dom(A_i)\} \subseteq \prod_{i=1}^n dom(A_i)$$

- **Quan hệ tương thích**: 2 quan hệ có cùng tập thuộc tính, các thuộc tính cùng tên có cùng miền giá trị.
- $\leadsto$ Ký hiệu:
 - $t(X)$: bộ t chứa các giá trị tương ứng với $X \subseteq U$.
 - $t(a_i)$: giá trị của bộ t tại $a_i \in U$.
 - $R \equiv$ **bảng** với mỗi cột - 1 thuộc tính, mỗi hàng - 1 bộ.

Khái niệm quan hệ: ví dụ

- Xét quan hệ *TKB* (thời khóa biểu) của 1 trường học:

<i>Ngày</i>	<i>Tiet</i>	<i>Mon</i>	<i>Phong</i>	<i>Giaovien</i>
1/9	1	Học máy	E201	Nguyễn Văn Hùng
1/9	2	Học máy	E201	Nguyễn Văn Hùng
1/9	1	Anh văn 1	H101	Lê Thị Hòa
1/9	2	Anh văn 1	H102	Trần Văn Hiệp
1/9	3	Anh văn 2	E201	Hoàng Thị Ngọc

- $\leadsto$

- $U = \{ \text{Ngày, Tiet, Mon, Phong, Giaovien} \}$
- $(\text{"1/9"}, 1, \text{"Anh Văn 1"}, \text{"H101"}, \text{"Lê Thị Hoà"})$: 1 bộ
- $X = \{ \text{Tiet, Mon, Phong} \} \subseteq U$, $t =$ bộ thứ 3:

$\Rightarrow$

$$t(X) = (1, \text{"Anh văn 1"}, \text{"H101"})$$

$$t(\text{Phong}) = \text{H101}$$

Khái niệm quan hệ: Định nghĩa- (2)

- **Lược đồ quan hệ:** cặp $S = (U, F)$ trong đó:
 - U : tập hữu hạn các thuộc tính
 - F : tập các ràng buộc giữa các thuộc tính.
- $\rightsquigarrow$
 - Quan hệ: có thể thay đổi theo thời gian
 - LĐQH: mô tả về cấu trúc và các ràng buộc toàn vẹn của 1 quan hệ (bất biến, ít thay đổi)
 - **Thể hiện:** quan hệ thỏa LĐQH.
- **Lược đồ CSDL quan hệ:** tập hữu hạn các LĐQH $S_1, S_2, \dots, S_k$.

Khoá quan hệ: Định nghĩa

- Cho quan hệ R trên U .
 - **Khóa** của R : tập con $K \subseteq U$ thoả
 - 1 Tính duy nhất:

$$\forall t_i, t_j \in R : t_i(K) \neq t_j(K)$$

- 2 Tính tối thiểu:

$$\forall K' \subset K, K' \text{ không thoả (1).}$$

- **Siêu khoá**: thoả (1).
- $\leadsto$
 - 1 $\text{Khoá} \Rightarrow \text{Siêu khoá}$.
 - 2 R luôn có ít nhất 1 khoá.
 - 3 Tính chất (hệ) Sperner: 2 khoá bất kỳ không chứa nhau.
 - 4 Ký hiệu tập tất cả khóa: $\text{Key}(R)$

Khoá quan hệ: ví dụ

- ① Khóa của quan hệ TKB : $\{Tiet, Monhoc\}$.
- ② Xét quan hệ $R = \{t_1, t_2, t_3\}$ trên $U = \{A, B, C\}$:

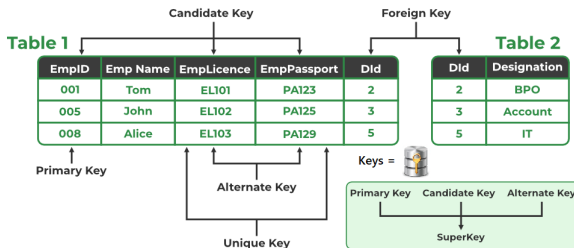
$$R =$$

	A	B	C
t_1	0	1	1
t_2	1	2	0
t_3	2	1	1

- $\hookrightarrow$
 - Khóa (duy nhất): $\{A\}$
 - Siêu khoá: $\{A, B\}$...
- (*Bài tập, thảo luận*)

Khoá quan hệ: Lưu ý

- ~ Trong thiết kế:
 - **Khoá dự tuyển**: $\equiv$ *khoá*
 - **Khoá chính**: 1 khoá dự tuyển được chọn (cài đặt)
 - **Khoá thay thế**: các khoá dự tuyển còn lại.
- ~ Lưu ý khi đọc: một số sách, tài liệu sử dụng thuật ngữ
 - **Khoá tối thiểu** $\equiv$ *khoá*
 - **Khoá** $\equiv$ *siêu khoá*



Khoá quan hệ: Hệ bằng nhau

- Cho quan hệ $R = \{t_1, t_2, \dots, t_m\}$ trên U .
- Hệ bằng nhau** của R :

$$E(R) = \{E_{ij} : 1 \leq i < j \leq m\}$$

với $E_{ij} = \{A \in U : t_i(A) = t_j(A)\}$.

- Hệ bằng nhau cực đại** của R :

$$\begin{aligned} \max(E(R)) &= \{E_{ij} \in E(R) : \nexists E_{pq} \in E(R), E_{pq} \neq E_{ij}, \\ &\Rightarrow E_{pq} \supset E_{ij}\} \end{aligned}$$

$R =$

	A	B	C
t_1	0	1	1
t_2	1	2	0
t_3	2	1	1

$E_{13} = \{B, C\}$

Khoá quan hệ: Thuật toán

- Đầu vào: quan hệ $R = \{t_1, \dots, t_m\}$ trên $U = \{A_1, \dots, A_n\}$.
- Đầu ra: Một khóa K của R .
- Phương pháp:
- **Bước 1.** Tính $E(R)$.
- **Bước 2.** Tính $\max(R)$.
- **Bước $i + 2$.** Xây dựng dãy $K_0, K_1, \dots, K_n$:
 - ① $K_0 = U$
 - ② $K_i = \begin{cases} K_{i-1} - \{A_i\} & \text{nếu } \forall E_{ij} \in \max(R) : K_{i-1} - \{A_i\} \not\subseteq E_{ij} \\ K_{i-1} & \text{ngược lại} \end{cases}$
 $\leadsto i = 1, 2, \dots, n$
- **Bước cuối.** Đặt $K = K_n$.

Khoá quan hệ: ví dụ

- Xét lại:

$$R =$$

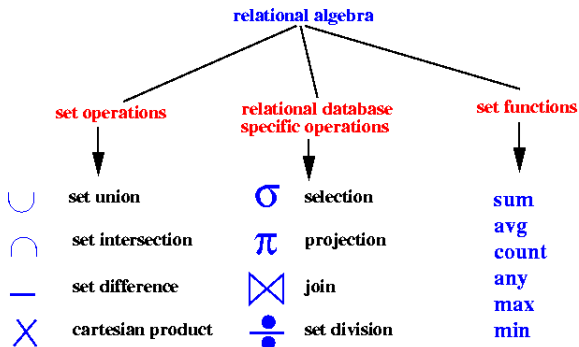
	A	B	C
t_1	0	1	1
t_2	1	2	0
t_3	2	1	1

- Ta có:
 - $E_{12} = \emptyset, E_{13} = \{B, C\}, E_{23} = \emptyset$
- $\Rightarrow$
 - $E(R) = \{\emptyset, \{B, C\}\}$
 - $\max(R) = \{\{B, C\}\}$
- Tính dãy $K_0, \dots, K_3$:
 - $K_0 = \{A, B, C\}$
 - $K_1 = K_0$
 - $K_2 = \{A, C\}$
 - $K_3 = \{A\}$

$\Rightarrow$ Một khoá của R : $\{A\}$.

Đại số quan hệ: Giới thiệu

- Đại số quan hệ: truy vấn, thao tác trên CSDL quan hệ.



Đại số quan hệ: Định nghĩa- (1)

- **Nhóm 1:** xét 2 quan hệ tương thích R_1, R_2 .

① Phép hợp:

$$R_1 \cup R_2 = \{t : t \in R_1 \text{ hoặc } t \in R_2\}$$

② Phép giao:

$$R_1 \cap R_2 = \{t : t \in R_1, t \in R_2\}$$

③ Phép hiệu:

$$R_1 - R_2 = \{t : t \in R_1, t \notin R_2\}$$

Course_1

C_id	C_name
11	Foundation C
21	C++
31	JAVA

Course_2

C_id	C_name
12	Python
21	C++

=>

Course_1 U Course_2

C_id	C_name
11	Foundation C
21	C++
31	JAVA
12	Python

Đại số quan hệ: ví dụ- (1)

Ví dụ: (Hợp, giao, hiệu)

$$R_1 = \begin{array}{|c|c|c|} \hline A & B & C \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 3 \\ \hline 0 & 2 & 3 \\ \hline \end{array}, R_2 = \begin{array}{|c|c|c|} \hline A & B & C \\ \hline 0 & 1 & 2 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

$\Rightarrow$

$$R_1 \cup R_2 = \begin{array}{|c|c|c|} \hline A & B & C \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 3 \\ \hline 0 & 2 & 3 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

$$R_1 \cap R_2 = \begin{array}{|c|c|c|} \hline A & B & C \\ \hline 0 & 1 & 2 \\ \hline \end{array}, R_1 - R_2 = \begin{array}{|c|c|c|} \hline A & B & C \\ \hline 0 & 1 & 3 \\ \hline 0 & 2 & 3 \\ \hline \end{array}$$

Đại số quan hệ: Định nghĩa- (2)

- Cho quan hệ R_1 trên U_1 , quan hệ R_2 trên U_2 .

① **Phép tích Descartes:** điều kiện $U_1 \cap U_2 = \emptyset$

$$R_1 \times R_2 = \{t_1 \star t_2 : t_1 \in R_1, t_2 \in R_2\}$$

$\leadsto$ **phép dán bộ:**

$$t_1 \star t_2 = \{t(U_1 \cup U_2) : t(U_1) = t_1, t(U_2) = t_2\}$$

② **Phép chia:** điều kiện $R_2 \neq \emptyset, U_2 \subseteq U_1$

$$\begin{aligned} R_1 : R_2 &= \{t(U_1 - U_2) : t \in R_1, (\forall t_2 \in R_2 \Rightarrow \\ &\Rightarrow t(U_1 - U_2) \star t_2 \in R_1)\} \end{aligned}$$

Đại số quan hệ: ví dụ- (2)

Ví dụ: (Tích Descartes)

$$R_1 = \begin{array}{|c|c|c|} \hline A & B & C \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 3 \\ \hline 0 & 2 & 3 \\ \hline \end{array}, R_2 = \begin{array}{|c|c|} \hline D & E \\ \hline 0 & 1 \\ \hline 1 & 1 \\ \hline \end{array}$$

$\Rightarrow$

$$R_1 \times R_2 = \begin{array}{|c|c|c|c|c|} \hline A & B & C & D & E \\ \hline 0 & 1 & 2 & 0 & 1 \\ \hline 0 & 1 & 2 & 1 & 1 \\ \hline 0 & 1 & 3 & 0 & 1 \\ \hline 0 & 1 & 3 & 1 & 1 \\ \hline 0 & 2 & 3 & 0 & 1 \\ \hline 0 & 2 & 3 & 1 & 1 \\ \hline \end{array}$$

Đại số quan hệ: ví dụ- (3)

Ví dụ: (Chia)

$$R_1 = \begin{array}{|c|c|} \hline A & B \\ \hline 1 & 0 \\ \hline 1 & 1 \\ \hline 2 & 0 \\ \hline 2 & 2 \\ \hline 3 & 0 \\ \hline 3 & 1 \\ \hline \end{array}, R_2 = \begin{array}{|c|} \hline B \\ \hline 0 \\ \hline 1 \\ \hline \end{array}$$

⇒

$$R_1 : R_2 = \begin{array}{|c|} \hline A \\ \hline 1 \\ \hline 3 \\ \hline \end{array}$$

Đại số quan hệ: Định nghĩa- (3)

- **Nhóm 2:** Cho quan hệ R trên U .

- ① **Phép chọn (lọc)** theo điều kiện F : F là biểu thức logic

$$\sigma_F(R) = \{t \in R : F(t) \text{ đúng}\}$$

- ② **Phép chiếu** (trên $X \subseteq U$):

$$\Pi_X(R) = \{t(X) : t \in R\}$$

- ③ **Phép kết nối** (tự nhiên):

$$R_1 \bowtie R_2 = \{t_1 \star t_2 : t_1 \in R_1, t_2 \in R_2\}$$

- ④ **Phép kết nối θ** (điều kiện): $U_1 \cap U_2 = \emptyset$,

$$A \in U_1, B \in U_2$$

$$R_1 \bowtie_{A\theta B} R_2 = \{t_1 \star t_2 : t_1 \in R_1, t_2 \in R_2, t_1(A)\theta t_2(B)\}$$

Đại số quan hệ: Lưu ý



- ① F : còn gọi **biểu thức chọn**
- ② $U_1 \cap U_2 = \emptyset \Rightarrow R_1 \bowtie R_2 = R_1 \times R_2$
- ③ θ : $=, \neq, <, \leq, >, \geq$
- ④ $\theta \equiv$ phép toán " $=$ " $\Rightarrow$ gọi **kết nối bằng**
- ⑤ $R_1 \bowtie_{A\theta B} R_2 = \sigma_{t_1(A)\theta t_2(B)}(R_1 \times R_2)$.

Đại số quan hệ: ví dụ- (4)

Ví dụ: (Chọn, chiếu, kết nối)

$$R = \begin{array}{|c|c|c|} \hline A & B & C \\ \hline 0 & 1 & 1 \\ \hline 1 & 2 & 0 \\ \hline 2 & 1 & 1 \\ \hline \end{array}, S = \begin{array}{|c|c|c|} \hline B & C & D \\ \hline 2 & 1 & 0 \\ \hline 1 & 1 & 1 \\ \hline \end{array}, T = \begin{array}{|c|c|} \hline D & E \\ \hline 0 & 1 \\ \hline 1 & 1 \\ \hline \end{array}$$

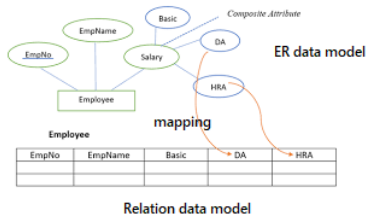
$\Rightarrow$

$$\sigma_{F=\{B='1' \wedge C='1'\}}(R) = \begin{array}{|c|c|c|} \hline A & B & C \\ \hline 0 & 1 & 1 \\ \hline 2 & 1 & 1 \\ \hline \end{array}, \Pi_{\{B,C\}}(R) = \begin{array}{|c|c|} \hline B & C \\ \hline 1 & 1 \\ \hline 2 & 0 \\ \hline \end{array}$$

$$R \bowtie S = \begin{array}{|c|c|c|c|} \hline A & B & C & D \\ \hline 0 & 1 & 1 & 1 \\ \hline 2 & 1 & 1 & 1 \\ \hline \end{array}, R \underset{C>D}{\bowtie} T = \begin{array}{|c|c|c|c|c|} \hline A & B & C & D & E \\ \hline 0 & 1 & 1 & 0 & 1 \\ \hline 2 & 1 & 1 & 0 & 1 \\ \hline \end{array}$$

Chuyển đổi mô hình: Quy tắc - (1)

- Chuyển đổi MHDL ER sang MHDL quan hệ:



- Chuyển đổi các thực thể:

① Thực thể mạnh: $\mapsto$ 1 bảng (quan hệ)

- Các thuộc tính $\mapsto$ các cột
- Khóa chính (thực thể) $\mapsto$ khóa chính của bảng

② Thực thể yếu: $\mapsto$ 1 bảng riêng

- Các thuộc tính $\mapsto$ các cột
- Thêm 1 cột mới: làm khóa ngoại để liên kết với thực thể mạnh mà nó phụ thuộc
- Khóa chính (thực thể mạnh) và khóa phân biệt (thực thể yếu) $\mapsto$ khóa chính của bảng riêng.

Chuyển đổi mô hình: Quy tắc - (2)

- Chuyển đổi các mối quan hệ:

① **Quan hệ 1-1:** thêm *khóa chính* của 1 thực thể (nào đó) làm *khóa ngoại* trong bảng của thực thể còn lại cùng với ràng buộc duy nhất (UNIQUE).

② **Quan hệ 1- n :**

- Thêm *khóa chính* của thực thể phía "1" làm *khóa ngoại* vào bảng của thực thể phía " n ".
- Nếu mỗi quan hệ có thuộc tính $\Rightarrow$ thêm vào bảng thực thể phía " n ".

③ **Quan hệ m - n :**

- $\mapsto$ 1 *bảng trung gian*
- Bảng trung gian: chứa *khóa chính* của cả 2 thực thể tham gia làm *khóa ngoại*
- Nếu mỗi quan hệ có thuộc tính $\Rightarrow$ thêm vào bảng trung gian

Chuyển đổi mô hình: Quy tắc - (3)

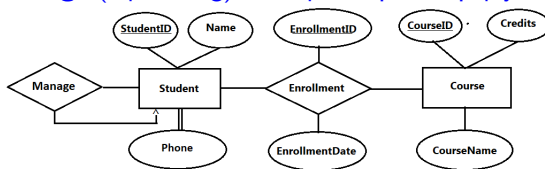
- Chuyển đổi quan hệ đệ quy:
 - Thuộc quan hệ 1-n:
 - Thêm *khóa chính* của thực thể đó làm *khóa ngoại* vào chính bảng đó, và thường đặt tên khác đi để phân biệt vai trò.
 - Khóa ngoại này cho phép 1 bản ghi tham chiếu đến 1 bản ghi khác trong cùng 1 bảng.

Chuyển đổi mô hình: Quy tắc - (4)

- **Chuyển đổi mỗi quan hệ *isa*:**
 - **Cách 1:** mỗi thực thể (phân cấp) $\mapsto$ 1 *bảng riêng*
 - **Cách 2:** các thực thể cha-con $\mapsto$ 1 *bảng duy nhất*
- **Chuyển đổi các thuộc tính đa trị:** $\mapsto$ 1 *bảng riêng*
 - *Khóa chính* (thực thể sở hữu thuộc tính đa trị) $\mapsto$ *khoá ngoại* của bảng.
 - *Khoá chính* của bảng: gồm *khóa ngoại* và thuộc tính đa trị đó.

Chuyển đổi mô hình: Ví dụ - (1)

- Xét MHDL ER với 2 thực thể và 2 mối quan hệ:
 - Student**: StudentID (PK), Name, Phone (đa trị)
 - Course**: CourseID (PK), CourseName, Credits
 - Enrollment** (mối quan hệ $n-m$): EnrollmentID (PK), EnrollmentDate
 - Manage** (lớp trưởng): mối quan hệ $1-n$ đệ quy.



- Chuyển đổi sang MHDL quan hệ: biểu diễn bằng LĐQH
 - Course**(CourseID, CourseName, Credits)
 - Student**(StudentID, Name, *ClassMonitorID*)
 - Student_P**(StudentID, Phone)
 - Enrollment**(EnrollmentID, EnrollmentDate, *StudentID*, *CourseID*)

Chuyển đổi mô hình: Ví dụ - (2)

- MHDL quan hệ biểu diễn dưới dạng bảng:

1. Bảng Student :

StudentID (PK)	Name	DateOfBirth	ClassMonitorID(FK)
S001 (ClassMonitor)	Alice	2000-01-01	NULL
S002	Tom	2000-10-05	S001

(S001 lớp trưởng của S002)

2. Bảng Course :

CourseID (PK)	CourseName	Credits
101	Database	3
...		

3. Bảng Enrollment :

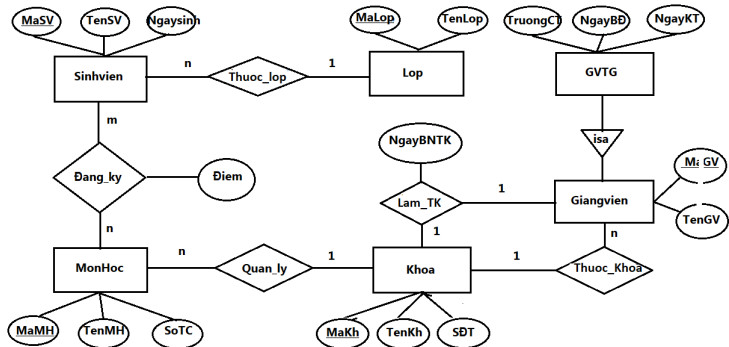
EnrollmentID (PK)	StudentID (FK)	CourseID (FK)	EnrollmentDate
1	S002	101	2025-01-01
...			

Chuyển đổi mô hình: Ví dụ - (3)

- Cho MHDL ER:
 - **Thực thể:**
 - ① *Sinh viên*: MaSV (PK), TenSV, Ngaysinh
 - ② *Lớp*: MaLop (PK), TenLop
 - ③ *Môn học*: MaMH (PK), TenMH, SoTC
 - ④ *Khoa*: MaKh (PK), TenKh, SĐT
 - ⑤ *Giảng viên*: MaGV (PK), TenGV
 - ⑥ *Giảng viên thỉnh giảng*: TruongCT, NgayBD, NgayKT
 - **Mối quan hệ:**
 - ① *Thuộc lớp*: quan hệ 1- n từ Lớp đến Môn học
 - ② *Đăng ký*: quan hệ n - m giữa Sinh viên và Môn học, có thêm thuộc tính Diem
 - ③ *Quản lý*: quan hệ 1- n từ Khoa đến Môn học
 - ④ *Thuộc Khoa*: quan hệ 1- n từ Khoa đến Giảng viên
 - ⑤ *Làm Trưởng Khoa*: quan hệ 1-1 giữa Giảng viên và Trưởng Khoa, có thêm thuộc tính NgayBNTK
 - ⑥ *isa*: Giảng viên thỉnh giảng *isa* với Giảng viên.

Chuyển đổi mô hình: Ví dụ - (4)

- Biểu diễn bằng sơ đồ ER:



Chuyển đổi mô hình: Ví dụ - (5)

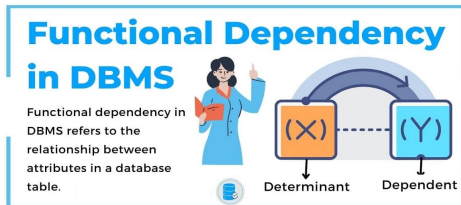
- Chuyển đổi sang MHDL quan hệ:
 - ① **SINHVIEN**(MaSV, TenSV, Ngaysinh, *MaLop*)
 - ② **LOP**(MaLop, TenLop)
 - ③ **MONHOC**(MaMH, TenMH, SoTC, *MaKh*)
 - ④ **KHOA**(MaKh, TenKh, SDT, *MaGV_TK*, NBNTK)
 - ⑤ **GIANGVIEN**(MaGV, TenGV, *MaKh*)
 - ⑥ **GVTG**(MaGV, TruongCT, NgayBD, NgayKT)
 - ⑦ **DANGKY**(MaSV, MaMH, Diem)
- ~→
 - MaGV trong **GVTG** vừa là *khoá chính*, vừa là *khoá ngoại* trở về MaGV trong **GIANGVIEN**
 - MaGV_TK là khóa ngoại trong **KHOA** trở về bảng **GIANGVIEN** để thực hiện quan hệ Lam_TK. Trong SQL, MaGV_TK này bắt buộc phải có ràng buộc UNIQUE.

Nội dung môn học

- ① Tổng quan
- ② Các mô hình dữ liệu
- ③ Phụ thuộc hàm**
- ④ Khoá lược đồ quan hệ
- ⑤ Chuẩn hoá lược đồ quan hệ
- ⑥ Tài liệu tham khảo

Chương 3. Phụ thuộc hàm

- 1 Khái niệm phụ thuộc hàm
- 2 Suy diễn PTH
- 3 Hệ tiên đề Armstrong
- 4 Bao đóng tập thuộc tính
- 5 Phủ tối thiểu



Khái niệm phụ thuộc hàm: Định nghĩa- (1)

- Cho tập thuộc tính $U \neq \emptyset$, và $R = \{t_1, t_2, \dots, t_m\}$.
- **Phụ thuộc hàm**: là ký hiệu $X \rightarrow Y$ với $X, Y \subseteq U$ (trên U). PTH $X \rightarrow Y$ **đúng** trên quan hệ R , nếu:

$$\forall t_i, t_j \in R: t_i(X) = t_j(X) \Rightarrow t_i(Y) = t_j(Y).$$

Ví dụ: Xét quan hệ R trên $U = \{A, B, C\}$

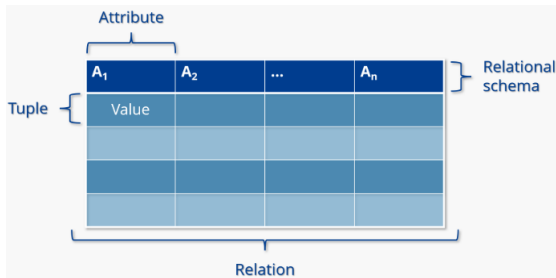
$$R =$$

	A	B	C
t_1	0	0	1
t_2	1	0	1
t_3	2	0	0

- 1 Các PTH đúng trên R : $A \rightarrow B, C \rightarrow B \dots$
- 2 Các PTH không đúng trên R : $B \rightarrow A, C \rightarrow A \dots$

Khái niệm phụ thuộc hàm: Định nghĩa- (2)

- **Lược đồ quan hệ**: là cặp $S = (U, F)$ trong đó
 - $U = \{A_1, A_2, \dots, A_n\} \neq \emptyset$: tập các thuộc tính
 - $F = \{X \rightarrow Y : X, Y \subseteq U\}$: tập các PTH trên U .



- $\leadsto$
 - LQKH được xem như cái khung (các cột) và mối quan hệ (PTH) giữa các cột.
 - Trong F người ta không quan tâm các PTH: $X \rightarrow \emptyset$, $\emptyset \rightarrow Y$.

Khái niệm phụ thuộc hàm: Định nghĩa- (3)

- $\leadsto$ Gọi/đọc:
- $X \rightarrow Y$: X xác định hàm Y , hay Y PTH vào X .
 - X : vế trái; Y : vế phải.
 - Các tập thuộc tính X, Y có thể viết dưới dạng xâu ký tự: $\{A, B, C\} \rightarrow \{D\}$ viết lại $ABC \rightarrow D$.
- $\leadsto$ Ký hiệu:
 - ① $R(X \rightarrow Y)$: khi $X \rightarrow Y$ đúng trên R .
 - ② $R(F)$: nếu $R(X \rightarrow Y)$ với mọi $X \rightarrow Y$ trong tập PTH F .

Suy diễn PTH: Định nghĩa

- **Suy diễn PTH** (theo quan hệ): $F \models X \rightarrow Y$, nếu với mọi quan hệ R :

$$R(F) \Rightarrow R(X \rightarrow Y)$$

- Đặt

$$F^* = \{X \rightarrow Y : F \models X \rightarrow Y\}$$

- Khi F không suy diễn được $X \rightarrow Y$ thì ký hiệu $F \not\models X \rightarrow Y$.

Hệ tiên đề Armstrong: Định nghĩa

- ① (F1) Quy tắc phản xạ:

$$Y \subseteq X \Rightarrow X \rightarrow Y.$$

- ② (F2) Quy tắc gia tăng:

$$X \rightarrow Y \Rightarrow X \cup Z \rightarrow Y \cup Z.$$

- ③ (F3) Quy tắc bắc cầu:

$$X \rightarrow Y, Y \rightarrow Z \Rightarrow X \rightarrow Z.$$

Ví dụ:

Cho $F = \{A \rightarrow B, B \rightarrow C\}$. Từ F suy ra được các PTH mới:

① $A \rightarrow B$ (gt) $\xRightarrow{(F2)}$ **$A \rightarrow AB$** .

② $B \rightarrow C$ (gt) $\xRightarrow{(F2)}$ **$AB \rightarrow ABC$** .

③ Từ $A \rightarrow AB$ và $AB \rightarrow ABC$ $\xRightarrow{(F3)}$ **$A \rightarrow ABC$** .

Hệ tiên đề Armstrong: Suy diễn theo tiên đề

- **Suy diễn theo tiên đề:** các PTH được suy diễn từ F dựa vào HTĐ Armstrong.
- **Bao đóng F^+ :**

$$F^+ = \{X \rightarrow Y : X \rightarrow Y \text{ được SD từ } F \text{ dựa vào (F1)-(F3)}\}.$$

Định lý 3.1 (Tính đúng đắn và đầy đủ của HTĐ)

$$F^+ = F^*.$$

Suy diễn (theo quan hệ) $\iff$ Suy diễn theo tiên đề.

$$X \rightarrow Y \in F^+ \iff F \models X \rightarrow Y$$

Hệ tiên đề Armstrong: Các quy tắc suy diễn MR

Mệnh đề 3.1

(F4) **Quy tắc cộng tính phải:**

$$X \rightarrow Y, X \rightarrow Z \Rightarrow X \rightarrow Y \cup Z$$

(F5) **Quy tắc thu hẹp phải (tách):**

$$X \rightarrow Y \Rightarrow X \rightarrow Z \text{ với mọi } Z \subseteq Y$$

(F6) **Quy tắc giả bắc cầu:**

$$X \rightarrow Y \text{ và } Y \cup W \rightarrow Z \Rightarrow X \cup W \rightarrow Z$$

(F7) **Quy tắc cộng tính đầy đủ:**

$$X \rightarrow Y \text{ và } W \rightarrow Z \Rightarrow X \cup W \rightarrow Y \cup Z$$

(F8) **Quy tắc mở rộng trái và thu hẹp phải:**

$$X \rightarrow Y \Rightarrow X \cup W \rightarrow Y - Z$$

(F9) **Hệ quả của (F4) và (F5):**

$$X \rightarrow A_1 A_2 \dots A_n \iff X \rightarrow A_i, \quad \forall i = 1, \dots, n.$$

Bao đóng tập thuộc tính: Định nghĩa

- Xét LĐQH $S = (U, F)$ và $X \subseteq U$.
- **Bao đóng** của X (ứng với F):

$$X_F^+ = \{A \in U : F \models X \rightarrow A\} \quad (\equiv X^+)$$

Ví dụ:

- Cho LĐQH $S = (U, F)$ với:
 - $U = \{A, B, C\}$
 - $F = \{A \rightarrow B, B \rightarrow C\}$.
- Ta có:
 - ① $\{B\}^+ = \{B, C\}$
 - ② $\{A\}^+ = \{A, B, C\}$
 - ③ $\{B, C\}^+ = \{B, C\}$.

Bao đóng tập thuộc tính: Tính chất

Mệnh đề 3.2 (Tính chất của bao đóng)

- (1) *Tính phản xạ*: $X \subseteq X^+$
- (2) *Tính đơn điệu*: $X \subseteq Y \Rightarrow X^+ \subseteq Y^+$
- (3) *Tính lũy đẳng*: $(X^+)^+ = X^+$
- (4) $(X \cup Y)^+ \supseteq X^+ \cup Y^+$
- (5) $(X \cap Y)^+ \subseteq X^+ \cap Y^+$
- (6) $(X^+ \cup Y)^+ = (X \cup Y^+)^+ = (X \cup Y)^+$
- (7) $F \models X \rightarrow Y \Leftrightarrow Y \subseteq X^+$.

Bao đóng tập thuộc tính: Thuật toán

- Đầu vào: LĐQH $S = (U, F)$ và $X \subseteq U$.

- Đầu ra: Bao đóng X^+ .

- Phương pháp:

Xây dựng dãy $X_0 \subseteq X_1 \subseteq \dots \subseteq X_k$ theo qui tắc:

- **Bước 1.** (*Khởi tạo*) Đặt $X_0 := X$.
- **Bước 2.** (*Lặp, $i = 0, 1, \dots$*) Tính:

$$X_{i+1} := X_i \cup Z_i$$

với

$$Z_i = \{A \notin X_i : \exists Z \rightarrow W \in F, A \in W, Z \subseteq X_i\}$$

- Thuật toán dừng: $\exists k \in \mathbb{N}$ nhỏ nhất sao cho $X_{k+1} = X_k$.
- **Bước 3.** (*Kết luận*) Đặt $X^+ := X_k$.

Bao đóng tập thuộc tính: ví dụ

Ví dụ:

- Cho LĐQH $S = (U, F)$ với
 - $U = \{A, B, C\}$
 - $F = \{A \rightarrow B, B \rightarrow C\}$.

1 Tính $\{A\}^+$:

- $X_0 = \{A\}$
- $X_1 = \{A, B\}$
- $X_2 = \{A, B, C\}$
- $X_3 = X_2$

$$\Rightarrow \{A\}^+ = \{A, B, C\}.$$

2 Tính tương tự:

- $\{C\}^+ = \{C\}$
- $\{B, C\}^+ = \{B, C\}$.

Bài toán thành viên: Phát biểu, thuật toán

Bài toán 3.1 (Bài toán thành viên)

Cho F và $X \rightarrow Y$. Xác định xem F có suy diễn được $X \rightarrow Y$?

$$F \text{ suy diễn được } X \rightarrow Y \iff Y \subseteq X^+.$$

- **Thuật toán.**
- **Đầu vào:** LĐQH $S = (U, F)$ và PTH $X \rightarrow Y$.
- **Đầu ra:** Cho biết $F \models X \rightarrow Y$?
- **Phương pháp:**
- **Bước 1.** (Tính bao đóng) Tính X^+ theo thuật toán tính bao đóng.
- **Bước 2.** (Kiểm tra và kết luận)
 - ① Nếu $Y \subseteq X^+$ thì $F \models X \rightarrow Y$
 - ② Ngược lại: $F \not\models X \rightarrow Y$.

Bài toán thành viên: Ví dụ

Ví dụ:

- Cho LĐQH $S = (U, F)$ với:
 - $U = \{A, B, C, D\}$
 - $F = \{A \rightarrow B, B \rightarrow C, A \rightarrow D\}$.
- ① Xác định xem: $F \models B \rightarrow ABC$?
 - $\{B\}^+ = \{B, C\} \Rightarrow \{A, B, C\} \not\subseteq \{B\}^+$
 - $\Rightarrow F \not\models B \rightarrow ABC$.
- ② Xác định xem: $F \models A \rightarrow U$?
 - $\{A\}^+ = \{A, B, C, D\} \Rightarrow U = \{A\}^+$
 - $\Rightarrow F \models A \rightarrow U$.

Phủ tối thiểu: Tương đương, phủ

- Cho F và G là 2 tập PTH.
- ① **Suy diễn tập PTH:** $F \models G$, nếu

$$\forall X \rightarrow Y \in G \Rightarrow F \models X \rightarrow Y.$$

- ② **PTH tương đương:** $F \equiv G$, nếu $F \models G$ và $G \models F$.
- ③ **Phủ PTH:** Khi $G \equiv F$, nói G là 1 **phủ** của F .

Ví dụ:

Xét $F = \{A \rightarrow B, B \rightarrow C, A \rightarrow C\}$, $G = \{A \rightarrow B, B \rightarrow C\}$.

$\hookrightarrow$

- $F \models G$
- $G \models F$
- $\Rightarrow F \equiv G$, do đó G là 1 phủ của F .

Phủ tối thiểu: PTH dư

- **PTH dư:** PTH $X \rightarrow Y \in F$ được gọi là **dư**, nếu:

$$F - \{X \rightarrow Y\} \equiv F.$$

Mệnh đề 3.3

$$PTH X \rightarrow Y \in F \text{ dư} \iff X_{F - \{X \rightarrow Y\}}^+ \supseteq Y.$$

Ví dụ:

Cho $F = \{A \rightarrow B, B \rightarrow C, A \rightarrow C\}$.

- Xét PTH $A \rightarrow C$: $\{A\}_{\{A \rightarrow B, B \rightarrow C\}}^+ = \{A, B, C\} \supseteq \{C\}$.
 $\Rightarrow$ PTH $A \rightarrow C$ dư.
- Xét PTH $A \rightarrow B$: $\{A\}_{\{B \rightarrow C, A \rightarrow C\}}^+ = \{A, C\} \not\supseteq \{B\}$.
 $\Rightarrow$ PTH $A \rightarrow B$ không dư.

Phủ tối thiểu: Thuộc tính dư trái

- **Thuộc tính dư trái:** Thuộc tính $A \in X$ được gọi là **dư trái** $X \rightarrow Y \in F$, nếu:

$$F - \{X \rightarrow Y\} \cup \{X - \{A\} \rightarrow Y\} \equiv F.$$

Mệnh đề 3.4

$A \in X$ là dư trái trong $X \rightarrow Y \iff (X - \{A\})^+ \supseteq Y$.

Ví dụ:

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

- Xét B trong $AB \rightarrow D$: $\{A\}^+ = \{A, B, C, D\} \supseteq \{D\}$.
 $\Rightarrow B$ là thuộc tính dư trái trong $AB \rightarrow D$.
- Xét A trong $AB \rightarrow D$: $\{B\}^+ = \{B, C\} \not\supseteq \{D\}$.
 $\Rightarrow A$ không phải thuộc tính dư trái trong $AB \rightarrow D$.

Phủ tối thiểu: Định nghĩa phủ tối thiểu

- **Phủ tối thiểu:** Tập PTH G được gọi là **phủ tối thiểu** của tập PTH F , nếu:
 - ① G là 1 phủ của F
 - ② Vế phải của mỗi PTH trong G chỉ có đúng 1 thuộc tính
 - ③ G không có TTTD
 - ④ G không có PTH dư.

Ví dụ:

Cho $F = \{A \rightarrow BC, B \rightarrow C\}$ và $G = \{A \rightarrow B, B \rightarrow C\}$.

- G là 1 phủ của F
- Vế phải mỗi PTH trong G chỉ có đúng 1 thuộc tính
- G không có TTTD
- G không có PTH dư
- $\Rightarrow G$ phủ tối thiểu của F .

Phủ tối thiểu không duy nhất

Phủ tối thiểu: Thuật toán tìm 1 phủ tối thiểu

- **Đầu vào:** Tập PTH F .
- **Đầu ra:** Phủ tối thiểu G của F .
- **Phương pháp:**
- **Bước 1.** (Khởi tạo) $G := \emptyset$.
- **Bước 2.** (Chuẩn hoá VP) Với mỗi $X \rightarrow Y \in F$, với mỗi $A \in Y - X$, nếu $X \rightarrow A \notin G$ thì

$$G := G \cup \{X \rightarrow A\}$$

- **Bước 3.** (Loại bỏ TTDT) Với mỗi $X \rightarrow Y \in G$, với mỗi $A \in X$, nếu $(X - \{A\})^+ \supseteq Y$ thì

$$G := G - \{X \rightarrow Y\} \cup \{X - \{A\} \rightarrow Y\}.$$

- **Bước 4.** (Loại bỏ PTH dư) Với mỗi $X \rightarrow Y \in G$, nếu $X_{G - \{X \rightarrow Y\}}^+ \supseteq Y$ thì

$$G := G - \{X \rightarrow Y\}$$

- **Bước 5.** (Kết luận) Tập G là 1 phủ tối thiểu cần tìm.

Phủ tối thiểu: ví dụ- (1)

Ví dụ:

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

- $G := \emptyset$
- **Chuẩn hoá VP:**

$$G = \{A \rightarrow C, B \rightarrow C, A \rightarrow B\}$$

- **Loại bỏ TTDT:** không
- **Loại bỏ PTH dư:**
 - Xét $A \rightarrow C$: $\{A\}_{G - \{A \rightarrow C\}}^+ = \{A, B, C\} \supseteq \{C\}$
 $\Rightarrow G := G - \{A \rightarrow C\} = \{B \rightarrow C, A \rightarrow B\}.$
- $\Rightarrow$ Một phủ tối thiểu của F :

$$G = \{B \rightarrow C, A \rightarrow B\}.$$

Phủ tối thiểu: ví dụ- (2)

Ví dụ:

Cho $F = \{A \rightarrow BCE, AB \rightarrow DE, BI \rightarrow J\}$.

- $G := \emptyset$.
- **Tách VP:**

$$G = \{A \rightarrow B, A \rightarrow C, A \rightarrow E, \\ AB \rightarrow D, AB \rightarrow E, \\ BI \rightarrow J\}$$

- **Loại bỏ TTDT:**
 - Xét $AB \rightarrow D$: $\{A\}^+ = \{A, B, C, E, D\} \supseteq \{D\} \Rightarrow B$ TTDT, cập nhật lại G :
 $G := \{A \rightarrow B, A \rightarrow C, A \rightarrow E, A \rightarrow D, AB \rightarrow E, BI \rightarrow J\}$

Phủ tối thiểu: ví dụ- (3)

Ví dụ: (tiếp theo)

- **Loại bỏ TTDT (tiếp theo):**
 - Xét $AB \rightarrow E$: $\{A\}^+ = \{A, B, C, E, D\} \supseteq \{E\} \Rightarrow B$ TTDT, cập nhật lại G :
 $G := \{A \rightarrow B, A \rightarrow C, A \rightarrow E, A \rightarrow D, BI \rightarrow J\}$
- **Loại bỏ PTH dư: không**
- $\Rightarrow$ Một phủ tối thiểu của F :

$$G = \{A \rightarrow B, A \rightarrow C, A \rightarrow E, A \rightarrow D, BI \rightarrow J\}.$$

Phủ tối thiểu: Lưu ý

Phủ tối thiểu không duy nhất

Ví dụ:

Xét $F = \{A \rightarrow B, B \rightarrow A, A \rightarrow C, C \rightarrow A, B \rightarrow C\}$.

- Tập F có 2 phủ tối thiểu:
 - ① $G_1 = \{A \rightarrow B, C \rightarrow A, B \rightarrow C\}$
 - ② $G_2 = \{A \rightarrow B, B \rightarrow A, A \rightarrow C, C \rightarrow A\}$.

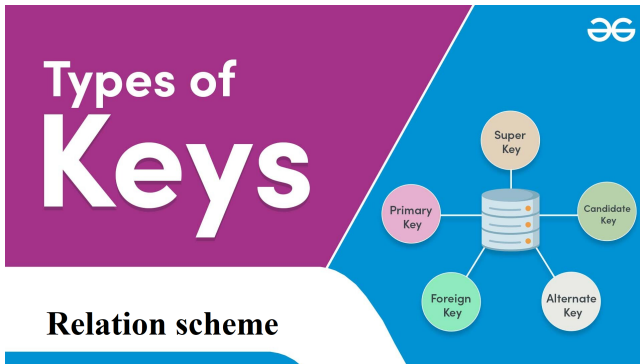
- **Bài tập**

Nội dung môn học

- ① Tổng quan
- ② Các mô hình dữ liệu
- ③ Phụ thuộc hàm
- ④ Khoá lược đồ quan hệ**
- ⑤ Chuẩn hoá lược đồ quan hệ
- ⑥ Tài liệu tham khảo

Chương 4. Khoá lược đồ quan hệ

- 1 Khái niệm khoá
- 2 Tính chất khoá
- 3 Thuật toán tìm khoá



Khái niệm khoá: Định nghĩa- (1)

- **Khoá** của LĐQH $S = (U, F)$ là tập con thuộc tính $K \subseteq U$ thoả 2 điều kiện:
 - ① **Tính đầy đủ:** $K^+ = U$ (tức $F \models K \rightarrow U$).
 - ② **Tính tối thiểu:** $\nexists K' \subset K : (K')^+ = U$.
- Nếu K thoả (1) thì K được gọi là **siêu khoá** của S .
- Ký hiệu: $\text{Key}(S)$ là tập tất cả khoá của S .

Ví dụ:

- Cho LĐQH $S = (U, F)$ với:
 - $U = \{A, B, C\}$
 - $F = \{A \rightarrow B, B \rightarrow C\}$.
- Ta có:
 - ① $\{A\}^+ = \{A, B, C\} = U$.
 - ② $\{A\}$ là tập nhỏ nhất có bao đóng bằng U .
 - $\Rightarrow \{A\}$: 1 khoá của S .
- Các siêu khoá của S : $\{A, B\}, \{A, C\} \dots$

Khái niệm khoá: Định nghĩa- (2)

- **Thuộc tính khoá:** thuộc tính thuộc 1 khóa nào đó.
- **Thuộc tính không khoá:** ngược lại
 - Ký hiệu: U_N là tập tất cả thuộc tính không khoá.

$$U_N = U - \bigcup \text{Key}(S)$$

- $\leadsto$
 - ① *Khoá $\Rightarrow$ siêu khoá.*
 - ② *U là 1 siêu khoá.*
 - ③ *LĐQH luôn có ít nhất 1 khoá.*

Khái niệm khoá: Ví dụ

Ví dụ:

- Tìm các khoá, thuộc tính khoá của LĐQH $S = (U, F)$ với:
 - $U = \{C, A, Z\}$
 - $F = \{CA \rightarrow Z, Z \rightarrow C\}$.
- Ta có:
 - $\{C, A\}^+ = \{C, A, Z\} = U$
 - $\{A, Z\}^+ = \{A, Z, C\} = U$
- Và
 - $\{C\}^+ = \{C\} \neq U$
 - $\{A\}^+ = \{A\} \neq U$
 - $\{Z\}^+ = \{Z, C\} \neq U$
- $\Rightarrow$
 - Các khoá: $\{C, A\}, \{A, Z\}$
 - Các thuộc tính khoá: C, A, Z , và $U_N = \emptyset$.

Tính chất khoá

- Xét LĐQH $S = (U, F)$:
 - $F = \{X_i \rightarrow Y_i | X_i, Y_i \subseteq U, i = 1, 2, \dots, m\}$
 - Tập tất cả thuộc tính bên trái: $U_L = \bigcup_{i=1}^m X_i$
 - Tập tất cả thuộc tính bên phải: $U_R = \bigcup_{i=1}^m Y_i$

Định lý 4.1 (Điều kiện cần của khoá)

Nếu $K \in \text{Key}(S)$ thì

$$(U - U_R) \subseteq K \subseteq (U - U_R) \cup (U_L \cap U_R).$$

Định lý 4.2 (Điều kiện có khoá duy nhất)

LĐQH S có 1 khoá duy nhất $\iff (U - U_R)^+ = U$.

Thuật toán tìm khoá: Tìm 1 khoá

- Đầu vào: LĐQH $S = (U, F)$.
- Đầu ra: Một khoá $K \in \text{Key}(S)$.
- Phương pháp:
- Bước 1.
 - ① Tính $U - U_R$.
 - ② Nếu $(U - U_R)^+ = U$ thì $K := U - U_R$, và đến bước cuối.
- Bước 2. Tính $U_L \cap U_R = \{B_1, B_2, \dots, B_n\}$. Xây dựng dãy $K_0, K_1, \dots, K_n$:

$$K_0 := (U - U_R) \cup \{B_1, B_2, \dots, B_n\}$$

$$K_i := \begin{cases} K_{i-1} - \{B_i\} & \text{nếu } (K_{i-1} - \{B_i\})^+ = U \\ K_{i-1} & \text{ngược lại} \end{cases}$$

$\forall i = 1, 2, \dots, n$. Đặt $K := K_n$.

- Bước 3. (*Kết luận*) K là 1 khoá cần tìm.

Thuật toán tìm khoá: ví dụ- (1)

- Tìm một khoá của LĐQH $S = (U, F)$ với
 - $U = \{C, A, Z\}$
 - $F = \{CA \rightarrow Z, Z \rightarrow C\}$.
- Ta có:
 - $U_R = \{C, Z\}$
 - $U - U_R = \{A\}$
 - $(U - U_R)^+ = \{A\}^+ = \{A\}$.
- Ta có
 - $U_L = \{C, A, Z\}$
 - $U_L \cap U_R = \{C, Z\}$
 - $K_0 = \{C, A, Z\}$
 - $K_1 = \{A, Z\}$
 - $K_2 = K_1$.
- $\Rightarrow \{A, Z\}$: 1 khoá của S .
- $\leadsto$

$$\text{Key}(S) = \{\{A, Z\}, \{A, C\}\}.$$

Thuật toán tìm khoá: ví dụ- (2)

- Tìm một khoá của LĐQH $S = (U, F)$ với:
 - $U = \{A, B, C\}$
 - $F = \{A \rightarrow B, B \rightarrow C\}$.
- Ta có:
 - $U_R = \{B, C\}$
 - $U - U_R = \{A\}$
 - $(U - U_R)^+ = \{A\}^+ = U$.
- $\Rightarrow \{A\}$: khoá duy nhất của S .
- **Bài tập**

Thuật toán tìm khoá: Tìm tất cả khoá- (1)

Định lý 4.3 (Phát hiện thêm khoá mới)

Giả sử $K(S)$ là 1 tập khác rỗng một số khóa của LĐQH $S = (U, F)$. Điều kiện cần và đủ để có thêm 1 khóa chưa có trong $K(S)$ là:

$$\exists X_i \rightarrow Y_i \in F, \exists K_j \in K(S) \text{ sao cho}$$

$$\forall K_p \in K(S) : X_i \cup (K_j - Y_i) \not\subseteq K_p.$$

Thuật toán tìm khoá: Tìm tất cả khoá- (2)

- Đầu vào: LĐQH $S = (U, F)$.
- Đầu ra: Tất cả khoá $Key(S)$.
- Phương pháp:
- **Bước 1.** (*Khởi tạo*) Tìm 1 khóa K , đặt $K(S) := \{K\}$.
- **Bước 2.** (*Lặp*)
 - Với mỗi $X_i \rightarrow Y_i \in F$, với mỗi $K_j \in K(S)$, nếu

$$X_i \cup (K_j - Y_i) \not\subseteq K_p, \forall K_p \in K(S) \quad (1)$$

thì tìm 1 khóa mới K' từ siêu khóa $X_i \cup (K_j - Y_i)$.

- Cập nhật $K(S) := K(S) \cup \{K'\}$.
- Lặp lại quá trình trên cho đến khi (1) không thỏa mãn.
- **Bước 3.** (*Kết luận*) $Key(S) := K(S)$.

Thuật toán tìm khoá: ví dụ

- Tìm tất cả khoá của LĐQH $S = (U, F)$ với:
 - $U = \{A, B, C, D, E, G\}$
 - $F = \{B \rightarrow C, C \rightarrow B, A \rightarrow GD\}$.
- Ta có khoá đầu tiên: $\{A, C, E\}$. Đặt

$$K(S) := \{\{A, C, E\}\}.$$

① Xét PTH $B \rightarrow C \in F$:

$$\{B\} \cup (\{A, C, E\} - \{C\}) = \{A, B, E\} \not\supseteq K_p, \forall K_p \in K(S)$$

- $\Rightarrow S$ có thêm 1 khóa mới $\subseteq \{A, B, E\}$:
- $K_0 = \{A, B, E\}$
- $K_1 = K_0$
- $K_2 = K_1$
- $K_3 = K_2$
- $\Rightarrow \{A, B, E\}$ là khoá mới.
- Cập nhật: $K(S) := \{\{A, C, E\}, \{A, B, E\}\}$.

② Không tìm được thêm PTH nào nữa. Thuật toán dừng.

- $\Rightarrow$ Tất cả khoá của S :

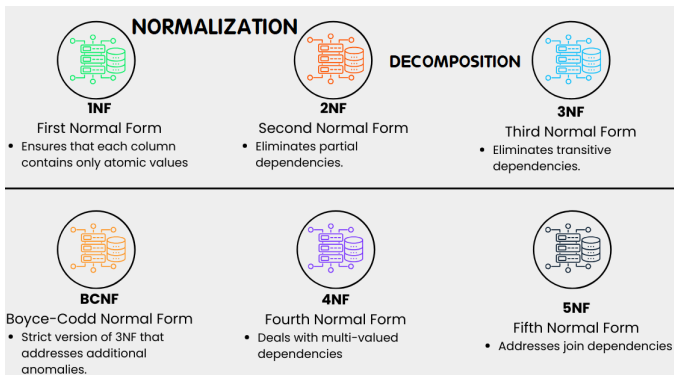
$$Key(S) = \{\{A, C, E\}, \{A, B, E\}\}.$$

Nội dung môn học

- ① Tổng quan
- ② Các mô hình dữ liệu
- ③ Phụ thuộc hàm
- ④ Khoá lược đồ quan hệ
- ⑤ Chuẩn hoá lược đồ quan hệ**
- ⑥ Tài liệu tham khảo

Chương 5. Chuẩn hoá lược đồ quan hệ

- ① Phép tách lược đồ quan hệ
- ② Các dạng chuẩn của lược đồ quan hệ
- ③ Chuẩn hóa lược đồ quan hệ qua phép tách



Phép tách LĐQH: Định nghĩa

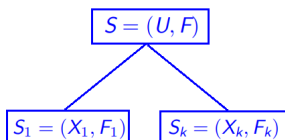
- **Tách LĐQH** $S = (U, F)$: chia S thành các lược đồ con:

$$S_1 = (X_1, F_1), \dots, S_k = (X_k, F_k)$$

sao cho:

- Điều kiện phủ: $\bigcup_{i=1}^k X_i = U, X_i \subseteq U$.
- Chiều: $\Pi_{X_i}(F) \equiv F_i$ - tập các PTH suy diễn từ F trên X_i :

$$F_i = \{X \rightarrow Y : F \models X \rightarrow Y, X \cup Y \subseteq X_i\}$$



- $\rho = \{X_1, \dots, X_k\}$: phép tách (dọc) của S .

Phép tách LĐQH: ví dụ

Ví dụ:

Cho LĐQH $S = (U, F)$ với

- $U = \{A, B, C\}$
- $F = \{A \rightarrow B, A \rightarrow C\}$.

① $\rho_1 = \{\{A\}, \{C\}\}$: không phải phép tách của S .

② $\rho_2 = \{\{A, B\}, \{A, C\}\}$: phép tách của S

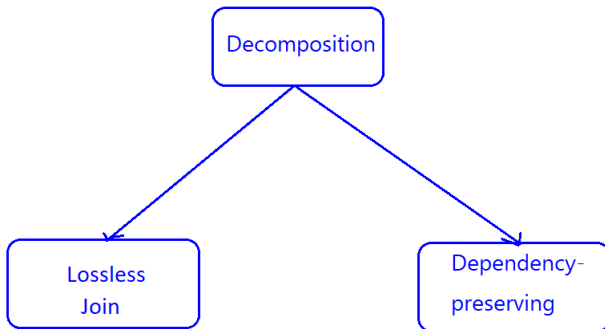
- Ta có:
 - Các lược đồ con: $S_1 = (\{A, B\}, F_1)$, $S_2 = (\{A, C\}, F_2)$
 - Các phép chiếu: $F_1 = \{A \rightarrow B\}$, $F_2 = \{A \rightarrow C\}$.

③ $\rho_3 = \{\{A, B\}, \{B, C\}\}$: phép tách của S .

- Ta có:
 - Các lược đồ con: $S_1 = (\{A, B\}, F_1)$, $S_2 = (\{B, C\}, F_2)$
 - Các phép chiếu: $F_1 = \{A \rightarrow B\}$, $F_2 = \emptyset$.

Phép tách LĐQH: Điều kiện của phép tách

- Một phép tách được xem là “tốt ” cần thoả 2 điều kiện:
 - ① Kết nối không mất thông tin (bảo toàn thông tin).
 - ② Bảo toàn các PTH.

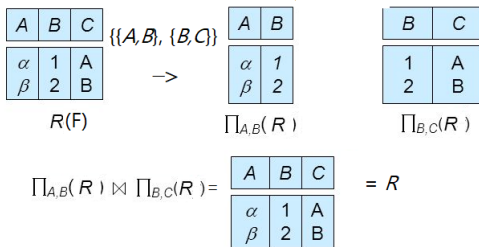


Phép tách LĐQH: Kết nối không mất thông tin

- Phép tách $\rho = \{X_1, \dots, X_k\}$ được gọi là (kết nối) **không mất thông tin**, nếu:

$$\forall R(F) : \Pi_{X_1}(R) \bowtie \Pi_{X_2}(R) \bowtie \dots \bowtie \Pi_{X_k}(R) = R$$

Ví dụ:



$\rho = \{\{A, B\}, \{B, C\}\}$ không mất thông tin

Phép tách LĐQH: Thuật toán- (1)

- Đầu vào: LĐQH $S = (U, F)$ với $U = \{A_1, \dots, A_n\}$, phép tách $\rho = \{X_1, \dots, X_k\}$.
- Đầu ra: Cho biết ρ có mất thông tin?
- Phương pháp:
- **Bước 1.** (*Khởi tạo*) Lập bảng k hàng, n cột:

$$(i, j) = \begin{cases} x_j & \text{nếu } A_j \in X_i \\ y_{ij} & \text{ngược lại} \end{cases}$$

Với

- hàng i : tương ứng X_i ($1 \leq i \leq k$)
- cột j : tương ứng A_j ($1 \leq j \leq n$).

Phép tách LĐQH: Thuật toán- (2)

- **Bước 2.** (*Biến đổi bảng*) Với mỗi $X \rightarrow Y \in F$:
 - ① Tìm các hàng có giá trị giống nhau trên X , rồi làm bằng nhau trên Y :
 - ① Nếu 1 ô có $x_j \Rightarrow$ gán x_j cho các ô tương ứng.
 - ② Nếu các ô đều là $y \Rightarrow$ chọn 1 trong chúng để thay thế (thường chọn y với chỉ số nhỏ hơn).
 - ② Quá trình biến đổi cứ thế, kể cả việc lặp lại PTH đã áp dụng, cho đến khi không thể biến đổi bảng được.
- **Bước 3.** (*Kết luận*)
 - ① Nếu bảng cuối có ít nhất 1 hàng chứa toàn ký hiệu $x_j \Rightarrow \rho$ không mất thông tin.
 - ② Ngược lại, ρ mất thông tin.

Phép tách LĐQH: ví dụ- (1)

Ví dụ:

Cho LĐQH $S = (U, F)$ với

- $U = \{A, B, C\}$
- $F = \{A \rightarrow B, B \rightarrow C\}$.

(1) Xét phép tách $\rho = \{\{A, C\}, \{B\}\}$:

- Lập bảng:

	A	B	C
$\{A, C\}$	x_1	y_{12}	x_3
$\{B\}$	y_{21}	x_2	y_{23}

- Biến đổi bảng:
 - $A \rightarrow B$: không
 - $B \rightarrow C$: không
- $\Rightarrow \rho$ mất thông tin.

Phép tách LĐQH: ví dụ- (2)

Ví dụ:

(2) Xét phép tách $\rho = \{\{A, B\}, \{B, C\}\}$:

- Lập bảng:

	A	B	C
$\{A, B\}$	x_1	x_2	y_{13}
$\{B, C\}$	y_{21}	x_2	x_3

- Biến đổi bảng: $A \rightarrow B$: không; $B \rightarrow C$: $y_{13} \Rightarrow x_3$.
- Bảng cuối:

	A	B	C
$\{A, B\}$	x_1	x_2	~~y_{13}~~ x_3
$\{B, C\}$	y_{21}	x_2	x_3

- Hàng thứ 1: (x_1, x_2, x_3) . $\Rightarrow \rho$ không mất thông tin.

Phép tách LĐQH: ví dụ- (3)

Ví dụ:

- Cho LĐQH $S = (U, F)$ với:
 - $U = \{A, B, C, D, E\}$
 - $F = \{A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A\}$
 - $\rho = \{\{A, D\}, \{A, B\}, \{B, E\}, \{C, D, E\}, \{A, E\}\}$.
- Lập bảng:

	A	B	C	D	E
$\{A, D\}$	x_1	y_{12}	y_{13}	x_4	y_{15}
$\{A, B\}$	x_1	x_2	y_{23}	y_{24}	y_{25}
$\{B, E\}$	y_{31}	x_2	y_{33}	y_{34}	x_5
$\{C, D, E\}$	y_{41}	y_{42}	x_3	x_4	x_5
$\{A, E\}$	x_1	y_{52}	y_{53}	y_{54}	x_5

Phép tách LĐQH: ví dụ- (4)

Ví dụ: (tiếp theo)

- Biến đổi bảng:
 - $A \rightarrow C: y_{23}, y_{53} \Rightarrow y_{13}$
 - $B \rightarrow C: y_{33} \Rightarrow y_{13}$
 - $C \rightarrow D: y_{24}, y_{34}, y_{54} \Rightarrow x_4$
 - $DE \rightarrow C: y_{13} \Rightarrow x_3$
 - $CE \rightarrow A: y_{31}, y_{41} \Rightarrow x_1$

	A	B	C	D	E
$\{A, D\}$	x_1	y_{12}	~~y_{13}~~ ^{x_3}	x_4	y_{15}
$\{A, B\}$	x_1	x_2	~~y_{23}~~ ^{x_3}	~~y_{24}~~ ^{x_4}	y_{25}
$\{B, E\}$	~~y_{31}~~ ^{x_1}	x_2	~~y_{33}~~ ^{x_3}	~~y_{34}~~ ^{x_4}	x_5
$\{C, D, E\}$	~~y_{41}~~ ^{x_1}	y_{42}	x_3	x_4	x_5
$\{A, E\}$	x_1	y_{52}	~~y_{53}~~ ^{x_3}	~~y_{54}~~ ^{x_4}	x_5

Phép tách LĐQH: ví dụ- (5)

Ví dụ: (tiếp theo)

- Bảng cuối:

	A	B	C	D	E
$\{A, D\}$	x_1	y_{12}	x_3	x_4	y_{15}
$\{A, B\}$	x_1	x_2	x_3	x_4	y_{25}
$\{B, E\}$	x_1	x_2	x_3	x_4	x_5
$\{C, D, E\}$	x_1	y_{42}	x_3	x_4	x_5
$\{A, E\}$	x_1	y_{52}	x_3	x_4	x_5

- Hàng thứ 3: $(x_1, x_2, x_3, x_4, x_5)$.
- $\Rightarrow \rho$ không mất thông tin.

Phép tách LĐQH: Điều kiện đủ

Định lý 5.1 (Đối với phép tách có 2 thành phần)

Phép tách $\rho = \{X_1, X_2\}$ không mất thông tin, nếu

$$X_1 \cap X_2 \rightarrow X_1 - X_2 \in F$$

hoặc

$$X_1 \cap X_2 \rightarrow X_2 - X_1 \in F.$$

Ví dụ:

- Xét LĐQH $S = (U, F)$ với
 - $U = \{A, B, C\}$, $F = \{A \rightarrow B, B \rightarrow C\}$
 - Phép tách: $\rho = \{\{A, B\}, \{B, C\}\}$.
- Ta có: $\{B, C\} \cap \{A, B\} \rightarrow \{B, C\} - \{A, B\}$

Phép tách LĐQH: Bảo toàn PTH

- Phép tách $\rho = \{X_1, \dots, X_k\}$ được gọi là **bảo toàn PTH**, nếu:

$$\bigcup_{i=1}^k \Pi_{X_i}(F) \models F$$

- $\leadsto$ Luôn có $F \models \bigcup_{i=1}^k \Pi_{X_i}(F)$. Hai tính chất kết nối không mất thông tin và bảo toàn PTH là độc lập nhau.

Ví dụ:

- Cho LĐQH $S = (U, F)$ với
 - $U = \{A, B, C\}$, $F = \{A \rightarrow B, B \rightarrow C\}$
 - Phép tách: $\rho = \{\{A, B\}, \{B, C\}\}$
- Ta có:
 - $\Pi_{\{A, B\}}(F) = \{A \rightarrow B\}$
 - $\Pi_{\{B, C\}}(F) = \{B \rightarrow C\}$
- $\Rightarrow \rho$ bảo toàn F .

Phép tách LĐQH: ví dụ

Ví dụ: (tiếp theo)

- Cho LĐQH $S = (U, F)$ với
 - $U = \{A, B, C\}$
 - $F = \{A \rightarrow B, B \rightarrow C\}$
 - Phép tách: $\rho = \{\{A, B\}, \{A, C\}\}$
- Ta có:
 - $\Pi_{\{A, B\}}(F) = \{A \rightarrow B\}$
 - $\Pi_{\{A, C\}}(F) = \{A \rightarrow C\}$
- $\Rightarrow \rho$ không bảo toàn F .

Phép tách LĐQH: Thuật toán

- Đầu vào: $S = (U, F)$ và phép tách $\rho = \{X_1, \dots, X_k\}$.
- Đầu ra: Cho biết ρ có bảo toàn PTH?
- Phương pháp:
- **Bước 1.** (*Khởi tạo*) Đặt

$$G := \bigcup_{i=1}^k \Pi_{X_i}(F)$$

- **Bước 2.** (*Kiểm tra suy diễn*) Xét $X \rightarrow Y \in F$, tính
 - ① $Z := X$. Với mỗi X_i , tính $Z := Z \cup \left[(Z \cap X_i)_F^+ \cap X_i \right]$.
 - ② Đặt $X_G^+ := Z$.
 - ③ Nếu $X_G^+ \supseteq Y \Rightarrow G \models X \rightarrow Y$. Ngược lại, $G \not\models X \rightarrow Y$.
- **Bước 3.** (*Kết luận*)
 - ① Với mỗi $X \rightarrow Y \in F$: $G \models X \rightarrow Y \Rightarrow \rho$ bảo toàn F .
 - ② Ngược lại, ρ không bảo toàn F .

Phép tách LĐQH: ví dụ- (1)

Ví dụ:

- Cho LĐQH $S = (U, F)$ với:
 - $U = \{A, B, C\}$, $F = \{A \rightarrow B, B \rightarrow C\}$
 - $\rho = \{\{A, B\}, \{B, C\}\}$.
- Đặt: $G := \Pi_{\{A, B\}}(F) \cup \Pi_{\{B, C\}}(F)$.
 - Xét $A \rightarrow B$:
 - $Z := \{A\}$
 - $\{A, B\}: Z := \{A\} \cup [\{A\}_F^+ \cap \{A, B\}] = \{A, B\}$
 - $\{B, C\}: Z := \{A, B\} \cup [\{B\}_F^+ \cap \{B, C\}] = \{A, B, C\}$
 - $\{A\}_G^+ := \{A, B, C\} \supseteq \{B\} \Rightarrow G \models A \rightarrow B$.
 - $B \rightarrow C$:
 - $Z := \{B\}$
 - $\{A, B\}: Z := \{B\} \cup [\{B\}_F^+ \cap \{A, B\}] = \{B\}$
 - $\{B, C\}: Z := \{B\} \cup [\{B\}_F^+ \cap \{B, C\}] = \{B, C\}$
 - $\{B\}_G^+ := \{B, C\} \supseteq \{C\} \Rightarrow G \models B \rightarrow C$.

Phép tách LĐQH: ví dụ- (2)

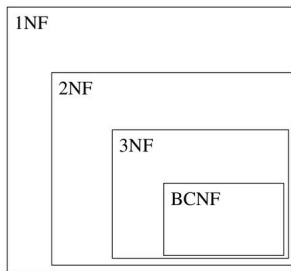
Ví dụ:

- Xét LĐQH $S = (U, F)$ với:
 - $U = \{A, B, C\}$, $F = \{A \rightarrow B, B \rightarrow C\}$
 - $\rho = \{\{A, B\}, \{A, C\}\}$.
- Đặt: $G := \Pi_{\{A, B\}}(F) \cup \Pi_{\{A, C\}}(F)$.
 - Xét $A \rightarrow B$:
 - $Z := \{A\}$
 - $\{A, B\}: Z := \{A\} \cup [\{A\}_F^+ \cap \{A, B\}] = \{A, B\}$
 - $\{A, C\}: Z := \{A, B\} \cup [\{A\}_F^+ \cap \{A, C\}] = \{A, B, C\}$
 - $\{A\}_G^+ := \{A, B, C\} \supseteq \{B\} \Rightarrow G \models A \rightarrow B$.
 - Xét $B \rightarrow C$:
 - $Z := \{B\}$
 - $\{A, B\}: Z := \{B\} \cup [\{B\}_F^+ \cap \{A, B\}] = \{B\}$
 - $\{A, C\}: Z := \{B\} \cup [\emptyset \cap \{A, C\}] = \{B\}$
 - $\{B\}_G^+ := \{B\} \not\supseteq \{C\} \Rightarrow G \not\models B \rightarrow C$.
 - $\Rightarrow \rho$ không bảo toàn F .

Các dạng chuẩn của LĐQH: Giới thiệu

- 1 Dạng chuẩn một (First Normal Form - 1NF)
- 2 Dạng chuẩn hai (Second Normal Form - 2NF)
- 3 Dạng chuẩn ba (Third Normal Form - 3NF)
- 4 Dạng chuẩn Boyce-Codd (Boyce-Codd Normal Form - BCNF)

Normalization



Các dạng chuẩn của LĐQH: Giá trị nguyên tử

- **Giá trị nguyên tử** (Atomic) trong CSDL: giá trị không thể chia nhỏ hơn được nữa (theo ngữ cảnh của ứng dụng).

Ứng dụng	Nguyên tử	Số điện thoại
Chỉ lưu/tìm kiếm cả số	 Có	02343529466
Cần truy vấn theo mã vùng	 Không	Tách 023 và ...9466

Các dạng chuẩn của LĐQH: Chuẩn một

- **Chuẩn một:** các thuộc tính của LĐQH $S = (U, F)$ chỉ chứa các giá trị nguyên tử.

• $\leadsto$

- ① Bảng sau không phải chuẩn một:

MaSV	TenSV	NgaySinh	Mon	Diem
SV0001	Lê Văn A	10-10-2000	Toán, Lý	9,9
SV0002	Lê Thị B	15-11-1999	Toán	10

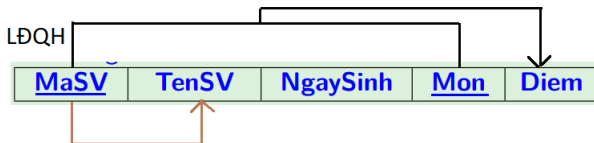
- ② Bảng sau là chuẩn một:

MaSV	TenSV	NgaySinh	Mon	Diem
SV0001	Lê Văn A	10-10-2000	Toán	9
SV0001	Lê Văn A	10-10-2000	Lý	9
SV0002	Lê Thị B	15-11-1999	Toán	10

- $\leadsto$ Giả thiết: *LĐQH tối thiểu là chuẩn một.*

Các dạng chuẩn của LĐQH: Chuẩn hai- (1)

- Xét LĐQH $S = (U, F)$:
 - ① ***Y phụ thuộc đầy đủ*** (hoàn toàn) vào X : nếu $F \models X \rightarrow Y$ và X không có TTDT.
 - ② ***Y phụ thuộc bộ phận*** (một phần) vào X : ngược lại.



$\{MaSV, Mon\} \rightarrow \{Diem\}$: Phụ thuộc đầy đủ

$\{MaSV, Mon\} \rightarrow \{TenSV\}$: Phụ thuộc bộ phận

MaSV	TenSV	NgaySinh	Mon	Diem
SV0001	Lê Văn A	10-10-2000	Toán	9
SV0001	Lê Văn A	10-10-2000	Lý	10
SV0002	Lê Thị B	15-1-1999	Toán	7
SV0002	Lê Thị B	15-1-1999	Hoá	9

Các dạng chuẩn của LĐQH: Chuẩn hai- (2)

Ví dụ:

- Xét LĐQH $S = (U, F)$ với:
 - $U = \{A, B, C, D\}$
 - $F = \{AB \rightarrow C, B \rightarrow CD, A \rightarrow C\}$.
- Ta có:
 - $B \rightarrow CD$ và $\{B\}$ không có TTDT
 $\Rightarrow \{C, D\}$ PTĐĐ vào $\{B\}$.
 - $AB \rightarrow C$ nhưng B là TTDT trong $\{A, B\}$
 $\Rightarrow \{C\}$ PTBP vào $\{A, B\}$.

Các dạng chuẩn của LĐQH: Chuẩn hai- (3)

- **Chuẩn hai:** LĐQH $S = (U, F)$ thỏa:
 - ① 1NF.
 - ② Các thuộc tính không khoá phải PTĐĐ vào khóa.

Định lý 5.2 (2 dấu hiệu nhận biết 2NF)

- (1) $\forall K \in \text{Key}(S) : |K| = 1 \Rightarrow S \text{ là } 2NF.$
- (2) $U_N = \emptyset \Rightarrow S \text{ là } 2NF.$

Các dạng chuẩn của LĐQH: Chuẩn hai- (4)

Ví dụ:

① Xét LĐQH $S = (U, F)$ với:

- $U = \{A, B, C, D\}$
- $F = \{AB \rightarrow C, B \rightarrow CD, A \rightarrow C\}$.

Ta có:

- $Key(S) = \{\{A, B\}\} \Rightarrow U_N = \{C, D\}$
- $AB \rightarrow C$ và $A \rightarrow C \Rightarrow$ TTKK C PTBP vào khoá $\{A, B\}$
- $\Rightarrow S$ không phải 2NF.

② Xét LĐQH $S = (U, F)$ với:

- $U = \{A, B, C\}$
- $F = \{A \rightarrow B, B \rightarrow C\}$.

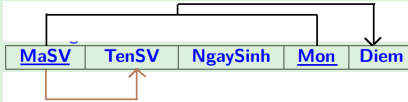
Ta có:

- $Key(S) = \{\{A\}\}$
- $\Rightarrow S$ 2NF (theo định lý).

Các dạng chuẩn của LĐQH: Chuẩn hai- (5)

Ví dụ:

- Xét LĐQH sau:



<u>MaSV</u>	<u>TenSV</u>	NgaySinh	<u>Mon</u>	<u>Diem</u>
SV0001	Lê Văn A	10-10-2000	Toán	9
SV0001	Lê Văn A	10-10-2000	Lý	10
SV0002	Lê Thị B	15-1-1999	Toán	7
SV0002	Lê Thị B	15-1-1999	Hoá	9

- Ta có: $\{MaSV, Mon\}$ là khoá $\Rightarrow$
 $U_N = \{TenSV, NgaySinh, Diem\}$
 - $MaSV Mon \rightarrow Ten$ và $MaSV \rightarrow TenSV \Rightarrow$ TTKK
 $TenSV$ PTBP khoá $MaSV Mon$.
- $\Rightarrow$ LĐQH không phải 2NF.

Các dạng chuẩn của LĐQH: Chuẩn ba- (1)

- Xét $S = (U, F)$ và PTH $X \rightarrow Z$:
 - Z **phụ thuộc bắc cầu** vào X , nếu tồn tại tập thuộc tính trung gian Y :
 - ① $F \models X \rightarrow Y, F \models Y \rightarrow Z$
 - ② $F \not\models Y \rightarrow X, Z \not\subseteq X \cup Y$.
 - Ngược lại: PTH $X \rightarrow Z$ được gọi là **phụ thuộc trực tiếp**.

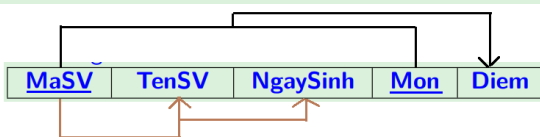
Ví dụ:

- Cho LĐQH $S = (U, F)$ với:
 - $U = \{A, B, C, D\}$
 - $F = \{AB \rightarrow C, C \rightarrow D\}$.
- Ta có:
 - $AB \rightarrow C$ và $C \rightarrow D$
 - $F \not\models C \rightarrow AB$ và $\{D\} \not\subseteq \{A, B, C\}$
 $\Rightarrow F \models AB \rightarrow D$: $\{D\}$ PTBC vào $\{A, B\}$.
 - PTH $AB \rightarrow C$: PTTT.

Các dạng chuẩn của LĐQH: Chuẩn ba- (2)

Ví dụ:

Xét LĐQH sau:



{MaSV} → {NgaySinh}: Phụ thuộc bắc cầu

<u>MaSV</u>	TenSV	NgaySinh	<u>Mon</u>	Diem
SV0001	Lê Văn A	10-10-2000	Toán	9
SV0001	Lê Văn A	10-10-2000	Lý	10
SV0002	Lê Thị B	15-1-1999	Toán	7
SV0002	Lê Thị B	15-1-1999	Hoá	9

Các dạng chuẩn của LĐQH: Chuẩn ba- (3)

- **Chuẩn ba:** LĐQH $S = (U, F)$ thoả:
 - ① 1NF
 - ② Các thuộc tính không khoá phải PTTT vào khóa.

Định lý 5.3 (Dấu hiệu nhận biết 3NF)

$U_N = \emptyset \Rightarrow S$ là 3NF.

Ví dụ: Xét bảng có LĐQH S :

<u>MaSV</u>	TenSV	MaLop	TenLop
SV0001	Nguyễn Văn A	L01	Toán
SV0002	Nguyễn Văn A	L02	Hoá
SV0003	Lê Thị B	L01	Toán

- Khoá: $\{MaSV\} \Rightarrow U_N = \{TenSV, MaLop, TenLop\}$.
- PTH: $MaSV \rightarrow MaLop, MaLop \rightarrow TenLop$
- $\Rightarrow MaSV \rightarrow TenLop$: PTBC $\Rightarrow S$ không phải chuẩn 3.

Các dạng chuẩn của LĐQH: Chuẩn ba- (4)

Ví dụ:

① Xét LĐQH $S = (U, F)$ với:

- $U = \{C, A, Z\}$
- $F = \{CA \rightarrow Z, Z \rightarrow C\}$.

Ta có:

- $\text{Key}(S) = \{\{C, A\}, \{A, Z\}\} \Rightarrow U_N = \emptyset$.
- $\Rightarrow S$ là chuẩn ba (theo định lý).

② Xét LĐQH $S = (U, F)$ với:

- $U = \{A, B, C\}$
- $F = \{A \rightarrow B, B \rightarrow C\}$.

Ta có:

- $\text{Key}(S) = \{\{A\}\} \Rightarrow U_N = \{B, C\}$.
- $A \rightarrow B, B \rightarrow C$ và $F \not\models B \rightarrow A$
- $\Rightarrow F \models A \rightarrow C$: PTBC giữa khoá $\{A\}$ và TTKK C .
- $\Rightarrow S$ không phải chuẩn 3.

Các dạng chuẩn của LĐQH: Chuẩn BCNF- (1)

- **Chuẩn Boyce-Codd:** LĐQH $S = (U, F)$ thỏa:

- ① 1NF
- ② Các thuộc tính phải PTTT vào khóa.

Định lý 5.4 (Điều kiện cần và đủ)

S là BCNF $\Leftrightarrow \forall F \models X \rightarrow A$ và $A \notin X$ thì X phải là siêu khoá.

Ví dụ:

- Xét LĐQH $S = (U, F)$ với:
 - $U = \{C, A, Z\}$, $F = \{CA \rightarrow Z, Z \rightarrow C\}$.
- Ta có:
 - $Key(S) = \{\{C, A\}, \{A, Z\}\}$.
 - Dễ thấy: $\{C\}$ PTBC vào khoá $\{C, A\}$
 - $\Rightarrow S$ không phải chuẩn Boyce-Codd.
- Nếu thay lại $F = \{CA \rightarrow Z, Z \rightarrow CA\}$ thì S là BCNF.

Các dạng chuẩn của LĐQH: Chuẩn BCNF- (2)

Ví dụ: Xét bảng có LĐQH S:

<u>MaSV</u>	TenSV	MaLop	MaGV
SV0001	Nguyễn Văn A	L01	GV01
SV0002	Nguyễn Văn A	L02	GV02
SV0003	Lê Thị B	L01	GV01

- Ta có:
 - Khoá: {MaSV}
 - PTH: $MaLop \rightarrow MaGV$ và {MaLop} không phải siêu khoá
 - $\Rightarrow S$ không phải chuẩn Boyce-Codd.

Các dạng chuẩn của LĐQH: Chuẩn BCNF- (3)

Ví dụ:

Tách LĐQH S trên thành các lược đồ con:

① Lược đồ SV:

<u>MaSV</u>	TenSV	MaLop
SV0001	Nguyễn Văn A	L01
SV0002	Nguyễn Văn A	L02
SV0003	Lê Thị B	L01

② Lược đồ lớp-giáo viên:

<u>MaLop</u>	MaGV
L01	GV01
L02	GV02

- $\Rightarrow$ Các lược đồ con trên đều là chuẩn Boyce-Codd.

Các dạng chuẩn của LĐQH: Mỗi quan hệ

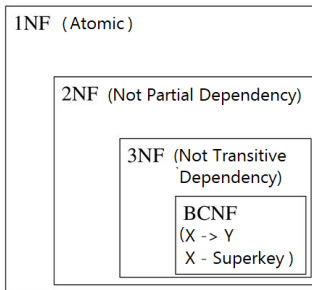
Định lý 5.5

Cho $S = (U, F)$ là LĐQH. Khi đó

(1) S là 2NF $\Rightarrow S$ là 1NF.

(2) S là 3NF $\Rightarrow S$ là 2NF.

(3) S là BCNF $\Rightarrow S$ là 3NF.



Chuẩn hoá LĐQH: 3NF-Thuật toán

- **Đầu vào:** $S = (U, F)$.
- **Đầu ra:** Phép tách $\rho = \{X_1, \dots, X_k\}$ thoả
 - ① Bảo toàn F và không mất thông tin.
 - ② Các X_i là 3NF ứng với $\Pi_{X_i}(F)$.
- **Phương pháp:**
- **Bước 1.** Tìm 1 PTT G của F . Tìm 1 khoá K của S .
- **Bước 2.** Nếu $\exists b \in U$ sao cho $b \notin X \cup \{a\}$, $\forall X \rightarrow a \in G$ thì $U := U - \{b\}$.
- **Bước 3.** (Tính toán và kết luận) Xét 2 trường hợp:
 - ① Nếu $\exists X \rightarrow a \in G$ sao cho $X \cup \{a\} = U$ thì $\rho := \{U\}$.
Thuật toán dừng. Kết luận ρ là phép tách cần tìm.
 - ② Ngược lại:
 - ① Khởi tạo $\rho := \emptyset$.
 - ② Với mỗi $X \rightarrow a \in G$ đặt $\rho := \rho \cup \{X \cup \{a\}\}$. Trường hợp, nếu có các $X \rightarrow a_1, \dots, X \rightarrow a_k \in G$ thì thay các $X \cup \{a_i\} \in \rho$ ($i = 1, \dots, k$) bằng $X \cup \{a_1, \dots, a_k\}$.
 - ③ $\rho := \rho \cup \{K\}$. Kết luận ρ là phép tách cần tìm.

Chuẩn hoá LĐQH: 3NF-Ví dụ - (1)

1. Xét $S = (U, F)$:

- $U = \{A, B, C, D\}$
- $F = \{AB \rightarrow C, C \rightarrow D, D \rightarrow A\}$.
- $\hookrightarrow$
 - ① Chuẩn hoá về phải: F thoả.
 - ② RGT: F thoả.
 - ③ Loại bỏ PTH dư: F thoả.
 $\Rightarrow F$ là tối thiểu.
- $\hookrightarrow$
 - Dễ thấy $K = \{A, B\}$ là 1 khoá.
 - Vận dụng thuật toán:
 - $\rho := \emptyset$
 - $AB \rightarrow C: \rho = \{\{A, B, C\}\}$
 - $C \rightarrow D: \rho = \{\{A, B, C\}, \{C, D\}\}$
 - $D \rightarrow A: \rho = \{\{A, B, C\}, \{C, D\}, \{D, A\}\}$
- $\Rightarrow$ Phép tách thành 3NF sao cho bảo toàn PTH và không mất thông tin:

Chuẩn hoá LĐQH: 3NF-Ví dụ - (2)

2. Xét $S = (U, F)$:

- $U = \{A, B, C, D, E\}$
- $F = \{AB \rightarrow CDE, C \rightarrow DE, D \rightarrow E\}$.

• $\hookrightarrow$

① Chuẩn hoá về phải: F thoả.

$$G = \{AB \rightarrow C, AB \rightarrow D, AB \rightarrow E, \\ C \rightarrow D, C \rightarrow E, \\ D \rightarrow E\}.$$

② RGT: G thoả.

③ Loại bỏ PTH dư:

- $AB \rightarrow D$ dư:

$$\{A, B\}_{G=\{AB \rightarrow D\}}^+ = \{A, B, C, E, D\} \supseteq \{D\} \\ \Rightarrow G := \{AB \rightarrow C, AB \rightarrow E, C \rightarrow D, C \rightarrow E, D \rightarrow E\}$$

- $AB \rightarrow E$ dư:

$$\{A, B\}_{G=\{AB \rightarrow E\}}^+ = \{A, B, C, D, E\} \supseteq \{E\} \\ \Rightarrow G := \{AB \rightarrow C, C \rightarrow D, C \rightarrow E, D \rightarrow E\}$$

- $C \rightarrow E$ dư: $\{C\}_{G=\{C \rightarrow E\}}^+ = \{C, D, E\} \supseteq \{E\}$

$$\Rightarrow G := \{AB \rightarrow C, C \rightarrow D, D \rightarrow E\}.$$

- **1 PTT của F :** $G := \{AB \rightarrow C, C \rightarrow D, D \rightarrow E\}.$

Chuẩn hoá LĐQH: 3NF-Ví dụ - (3)

- $\hookrightarrow$
 - Dễ thấy $Key(S) = \{\{A, B\}\}$ (tìm trên G nhanh hơn).
 - Vận dụng thuật toán:
 - $\rho := \emptyset$
 - $AB \rightarrow C: \rho = \{\{A, B, C\}\}$
 - $C \rightarrow D: \rho = \{\{A, B, C\}, \{C, D\}\}$
 - $D \rightarrow E: \rho = \{\{A, B, C\}, \{C, D\}, \{D, E\}\}.$
- $\Rightarrow$ Phép tách thành 3NF sao cho bảo toàn PTH và không mất thông tin:

$$\rho = \{\{A, B, C\}, \{C, D\}, \{D, E\}\}.$$

Chuẩn hoá LĐQH: BCNF-Thuật toán

- Đầu vào: LĐQH $S = (U, F)$.
- Đầu ra: Phép tách $\rho = \{X_1, \dots, X_k\}$ thoả
 - ① Không mất thông tin.
 - ② Các X_i là BCNF ứng với $\Pi_{X_i}(F)$.
- Phương pháp:
 - **Bước 1.** (Khởi tạo) $Z := U, \rho := \emptyset$.
 - **Bước 2.** (Lặp: cho đến khi Z là BCNF)
 - ① Tìm 1 $X \rightarrow A$ trên Z vi phạm BCNF.
 - ② Tách Z thành $X \cup \{A\}$ và $Z - \{A\}$.
 - ③ $\rho := \rho \cup \{X \cup \{A\}\}$.
 - ④ $Z := Z - \{A\}$.
 - **Bước 3.** (Kết luận) ρ là phép tách cần tìm.

Chuẩn hoá LĐQH: BCNF-Ví dụ - (1)

1. Xét $S = (U, F)$:

- $U = \{C, T, H, R, S', G\}$
- $F = \{C \rightarrow T, HR \rightarrow C, HT \rightarrow R, CS' \rightarrow G, HS' \rightarrow R\}$
- $\leadsto$

① U :

- C : Lớp học phần (Course)
- T : Giáo viên (Teacher)
- H : Giờ học (Hour)
- R : Phòng học (Room)
- S' : Sinh viên (Student)
- G : Điểm học phần (Grade)

② F :

- $C \rightarrow T$: mỗi LHP do 1 GV phụ trách
- $HR \rightarrow C$: tại mỗi PH, trong mỗi GH chỉ có 1 LHP
- $HT \rightarrow R$: tại mỗi GH, mỗi GV chỉ dạy trong 1 PH
- $CS' \rightarrow G$: đối với mỗi LHP, mỗi SV chỉ có 1 ĐHP
- $HS' \rightarrow R$: tại mỗi GH, mỗi SV chỉ có mặt trong 1 PH.

Chuẩn hoá LĐQH: BCNF-Ví dụ - (2)

- $\hookrightarrow$
 - $Key(S) = \{\{H, S'\}\}$
 - $Z := U, \rho := \emptyset$
 - **Vòng lặp 1:**
 - $CS' \rightarrow G$ vi phạm BCNF trên $Z = U$
 - $\Rightarrow Z := \{C, T, H, R, S'\}$ và $\rho := \{\{C, S', G\}\}$
 - **Vòng lặp 2:**
 - $C \rightarrow T$ vi phạm BCNF trên $Z = \{C, T, H, R, S'\}$
 - $\Rightarrow Z := \{C, H, R, S'\}$ và $\rho := \{\{C, S', G\}, \{C, T\}\}$
 - **Vòng lặp 3:**
 - $HR \rightarrow C$ vi phạm BCNF trên $Z = \{C, H, R, S'\}$
 - $\Rightarrow Z := \{C, H, S'\}$ và
 $\rho := \{\{C, S', G\}, \{C, T\}, \{C, H, R\}\}$
 - Z là BCNF.
 - $\Rightarrow$ Phép tách S thành BCNF không mất thông tin:
- $$\rho := \{\{C, S', G\}, \{C, T\}, \{C, H, R\}, \{C, H, S'\}\}.$$

Chuẩn hoá LĐQH: BCNF-Ví dụ - (3)

2. Xét LĐQH quản lý điểm SV $S = (U, F)$:

- $U = \{MaSV, TenSV, MaMH, TenMH, Diem\}$
- $F = \{MaSV \rightarrow TenSV, MaMH \rightarrow TenMH, \{MaSV, MaMH\} \rightarrow Diem\}$.
- $\hookrightarrow$
 - $Key(S) = \{\{MaSV, MaMH\}\}$
 - $Z := U, \rho := \emptyset$
 - **Vòng lặp 1:**
 - $MaSV \rightarrow TenSV$ vi phạm BCNF trên $Z = U$
 - $\Rightarrow Z := \{MaSV, MaMH, TenMH, Diem\}, \rho := \{\{MaSV, TenSV\}\}$
 - **Vòng lặp 2:**
 - $MaMH \rightarrow TenMH$ vi phạm BCNF trên $Z = \{MaSV, MaMH, TenMH, Diem\}$
 - $\Rightarrow Z := \{MaSV, MaMH, Diem\}, \rho := \{\{MaSV, TenSV\}, \{MaMH, TenMH\}\}$
 - Z đã là BCNF.

Chuẩn hoá LĐQH: BCNF-Ví dụ - (4)

- $\Rightarrow$ Phép tách S thành BCNF không mất thông tin:

$$\rho = \{ \{MaSV, TenSV\}, \{MaMH, TenMH\}, \\ \{MaSV, MaMH, Diem\} \}$$

- $\leadsto$ PTH:

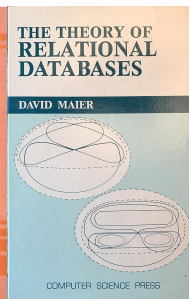
- ① $MaSV \rightarrow TenSV$: xác định duy nhất 1 SV
- ② $MaMH \rightarrow TenMH$: xác định duy nhất 1 môn học.

Nội dung môn học

- ① Tổng quan
- ② Các mô hình dữ liệu
- ③ Phụ thuộc hàm
- ④ Khoá lược đồ quan hệ
- ⑤ Chuẩn hoá lược đồ quan hệ
- ⑥ Tài liệu tham khảo**

Tài liệu tham khảo

-  Nguyễn Hoàng Sơn, *Giáo trình CSDL*, NXB ĐH Huế, 2018 (Download: **1**, **2**).
-  Nguyễn Xuân Huy, *Các phụ thuộc logic trong CSDL*, NXB Thống kê, 2006.
-  David Maier, *The theory of relational databases*, Computer Science Press, 1983.



Bảng ký hiệu, viết tắt - (1)

Viết tắt	Ý nghĩa
PTH	Phụ thuộc hàm
DL	Dữ liệu
CSDL	Cơ sở dữ liệu
ĐPT	Độ phức tạp
TTDT	Thuộc tính dư thừa
LĐQH	Lược đồ quan hệ
KD	không dư
PKD	Phủ không dư
RG	Rút gọn
RGT	Rút gọn trái
RGP	Rút gọn phải
PRG	Phủ rút gọn
PRGT	Phủ rút gọn trái
PTT	Phủ tối thiểu

Bảng ký hiệu, viết tắt - (2)

Viết tắt	Ý nghĩa
TTCB	Thuộc tính cơ bản
TTKCB	Thuộc tính không cơ bản
PTHT	Phụ thuộc hoàn toàn
PTBP	Phụ thuộc bộ phận
LĐKN	Lược đồ khái niệm
CTUD	Chương trình ứng dụng
XLGD	Xử lý giao dịch
TVDL	Truy vấn dữ liệu
TTH	Thực thể