

# CƠ SỞ DỮ LIỆU

ThS.Lê Văn Hạnh



# Mục tiêu môn học

---

- Thời lượng: 45 tiết lý thuyết + 45 tiết thực hành
- Các hệ quản trị cơ sở dữ liệu
- Các mô hình dữ liệu
  - Mô hình thực thể - kết hợp
  - Mô hình quan hệ
- Đại số quan hệ
- Ngôn ngữ truy vấn có cấu trúc
- Chuẩn hóa lược đồ CSDL



# Nội dung môn học

---

1. Mục đích
2. Thiết kế mô hình xử lý bên ngoài – giao diện
3. Thiết kế mô hình xử lý bên trong
4. Các sơ liệu cho mô xử lý mức logic



# Chương 1

---

# DẪN NHẬP



# Nội dung

---

- 1. Cơ Sở Dữ Liệu Là Gì?
- 2. Hệ Quản Trị CSDL
- 3. Các Mô Hình Dữ Liệu
- 4. Quản Lý Dữ Liệu Phân Tán
- 5. Mô Hình Thực Thể Kết Hợp
- 6. Bài tập



# \$1.1. Cơ Sở Dữ Liệu Là Gì?

---

- Nhu cầu: tin học hóa công tác quản lý
  - Dữ liệu: là những thông tin ta muốn lưu trữ và sử dụng lại
  - Quản lý dữ liệu trong thực tế:
    - Việc tin học hóa trong các đơn vị thường thực hiện độc lập với nhau về thời gian, công cụ và nội dung
    - Do yêu cầu quản lý, mỗi bộ phận đều tổ chức lưu trữ hồ sơ riêng cho mình dựa trên hồ sơ gốc ban đầu có bổ sung thêm 1 số thông tin cần thiết riêng



# \$1.1. Cơ Sở Dữ Liệu Là Gì?

---

- Tồn tại nhiều bất cập trong quản lý
  - Phương pháp cũ quản lý theo hướng công việc dẫn đến
    - Dữ liệu chứa giá trị trống (NULL) nên dễ dẫn đến sự nhập nhằng về mặt ngữ nghĩa
    - Khó mở rộng hoặc kết nối các hệ thống
    - Trùng lặp thông tin
    - Nhu cầu truy xuất đồng thời
    - Vấn đề phân quyền

# Bất cập trong quản lý

- Dữ liệu chứa giá trị trống (NULL) nên dễ dẫn đến sự nhập nhằng về mặt ngữ nghĩa

<i>Tên thọ</i>	<i>Phá i</i>	<i>Số con</i>
Trang	Nữ	2
Hung	Nữ	0
Trang	Nữ	3
Quoc	Na	0
Hau	Na	
Chau	Nữ	



# Bất cập trong quản lý

- Khó mở rộng hoặc kết nối các hệ thống
  - Thông tin đã lưu trữ không thể chia sẻ giữa các hệ thống. Vì vậy không có quản lý giao dịch và xử lý đồng thời giữa nhiều người dùng

<i>Tên thợ</i>	<i>Phái</i>	<i>Năm sinh</i>	<i>Quê quán</i>	<i>Các thông tin khác</i>
...	...	...	...	...
...	...	...	...	...

<i>Tên thợ</i>	<i>Trình độ chuyên môn</i>	<i>Bậc lương</i>	<i>Thâm niên</i>	<i>Các thông tin khác</i>
...	...	...	...	...
...	...	...	...	...

# Bất cập trong quản lý

- Trùng lặp thông tin nên gây ra :
  - ① Lãng phí:
    - Cùng 1 thông tin phải nhập nhiều lần.
    - Vấn đề bảo trì khi có thay đổi.
    - Lưu trữ .
  - ② Thiếu nhất quán của dữ liệu: tại 1 thời điểm, thông tin về một người có thể không

Tên NV	Phòng	Tháng	Lương
Thúy	QTKD	1/2007	32000
An	QYKD	1/2007	35000
Thúy	QTKD	13/2007	3200

# Định nghĩa CSDL

- Là một tập hợp dữ liệu được tổ chức và lưu trữ theo một cấu trúc chặt chẽ nhằm phục vụ cho nhiều đối tượng với các mục đích khác nhau

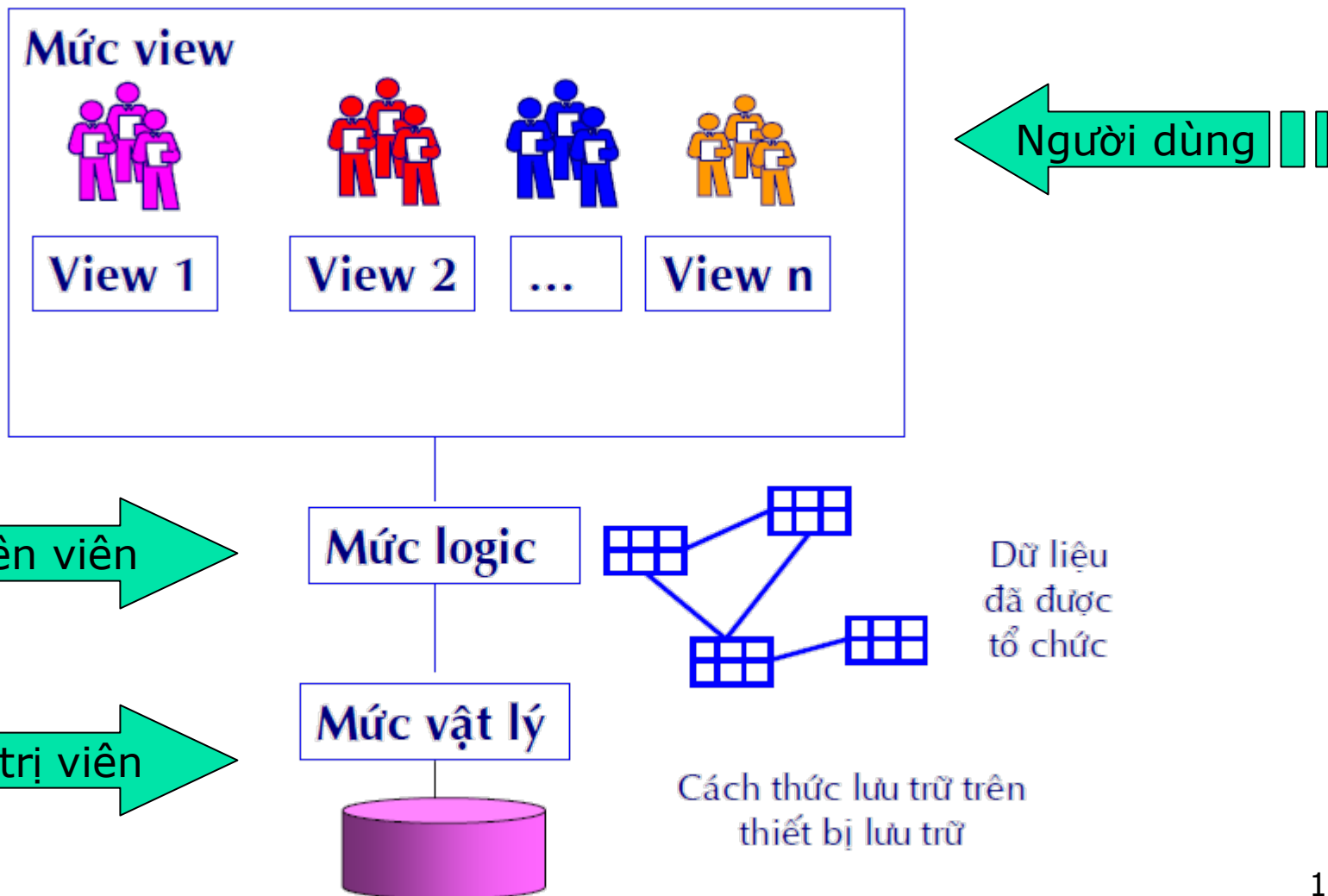


# Các đối tượng sử dụng CSDL

---

- Người sử dụng không chuyên về lĩnh vực tin học và CSDL
- Chuyên viên tin học biết khai thác CSDL
- Người quản trị CSDL

# Ba cấp độ thể hiện dữ liệu



# Như thế nào là 1 CSDL kém?

- Dư thừa (redundancy)
- Mâu thuẫn tiềm ẩn (potential inconsistency)
- Bất thường khi chèn (insertion anomaly)
- Bất thường khi xóa (deletion anomaly)

Hàng	ĐVT	SL	Đơn giá	Ngày nhập	NCC	Khu vực
Nokia N82	bộ	50	xxx	20/12/2007	Công ty FPT	TpHCM
Nokia N73	bộ	10	yyy	20/12/2007	Thành Cong Mobile	Tp HCM
Nokia N82	bộ	20	xx	03/05/2008	Cong ty FPT	Tp.HCM



# Nội dung

---

- 1. Cơ Sở Dữ Liệu Là Gì?
- 2. Hệ Quản Trị CSDL
- 3. Các Mô Hình Dữ Liệu
- 4. Quản Lý Dữ Liệu Phân Tán
- 5. Mô Hình Thực Thể Kết Hợp
- 6. Bài tập



## \$1.2. Hệ Quản Trị CSDL

---

- Khả năng quản lý những dữ liệu cố định
- Khả năng truy xuất có hiệu quả một số lượng lớn dữ liệu
- Hỗ trợ ít nhất một mô hình dữ liệu (data model): giúp người sử dụng có thể xem được dữ liệu nhanh chóng, dễ dàng.
- Hỗ trợ ngôn ngữ vấn tin (query language): cho phép người sử dụng định nghĩa các cấu trúc dữ liệu, truy xuất dữ liệu và thao tác dữ liệu.
- Quản lý các giao dịch (transaction): cho phép nhiều người sử dụng truy xuất đồng thời và chính xác đến một CSDL.





## \$1.2. Hệ Quản Trị CSDL

---

- Điều khiển các quá trình truy xuất: là khả năng giới hạn các quá trình truy xuất dữ liệu của những người không được phép và khả năng kiểm tra độ tin cậy của dữ liệu.
- Có đặc tính tự thích ứng: là khả năng tự phục hồi lại dữ liệu của hệ thống khi gặp sự cố mà không làm mất dữ liệu.
- Các hệ quản trị CSDL phổ biến:
  - Access
  - SQL Server
  - Oracle



# Nội dung

---

- 1. Cơ Sở Dữ Liệu Là Gì?
- 2. Hệ Quản Trị CSDL
- 3. Các Mô Hình Dữ Liệu
- 4. Quản Lý Dữ Liệu Phân Tán
- 5. Mô Hình Thực Thể Kết Hợp
- 6. Bài tập



## \$1.3. Các mô hình dữ liệu

---

- Mô hình dữ liệu là một hệ hình thức toán học gồm 2 phần:
  - Khả năng quản lý những dữ liệu cố định
  - Một tập hợp các phép toán thao tác trên dữ liệu đó
- Nói cách khác, mô hình dữ liệu là một tập hợp ký hiệu và quy tắc cho phép mô tả dữ liệu, mối liên hệ trên dữ liệu, ngữ nghĩa và các ràng buộc trên dữ liệu
- Mỗi loại mô hình đặc trưng cho một phương pháp tiếp cận dữ liệu của người phân tích-thiết kế



# \$1.3. Các mô hình dữ liệu

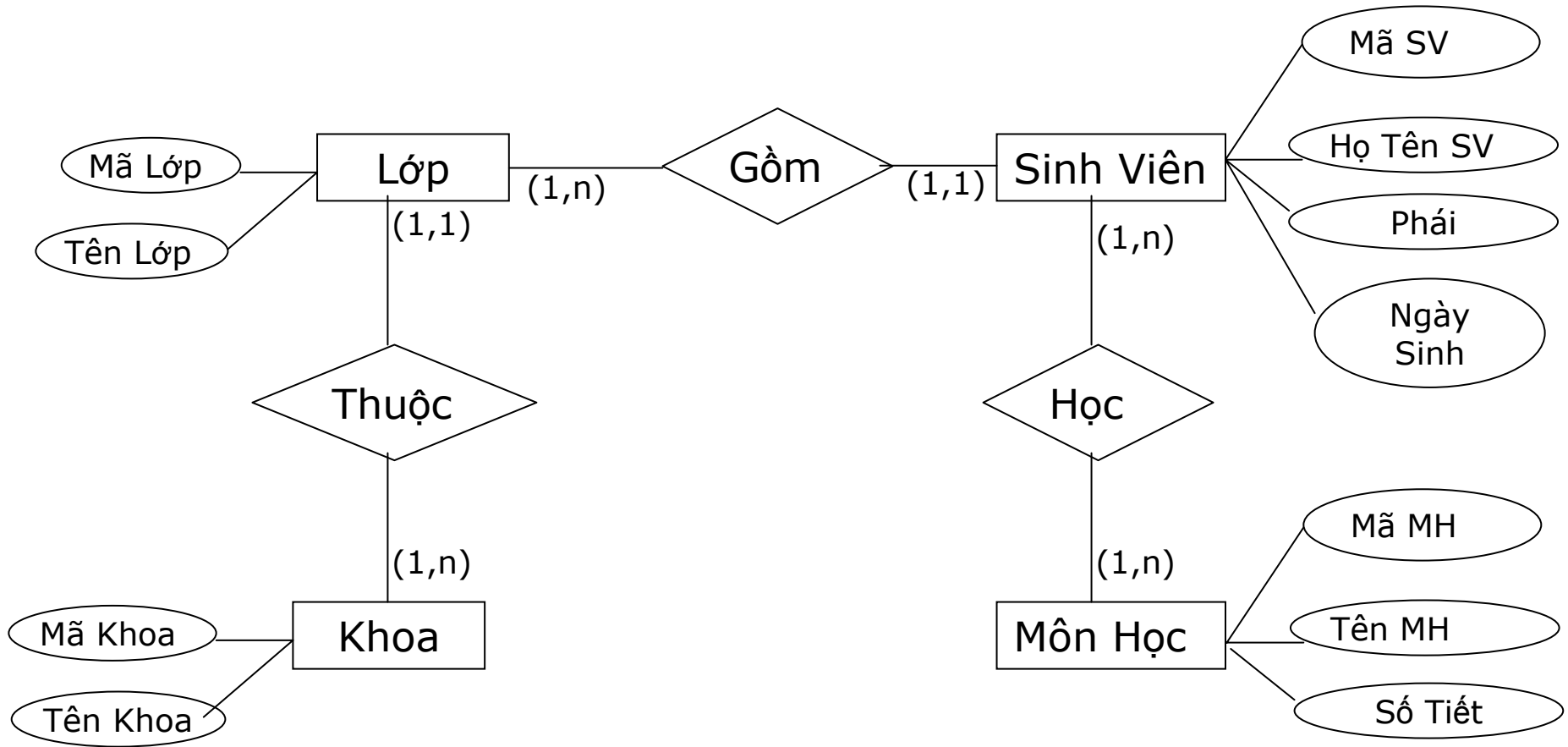
---

- Mô hình logic trên cơ sở đối tượng (Object-based logical models)
  - Mô hình Thực thể – Kết hợp (Entity-Relation)
  - Mô hình hướng đối tượng (Object-Oriented Model)
- Mô hình logic trên cơ sở mẫu tin (Record-based logical models)
  - Mô hình quan hệ
  - Mô hình mạng
  - Mô hình phân cấp
- Mô hình vật lý (Physical Models)

# Mô hình thực thể - kết hợp

- Mục đích:
  - Mô tả lược đồ khái niệm của 1 tổ chức mà không cần chú ý đến tính hiệu quả hoặc thiết kế vật lý như những mô hình khác.
  - Thường được chuyển thành lược đồ khái niệm trong mô hình khác.
- Các thành phần:
  - Loại thực thể
  - Mối kết hợp

# Mô hình thực thể - kết hợp



# Mô hình hướng đối tượng

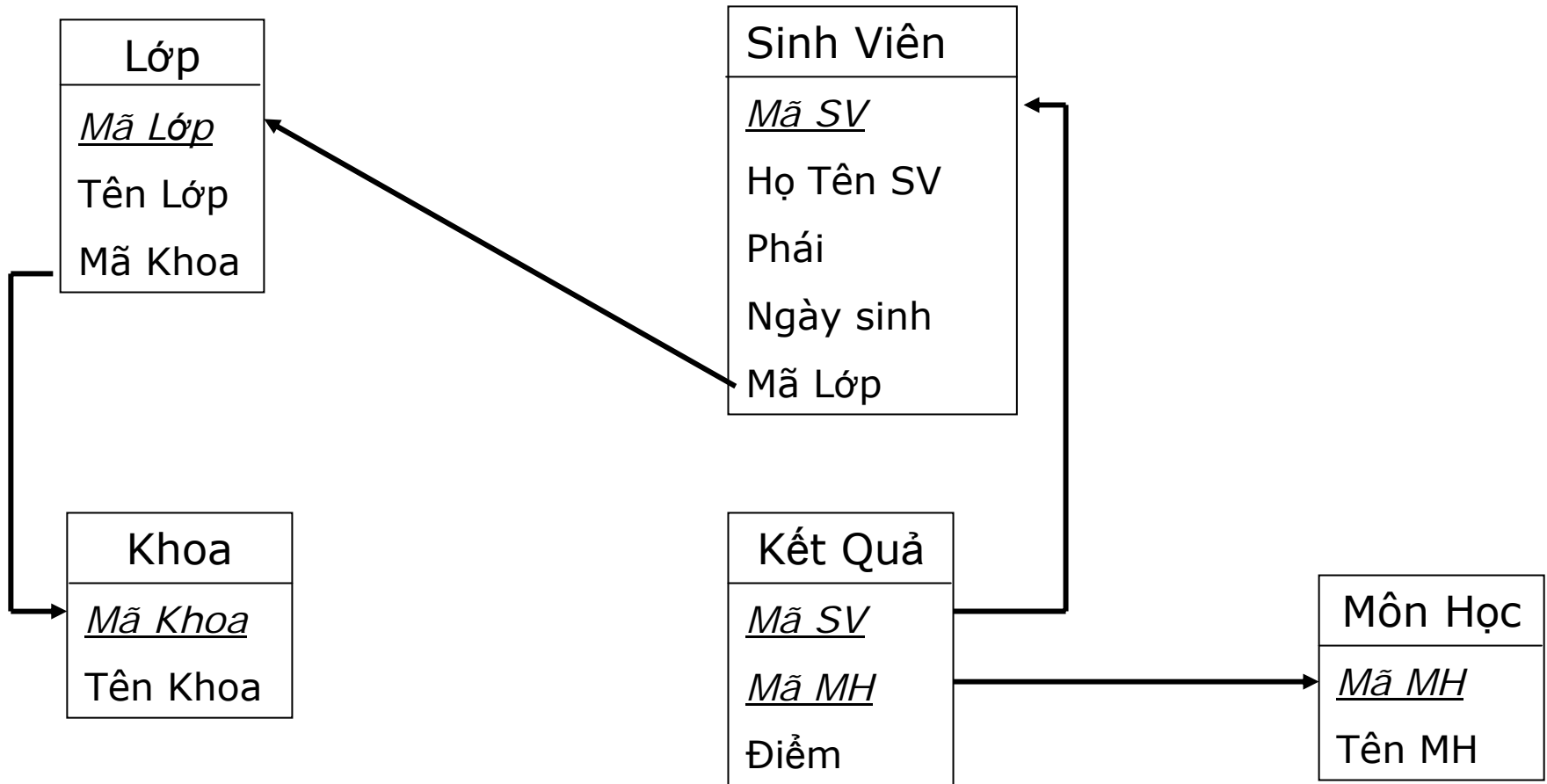
- Mục đích: tiếp cận hướng đối tượng (trong các phương pháp lập trình hướng đối tượng).
- Đặc trưng:
  - Tính đóng gói (encapsulation)
  - Tính đa hình (polymorphism)
  - Tính tái sử dụng (reusability).
- Các khái niệm:
  - Lớp (class)
  - Sự kế thừa (inheritance)
  - Kế thừa bội (tức là kế thừa từ nhiều lớp cơ sở - multi-inheritance)

# Mô hình quan hệ

- Mục đích: mô tả được dữ liệu dạng vật lý.
- Ưu điểm:
  - Đơn giản, hiệu quả.
  - Hỗ trợ các ngôn ngữ khai báo, các phép toán trên dữ liệu.
- Các thành phần:
  - Thuộc tính - Quan hệ
  - Lược đồ quan hệ - Bộ - Khóa



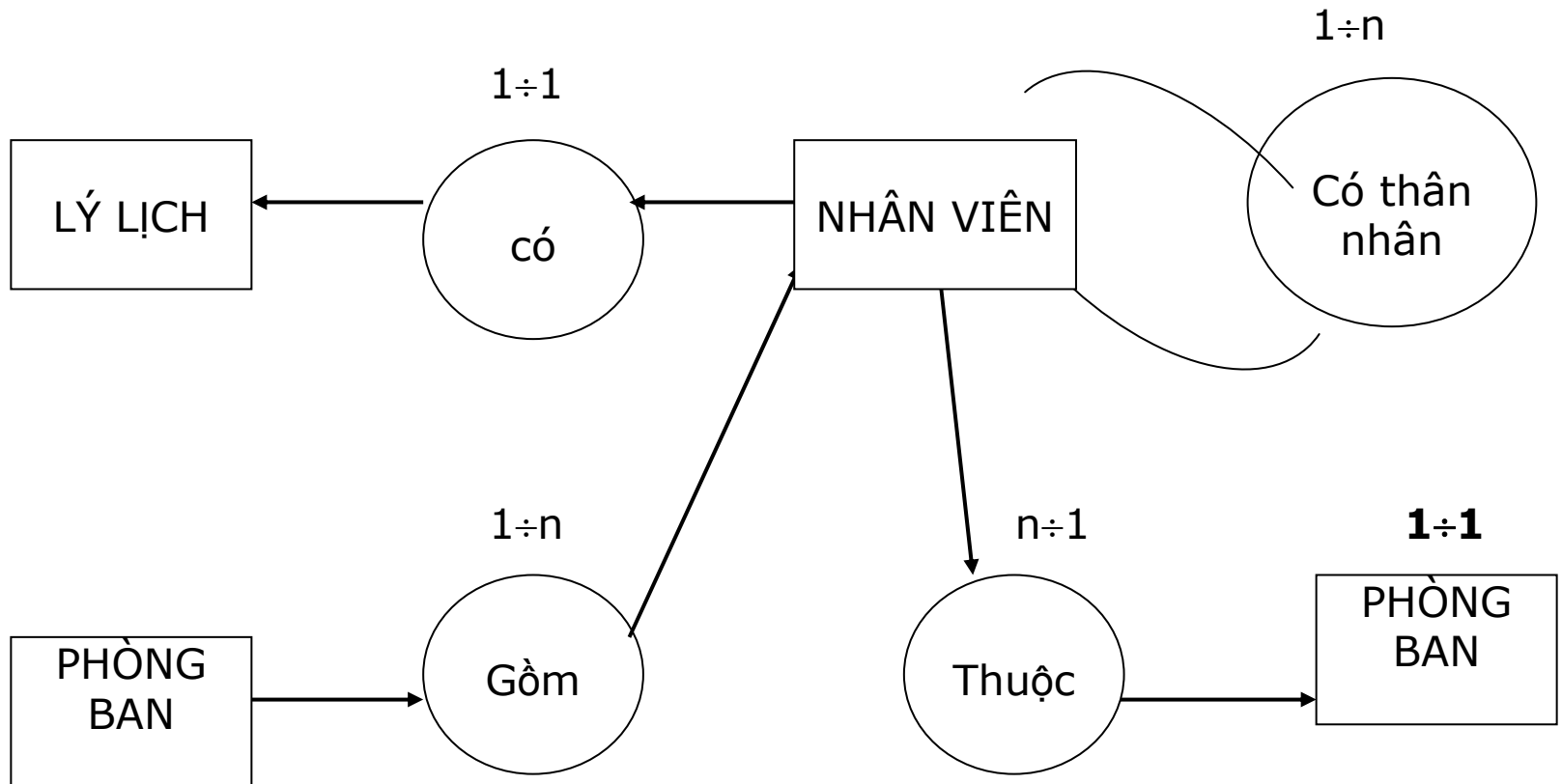
# Mô hình quan hệ



# Mô hình dữ liệu mạng

- Mục đích: dùng đồ thị có hướng cho các dữ liệu.
- Ưu điểm: đưa ra một cách ngắn gọn mô tả các mối liên hệ.
- Khuyết điểm: khi CSDL quá lớn, việc cài đặt các mối liên hệ nhiều nhiều làm cho hiệu năng giảm đến mức không thể chấp nhận được.
- Thành phần:
  - Loại mẫu tin (Record Type)
  - Mẫu tin (Record)
  - Loại liên hệ (Set Type)

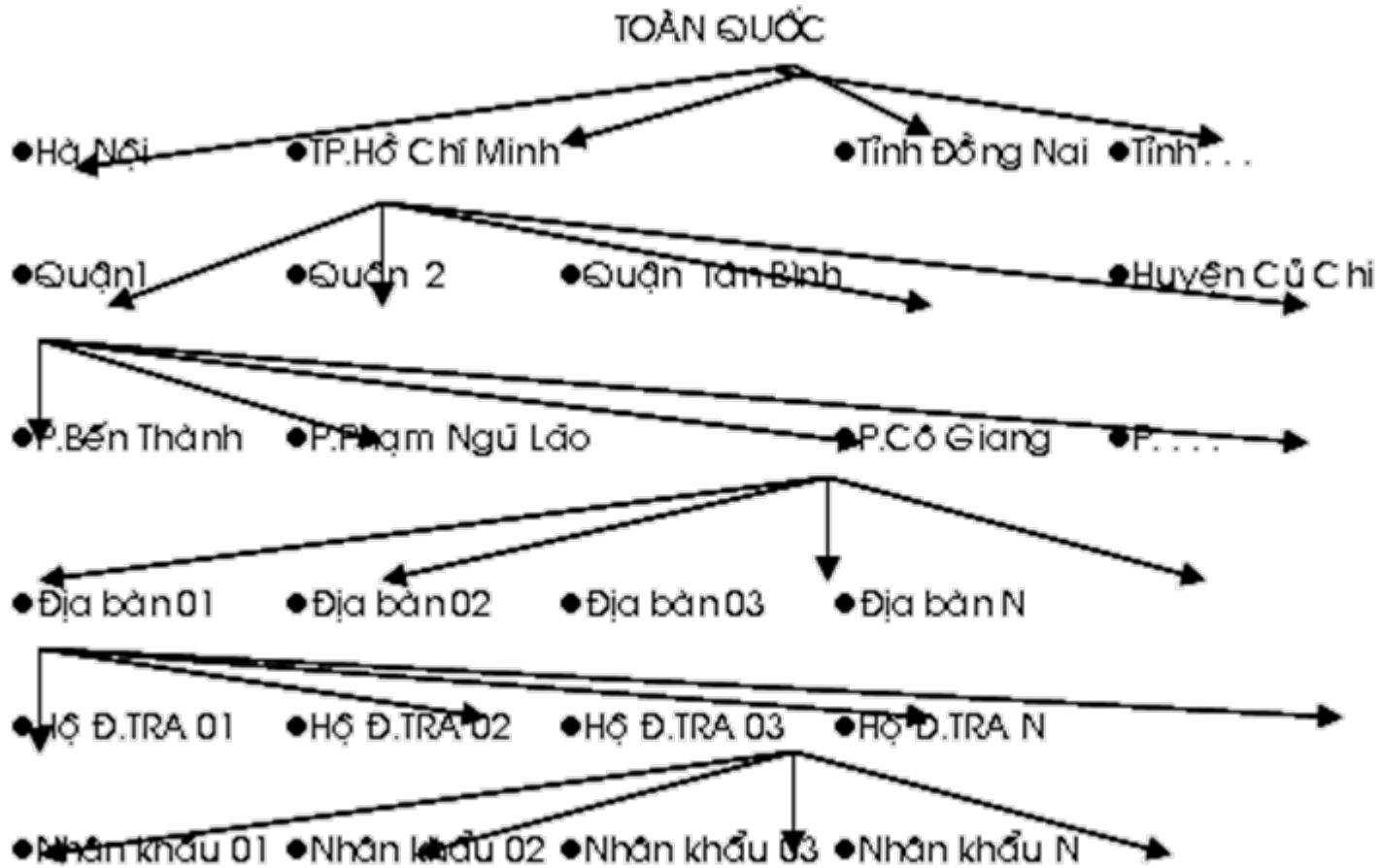
# Mô hình dữ liệu mạng



# Mô hình dữ liệu phân cấp

- Mục đích: biểu diễn dữ liệu dưới dạng hình cây. Trong đó các đường nối chỉ đi theo hướng từ con đến cha.
- Ưu điểm: đưa ra một cách ngắn gọn mô tả các mối liên hệ.
- Khuyết điểm: chỉ biểu diễn “kiểu mẫu tin ảo”.
- Thành phần:
  - Loại mẫu tin: giống khái niệm mẫu tin trong mô hình dữ liệu mạng.
  - Loại mối liên hệ: Kiểu liên hệ là phân cấp, theo cách: Giữa 2 loại mẫu tin chỉ tồn tại 1 mối liên hệ duy nhất

# Mô hình dữ liệu phân cấp



# Mô hình dữ liệu tốt nhất?

- Không có mô hình nào được xem là tốt nhất cho các hệ thống CSDL
- Nên xem xét theo kiểu: Chúng được sử dụng tốt nhất ở đâu và vào lúc nào?
- Phần lớn các mô hình:
  - Mục đích dùng như hệ thống ký hiệu cho dữ liệu, và
  - Làm hệ thống ký hiệu nền tảng cho ngôn ngữ thao tác dữ liệu
- E-R dùng thiết kế lược đồ khái niệm
- Mô hình hướng đối tượng giải quyết dư thừa tốt



# Nội dung

---

- 1. Cơ Sở Dữ Liệu Là Gì?
- 2. Hệ Quản Trị CSDL
- 3. Các Mô Hình Dữ Liệu
- 4. Quản Lý Dữ Liệu Phân Tán
- 5. Mô Hình Thực Thể Kết Hợp
- 6. Bài tập



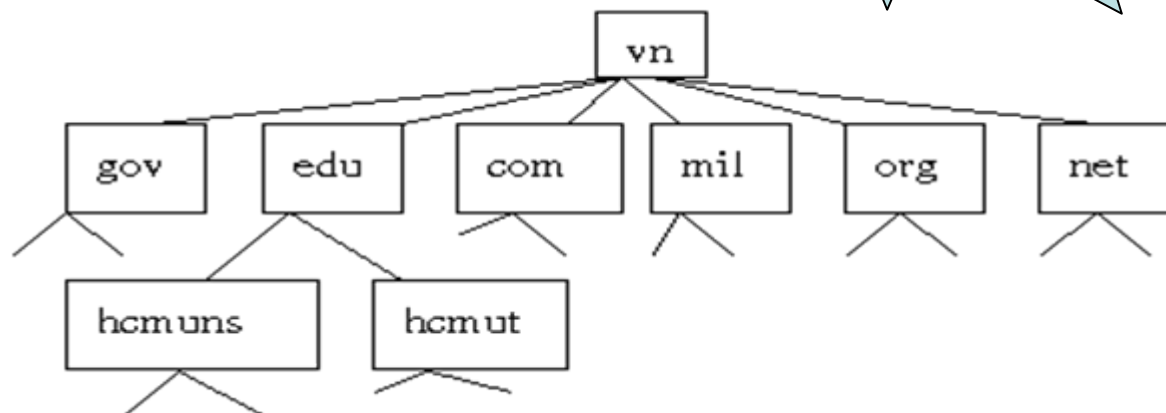
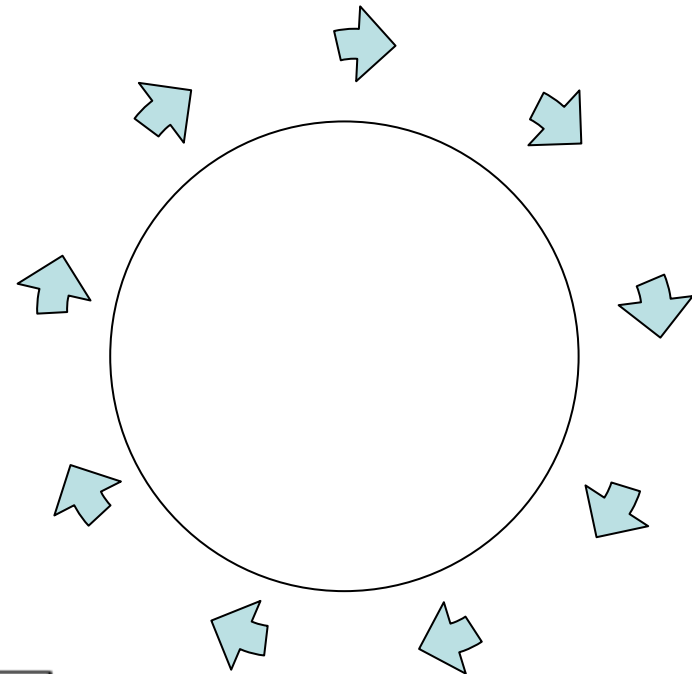
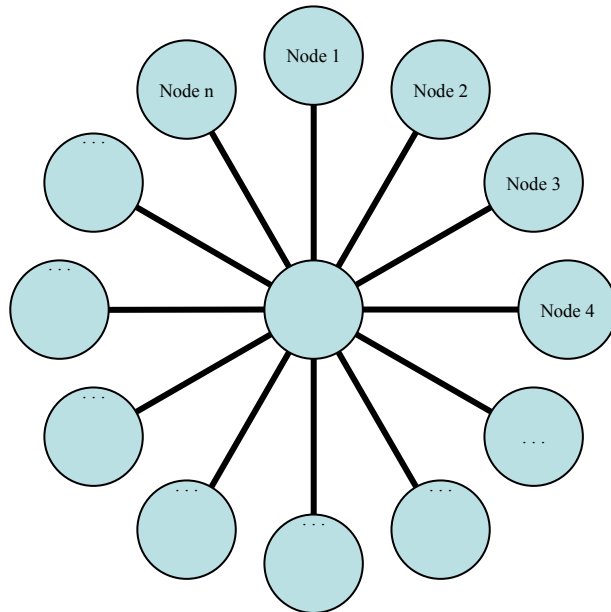
## \$1.4. Quản lý dữ liệu phân tán

---

- CSDL phân tán bao gồm một tập các nút (node) liên kết (link) với nhau, mỗi nút biểu diễn cho 1 máy tính và 1 thiết bị lưu trữ thứ cấp kèm theo. Các nút có thể chỉ là thiết bị lưu trữ nhỏ hoặc không có thiết bị lưu trữ (chỉ thuần có chức năng tính toán).
- Trong một mạng LAN (Local Area Network), mỗi máy trạm (workstation) đóng vai trò một node. Mỗi node sẽ lưu trữ dữ liệu của riêng mình.



# \$1.4. Quản lý dữ liệu phân tán



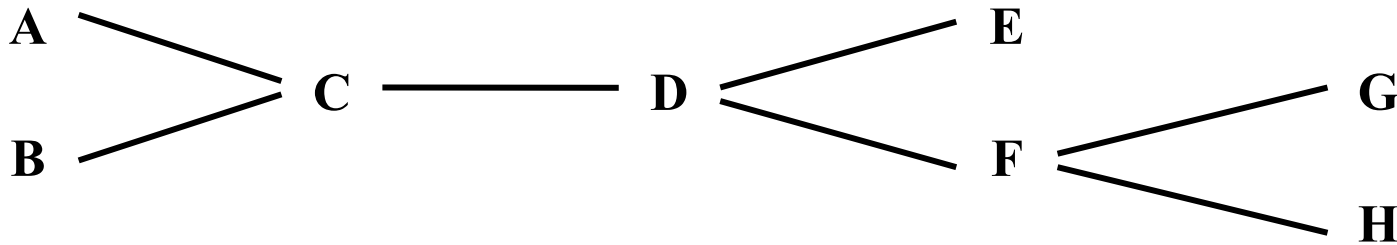


## \$1.4. Quản lý dữ liệu phân tán

---

- Dù tổ chức theo mạng cục bộ hay mạng phân tán thì “chi phí” phải trả cho việc truyền thông cho các nút là không nhỏ.
- CSDL phân tán sẽ bị “thương tổn” do rất nhiều sự cố như: đường liên lạc bị hỏng, máy tại 1 nút bị hư hoặc bị mất điện, ... Vì vậy hệ thống phải có khả năng tự thích ứng khi có sự cố về mạng.

# \$1.4. Quản lý dữ liệu phân tán



- Nếu sự cố xảy ra:
  - Trên đường truyền CD: cây sẽ bị tách thành 2 phần  $\{A, B, C\}$  và  $\{D, E, F, G, H\}$
  - Tại nút D: cây sẽ tách thành 3 phần  $\{A, B, C\}$ ,  $\{E\}$  và  $\{F, G, H\}$
- → Cần một thiết kế tốt



## \$1.4. Quản lý dữ liệu phân tán

---

- Để tăng cường sự thích ứng, người ta cần tạo ra nhiều bản sao và phải đảm bảo rằng các bản sao này đều phải giống nhau.
  - Khi có thao tác thay đổi trên một mục tại nút A thì những thay đổi đó phải được “cập nhật” trên tất cả các bản sao.
  - Khi nút A gặp sự cố cập nhật dữ liệu các nút khác vẫn cần được cập nhật cho nhau.
  - Sau đó khi nút A được khôi phục, thì điều cần thiết là bản sao mới nhất (có thể do nút khác cập nhật) phải được cập nhật lại trên nút A.



# Nội dung

---

- 1. Cơ Sở Dữ Liệu Là Gì?
- 2. Hệ Quản Trị CSDL
- 3. Các Mô Hình Dữ Liệu
- 4. Quản Lý Dữ Liệu Phân Tán
- 5. Mô Hình Thực Thể Kết Hợp
- 6. Bài tập

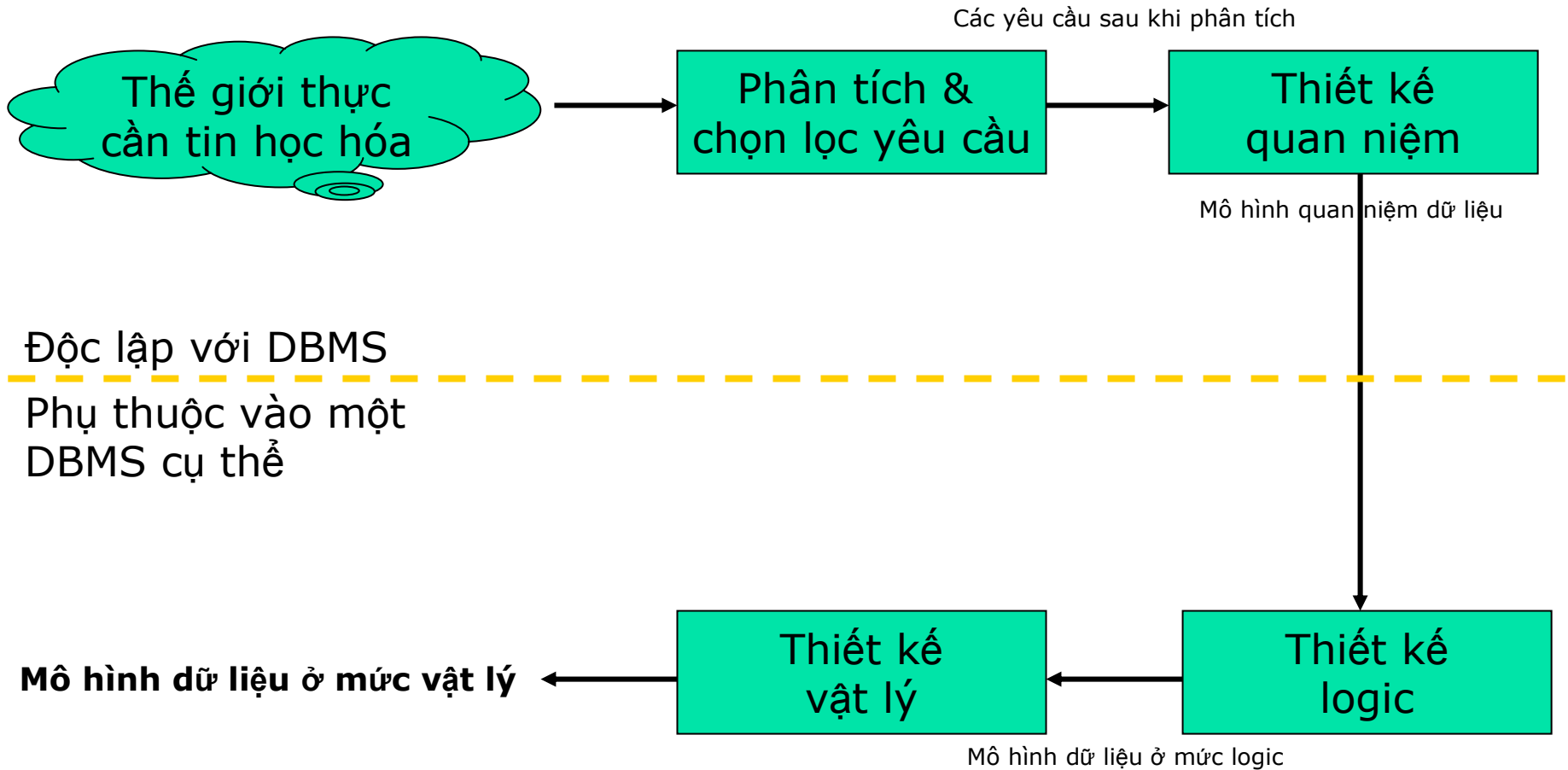


# \$1.5. Mô hình thực thể - kết hợp

---

- Các giai đoạn thiết kế CSDL
- Một ví dụ
- Các khái niệm về mô hình Thực thể – Kết hợp
- Các ký hiệu trong mô hình ER
- Hoàn chỉnh thiết kế cho ví dụ

# Các giai đoạn thiết kế CSDL



# Tin học hóa Quản lý phân công đề án tại một công ty

- Công ty có nhiều **phòng ban**: **tên**, **mã số** và một **nhân viên** làm **trưởng phòng**. **Ngày** mà nhân viên đó bắt đầu làm trưởng phòng cũng được quan tâm. Một **phòng ban** có thể định vị ở nhiều **địa điểm** khác nhau.
- Một **phòng ban** phụ trách một số **đề án**. Mỗi **đề án** có **tên**, **mã số** và **nơi thực hiện đề án**.
- Về **nhân viên**, cần quan tâm: **mã**, **tên**, **địa chỉ**, **mức lương**, **giới tính** và **ngày sinh**. Mỗi nhân viên thuộc một **phòng ban** nhưng có thể làm việc cho nhiều **đề án**. Mỗi **đề án** do một **phòng ban** phụ trách.



# Tin học hóa Quản lý phân công đề án tại một công ty

- Cần lưu lại **giờ làm việc** của một **nhân viên** làm cho một **đề án**. Ngoài ra cũng cần biết được **người phụ trách trực tiếp của một nhân viên**
- Để quản lý các thông tin có liên quan đến bảo hiểm, người sử dụng cũng có nhu cầu về thông tin của **mỗi quan hệ thân nhân** (vợ, chồng, con) của **nhân viên**. Các thông tin mà người sử dụng quan tâm bao gồm **tên thân nhân, giới tính và mỗi quan hệ** với nhân viên.

# Các khái niệm trong mô hình E-R

- **Thực thể (Entity):** Là 1 đối tượng tồn tại trong thế giới thực, có thể là cụ thể hoặc trừu tượng và có thể nhận biết.
  - Ví dụ: 1 nhân viên có mã là NV010, 1 sinh viên có mã là TH98020
- **Thuộc tính (Attribute):** Mỗi thực thể có nhiều đặc trưng, mỗi đặc trưng được gọi là một thuộc tính.
  - Ví dụ: Mã NV, Tên, Địa chỉ, Năm sinh, Điện thoại là các thuộc tính của thực thể nhân viên, dùng để nhận biết một đối tượng nhân viên cụ thể.
- Từng thực thể riêng biệt có giá trị riêng biệt cho mỗi thuộc tính.

# Các khái niệm trong mô hình E-R

- Phân loại thuộc tính:
  - Thuộc tính đơn (Single attribute)
  - Thuộc tính hợp (Composite attribute)
- Giá trị của thuộc tính:
  - Đơn trị (single-valued)
  - Đa trị (multi-valued)
  - Null
- Thuộc tính dẫn xuất.

# Các khái niệm trong mô hình E-R

- **Loại thực thể (Entity set):** Là tập hợp các thực thể có cùng các thuộc tính.
  - Ví dụ: Loại thực thể NHAN VIEN, loại thực thể ĐEAN.
- **Khóa của loại thực thể:** Tập hợp nhỏ nhất các thuộc tính của một loại thực thể mà giá trị của tập hợp này là duy nhất đối với mọi thực thể được gọi là khóa của loại thực thể.
  - Khóa dùng để nhận biết từng thực thể.

# Các khái niệm trong mô hình E-R

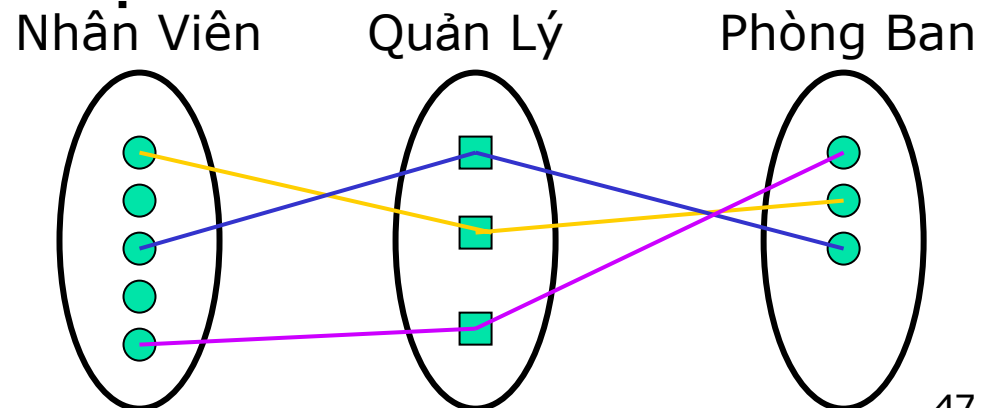
- **Mối kết hợp:** Một mối kết hợp (relationship) là mối quan hệ giữa hai hay nhiều thực thể mang một ý nghĩa nào đó. Ví dụ:
  - Nguyễn Văn A trực thuộc phòng Nghiên cứu.
  - Trần Thị B trực thuộc phòng Kinh doanh.
- **Loại mối kết hợp:** Một loại mối kết hợp là tập tất cả các mối kết hợp cùng loại.
  - Ví dụ: Tập tất cả những mối quan hệ trực\_thuộc trên được diễn đạt bởi loại mối kết hợp TRUC\_THUOC.

# Các khái niệm trong mô hình E-R

- **Bậc của loại mỗi kết hợp:** Là số lượng loại thực thể tham gia vào mỗi kết hợp
  - Loại mỗi kết hợp bậc 2 còn gọi là mỗi kết hợp nhị phân. Hầu hết các loại mỗi kết hợp trong CSDL là nhị phân.
- **Thuộc tính của loại mỗi kết hợp**
  - Giờ làm việc của nhân viên khi nhân viên tham gia làm đề án.

# Các khái niệm trong mô hình E-R

- Các ràng buộc trên loại mỗi kết hợp:  
gồm 2 loại:
  - Ràng buộc tỷ lệ lực lượng (Cardinality ratio constraint): ký hiệu (min, max).
  - Ràng buộc tham gia (Participation constraint): gồm 2 loại:
    - Toàn phần
    - Bán phần



# Các khái niệm trong mô hình E-R

- Loại thực thể yếu: Loại thực thể không có tập thuộc tính khóa gọi là loại thực thể yếu (weak entity type).
  - Loại thực thể yếu phải tham gia trong một loại mối kết hợp xác định (Identifying relationship type) trong đó có một loại thực thể chủ hay loại thực thể xác định (Identifying entity).



# Các khái niệm trong mô hình E-R

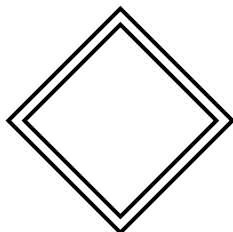
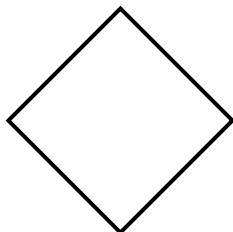
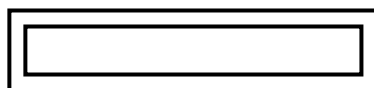
- Một thực thể yếu được xác định bằng sự kết hợp của:
  - Khóa riêng phần (partial key) của một thực thể yếu.
  - Khóa của một thực thể chủ xác định thực thể yếu này.

- Ví dụ:

Thực thể yếu	Thực thể nó phụ thuộc
Chi tiết hóa đơn	Hóa đơn
Các thân nhân	Nhân viên

# Các khái niệm trong mô hình E-R

- Ký hiệu



- Ý nghĩa

Loại thực thể

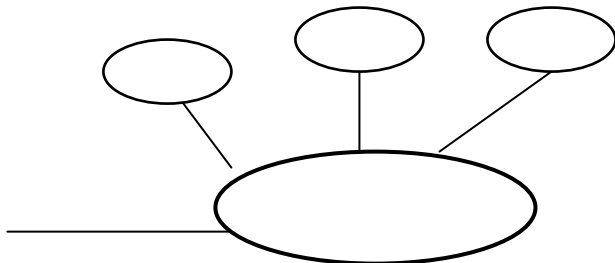
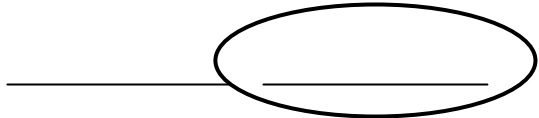
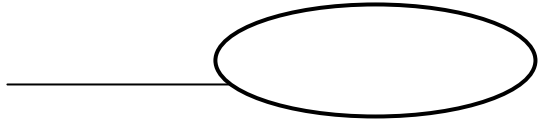
Loại thực thể yếu

Loại mối kết hợp

Loại mối kết hợp xác định

# Các khái niệm trong mô hình E-R

- Ký hiệu



- Ý nghĩa

Thuộc tính

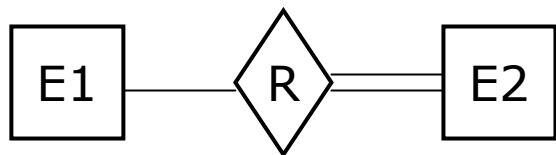
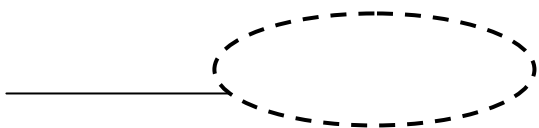
Thuộc tính khóa

Thuộc tính đa trị

Thuộc tính đa hợp

# Các khái niệm trong mô hình E-R

- Ký hiệu



- Ý nghĩa

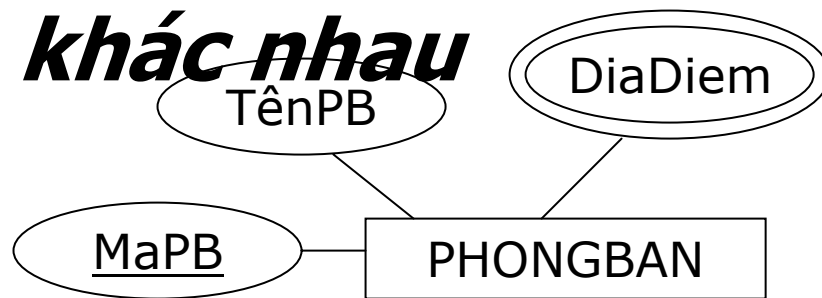
Thuộc tính dẫn xuất

E2 tham gia toàn phần vào R

E1 tham gia vào R với tỷ lệ (min,max)

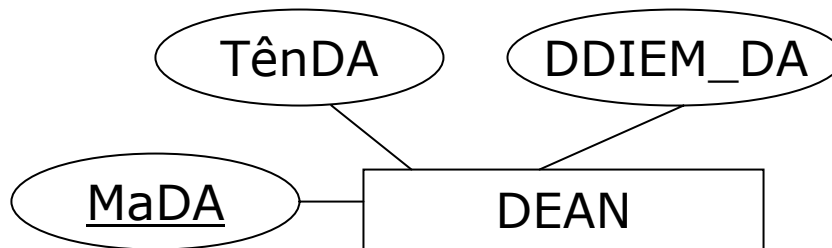
# Thiết kế mô hình E-R cho ví dụ

- **Công ty có nhiều phòng ban: tên, mã số** và một nhân viên làm trưởng phòng. Ngày mà nhân viên đó bắt đầu làm trưởng phòng cũng được quan tâm. **Một phòng ban có thể định vị ở nhiều địa điểm khác nhau**



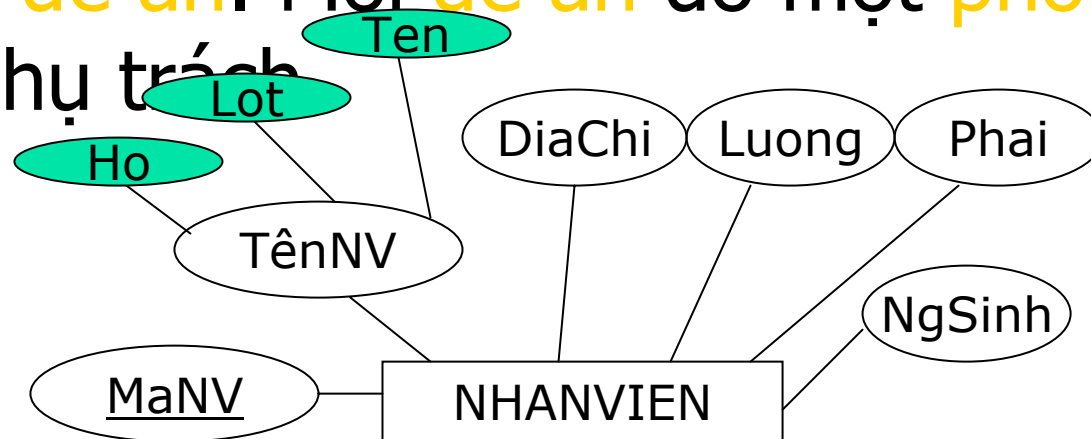
# Thiết kế mô hình E-R cho ví dụ

- Một phòng ban phụ trách một số đề án.  
**Mỗi đề án có tên, mã số và nơi thực hiện đề án.**



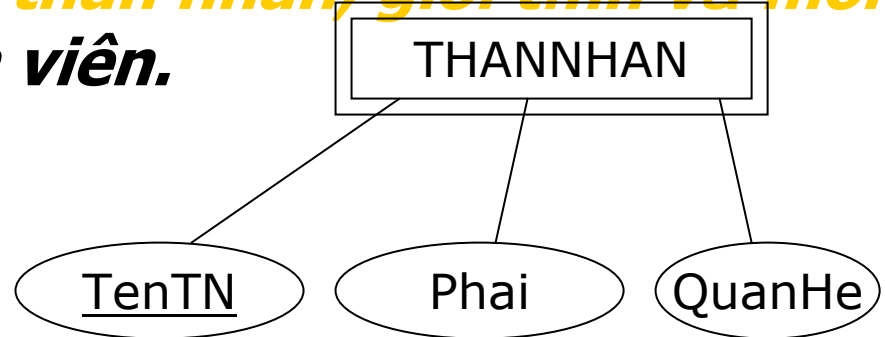
# Thiết kế mô hình E-R cho ví dụ

- **Về nhân viên, cần quan tâm: mã, tên, địa chỉ, mức lương, giới tính và ngày sinh.** Mỗi nhân viên thuộc một phòng ban nhưng có thể làm việc cho nhiều đề án. Mỗi đề án do một phòng ban phụ trách.



# Thiết kế mô hình E-R cho ví dụ

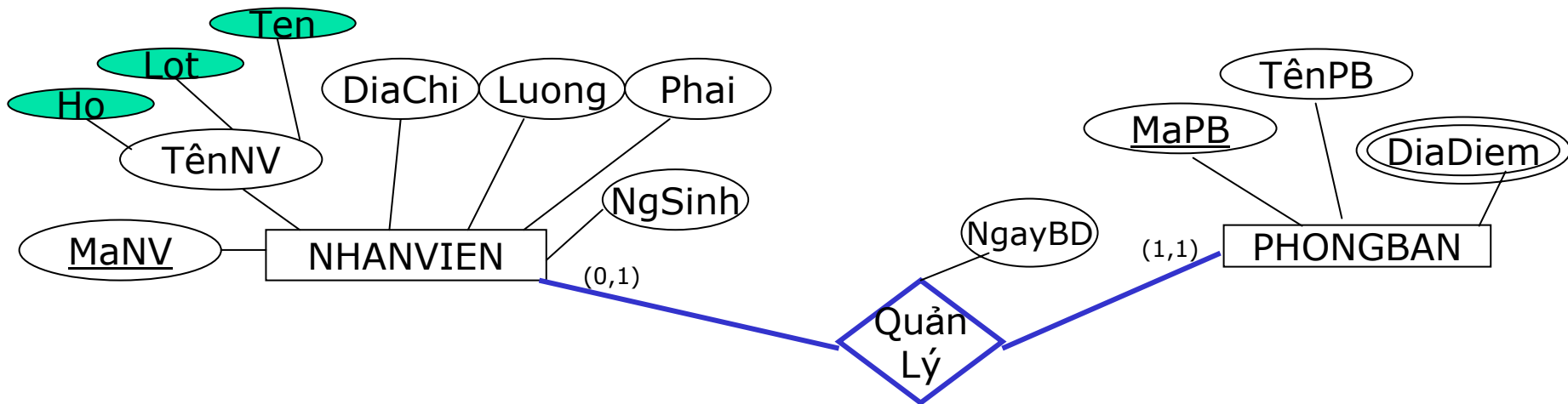
- Cần lưu lại **giờ làm việc** của một **nhân viên** làm cho một **đề án**. Ngoài ra cũng cần biết được **người phụ trách trực tiếp** của một **nhân viên**
- Để quản lý các thông tin có liên quan đến bảo hiểm, người sử dụng cũng có nhu cầu về thông tin của **mối quan hệ thân nhân (vợ, chồng, con) của nhân viên**. Các thông tin mà người sử dụng quan tâm bao gồm **tên thân nhân, giới tính và mối quan hệ với nhân viên**.





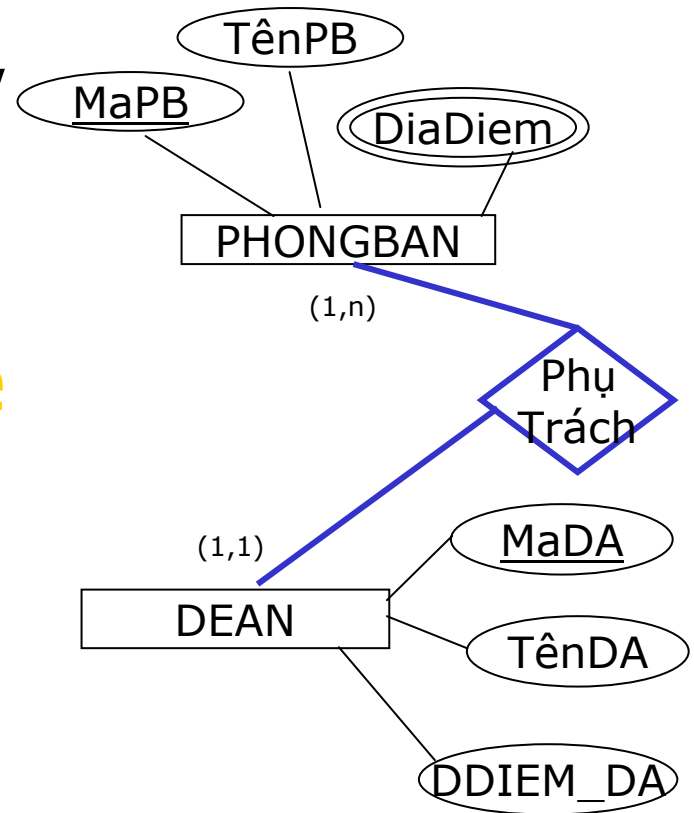
# Thiết kế mô hình E-R cho ví dụ

- Công ty có nhiều phòng ban: tên, mã số và **một nhân viên làm trưởng phòng.** Ngày mà nhân viên đó bắt đầu làm trưởng phòng cũng được quan tâm



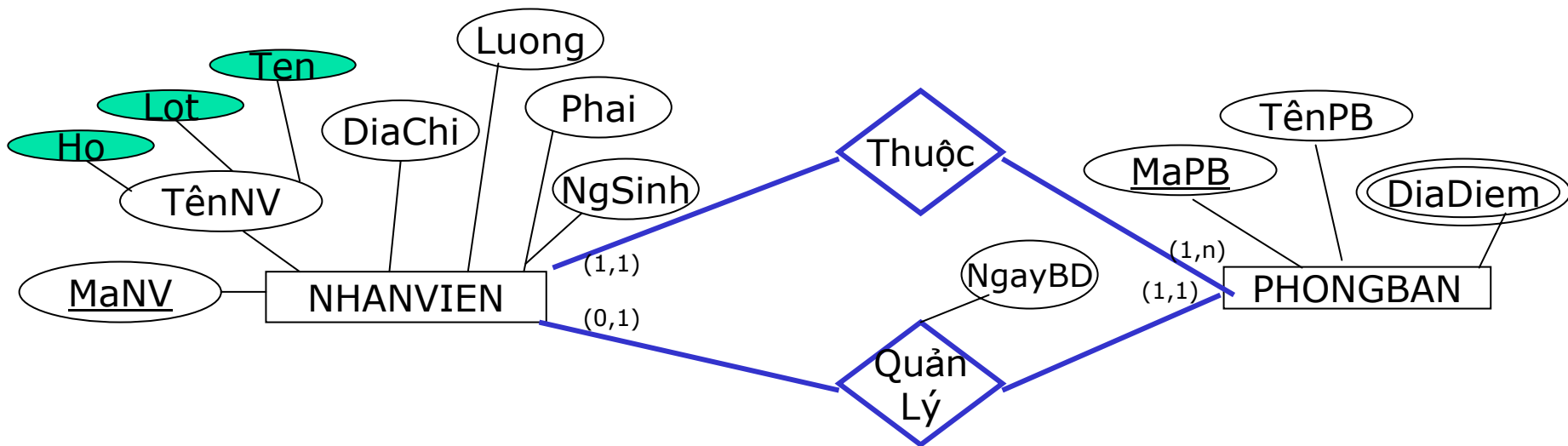
# Thiết kế mô hình E-R cho ví dụ

- **Một phòng ban phụ trách một số đề án.**  
Mỗi đề án có tên, mã số và nơi thực hiện đề án.
- Mỗi đề án do một phòng ban phụ trách.



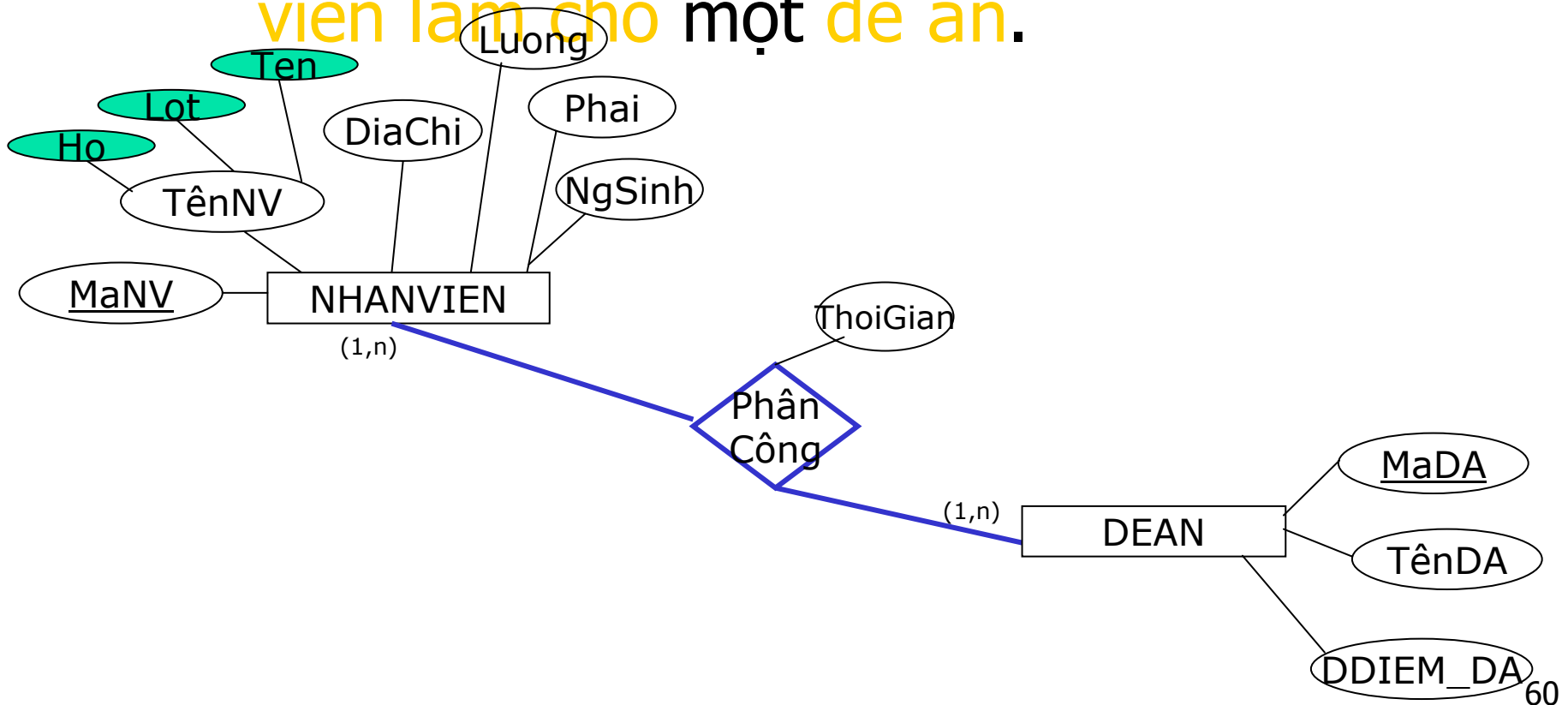
# Thiết kế mô hình E-R cho ví dụ

- Mỗi nhân viên thuộc một phòng ban nhưng có thể làm việc cho nhiều đề án



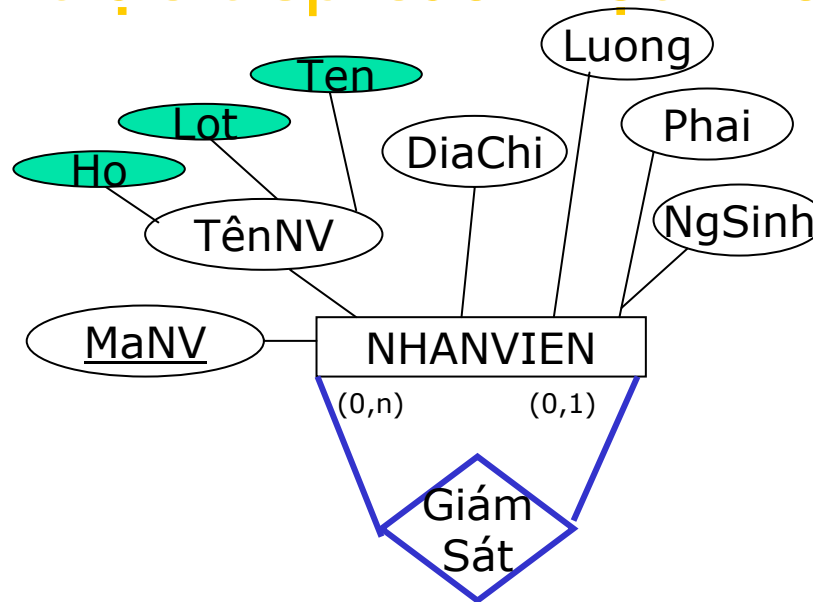
# Thiết kế mô hình E-R cho ví dụ

- Cần lưu lại **giờ làm việc** của một **nhân viên làm cho một đề án**.



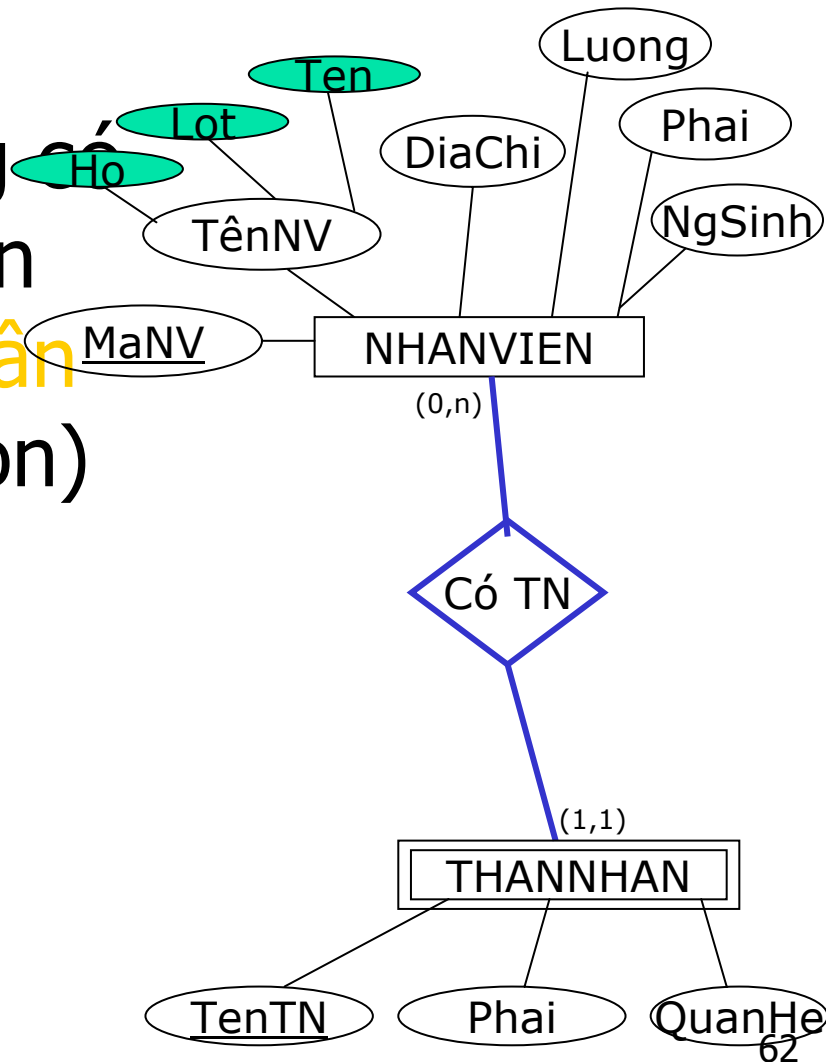
# Thiết kế mô hình E-R cho ví dụ

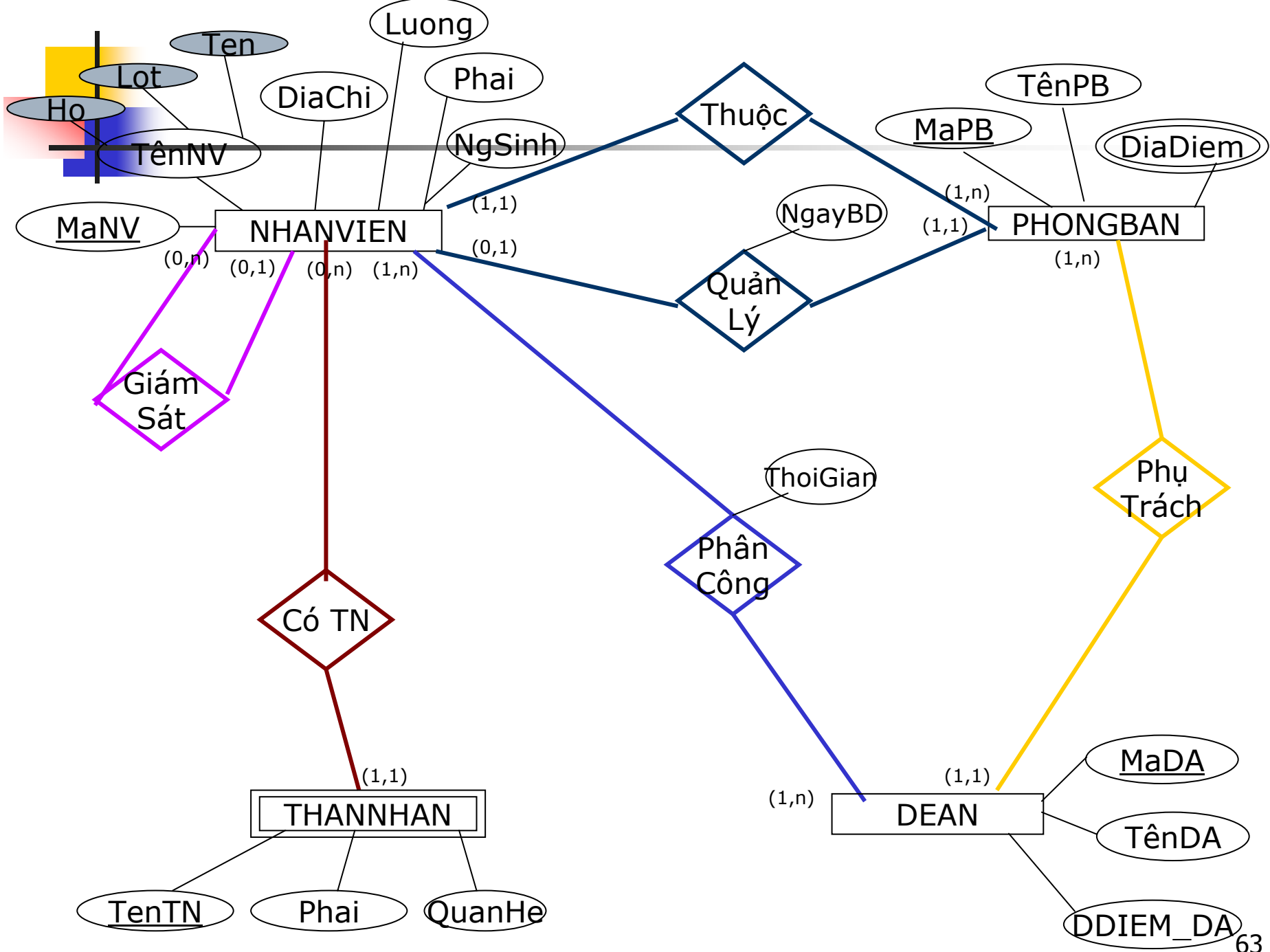
- Ngoài ra cũng cần biết được người phụ trách trực tiếp của một nhân viên



# Thiết kế mô hình E-R cho ví dụ

- người sử dụng cũng có nhu cầu về thông tin của **mỗi quan hệ thân nhân** (vợ, chồng, con) của **nhân viên**







# Nội dung

---

- 1. Cơ Sở Dữ Liệu Là Gì?
- 2. Hệ Quản Trị CSDL
- 3. Các Mô Hình Dữ Liệu
- 4. Quản Lý Dữ Liệu Phân Tán
- 5. Mô Hình Thực Thể Kết Hợp
- 6. Bài tập





## \$1.6. Bài tập làm ngay

---

- Tin học hóa Quản lý sinh viên tại một trường trung cấp
  - Các thông tin cần quản lý về một sinh viên gồm: họ tên, phái, địa chỉ, điện thoại. Ngoài ra, ta cũng cần biết sinh viên này đang theo học ngành nào.
  - Mỗi ngành sẽ có một tên ngành và số lượng giáo viên thuộc ngành đó.
  - Cần quản lý họ tên của giáo viên và thông tin về ngành mà giáo viên trực thuộc



## \$1.6. Bài tập làm ngay

---

- Tin học hóa Quản lý sinh viên tại một trường trung cấp
  - Trường cần quản lý thêm thông tin về các môn học gồm: tên và số tín chỉ
  - Công tác giảng dạy của trường được quản lý theo từng khóa học. Mỗi khóa sẽ được tổ chức theo từng học kỳ của từng năm do một giáo viên phụ trách một môn học cụ thể
  - Sinh viên sẽ có điểm khi tham gia các khóa học

# CƠ SỞ DỮ LIỆU

ThS.Lê Văn Hạnh

# MÔ HÌNH DỮ LIỆU QUAN HỆ (Relational Data Model)



# Nội Dung

---

- Được đề xuất bởi TS. E. F. Codd năm 1970.
- Nền tảng của mô hình này là lý thuyết tập hợp trên các quan hệ.
- **Nội dung chính**
  1. **Thuộc tính**
  2. Quan hệ
  3. Bộ giá trị
  4. Thể hiện của quan hệ
  5. Lược đồ quan hệ
  6. Lược đồ CSDL
  7. Khóa – Keys
  8. Chuyển ER sang RDM



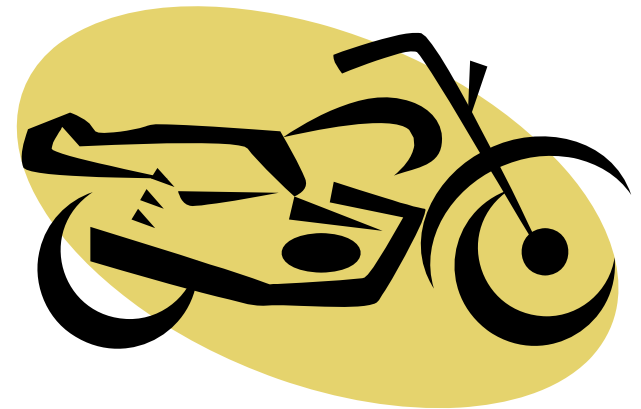
## \$2.1. Thuộc tính

---

- ***Đối tượng*** là một thực thể tồn tại khách quan hay một sự trừu tượng hóa (nhân viên, xe máy, hàng hóa, ...).
- **Thuộc tính** là đặc tính của *đối tượng* cần được phản ánh trong CSDL.

## \$2.1. Thuộc tính

- *Đối tượng xe máy* có các *thuộc tính* như :  
Số xe, Số máy, Chủ xe, ...
- *Đối tượng nhân viên* có các *thuộc tính* như : Mã nhân viên, Họ tên, Phái, ...





## \$2.1. Thuộc tính

---

- Thuộc tính được đặc trưng bởi 3 yếu tố:
  1. **Tên gọi:** Một dãy ký tự bất kỳ (gợi nhớ)
  2. **Kiểu dữ liệu:** Số, Chuỗi, Ngày, Logic, OLE.
  3. **Miền giá trị** của thuộc tính: Xác định tập giá trị mà thuộc tính có thể nhận. Ký hiệu miền giá trị của thuộc tính A là  $Dom(A)$ .





# Nội Dung

---

- Được đề xuất bởi TS. E. F. Codd năm 1970.
- Nền tảng của mô hình này là lý thuyết tập hợp trên các quan hệ.
- **Nội dung chính**
  1. Thuộc tính
  - 2. Quan hệ**
  3. Bộ giá trị
  4. Thể hiện của quan hệ
  5. Lược đồ quan hệ
  6. Lược đồ CSDL
  7. Khóa – Keys
  8. Chuyển ER sang RDM



## \$2.2. Quan hệ

- Quan hệ là một **bảng** (table) 2 chiều được định nghĩa trên ***một tập thuộc tính***.
- Tập toàn bộ thuộc tính của một quan hệ  $Q$  được ký hiệu là :  $Q^+$ .
- Ví dụ : quan hệ NhanVien với 5 thuộc tính

<b>NhanVien</b>				
<b>MaNV</b>	<b>HoTen</b>	<b>Phai</b>	<b>Luong</b>	<b>PHG</b>
123	N T A	Nữ	2000000	NC
124	L V M	Nam	2100000	NC

## \$2.2. Quan hệ

- Quan hệ PhongBan với 4 thuộc tính

Tên quan hệ

Tên thuộc tính

<b>PhongBan</b>			
<b>MAPH</b>	<b>TENPHG</b>	<b>TRPHG</b>	<b>NG_NHANCHUC</b>
1	Quan ly	888665555	19/06/1981
4	Dieu nang	987987987	1/1/1995
5	Nghien cuu	333445555	22/05/1968



# Nội Dung

---

- Được đề xuất bởi TS. E. F. Codd năm 1970.
- Nền tảng của mô hình này là lý thuyết tập hợp trên các quan hệ.
- **Nội dung chính**
  1. Thuộc tính
  2. Quan hệ
  - 3. Bộ giá trị**
  4. Thể hiện của quan hệ
  5. Lược đồ quan hệ
  6. Lược đồ CSDL
  7. Khóa – Keys
  8. Chuyển ER sang RDM

## \$2.3. Bộ giá trị

- Bộ là một dòng dữ liệu trong một quan hệ.
- Bộ thường được gọi là mẫu tin, bản ghi hay record.

Bộ

<b>NhanVien</b>				
<b>MaNV</b>	<b>HoTen</b>	<b>Phai</b>	<b>Luong</b>	<b>PHG</b>
123	N T A	Nữ	2000000	NC
124	L V M	Nam	2100000	NC

- Không có 2 bộ giống nhau trong cùng một quan hệ



# Nội Dung

---

- Được đề xuất bởi TS. E. F. Codd năm 1970.
- Nền tảng của mô hình này là lý thuyết tập hợp trên các quan hệ.
- **Nội dung chính**
  1. Thuộc tính
  2. Quan hệ
  3. Bộ giá trị
  - 4. Thể hiện của quan hệ**
  5. Lược đồ quan hệ
  6. Lược đồ CSDL
  7. Khóa – Keys
  8. Chuyển ER sang RDM

## \$2.4. Thể hiện của quan hệ

- Thể hiện của một quan hệ là tập hợp các bộ giá trị cụ thể của một quan hệ tại một thời điểm nhất định.
- Ký hiệu thể hiện của quan hệ Q là  $\tau_Q$
- $\tau_{\text{NhanVien1}}$  là thể hiện của quan hệ NhanVien vào ngày 1/10/2007

<b>NhanVien</b>				
<b>MaNV</b>	<b>HoTen</b>	<b>Phai</b>	<b>Luong</b>	<b>PHG</b>
123	N T A	Nữ	2000000	NC
124	L V M	Nam	2100000	NC

## \$2.4. Thể hiện của quan hệ

- **NhanVien2** là thể hiện của quan hệ **NhanVien** vào ngày 5/10/2007

<b>NhanVien</b>				
<b>MaNV</b>	<b>HoTen</b>	<b>Phai</b>	<b>Luong</b>	<b>PHG</b>
123	N T A	Nữ	2800000	NC
124	L V M	Nam	2100000	NC
125	T V B	Nam	1500000	Mang





# Nội Dung

---

- Được đề xuất bởi TS. E. F. Codd năm 1970.
- Nền tảng của mô hình này là lý thuyết tập hợp trên các quan hệ.
- **Nội dung chính**
  1. Thuộc tính
  2. Quan hệ
  3. Bộ giá trị
  4. Thể hiện của quan hệ
  - 5. Lược đồ quan hệ**
  6. Lược đồ CSDL
  7. Khóa – Keys
  8. Chuyển ER sang RDM



## \$2.5. Lược đồ quan hệ

---

- **Cấu trúc của một quan hệ** là tập thuộc tính hình thành nên quan hệ đó.
- **Lược đồ quan hệ** nhằm mục đích **mô tả cấu trúc** của một quan hệ và **các mối liên hệ** giữa các thuộc tính trong quan hệ đó.
- Một lược đồ quan hệ gồm:
  - một tập thuộc tính của quan hệ, kèm theo
  - một mô tả để xác định ý nghĩa và mối liên hệ giữa các thuộc tính



## \$2.5. Lược đồ quan hệ

---

- Lược đồ quan hệ được đặc trưng bởi:
  - Một tên phân biệt
  - Một tập hợp hữu hạn các thuộc tính  $(A_1, \dots, A_n)$
- Ký hiệu của lược đồ quan hệ  $Q$  gồm  $n$  thuộc tính  $(A_1, A_2, \dots, A_n)$  là :
  - $Q(A_1, A_2, \dots, A_n)$ , hoặc
  - $Q^+ = \{A_1, A_2, \dots, A_n\}$

## \$2.5. Lược đồ quan hệ

- Lược đồ quan hệ PhongBan
  - PhongBan(MaPHG, TenPHG, TrPHG, NG\_NhanChuc)
  - Mô tả: Mỗi phòng ban được cấp một mã số duy nhất để phân biệt với các phòng ban khác trong công ty, và có một tên phòng ban, một trưởng phòng cùng ngày nhận chức của trưởng phòng.

MAPHG	TENPHG	TRPHG	NG_NHANCHUC
1	Quan ly	888665555	19/6/1981
4	Dieu nang	987987987	1/1/1995
5	Nghien cuu	333445555	22/5/1968

Quan hệ  
PhongBan



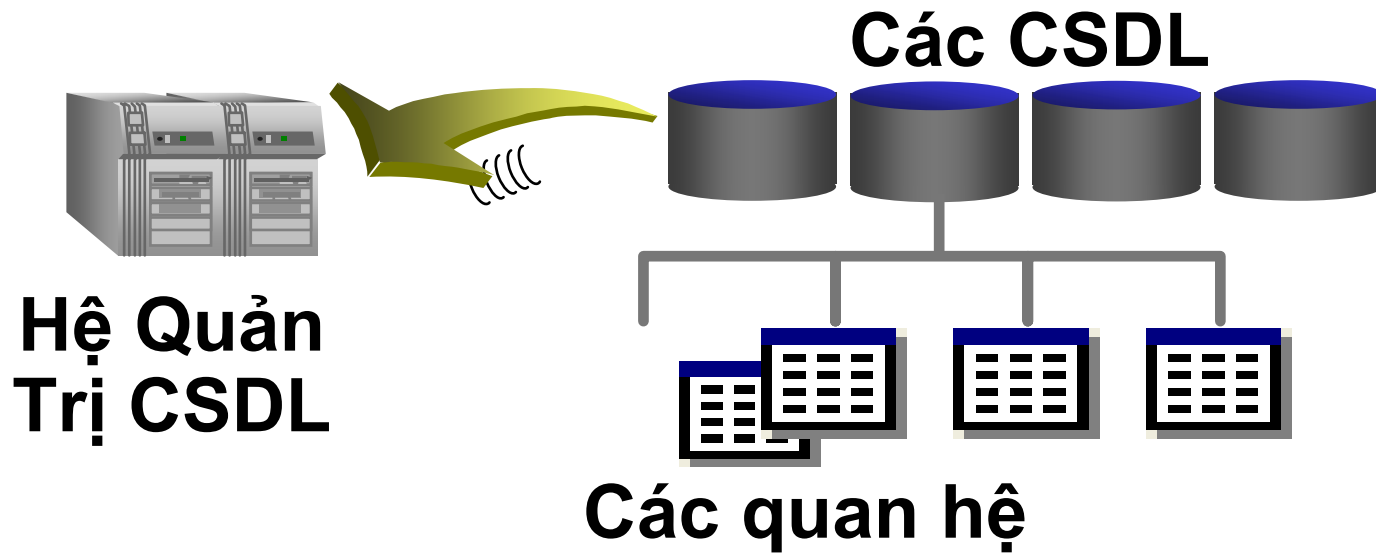
# Nội Dung

---

- Được đề xuất bởi TS. E. F. Codd năm 1970.
- Nền tảng của mô hình này là lý thuyết tập hợp trên các quan hệ.
- **Nội dung chính**
  1. Thuộc tính
  2. Quan hệ
  3. Bộ giá trị
  4. Thể hiện của quan hệ
  5. Lược đồ quan hệ
  - 6. Lược đồ CSDL**
  7. Khóa – Keys
  8. Chuyển ER sang RDM

## \$2.6. Lược đồ CSDL

- Là tập hợp gồm các lược đồ quan hệ và các mối liên hệ giữa chúng trong cùng một hệ thống quản lý.





## \$2.6. Lược đồ CSDL

---

Lược đồ CSDL “quản lý đề án cty”

**NHANVIEN** (MaNV, HoNV, TenLot, TenNV, Phai, Luong, PHG, NGSinh, DCHI, Ma\_NQL)

**PHONGBAN** (MaPHG, TenPHG, TrPHG, NG\_NhanChuc)

**DEAN** (MaDA, TenDA, DDIEM\_DA, Phong)

**PHANCONG** (Ma\_NVien, MaDA, ThoiGian)

**DIADIEM\_PHG** (MaPHG, DIADIEM)

**THANNHAN** (Ma\_NVien, TenTN, Phai, NGSinh, QuanHe)



# Nội Dung

---

- Được đề xuất bởi TS. E. F. Codd năm 1970.
- Nền tảng của mô hình này là lý thuyết tập hợp trên các quan hệ.
- **Nội dung chính**
  1. Thuộc tính
  2. Quan hệ
  3. Bộ giá trị
  4. Thể hiện của quan hệ
  5. Lược đồ quan hệ
  6. Lược đồ CSDL
  - 7. Khóa – Keys**
  8. Chuyển ER sang RDM



## \$2.7. Khóa

- Siêu khóa
- Khóa
- Khóa Chính
- Khóa Ngoại

Nguyễn Văn Tèo?



# Siêu khóa – Super Keys

- **Siêu khóa** : là một tập con các thuộc tính của  $Q^+$  mà nhờ vào đó chúng ta có thể phân biệt 2 bộ khác nhau trong cùng một thể hiện  $\mathcal{TQ}$  bất kỳ.
- $\forall t1, t2 \in \mathcal{TQ}, t1[K] \neq t2[K] \Leftrightarrow K$  là siêu khóa của  $Q$ .
- Một quan hệ có ít nhất một siêu khóa ( $Q^+$ ) và có thể có nhiều siêu khóa.

# Siêu khóa – Super Keys

- Ví dụ: các siêu khóa của quan hệ NhanVien là: {HoTen,Phai}; {HoTen,Phai,Luong,PHG}; ...

<b>NhanVien</b>				
Siêu Khóa?	<b>HoTen</b>	<b>Phai</b>	<b>Luong</b>	<b>PHG</b>
	N T A	Nữ	2000	NC
	L V M	Nam	2100	NC
	N T A	Nam	2500	KH

- Tổ hợp các thuộc tính có thể phân biệt các bộ trong 1 quan hệ

# Khóa – Keys

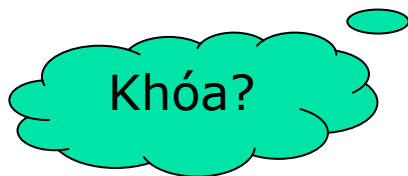
- Khóa: là một siêu khóa "*nhỏ nhất*" (chứa ít thuộc tính nhất và khác rỗng).
- Cho  $Q^+ = \{A_1, \dots, A_n\}$ ,  $K = \{A_1, \dots, A_m\}$ ,  $m \leq n$ ,  $K \subseteq Q^+$ ,  $K \neq \emptyset$ ,  $K$  là khóa của  $Q$  nếu:
  - $K$  là siêu khóa của  $Q$ .
  - nếu  $\exists K_1 \subset K$ ,  $K_1$  không là siêu khóa.

# Khóa – Keys

- Thuộc tính tham gia vào một khóa gọi là *thuộc tính khóa*, ngược lại là *thuộc tính không khóa*.
- Ví dụ:
  - MaNV là thuộc tính khóa của quan hệ NhanVien.
  - Trong quan hệ PhanCong có khóa  $K = \{MANV, MADA\}$ . Như vậy thuộc tính khóa sẽ là: MANV, MADA.

# Khóa Chính – Primary Keys

- Khi cài đặt trên một HQTCSDL cụ thể, nếu quan hệ có nhiều hơn một **khóa**, ta chỉ được chọn một khóa và gọi là **khóa chính**.
- Các *thuộc tính nằm trong khóa chính* khi liệt kê trong quan hệ phải được gạch dưới.



<b>HoTen</b>	<b>Phai</b>	<b>Luong</b>	<b>PHG</b>
N T A	Nữ	2000	NC
L V M	Nam	2100	NC
N T A	Nam	2500	KH

# Khóa Chính – Primary Keys

- Ví dụ 1 : Khóa chính gồm 1 thuộc tính
  - Khóa chính của quan hệ NhanVien là: {MANV}
  - NhanVien(MaNV, HoTen, Phai, Luong, PHG)
- Ví dụ 2 : Khóa chính gồm hơn 1 thuộc tính
  - Khóa chính của quan hệ PhanCong là: {MANV, MADA}
  - PhanCong=(MaNV, MaDA, ThoiGian)
  - PhanCong=(MaNV, ThoiGian, MaDA)

# Khóa Ngoại – Foreign Keys

- Cho  $Q1(U)$ ,  $Q2(V)$  là 2 quan hệ, với  $U$ ,  $V$  lần lượt là tập thuộc tính của các quan hệ  $Q1$  và  $Q2$ .
- $K1$  là khóa chính của  $Q1$ ,  $K2 \subseteq Q2^+$
- Nếu điều kiện sau được thỏa:
  - $K1$  và  $K2$  có cùng số thuộc tính và ngữ nghĩa của các thuộc tính.
  - Giữa  $Q1$  và  $Q2$  tồn tại mỗi quan hệ 1-n trên  $K1$  và  $K2$ , mỗi bộ  $q2 \in Q2$ ,  $\exists$  duy nhất 1 bộ  $q1 \in Q1$ , sao cho  $q1.K1 = q2.K2$
- Khi đó,  ***$K2$  là khóa ngoại*** của  $Q2$  ***tham chiếu đến khóa chính  $K1$***  của  $Q1$



# Khóa Ngoại – Foreign Keys

- Ví dụ, cho 2 quan hệ
  - PhongBan(MaPB, TenPB), với  $U = \text{PhongBan}^+$
  - NhanVien(MaNV, HoTen, Phai, Phg), với  $V = \text{NhanVien}^+$
- $K1 = \{\text{MaPB}\}$  là khóa chính của quan hệ PhongBan,  $K2 = \{\text{Phg}\} \subseteq V$ 
  - $K1$  và  $K2$  cùng có 1 thuộc tính **biểu thị** mã phòng ban.
  - $\text{Dom}(K2) \subseteq \text{Dom}(K1)$ , nghĩa là, quan hệ 1-n
- $K2$  là khóa ngoại của NhanVien tham chiếu đến khóa chính  $K1$  của PhongBan.

# Khóa Ngoại – Foreign Keys

- Ví dụ, cho 2 quan hệ
  - NhanVien(MaNV, HoTen, Phai, MaPHG), với  $U = \text{NhanVien}^+$
  - PhanCong(MSNV, MaDA, ThoiGian), với  $V = \text{PhanCong}^+$
- $K1 = \{\text{MaNV}\}$  là khóa chính của quan hệ NhanVien,  $K2 = \{\text{MSNV}\} \subseteq V$ 
  - $K1$  và  $K2$  cùng có 1 thuộc tính biểu thị mã nhân viên.
  - $\text{Dom}(K2) \subseteq \text{Dom}(K1)$ , nghĩa là, quan hệ 1-n
- $K2$  là khóa ngoại của PhanCong tham chiếu đến khóa chính  $K1$  của NhanVien.



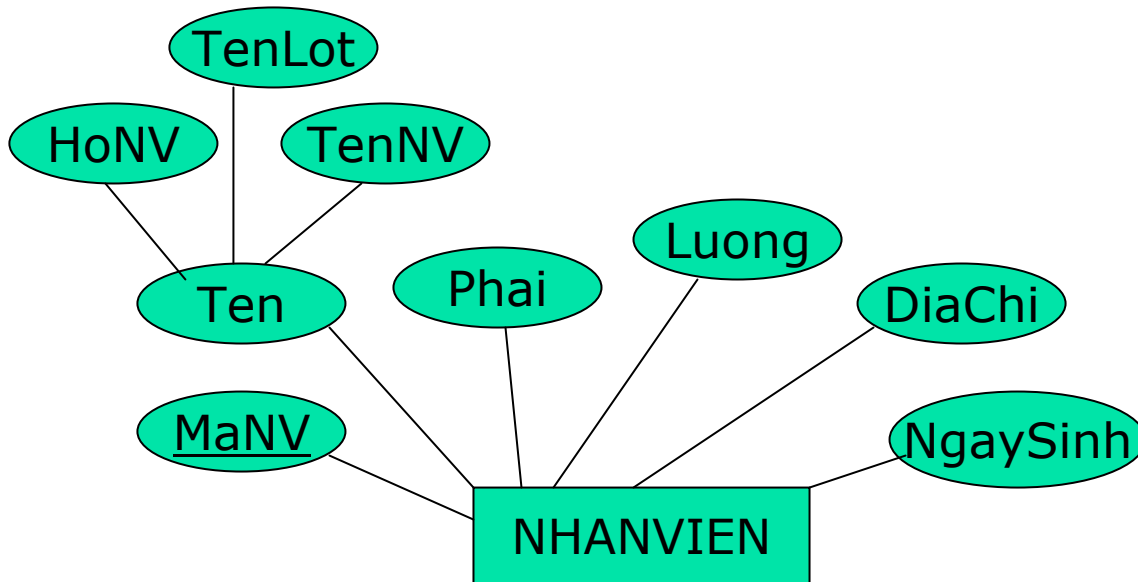
# Nội Dung

---

- Được đề xuất bởi TS. E. F. Codd năm 1970.
- Nền tảng của mô hình này là lý thuyết tập hợp trên các quan hệ.
- **Nội dung chính**
  1. Thuộc tính
  2. Quan hệ
  3. Bộ giá trị
  4. Thể hiện của quan hệ
  5. Lược đồ quan hệ
  6. Lược đồ CSDL
  7. Khóa – Keys
  - 8. Chuyển ER sang RDM**

## \$2.8. Chuyển từ ER sang RDM

- Bước 1. Mỗi loại thực thể → Quan hệ



- NHANVIEN(MaNV, HoNV, TenLot, TenNV, Phai, Luong, DiaChi, NgaySinh)



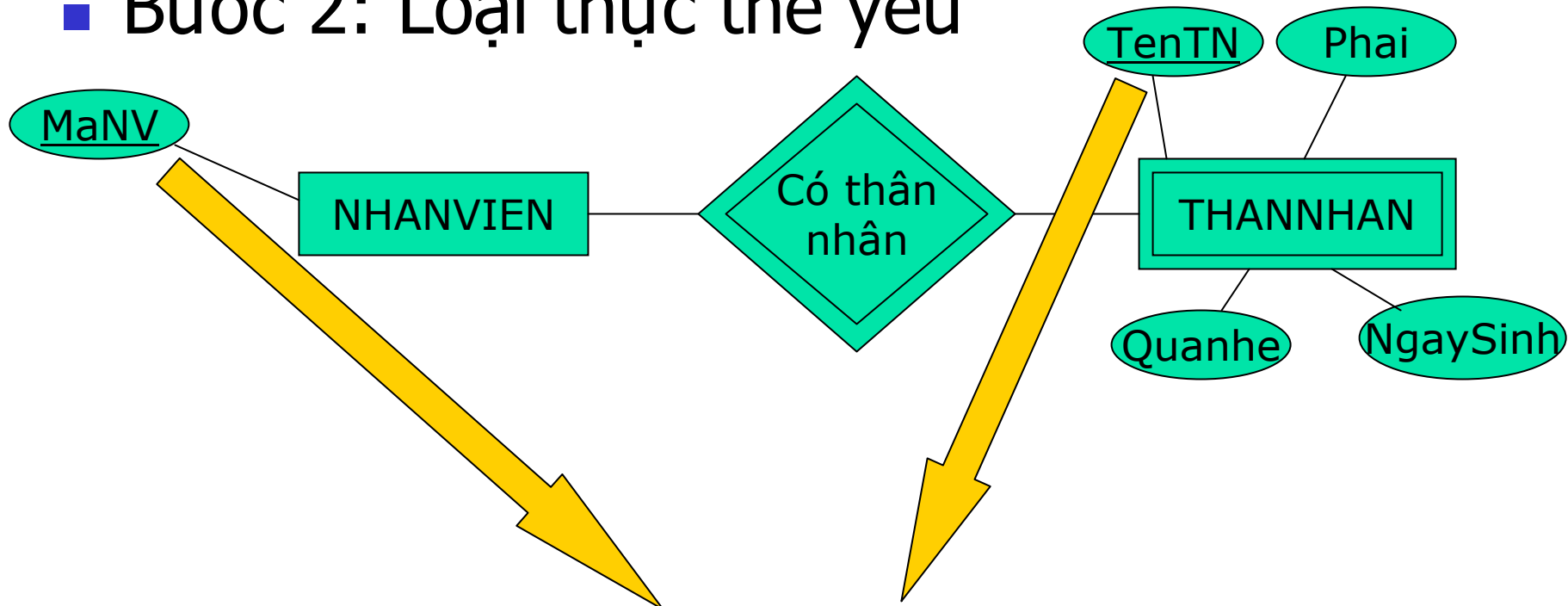
## \$2.8. Chuyển từ ER sang RDM

---

- Kết quả đạt được sau khi áp dụng Bước 1.
  - NHANVIEN(MaNV, HoNV, TenLot, TenNV, Phai, Luong, DiaChi, NgaySinh)
  - PHONGBAN(MaPHG, TenPHG)
  - DEAN(MaDA, TenDA, DDiem\_DA)

## \$2.8. Chuyển từ ER sang RDM

- Bước 2: Loại thực thể yếu



- THANNHAN(MaNV, TenTN, Phai, NgaySinh, Quanhe)



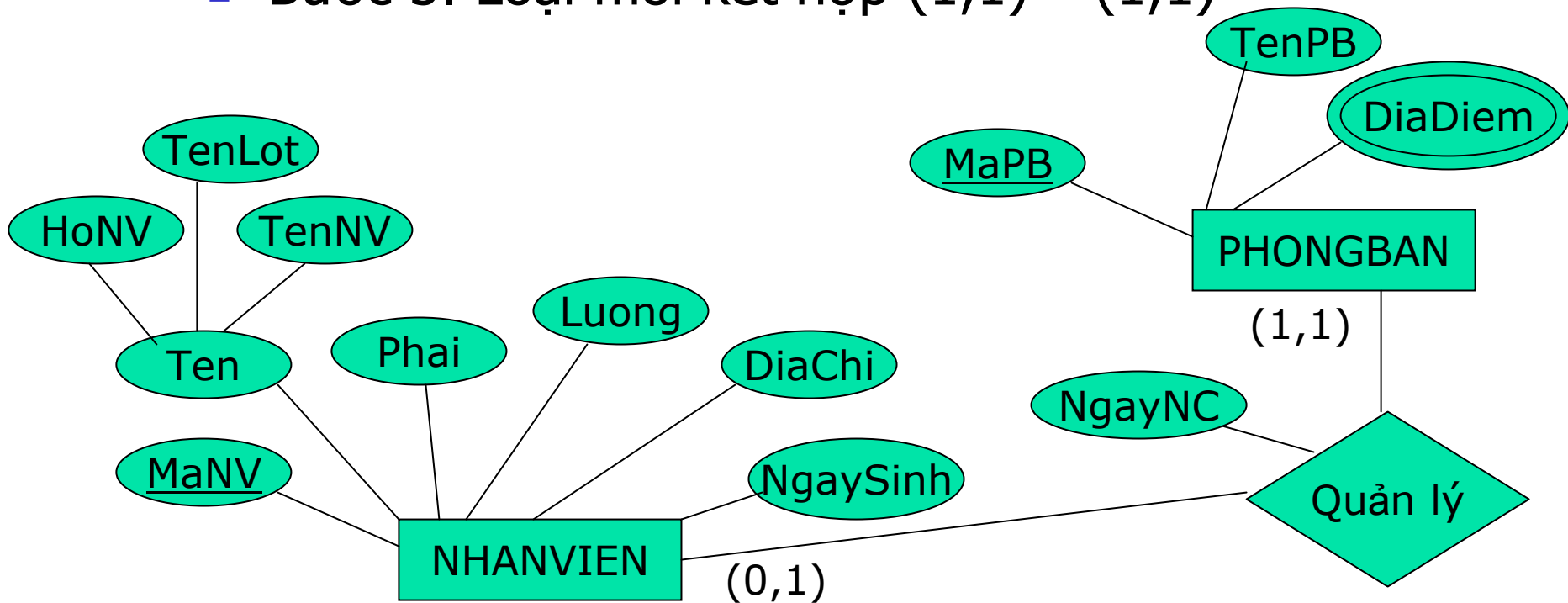
## \$2.8. Chuyển từ ER sang RDM

---

- Các lược đồ thu được từ các loại thực thể
  - NHANVIEN(MaNV, HoNV, TenLot, TenNV, Phai, Luong, DiaChi, NgaySinh)
  - PHONGBAN(MaPHG, TenPHG)
  - DEAN(MaDA, TenDA, DDiem\_DA)
  - THANNHAN(MaNV, TenTN, Phai, NgaySinh, Quanhe)

## \$2.8. Chuyển từ ER sang RDM

- Bước 3. Loại mỗi kết hợp (1,1) – (1,1)



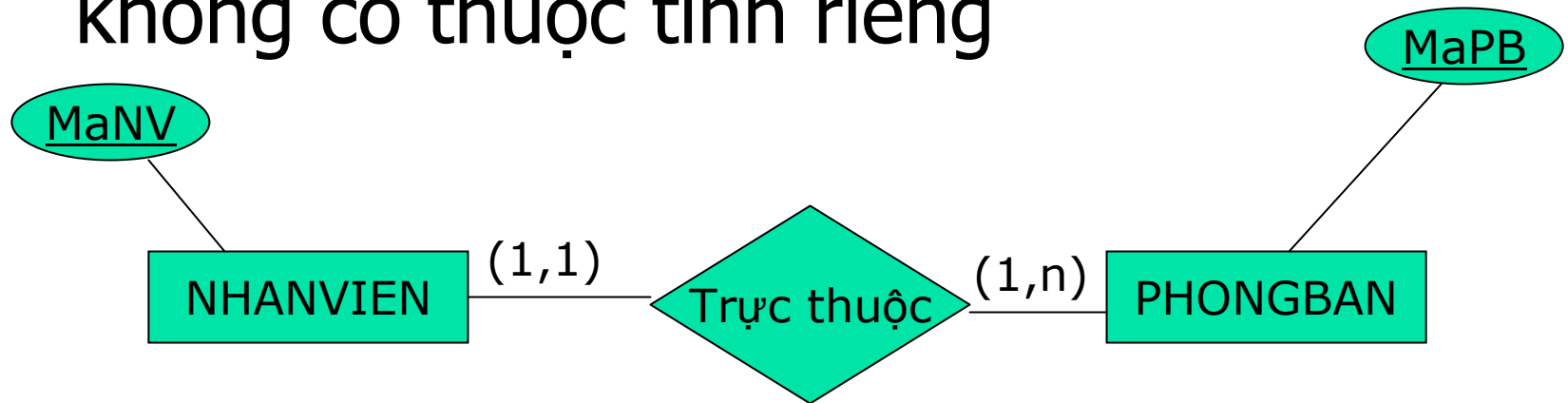
NHANVIEN(MaNV, HoNV, TenLot, TenNV, Phai, Luong, DiaChi, NgaySinh)

PHONGBAN(MaPB, TenPB, TrPHG, NgayNC)



## \$2.8. Chuyển từ ER sang RDM

- Bước 4. Loại mỗi kết hợp (1,1) – (1,n) không có thuộc tính riêng

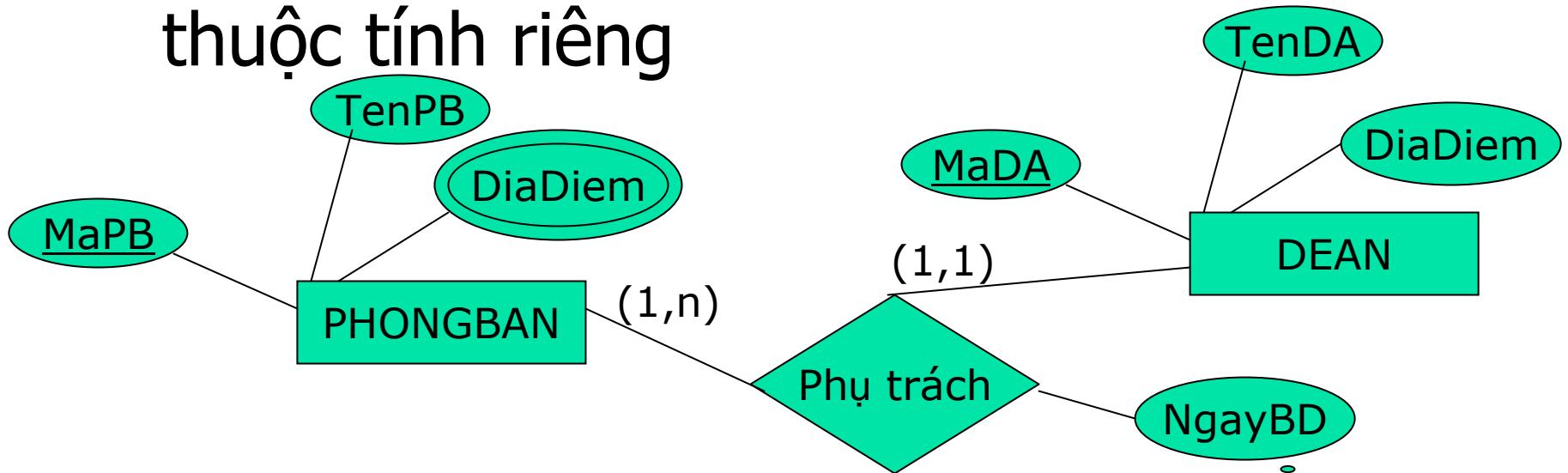


NHANVIEN(MaNV, HoNV, TenLot, TenNV, Phai, Luong, DiaChi, NgaySinh, MaPB)

PHONGBAN(MaPB, TenPB, TrPHG, NgayNC)

## \$2.8. Chuyển từ ER sang RDM

- Bước 5. Loại mỗi kết hợp (1,1) – (1,n) có thuộc tính riêng



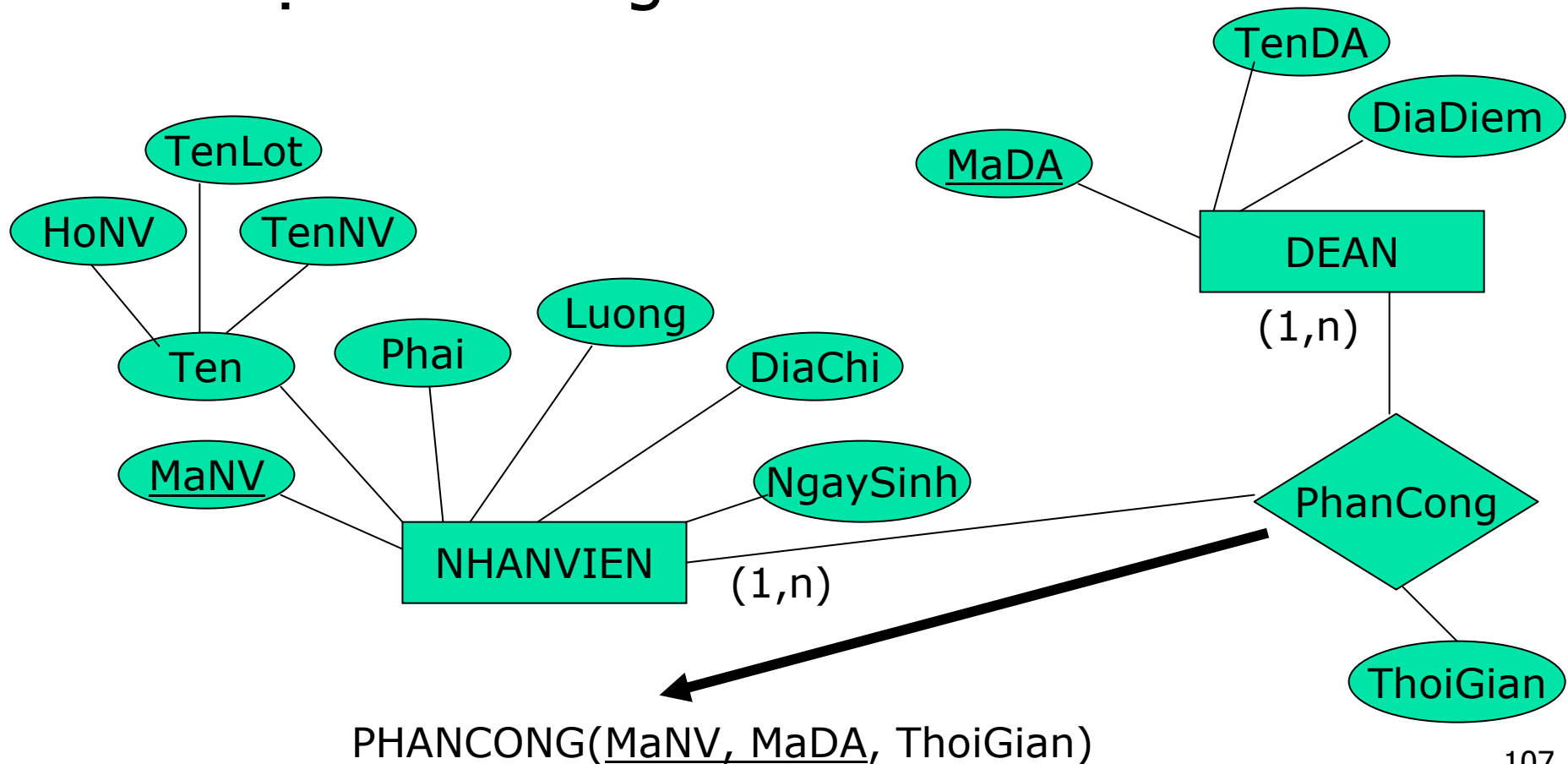
PHONGBAN(MaPB, TenPB, TrPhg, NgayNC)

DEAN(MaDA, TenDA, DiaDiem, MaPB, NgayBD)

Ví dụ

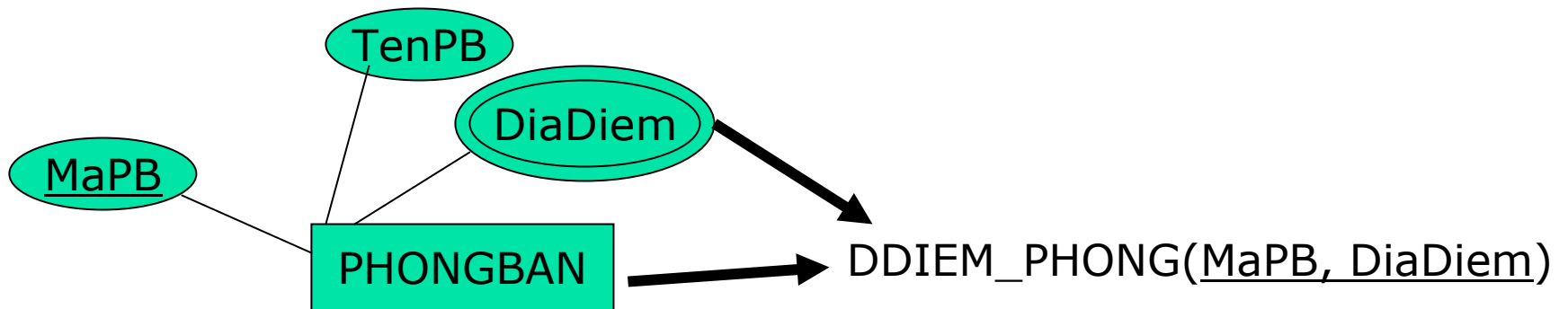
## \$2.8. Chuyển từ ER sang RDM

- Bước 6. Loại mỗi kết hợp (1,n) – (1,n) có thuộc tính riêng



# \$2.8. Chuyển từ ER sang RDM

- Bước 7. Thuộc tính đa trị



# \$2.8. Chuyển từ ER sang RDM

## ■ Lược đồ CSDL thu được

**THANNHAN**(MaNV, TenTN, Phai, NgaySinh, Quanhe)

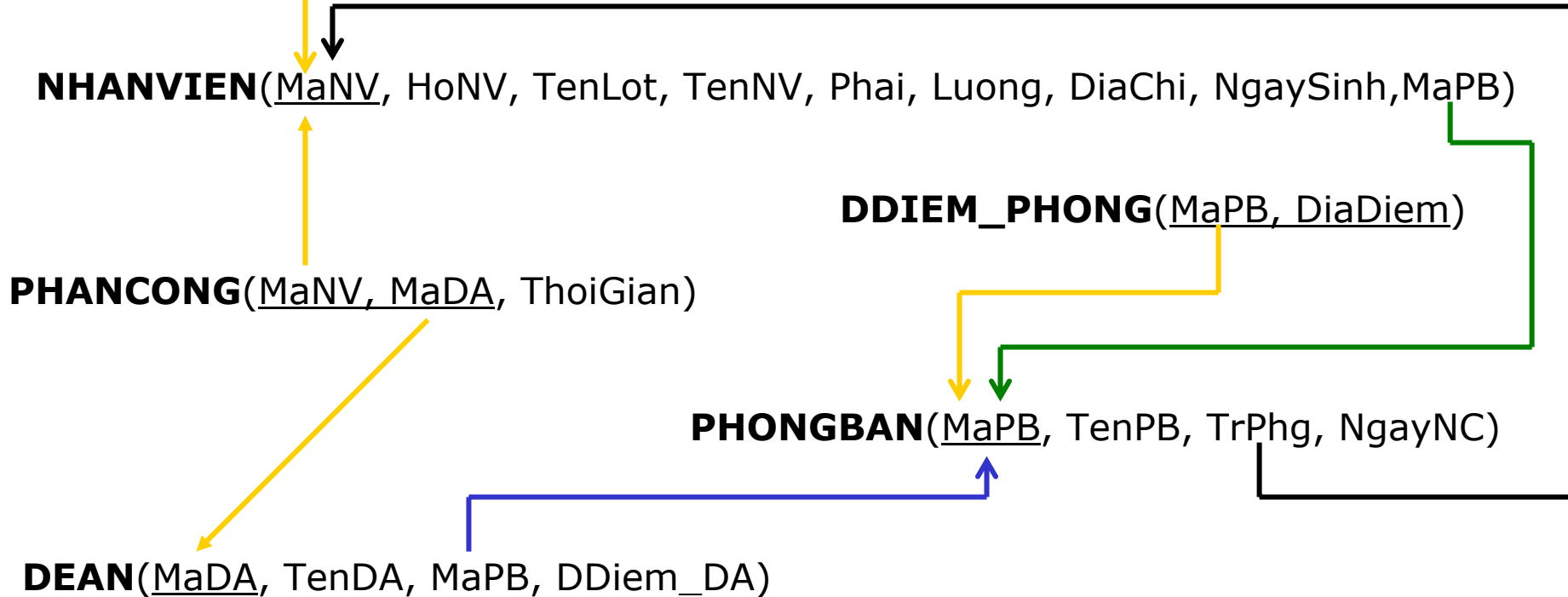
**NHANVIEN**(MaNV, HoNV, TenLot, TenNV, Phai, Luong, DiaChi, NgaySinh, MaPB)

**DDIEM\_PHONG**(MaPB, DiaDiem)

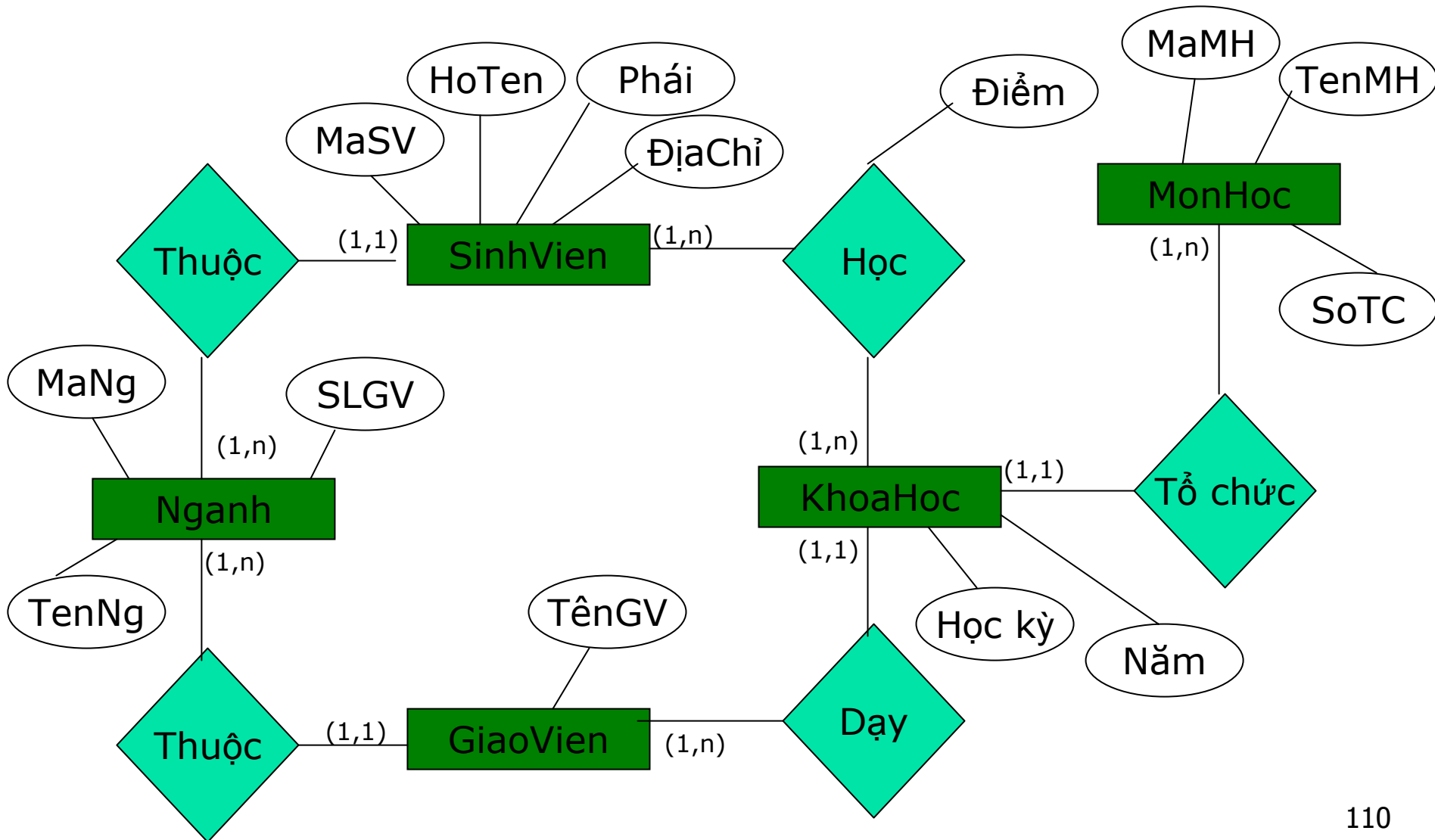
**PHANCONG**(MaNV, MaDA, ThoiGian)

**PHONGBAN**(MaPB, TenPB, TrPhg, NgayNC)

**DEAN**(MaDA, TenDA, MaPB, DDiem\_DA)



# \$2.8. Chuyển từ ER sang RDM



# CƠ SỞ DỮ LIỆU

ThS.Lê Văn Hạnh

# ĐẠI SỐ QUAN HỆ (Relational Algebra)





# Nội Dung

---

1. Giới thiệu
2. Các phép toán trên quan hệ
  - $\cup, \cap, -, \sigma, \Pi, \times$
  - $\bowtie$ :  $\theta$ -kết ( $\theta$ -join)
    - Kết bằng (equi-join), Kết tự nhiên (natural-join), Kết nội (inner-join), Kết trái (left-join), Kết phải (right-join), Kết ngoại (outer-join)
  - Phép chia
  - Các hàm kết hợp
3. Các thao tác trên dữ liệu quan hệ: thêm, xóa, sửa



## \$3.1. Giới thiệu

---

- Đại số quan hệ là ngôn ngữ dùng để đặc tả việc truy xuất dữ liệu trên quan hệ
- Gồm tập hợp các phép toán trên các quan hệ và cho kết quả là một quan hệ
- Đại số quan hệ có nền tảng toán học (cụ thể là lý thuyết tập hợp) để mô hình hóa CSDL quan hệ. Ngoài ra, ta có thể áp dụng các phép toán có sẵn của lý thuyết tập hợp để tối ưu quá trình rút trích dữ liệu từ CSDL quan hệ



## \$3.2. Các phép toán trên quan hệ

- Có năm phép toán cơ bản:
  - Chọn ( $\sigma$ ): Chọn tập con các dòng trong quan hệ.
  - Chiếu ( $\Pi$ ): Xóa các cột không mong muốn.
  - Tích Descartes ( $\times$ ): Kết hai quan hệ.
  - Hiệu ( $-$ ): Chứa các bộ của QH 1 nhưng không nằm trong QH 2.
  - Hợp ( $\cup$ ): Chứa các bộ của QH 1 và các bộ trong QH 2.
- Các phép toán khác: Giao ( $\cap$ ), kết, chia, đổi tên: không cơ bản, nhưng hữu ích.
- Kết quả thực hiện các phép toán trên cũng là các quan hệ, do đó có thể kết hợp giữa các phép toán này để tạo nên phép toán mới



## \$3.2. Các phép toán trên quan hệ

---

- Một số phép toán trên quan hệ đòi hỏi 2 quan hệ tham gia phải **khả hợp**
- Quan hệ khả hợp: hai quan hệ  $r$  và  $s$  được coi là khả hợp khi:
  - Số lượng thuộc tính phải bằng nhau (cùng bậc)
  - Miền giá trị của thuộc tính phải tương thích ( $\text{dom}(A_i) = \text{dom}(B_i)$ ) với  $1 \leq i \leq n$

# Phép hội – $\cup$

- Cho  $r$  và  $s$  là 2 quan hệ khả hợp  
 $r \cup s = \{ t / (t \in r) \vee (t \in s) \}$
- Quan hệ kết quả gồm tất cả các bộ có trong  $r$  hoặc có trong  $s$  (không trùng)

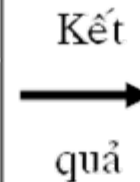
*DSLỚP ANH VÃN*

*DSLỚP TIN HỌC*

<i>MÃSV</i>	<i>HỌTÊN</i>
SV01	Trần Sửu
SV02	Phạm Tý
SV04	Hoàng Mão



<i>MÃSV</i>	<i>HỌTÊN</i>
SV02	Phạm Tý
SV05	Nguyễn Dần



<i>MÃSV</i>	<i>HỌTÊN</i>
SV01	Trần Sửu
SV02	Phạm Tý
SV04	Hoàng Mão
SV05	Nguyễn Dần

# Phép giao – $\cap$

- Cho  $r$  và  $s$  là 2 quan hệ khả hợp  
 $r \cap s = \{ t / (t \in r) \wedge (t \in s) \}$
- Quan hệ kết quả gồm tất cả các bộ có trong cả  $r$  lẫn  $s$

*ĐTỬ – TH*

<i>MÃ MÔN</i>	<i>TÊN MÔN</i>
07	Điện tử căn bản
08	Cấu trúc máy tính
09	Kỹ thuật số

*TH – ỨNG DỤNG*

<i>MÃ MÔN</i>	<i>TÊN MÔN</i>
08	Cấu trúc máy tính
10	Cơ sở dữ liệu
11	PTTK hệ thống

$\cap$

Kết  
quả  $\rightarrow$

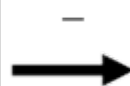
<i>MÃ MÔN</i>	<i>TÊN MÔN</i>
08	Cấu trúc máy tính

# Phép trừ - -

- Cho  $r$  và  $s$  là 2 quan hệ khả hợp  
 $r - s = \{ t / (t \in r) \wedge (t \notin s) \}$
- Quan hệ kết quả gồm tất cả các bộ có trong  $r$  nhưng không có trong  $s$

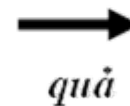
*ĐTỬ - TH*

MÃ MÔN	TÊN MÔN
07	Điện tử căn bản
08	Cấu trúc máy tính
09	Kỹ thuật số



*TH - ỨNG DỤNG*

MÃ MÔN	TÊN MÔN
08	Cấu trúc máy tính
10	Cơ sở dữ liệu
11	PTTK hệ thống



MÃ MÔN	TÊN MÔN
07	Điện tử căn bản
09	Kỹ thuật số

# Phép chọn – $\delta$

- Định nghĩa: Cho lược đồ quan hệ  $R (A_1, A_2, \dots, A_n)$ ,  $r(R)$ :
  - $\delta_E(r) = \{t, t \in r \text{ và } t \text{ thỏa } E\}$ , với
  - $E$  là một biểu thức cho biết điều kiện chọn
- Phép chọn được dùng để trích chọn các dòng thỏa điều kiện chọn  $E$  từ quan hệ ban đầu.



# Phép chọn – $\delta$

- Ví dụ, cho biết các sinh viên có học bổng

*SinhVien*

MÃ SV	HỌ TÊN	HỌC BỔNG
A01	Phạm Tý	200
B01	Trần Sứ	
A02	Nguyễn Dần	100
C01	Phùng Hợi	

$\delta_{HocBong>0}(SinhVien)$

*SV*

MÃ SV	HỌ TÊN	HỌC BỔNG
A01	Phạm Tý	200
A02	Nguyễn Dần	100

- Cho danh sách các nhân viên nữ của phòng KD có lương lớn hơn 1000000.

$\delta_{Phai='nu' \text{ and } Phg='KD' \text{ and } Luong>1000000}(\mathbf{NHANVIEN})$

# Phép chọn – $\delta$

- Đưa ra các phép chọn để có  $R'$  từ  $R$

$R$

<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
a	a	1	7
a	b	5	7
b	b	12	3
b	b	23	10

$R'$

<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
a	a	1	7
b	b	23	10

- Chọn những nhân viên thuộc Phòng có mã số là "P3" và phái là "Nam".
- Chọn những nhân viên có lương nằm trong khoảng từ 30000 đến 40000.
- Chọn những nhân viên phái là "Nữ" và lương >25000 hoặc những nhân viên có họ là "Đình"

# Phép chiếu - $\Pi$

- Cho trước một lược đồ quan hệ  $R(A_1, A_2, \dots, A_n)$ ,  $X \subset R$ ,  $r(R)$ :
  - $\Pi_X(r) = \{t[X] \mid t \in r\}$
  - Ký hiệu:  $\Pi_{A_1, A_2, \dots, A_k}(r)$
- Dùng để trích chọn các thuộc tính (cột) từ quan hệ  $r$  được chỉ ra trong danh sách thuộc tính
- Các dòng trùng nhau sẽ được loại bỏ khỏi quan hệ kết quả

# Phép chiếu - $\Pi$

- Phép chiếu lược đồ quan hệ SINHVIÊN (MaSV, HoTen, HocBong) lên tập thuộc tính gồm {HoTen, MaSV} sẽ cho kết quả như sau

*SinhVien*

MÃ SV	HỌ TÊN	HỌC BỔNG
A01	Phạm Tý	200.000
B01	Trần Sửu	
A02	Nguyễn Dân	100.000
C01	Phùng Hợi	

$\longrightarrow$   
 $\pi_{HoTen, MaSV}(SinhVien)$

*SV*

HỌ TÊN	MÃ SV
Phạm Tý	A01
Trần Sửu	B01
Nguyễn Dân	A02
Phùng Hợi	C01

# Phép chiếu - $\Pi$

- Cho danh sách nhân viên gồm các thông tin tên nhân viên, lương, ngày sinh, phái
  - $\Pi_{\text{TenNV, Luong, NgaySinh, Phai}}(\text{NHANVIEN})$
- Cho biết  $\text{MaNV}$ ,  $\text{TenNV}$ ,  $\text{Luong}$  của các nhân viên có  $\text{luong} > 20000$ 
  - $\Pi_{\text{MaNV, TenNV, Luong}}(\delta_{\text{Luong} > 20000}(\text{NHANVIEN}))$
- Cho biết tên nhân viên nữ phòng KD
  - $\Pi_{\text{TenNV}}(\delta_{\text{Phai} = \text{'nu'} \text{ and Phg} = \text{'KD'}}(\text{NHANVIEN}))$

# Phép gán ←

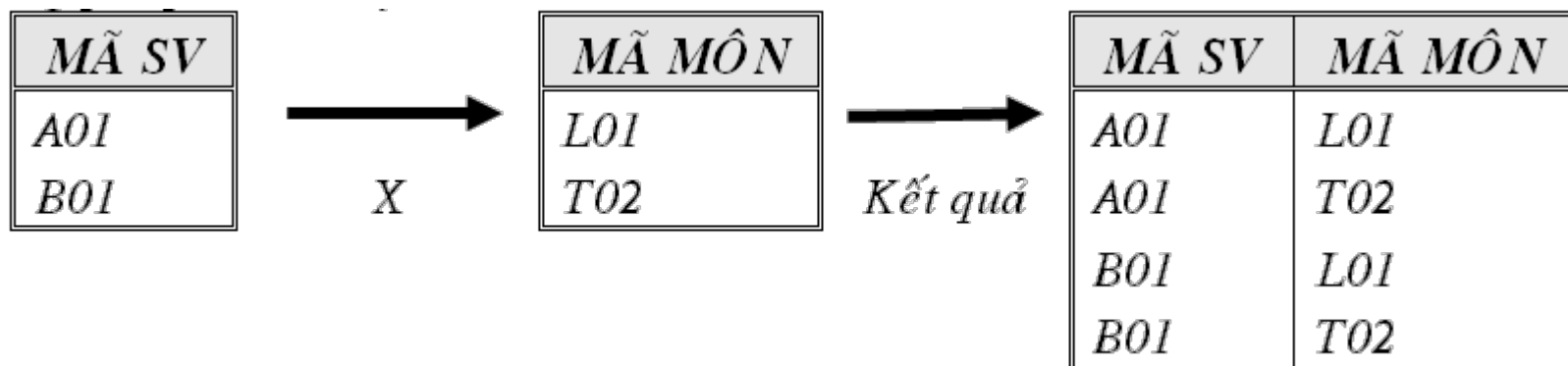
- Phép gán giúp thể hiện dễ dàng các câu truy vấn phức tạp.
- Ý tưởng: viết các câu truy vấn như là một chương trình tuần tự gồm một chuỗi các phép toán
- Ví dụ: Cho biết tên nhân viên nữ phòng KD
  - $KQ1 \leftarrow \delta_{\text{Phai}='nu' \text{ and } \text{Phg}='KD'} (\text{NHANVIEN})$
  - $KQ \leftarrow \Pi_{\text{TenNV}} (KQ1)$

# Phép đổi tên

- Cho phép tham chiếu tới một quan hệ bằng nhiều tên.
- Dễ dàng sử dụng trong trường hợp một quan hệ được sử dụng nhiều lần trong một biểu thức quan hệ
- Quan hệ: <tên mới> ← phép toán quan hệ
  - $SV \leftarrow \Pi_{HoTen, MaSV}(SinhVien)$
- Thuộc tính: <tên mới qh>(tên mới thuộc tính) ← phép toán quan hệ
  - $SV(Ten, Maso) \leftarrow \Pi_{HoTen, MaSV}(SinhVien)$

# Phép tích Cartersian - $\times$

- $r \times s := \{t \ q \mid t \in r \wedge q \in s\}$ 
  - Thuộc tính của  $r$  và  $s$  phải khác nhau ( $r \cap s = \emptyset$ )
  - Phép tích Cartersian là một phép toán vô nghĩa nếu đứng riêng. Do đó, người ta thường phải sử dụng phối hợp với các phép toán khác





# Phép tích Cartersian - $\times$

- Với mỗi phòng ban, cho biết thông tin về người trưởng phòng
  - $PB\_NV \leftarrow \text{PhongBan} \times \text{NhanVien}$
  - $PB\_TruongPhong \leftarrow \delta_{\text{MaNV}=\text{TrPhg}}(PB\_NV)$

NHANVIEN			
MAN	HONV	TENLO	TENN
001 <sup>V</sup>	Vuong	Ngoc <sup>T</sup>	Quy <sup>V</sup> en
002	Nguye	Thanh	Tung
003	Le	Thi	Nhan
004	Dinh	Ba	Tien

PHONGBAN		
MAPH	TENPHG	TRPH
QL <sup>G</sup>	Quan Ly	001 <sup>G</sup>
DH	Dieu Hanh	003
NC	Nghien	002

$\times$

MANV	HONV	TENLOT	TENNV	MAPHG	TENPHG	TRPHG
001	Vuong	Ngoc	Quy <sup>V</sup> en	QL	Quan Ly	001
001	Vuong	Ngoc	Quy <sup>V</sup> en	NC	Nghien Cuu	002
002	Nguye <sup>N</sup> n	Thanh	Tung	QL	Quan Ly	001
002	Nguye <sup>N</sup> n	Thanh	Tung	NC	Nghien Cuu	002
003	Le	Thi	Nhan	DH	Dieu Hanh	003
003	Le	Thi	Nhan	QL	Quan Ly	001
003	Le	Thi	Nhan	DH	Dieu Hanh	003
003	Le	Thi	Nhan	NC	Nghien Cuu	002
004	Dinh	Ba	Tien	QL	Quan Ly	001
004	Dinh	Ba	Tien	DH	Dieu Hanh	003
004	Dinh	Ba	Tien	NC	Nghien Cuu	002

# Phép kết - $\bowtie$

---

- $\theta$  – kết
- Kết bằng, kết tự nhiên
- Kết trái
- Kết phải
- Kết nội
- Kết ngoại

# Phép kết - $\bowtie$ - $\theta$ - kết

- Phép kết là sự kết hợp có thứ tự của phép tích Cartersian và phép chọn.
- Gọi  $\theta$  là một trong các phép so sánh  $\{=, >, \geq, <, \leq, \neq\}$ . Phép kết giữa hai lược đồ quan hệ  $Q1$  và  $Q2$  sẽ tạo thành một lược đồ quan hệ  $Q12$ :
  - $Q1 \bowtie_{A \theta B} Q2 = \{ q_1 \cap q_2 \mid q_1 \in r, u \in s \wedge q_1[A] \theta q_2[B] \}$
  - Điều kiện  $A \theta B$  được gọi là *điều kiện kết*.

# Phép kết - $\bowtie$ - $\theta$ - kết

- Phép kết bằng: khi  $\theta$  là phép so sánh "=" ta gọi là phép kết bằng hay còn gọi là equi-join
- Phép kết tự nhiên: Kết bằng hai quan hệ và một trong hai thuộc tính tham gia điều kiện kết đó được loại bỏ qua phép chiếu thì phép kết đó được gọi là kết tự nhiên (natural-join)

# Phép kết - ⋈ - Kết nội

- Phép kết nội (inner join): Thực chất là phép *kết nối bằng*.
  - Hai thuộc tính so sánh có cùng tên thì kết quả phép kết nối vẫn giữ lại 2 tên thuộc tính đó

*SINHVIÊN*

<i>MaSV</i>	<i>TênSinhViên</i>	<i>MãLớp</i>
S1	Tý	L1
S2	Sửu	L2
S3	Dần	L2
S4	Mẹo	L3
S5	Thìn	L4

*LỚP*

<i>MãLớp</i>	<i>TênLớp</i>
L1	Anh văn
L2	Tin học
L3	Toán
L5	Kế toán

*KQ1*

<i>MaSV</i>	<i>TênSinhViên</i>	<i>MãLớp</i>	<i>MãLớp</i>	<i>TênLớp</i>
S1	Tý	L1	L1	Anh văn
S2	Sửu	L2	L2	Tin học
S3	Dần	L2	L2	Tin học
S4	Mẹo	L3	L3	Toán

# Phép kết - Kết trái - $\bowtie$

- Phép kết trái (left join):

- $R \bowtie S = \{ v = (t, u) \mid (t \in R, u \in S \wedge t.A \in u.B) \text{ hoặc } (t \in R, u = u_{\text{NULL}} \text{ với } t.A \notin S[B]) \}$

*SINHVIÊN*

<i>MaSV</i>	<i>TênSinhViên</i>	<i>MãLớp</i>
S1	Tý	L1
S2	Sửu	L2
S3	Dần	L2
S4	Mẹo	L3
S5	Thìn	L4

*LỚP*

<i>MãLớp</i>	<i>TênLớp</i>
L1	Anh văn
L2	Tin học
L3	Toán
L5	Kế toán

*KQ1*

<i>MaSV</i>	<i>TênSinhViên</i>	<i>MãLớp</i>	<i>MãLớp</i>	<i>TênLớp</i>
S1	Tý	L1	L1	Anh văn
S2	Sửu	L2	L2	Tin học
S3	Dần	L2	L2	Tin học
S4	Mẹo	L3	L3	Toán
S5	Thìn	L4	-	-

# Phép kết - Kết phải - $\bowtie$

- Phép kết phải (right join):

- $R \bowtie S = \{ v = (t, u) \mid (t \in R, u \in S \wedge t.A \in u.B) \text{ hoặc } (u \in S, t = t_{\text{NULL}} \text{ với } u.B \notin R[A]) \}$

*SINHVIÊN*

<i>MaSV</i>	<i>TênSinhViên</i>	<i>MãLớp</i>
S1	Tý	L1
S2	Sửu	L2
S3	Dần	L2
S4	Mẹo	L3
S5	Thìn	L4

*LỚP*

<i>MãLớp</i>	<i>TênLớp</i>
L1	Anh văn
L2	Tin học
L3	Toán
L5	Kế toán

*KQ1*

<i>MaSV</i>	<i>TênSinhViên</i>	<i>MãLớp</i>	<i>MãLớp</i>	<i>TênLớp</i>
S1	Tý	L1	L1	Anh văn
S2	Sửu	L2	L2	Tin học
S3	Dần	L2	L2	Tin học
S4	Mẹo	L3	L3	Toán
-	-	-	L5	Kế toán

# Phép kết - Kết ngoại

- Phép kết ngoại (outer join):

■ Kết hợp hai phép kết trái và kết phải

*SINHVIÊN*

<i>MaSV</i>	<i>TênSinhViên</i>	<i>MãLớp</i>
S1	Tý	L1
S2	Sửu	L2
S3	Dần	L2
S4	Mẹo	L3
S5	Thìn	L4

*LỚP*

<i>MãLớp</i>	<i>TênLớp</i>
L1	Anh văn
L2	Tin học
L3	Toán
L5	Kế toán

*KQ1*

<i>MaSV</i>	<i>TênSinhViên</i>	<i>MãLớp</i>	<i>MãLớp</i>	<i>TênLớp</i>
S1	Tý	L1	L1	Anh văn
S2	Sửu	L2	L2	Tin học
S3	Dần	L2	L2	Tin học
S4	Mẹo	L3	L3	Toán
S5	Thìn	L4	-	-
-	-	-	L5	Kế toán



# Phép chia - $\div$

- $r \div s$ : lấy các bộ trong quan hệ  $r$  sao cho khớp với **tất cả** các bộ trong quan hệ  $s$ 
  - các thuộc tính của  $s$  phải là một tập thuộc tính con của  $r$ ,  $s^+ \subset r^+$ .
  - Thuộc tính của quan hệ kết quả chính là các thuộc tính của  $r - s$

# Phép chia - $\div$

<b>R</b>		<b>S</b>
A	B	B
$\alpha$	1	1
$\alpha$	2	2
$\alpha$	3	
$\beta$	1	
$\delta$	1	
$\delta$	2	
$\lambda$	1	
$\lambda$	3	

- Các bước thực hiện:

1. Xóa các thuộc tính không khóa trong R và S

2. Giả sử, gọi B là cột chung của cả 2 quan hệ R và S. Thực hiện xóa các dòng trong R sao cho giá trị của cột B trong R không có trong cột B của S

**B2:** Xóa dòng trong R

A	B
$\alpha$	1
$\alpha$	2
$\beta$	1
$\delta$	1
$\delta$	2
$\lambda$	1

# Phép chia - $\div$

**B2:** Xóa dòng trong  $R$

A	B
$\alpha$	1
$\alpha$	2
$\beta$	1
$\delta$	1
$\delta$	2
$\lambda$	1

- Các bước thực hiện:

3. Tạo  $t1$  bằng cách thực hiện phép chiếu trên cột không phải cột chung của  $R$  và  $S$
4. Tạo  $t2 = t1 \times S$

**B3:** Tạo  $t1$

A
$\alpha$
$\beta$
$\delta$
$\lambda$

$S$

B
1
2

**B4:**  $t2 = t1 \times S$

A	B
$\alpha$	1
$\alpha$	2
$\beta$	1
$\beta$	2
$\delta$	1
$\delta$	2
$\lambda$	1
$\lambda$	2

# Phép chia - $\div$

- Các bước thực hiện:

5. Tạo  $t3 = t2 - R$

**B4:**  $t2 = t1 \times S$

A	B
$\alpha$	1
$\alpha$	2
$\beta$	1
$\beta$	2
$\delta$	1
$\delta$	2
$\lambda$	1
$\lambda$	2

**R**

A	B
$\alpha$	1
$\alpha$	2
$\alpha$	3
$\beta$	1
$\delta$	1
$\delta$	2
$\lambda$	1
$\lambda$	3

**B5:**  $t3 = t2 - R$

A	B
$\alpha$	1
$\alpha$	2
$\beta$	1
<b><math>\beta</math></b>	<b>2</b>
$\delta$	1
$\delta$	2
$\lambda$	1
<b><math>\lambda</math></b>	<b>2</b>

# Phép chia - $\div$

- Các bước thực hiện:
  6. Chiếu để lấy các thuộc tính của  $t3$  có trong  $R$  và không có trong  $S$
  7. Tạo tập kết quả:  $tKq = t1 - t4$

*B6:  $t4 =$  Chiếu thuộc tính A của  $t3$*

$t4$

A
$\beta$
$\lambda$

*B7:  $tKq = t1 - t4$*

$t1 \quad - \quad t4 \quad = \quad tKq$

A
$\alpha$
$\beta$
$\delta$
$\lambda$

A
$\beta$
$\lambda$

A
$\alpha$
$\delta$

$R$		$S$
A	B	B
$\alpha$	1	1
$\alpha$	2	2
$\alpha$	3	
$\beta$	1	
$\delta$	1	
$\delta$	2	
$\lambda$	1	
$\lambda$	3	

# Phép chia - ÷

- Cho biết các nhân viên được phân công tham gia tất cả các đề án

PHANCONG		
MANV	MADA	THOIGIAN
001	TH001	30,0
001	TH002	12,5
002	TH001	10,0
002	TH002	10,0
002	DT001	10,0
002	DT002	10,0
003	TH001	37,5
004	DT001	22,5
004	DT002	10,0
006	DT001	30,5
007	TH001	20,0
007	TH002	10,0
008	TH001	10,0
008	DT002	12,5

DEAN	
MADA	TENDA
TH001	Tin hoc hoa 1
TH002	Tin hoc hoa 2
DT001	Dao tao 1
DT002	Dao tao 2

# Phép chia - ÷

- Cho biết các nhân viên được phân công tham gia tất cả các đề án có địa điểm tại TPHCM

PHANCONG		
MANV	MADA	THOIGIAN
001	TH001	30,0
001	TH002	12,5
002	TH001	10,0
002	TH002	10,0
002	DT001	10,0
002	DT002	10,0
003	TH001	37,5
004	DT001	22,5
004	DT002	10,0
006	DT001	30,5
007	TH001	20,0
007	TH002	10,0
008	TH001	10,0
008	DT002	12,5

DEAN		
MADA	TENDA	DDIEM_DA
TH001	Tin hoc hoa 1	HANOI
TH002	Tin hoc hoa 2	TPHCM
DT001	Dao tao 1	NHATRANG
DT002	Dao tao 2	HANOI

# Các hàm kết hợp – gom nhóm

- Max (thuộc tính kiểu số) : trả về giá trị lớn nhất.
- Min (thuộc tính kiểu số) : trả về giá trị nhỏ nhất.
- Count (tên thuộc tính), Count(\*): Đếm số phần tử, trùng vẫn đếm.
- Sum (thuộc tính kiểu số): Trả về tổng.
- Avg (thuộc tính kiểu số): Trả về giá trị trung bình.



# Các hàm kết hợp – gom nhóm

■  $G_1, G_2, \dots, G_n \rightsquigarrow F_1(A_1), F_2(A_2), \dots, F_n(A_n) (E)$

- E là biểu thức đại số quan hệ
- $G_i$  là tên thuộc tính gom nhóm (có thể tính toán không gom nhóm)
- $F_i$  là hàm gom nhóm, có thể là tên thuộc tính nếu muốn chọn thuộc tính này.
- $A_i$  là tên thuộc tính tính toán trong hàm gom nhóm  $F_i$

# Các hàm kết hợp – gom nhóm

MaNV	HoTen	Phai	MaPB
NV1	Bưởi	Nữ	KD
NV2	Cam	Nam	DH
NV3	Xoài	Nam	KD
NV4	Mít	Nữ	NC
NV5	Ổi	Nữ	DH

MaPB  $\mathfrak{Z}_{\text{COUNT}(\text{MaNV})}(\text{NhanVien})$

MaPB	No column name
KD	2
DH	2
NC	1

KQ(MaPB,SLNV)  $\leftarrow$  MaPB  $\mathfrak{Z}_{\text{COUNT}(\text{MaNV})}(\text{NhanVien})$

MaPB	SLNV
KD	2
DH	2
NC	1

# Các hàm kết hợp – gom nhóm

- Tính số nhân viên trong toàn công ty.

- $\tilde{\Sigma}_{\text{Count}(*)}$  (NHANVIEN)

No column name
20

- KQ(SLNV)  $\leftarrow \tilde{\Sigma}_{\text{Count}(*)}$  (NHANVIEN)

SLNV
20

# Các hàm kết hợp – gom nhóm

- Cho biết có bao nhiêu đề án ở TPHCM
- Đối với mỗi địa điểm, cho biết tên và số lượng đề án được thực hiện ở đó
- Cho biết số năm thâm niên (số năm làm trưởng phòng) của từng trưởng phòng
- Đối với từng đề án, cho biết số lượng nhân viên tham gia đề án, tổng số giờ làm việc của đề án

## §5.5. Các thao tác trên dữ liệu quan hệ

- Nội dung của CSDL có thể được cập nhật bằng cách dùng các thao tác thêm, xóa, sửa
- Tất cả các thao tác này được diễn đạt thông qua phép toán gán
  - R (kết quả) ← Các phép toán trên R (ban đầu)

# Thêm – Chèn (Insertion)

- Cú pháp:

- $R \leftarrow R \cup E$  ( với E là một biểu thức đại số quan hệ)

- Ví dụ: phân công NV 001 vào đề án X với thời gian là 25

- $PHANCONG \leftarrow PHANCONG \cup \{\mathbf{'001'}, \mathbf{'X'}, \mathbf{25}\}$

# Xóa (Delete)

- Cú pháp:
  - $R \leftarrow R - E$  ( với E là một biểu thức đại số quan hệ)
- Ví dụ: xóa toàn bộ phân công của NV 001
  - $T1 \leftarrow \delta_{MANV='001'}(PHANCONG)$
  - $PHANCONG \leftarrow PHANCONG - T1$

# Sửa (Update)

- Thường được diễn đạt bằng chuỗi các thao tác thêm và xóa. Tuy nhiên, có thể sử dụng trực tiếp các phép toán số học
- Ví dụ: Tăng lương gấp 1,5 cho nhân viên có mã số 001
  - $T1 \leftarrow \Pi_{MaNV, TenNV, Luong * 1.5} (\delta_{MaNV='001'} (NHANVIEN))$
  - $T2 \leftarrow \delta_{MaNV='001'} (NHANVIEN)$
  - $NHANVIEN \leftarrow NHANVIEN - T2$
  - $NHANVIEN \leftarrow NHANVIEN \cup T1$



# Sửa (Update)

- Tăng lương gấp 1,5 cho nhân viên có mã số 001

$$\text{NHANVIEN} \leftarrow \Pi_{\text{MANV}, \text{TENNV}, \text{LUONG} * 1.5} (\delta_{\text{MANV} = '001'} (\text{NHANVIEN})) \cup \Pi_{\text{MANV}, \text{TENNV}, \text{LUONG}} (\delta_{\text{MANV} \neq '001'} (\text{NHANVIEN}))$$

# Bài tập

- Xóa các nhân viên có số thân nhân  $\geq 3$
- Tăng 10% lương đối với các nhân viên nữ
- Tăng 10% lương đối với các nhân viên nam được phân công làm việc cho đề án ở Vũng Tàu hay Nha Trang

# CƠ SỞ DỮ LIỆU

ThS.Lê Văn Hạnh

# NGÔN NGỮ TRUY VẤN

(SQL – Structured Query Language)



# Nội dung

---

- **Giới thiệu**
- Ngôn ngữ định nghĩa dữ liệu  
(DDL – Data Definition Language)
- Ngôn ngữ thao tác dữ liệu  
(DML – Data Manipulation Language)
  - Tương quan giữa SQL và ngôn ngữ ĐSQH
- Ngôn ngữ điều khiển dữ liệu  
(DCL – Data Control Language)



## \$4.1. Giới thiệu

---

- Là ngôn ngữ chuẩn để truy vấn và thao tác trên CSDL quan hệ
- Là ngôn ngữ phi thủ tục
- Khởi nguồn của SQL là SEQUEL - Structured English Query Language, năm 1974)
- Các chuẩn SQL
  - SQL89
  - SQL92 (SQL2)
  - SQL99 (SQL3)



## \$4.1. Giới thiệu

---

- Đây là ngôn ngữ truy vấn dựa trên đại số quan hệ và là ngôn ngữ phi thủ tục. Các công dụng chính của SQL bao gồm:
  - Định nghĩa dữ liệu (data definition).
  - Truy vấn và Xử lý dữ liệu (data manipulation).
  - Điều khiển dữ liệu (data control)



## \$4.1. Giới thiệu

---

- Quy định về cú pháp của SQL:
  - Có thể được viết trên nhiều dòng (cuối mỗi lệnh nên có lệnh GO)
  - Từ khóa, tên hàm, tên thuộc tính, tên bảng, tên đối tượng (*Objects*) thì không được phép viết tách xuống hàng.
  - Không phân biệt chữ hoa, chữ thường trong câu lệnh.





# Nội dung

---

- Giới thiệu
- **Ngôn ngữ định nghĩa dữ liệu  
(DDL – Data Definition Language)**
- Ngôn ngữ thao tác dữ liệu  
(DML – Data Manipulation Language)
  - Tương quan giữa SQL và ngôn ngữ ĐSQH
- Ngôn ngữ điều khiển dữ liệu  
(DCL – Data Control Language)



## \$4.2. Ngôn ngữ định nghĩa dữ liệu

---

- **Lệnh tạo mới – CREATE**
  - Database – Table – View - ...
- **Lệnh sửa cấu trúc – ALTER**
  - Database – Table – View - ...
- **Lệnh xóa – DROP**
  - Database – Table – View - ...
- **Lệnh thêm – xóa – sửa dữ liệu**
  - INSERT – DELETE - UPDATE

# Lệnh tạo mới - CREATE

```
CREATE DATABASE < tên CSDL >
```

```
ON (
```

```
    NAME = ' Tên dữ liệu ,
```

```
    FILENAME = ' Tên file dữ liệu và đường dẫn ,
```

```
    SIZE = < kích thước ban đầu > ,
```

```
    MAXSIZE = < kích thước tối đa > ,
```

```
    FILEGROWTH = < kích thước tăng trưởng > )
```

```
LOG ON (
```

```
    NAME = ' Tên log ,
```

```
    FILENAME = ' Tên file nhật ký và đường dẫn ,
```

```
    SIZE = < kích thước ban đầu > ,
```

```
    MAXSIZE = < kích thước tối đa > ,
```

```
    FILEGROWTH = < kích thước tăng trưởng > )
```

# Lệnh tạo mới - CREATE

```
CREATE DATABASE QLDA
ON (
    NAME = 'QLDA_Data',
    FILENAME = 'D:|QLDA_Data.mdf',
    SIZE = 10MB,
    MAXSIZE = UNLIMITED,
    FILEGROWTH = 5MB)
LOG ON (
    NAME = 'QLDA_Log',
    FILENAME = 'D:|QLDA_Log.ldf',
    SIZE = 5MB,
    MAXSIZE = UNLIMITED,
    FILEGROWTH = 2MB)
```

# Lệnh tạo mới - CREATE

CREATE TABLE *< tên table >*

(

*< tên thuộc tính 1 >* kiểu\_dữ\_liệu [not null],

*< tên thuộc tính 2 >* kiểu\_dữ\_liệu [not null],

...

*< tên thuộc tính n >* kiểu\_dữ\_liệu [not null]

)

# Lệnh tạo mới - CREATE

- Các kiểu dữ liệu trong SQL Server

Các kiểu dữ liệu	SQL Server
Chuỗi ký tự	varchar(n), char(n), nvarchar(n), nchar(n) Text
Số	tinyint, smallint, int numeric, decimal, float, real
Ngày tháng	smalldatetime, datetime

# Lệnh tạo mới - CREATE

```
CREATE TABLE NhanVien
```

```
(
```

```
    MaNV      nvarchar(5) not null,
```

```
    HoNV      nvarchar(10) not null,
```

```
    TenLot    nvarchar(40) not null,
```

```
    TenNV     nvarchar(10) not null,
```

```
    PhaiNu    bit,
```

```
    NgaySinh  datetime not null,
```

```
    DiaChi    nvarchar(100),
```

```
    Luong     int,
```

```
    MaPB      nvarchar(2)
```

```
)
```

# Lệnh tạo mới - CREATE

- Cho lược đồ CSDL “quản lý đề án cty” như sau
  - **NHANVIEN** (MaNV, HoNV, TenLot, TenNV, Phai, Luong, Phong, NgaySinh, DiaChi, Ma\_NQL)
  - **PHONGBAN** (MaPHG, TenPHG, TrPHG, NG\_NhanChuc)
  - **DEAN** (MaDA, TenDA, DDIEM\_DA, Phong)
  - **PHANCONG** (MaNV, MaDA, ThoiGian)
  - **DIADIEM\_PHG** (MaPHG, DIADIEM)
  - **THANNHAN** (MaNV, TenTN, Phai, NGSinh, QuanHe)
- Viết lệnh tạo CSDL và các table trên



# Lệnh sửa cấu trúc - ALTER

- Thêm thuộc tính
- ALTER TABLE <*tên table*> ADD
  - <*tên thuộc tính 1*> kiểu\_dữ\_liệu,
  - <*tên thuộc tính 2*> kiểu\_dữ\_liệu,
  - ...
  - <*tên thuộc tính n*> kiểu\_dữ\_liệu
- Ví dụ: thêm cột Ghi\_chú vào bảng nhân viên  
ALTER TABLE NhanVien ADD GhiChu  
nvarchar(20)

# Lệnh sửa cấu trúc - ALTER

- Sửa kiểu dữ liệu thuộc tính
- ALTER TABLE <*tên table*> ALTER COLUMN
  - <*tên thuộc tính 1*> kiểu\_dữ\_liệu\_mới,
  - <*tên thuộc tính 2*> kiểu\_dữ\_liệu\_mới,
  - ...
  - <*tên thuộc tính n*> kiểu\_dữ\_liệu\_mới
- Ví dụ: sửa ngày sinh của nhân viên
  - ALTER TABLE NhanVien ALTER COLUMN NgaySinh SmallDateTime

# Lệnh sửa cấu trúc - ALTER

- Xóa thuộc tính
- ALTER TABLE *<tên table>* DROP COLUMN *<tên thuộc tính>*
- Ví dụ: xóa cột Ghi\_chú từ bảng nhân viên
  - ALTER TABLE NhanVien DROP COLUMN GhiChu

# Lệnh sửa cấu trúc - ALTER

## ■ Thêm ràng buộc toàn vẹn

```
ALTER TABLE <tên table>  
ADD CONSTRAINT <tên  
ràng buộc>
```

**UNIQUE** tên thuộc tính

**PRIMARY KEY** tên t\_tính

**FOREIGN KEY** tên t\_tính  
**REFERENCES** tên\_table  
(t\_tính khóa chính) [ON  
DELETE CASCADE] [ON  
UPDATE CASCADE]

**CHECK** (tên t\_tính  
điều\_kiện)

# Lệnh sửa cấu trúc - ALTER

- Tạo khóa chính cho table nhân viên
  - ALTER TABLE NhanVien ADD CONSTRAINT pk\_NV PRIMARY KEY (MaNV)
- Tạo khóa ngoại: nhân viên trực thuộc phòng ban
  - ALTER TABLE NhanVien ADD CONSTRAINT fk\_NV\_PB FOREIGN KEY (MaPB) REFERENCES PhongBan(MaPB)

# Lệnh sửa cấu trúc - ALTER

- Tạo khóa ngoại: phòng ban có một trường phòng
  - ALTER TABLE PhongBan ADD CONSTRAINT fk\_PB\_NV FOREIGN KEY (TrPhg) REFERENCES NhanVien(MaNv)
- Tạo ràng buộc lương phải > 0
  - ALTER TABLE NhanVien ADD CONSTRAINT chk\_Luong CHECK (Luong>0)

# Lệnh sửa cấu trúc - ALTER

- Tạo khóa ngoại: một đề án do một phòng ban quản lý
  - ALTER TABLE DeAn ADD CONSTRAINT fk\_DA\_PB FOREIGN KEY (MaPB) REFERENCES PhongBan(MaPB)
- Xóa khóa ngoại: một đề án do một phòng ban quản lý
  - ALTER TABLE DeAn DROP CONSTRAINT fk\_DA\_PB

# Lệnh xóa - DROP

- Cú pháp
  - DROP <DATABASE | TABLE | VIEW | ...> <tên>
  - Khi xóa table, phải xóa các table thuộc đầu nhiều trước hoặc phải xóa các khóa ngoại tham chiếu đến table cần xóa
- Ví dụ: xóa table nhân viên
  - Alter Table PhongBan Drop Constraint fk\_PB\_NV
  - Alter Table PhanCong Drop Constraint fk\_PC\_NV
  - Drop Table NhanVien





# Nội dung

---

- Giới thiệu
- Ngôn ngữ định nghĩa dữ liệu  
(DDL – Data Definition Language)
- **Ngôn ngữ thao tác dữ liệu  
(DML – Data Manipulation Language)**
  - Tương quan giữa SQL và ngôn ngữ ĐSQH
- Ngôn ngữ điều khiển dữ liệu  
(DCL – Data Control Language)

- Gồm các lệnh

- Thêm mới dữ liệu: INSERT INTO
- Xóa dữ liệu đang có: DELETE FROM
- Sửa dữ liệu: UPDATE
- Và lệnh truy vấn dữ liệu SELECT
  - Các lệnh truy vấn dùng để khai thác, rút trích dữ liệu của một CSDL.
  - Kết quả của 1 lệnh truy vấn là 1 bảng mới có thể được kết xuất ra màn hình, máy in hoặc lưu trữ trên các thiết bị trữ tin

# Lệnh thêm dữ liệu vào table

- Cú pháp
  - INSERT INTO *tên\_table* VALUES (*giá\_trị\_1*, *giá\_trị\_2*, ..., *giá\_trị\_n*)
  - INSERT INTO *tên\_bảng* (*t\_tính\_1*, *t\_tính\_2*) VALUES (*giá\_trị\_1*, *giá\_trị\_2*)
- Ví dụ: thêm một nhân viên mới
  - INSERT INTO NhanVien VALUES ('001', **N**'Vương Ngọc Quyên', '01/01/1977', **N**'450 Trần Hưng Đạo', **N**'TpHCM', 'QL')

# Lệnh sửa dữ liệu trong table

- Cú pháp
  - UPDATE *tên\_table* SET
    - t\_tính\_1 = giá\_trị\_1,
    - t\_tính\_2 = giá\_trị\_2
  - [WHERE điều\_kiện]
- Ví dụ: Sửa họ nhân viên có mã số '001' thành 'Nguyễn'
  - UPDATE NhanVien SET
    - HoNV = **N**'Nguyễn'
  - WHERE MaNV='001'

# Lệnh sửa dữ liệu trong table

- Ví dụ: Sửa họ tên của nhân viên có mã số '001' thành 'Nguyễn Thanh Tùng' và ngày sinh mới là 1/1/1978
  - UPDATE NhanVien SET
    - HoNV = **N**'Nguyen',
    - TenLot = **N**'Thanh',
    - TenNV = **N**'Tùng',
    - NgaySinh = '1/1/1978'
  - WHERE MaNV='001'

# Lệnh xóa dữ liệu trong table

- Cú pháp:
  - DELETE FROM *tên\_table*
  - [WHERE điều\_kiện]
- Ví dụ: xóa nhân viên có mã số '001'
  - DELETE FROM NhanVien
  - WHERE MaNV='001'
- Ví dụ: xóa toàn bộ nhân viên
  - DELETE FROM NhanVien

# Lệnh truy vấn dữ liệu

- Cú pháp:

**SELECT** [ALL|DISTINCT|TOP n [PERCENT]]

\*|<thuộc tính>|<biểu thức> | <hàm tập hợp>

[**INTO** <tên table mới>]

**FROM** <table>

[**WHERE** <điều kiện trên record>]

[**GROUP BY** <thuộc tính>]

[**HAVING** <điều kiện trên nhóm>]

[**ORDER BY** <thuộc tính> [ASC | DESC], . . . ]

# Lệnh truy vấn dữ liệu

- Tương quan giữa SQL và ĐSQH
  - Phép chọn:  $\delta_{\langle \text{điều kiện} \rangle}(\langle \text{tên quan hệ} \rangle)$  tương ứng mệnh đề WHERE trong SQL
  - Phép chiếu:  $\Pi_{\langle \text{thuộc tính} \rangle}(\langle \text{tên quan hệ} \rangle)$  tương ứng mệnh đề SELECT trong SQL
  - Các phép toán tập hợp: Hợp tương ứng mệnh đề UNION trong SQL
  - Tích Descartes ( $R \times S$ ) tương ứng mệnh đề FROM trong SQL
  - Kết ( $\theta$ ): là các điều kiện trong mệnh đề WHERE
    - Equi-join: Tất cả đều là phép so sánh bằng
    - Inner join, left join, right join: cùng tên



# Lệnh truy vấn dữ liệu

- Các toán tử sử dụng trong truy vấn
  - Toán tử so sánh: =, >, <, >=, <=, <>
  - Toán tử logic: AND, OR, NOT
  - Phép toán số học: +, -, \*, /
  - Khác:
    - BETWEEN
    - IS NULL
    - LIKE
    - IN
    - EXISTS

# Lệnh truy vấn dữ liệu

## ■ Toán tử so sánh

- Cho biết những nhân viên có lương trên 2000 (MaNV, TenNV)

- SELECT MaNV, TenNV
- FROM NhanVien
- WHERE Luong > 2000

- Cho biết tên nhân viên và tên phòng ban họ trực thuộc

- SELECT TenNV, TenPB
- FROM NhanVien nv, PhongBan pb
- WHERE nv.MaPB=pb.MaPB

# Lệnh truy vấn dữ liệu

## ■ Toán tử so sánh

- Cho biết tên nhân viên trưởng phòng và tên phòng ban họ quản lý

- SELECT      `Tên trưởng phòng`=TenNV, TenPB
- FROM        NhanVien nv, PhongBan pb
- WHERE       nv.MaNV=pb.TrPhg

- Cho biết tên trưởng phòng có lương > 2000

- SELECT      `Tên trưởng phòng`=TenNV
- FROM        NhanVien nv, PhongBan pb
- WHERE       nv.MaNV=pb.TrPhg AND  
                nv.Luong > 2000

# Lệnh truy vấn dữ liệu

## ■ Toán tử BETWEEN

- Cho biết tên và ngày sinh các nhân viên có năm sinh từ 1965 đến 1977
  - SELECT TenNV, NgaySinh
  - FROM NhanVien
  - WHERE (year(NgaySinh) BETWEEN 1965 AND 1977)
- Cho biết tên và ngày sinh các nhân viên sinh từ ngày 10 đến 20 tháng 8
  - SELECT TenNV, NgaySinh
  - FROM NhanVien
  - WHERE (day(NgaySinh) BETWEEN 10 AND 20) AND month(NgaySinh)=8

# Lệnh truy vấn dữ liệu

## ■ Toán tử LIKE

- So sánh chuỗi tương đối
- Cú pháp: `s LIKE 'p'`, `p` có thể chứa `%` hoặc `_`
  - `%` : thay thế bất kỳ chuỗi ký tự nào có chiều dài  $\geq 0$
  - `_` : thay thế **một** ký tự bất kỳ
- Ví dụ: cho biết các nhân viên có họ bắt đầu bằng chữ Ng
  - `SELECT MaNV, HoNV, TenLot, TenNV`
  - `FROM NhanVien`
  - `WHERE HoNV LIKE 'Ng%'`

# Lệnh truy vấn dữ liệu

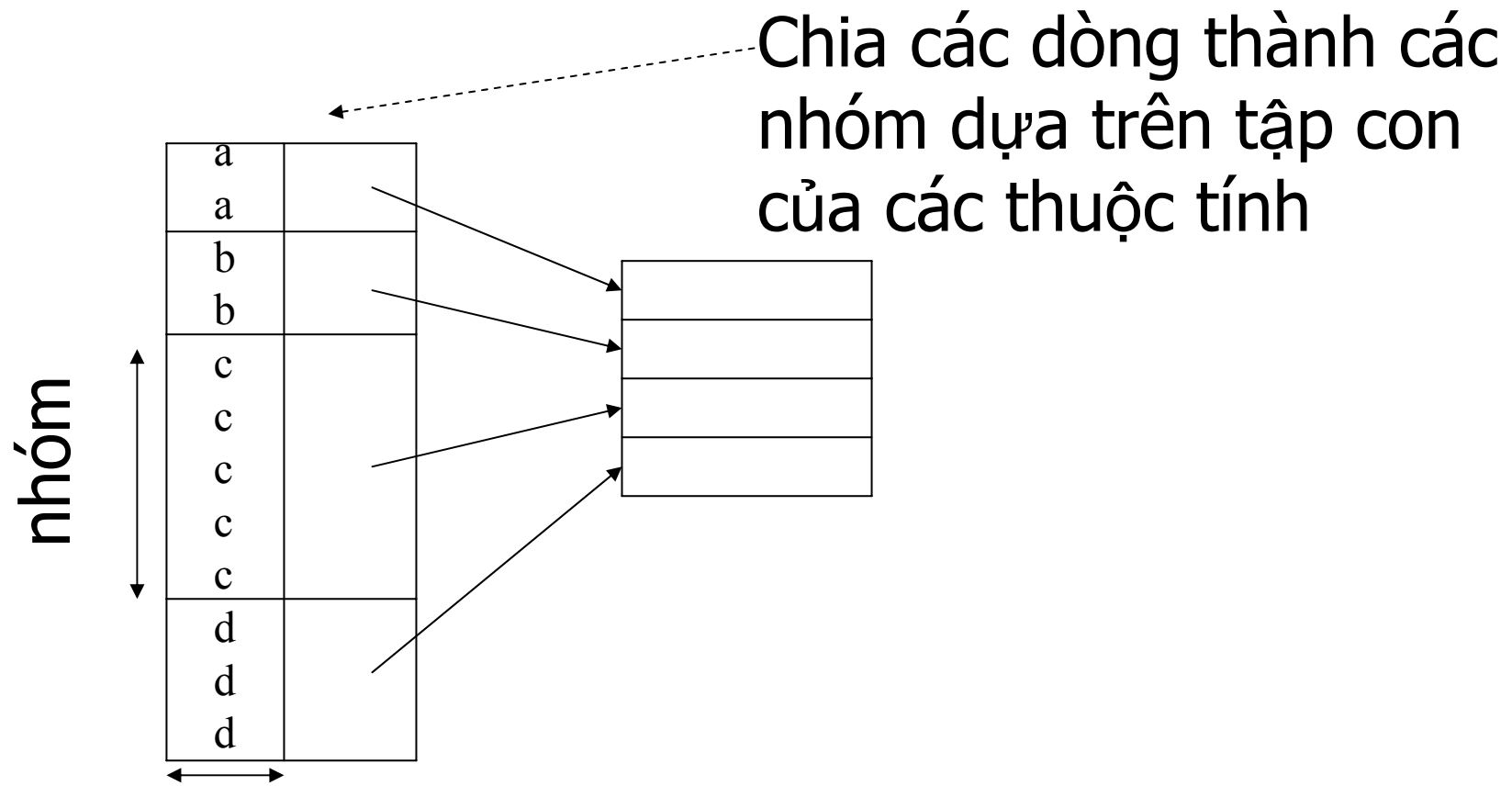
## ■ Toán tử IN

- Cho biết tên và ngày sinh các nhân viên sinh vào quý 3
  - SELECT TenNV, NgaySinh
  - FROM NhanVien
  - WHERE month(NgaySinh) in (7,8,9)
- Cho biết tên các nhân viên chưa hề được phân công đề án nào
  - SELECT TenNV
  - FROM NhanVien
  - WHERE MaNV NOT IN (
    - SELECT MaNV
    - FROM PhanCong )

# Lệnh truy vấn dữ liệu

- Mệnh đề GROUP BY
  - Gom nhiều dòng thành các nhóm dựa trên tập con của các thuộc tính
  - Tất cả các thành viên của nhóm đều thỏa các thuộc tính này.
  - Mỗi nhóm được mô tả bằng một dòng các thuộc tính, được giới hạn bởi:
    - Các thuộc tính chung của tất cả thành viên thuộc nhóm (được liệt kê trong mệnh đề GROUP BY).
    - Các phép toán trên nhóm.

# Lệnh truy vấn dữ liệu



Các thuộc tính GROUP BY



# Lệnh truy vấn dữ liệu

- Mệnh đề GROUP BY
  - Các hàm SQL cơ bản
    - COUNT: Đếm số bộ dữ liệu trong nhóm
    - MIN: Tính giá trị nhỏ nhất của thuộc tính trong nhóm
    - MAX: Tính giá trị lớn nhất của thuộc tính trong nhóm
    - AVG: Tính giá trị trung bình của thuộc tính trong nhóm
    - SUM: Tính tổng giá trị các bộ dữ liệu trong nhóm

# Lệnh truy vấn dữ liệu

- Mệnh đề GROUP BY: cho biết tổng lương từng phòng ban
  - SELECT pb.TenPB, 'TL'=SUM(nv.Luong)
  - FROM NhanVien nv, PhongBan pb
  - WHERE nv.MaPB = pb.MaPB
  - GROUP BY pb.TenPB

# Lệnh truy vấn dữ liệu

NHANVIEN			
MAN	TENN	MAP	LUONG
001 <sup>V</sup>	Quy <sup>V</sup> en	QL <sup>B</sup>	3.000.00
002	Tung	NC	2.500.00
003	Nhan	DH	2.500.00
004	Tien	NC	2.200.00
005	Vu	DH	2.200.00
006	Hung	NC	2.000.00
007	Tam	NC	2.200.00
008	Van	NC	1.800.00

nv.MaPB=pb.MaPB

PHONGBAN			
MAPH	TENPHG	TRPH	NGNC
QL <sup>G</sup>	Quan Ly	001 <sup>G</sup>	22/05/200
DH	Dieu Hanh	003	10/10/200
NC	Nghien	002	15/03/200

Cuu

2

TenPB	TenNV	Luong
Quan Ly	Quy <sup>V</sup> en	3000000
Nghien Cuu	Tung	2500000
Dieu Hanh	Nhan	2500000
Nghien Cuu	Tien	2200000
Dieu Hanh	Vu	2200000
Nghien Cuu	Hung	2000000
Nghien Cuu	Tam	2200000
Nghien Cuu	Van	1800000

# Lệnh truy vấn dữ liệu

TenPB	TenNV	Luong
Quan Ly	Quyên	3000000
Nghiên Cứu	Tung	2500000
Điều Hành	Nhan	2500000
Nghiên Cứu	Tien	2200000
Điều Hành	Vu	2200000
Nghiên Cứu	Hung	2000000
Nghiên Cứu	Tam	2200000
Nghiên Cứu	Van	1800000

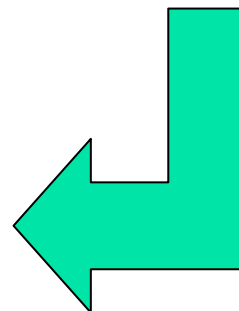
**Gom nhóm**



TenPB	TenN	Luong
Quan Ly	Quyên	3000000
Điều Hành	Nhan	2200000
Điều Hành	Vu	2500000
Nghiên Cứu	Tung	2200000
Nghiên Cứu	Tien	2000000
Nghiên Cứu	Hung	2200000
Nghiên Cứu	Tam	1800000
Nghiên Cứu	Van	0

TenPB	TL
Quan Ly	3000000
Điều Hành	4700000
Nghiên Cứu	0

**Tính toán trên từng nhóm**



# Lệnh truy vấn dữ liệu

---

- Mệnh đề HAVING
  - Lọc kết quả theo điều kiện, sau khi đã gom nhóm
  - Điều kiện của HAVING là các thuộc tính trong danh sách GROUP BY và các phép tính toán khác.

# Lệnh truy vấn dữ liệu

- Mệnh đề GROUP BY: cho biết các phòng ban có tổng lương > 3 triệu
  - SELECT pb.TenPB, 'TL'=SUM(nv.Luong)
  - FROM NhanVien nv, PhongBan pb
  - WHERE nv.MaPB = pb.MaPB
  - GROUP BY pb.TenPB
  - **HAVING SUM(nv.Luong) > 3000000**

# Lệnh truy vấn dữ liệu

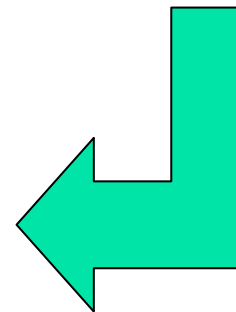
TenN		Luong
TenPB	V	
Quan Ly	Quyên	250000
Dieu Hanh	Nhan	220000
Dieu Hanh	Vu	250000
Nguyen	Tung	220000
Nguyen	Tien	200000
Nguyen	Hung	220000
Nguyen	Tam	180000
Cuu	Van	0

**Tính toán trên từng nhóm**



TenPB	TL
Quan Ly	3000000
Dieu Hanh	4700000
Nguyen	4700000
Cuu	0

TenPB	TL
Dieu Hanh	4700000
Nguyen	4700000
Cuu	0



**Lọc bỏ các nhóm không thỏa**

# Lệnh truy vấn dữ liệu

- Truy vấn lồng
  - Các mệnh đề SELECT được lồng vào nhau
  - Kết quả của câu SELECT bên trong sẽ được chuyển cho truy vấn bên ngoài.
  - Truy vấn ngoài cùng sẽ cho kết quả cuối cùng
  - Các mệnh đề SELECT được nối với nhau bằng các phép so sánh  $<$ ,  $>$ ,  $<>$ ,  $=$ ,  $>=$ ,  $<=$ , IN, EXISTS, ALL, ANY, SOME,...



# Lệnh truy vấn dữ liệu

## ■ Truy vấn lồng

- Mã số và tên các nhân viên có số giờ làm việc nhiều nhất
  - SELECT      nv.MaNV, nv.TenNV, SUM(pc.ThoiGian)
  - FROM        NhanVien nv, PhanCong pc
  - WHERE       nv.MaNV = pc.MaNV
  - GROUP BY   nv.MaNV, nv.TenNV
  - HAVING      SUM(pc.ThoiGian) >= ALL (  
              SELECT        SUM(ThoiGian)  
              FROM        PhanCong  
              GROUP BY    MaNV )

# Lệnh truy vấn dữ liệu

```
SELECT nv.TenNV, SUM(pc.ThoiGian)
FROM NhanVien nv, PhanCong pc
WHERE nv.MaNV = pc.MaNV
GROUP BY nv.TenNV
```

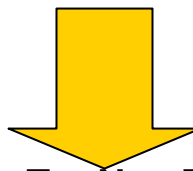
MaN	TenN	Tong
V	Quy	TG
001	n	42.5
002	Tung	40
003	Nhan	37.5
004	Tien	32.5
006	Hung	30.5
007	Tam	30
008	Van	22.5

```
SELECT
FROM
GROUP BY
```

```
SUM(ThoiGian)
PhanCong
MaNV
```

MaN	Tong
V	TG
001	42.5
002	40
003	37.5
004	32.5
006	30.5
007	30
008	22.5

HAVING SUM(pc.ThoiGian) >= ALL



MaN	TenN	Tong
V	Quy	TG
001	n	42.5



# Nhập Môn Cơ Sở Dữ Liệu

---

Minh họa ĐSQH & SQL



1. Cho biết tên, địa chỉ, điện thoại của tất cả các sinh viên

---

■  $KQ \leftarrow \Pi_{\text{Tên, ĐịaChị, ĐiệnThoai}} (\text{SinhVien})$

```
SELECT    Tên, ĐịaChị, ĐiệnThoai
FROM      SinhVien
```



## 2. Cho biết tên và số tín chỉ của các môn học

---

- $KQ \leftarrow \Pi_{\text{TenMH}, \text{SoTC}} (\text{MonHoc})$

```
SELECT    TenMH, SoTC  
FROM      MonHoc
```




### 3. Cho biết kết quả học tập của sinh viên có mã số SV03

---

- $KQ \leftarrow \Pi_{\text{MaKhoaHoc, Diem}} (\delta_{\text{MSSV}='SV03'}(\text{KetQua}))$

```
SELECT    MaKhoaHoc, Diem
FROM      KetQua
WHERE     MSSV='SV03'
```



4. Cho biết tên và số tín chỉ của các môn học có mã có dạng: ký tự thứ 1 là C, ký tự thứ 3 là D

- $KQ \leftarrow \Pi_{TenMH, SoTC} (\delta_{MaMH \text{ LIKE 'C?D*'}} (MonHoc))$

```
SELECT    TenMH, SoTC
FROM      MonHoc
WHERE     MaMH LIKE 'C?D*'
```



## 9. Cho biết danh sách các môn học được dạy trong năm 2002

---

- $KQ \leftarrow \Pi_{MaMH} (\delta_{Nam=2002} (GiangDay))$

```
SELECT    MaMH
FROM      GiangDay
WHERE     Nam=2002
```





## 10. Cho biết mã, tên, địa chỉ của các sv theo từng khoa, sắp A-Z theo tên sv

---

- $KQ \leftarrow \Pi_{\text{MaKhoa, MSSV, Ten, DiaChi}} (\text{SinhVien})$

```
SELECT    MaKhoa, MSSV, Ten, DiaChi
FROM      SinhVien
ORDER BY  MaKhoa ASC, Ten ASC
```



## 11. Cho biết điểm của các sv theo từng môn học

---

- $KQ \leftarrow \Pi_{MSSV, MaMH, Diem} (KetQua \bowtie GiangDay)$

```
SELECT    MSSV, MaMH, Diem
FROM      KetQua kq, GiangDay gd
WHERE     kq.MaKhoaHoc = gd.MaKhoaHoc
ORDER BY MaMH
```

## 12. Cho biết các sv học môn CSDL có điểm từ 8 đến 10

- $R1 \leftarrow \delta_{\text{MaMH}='CSDL'} (\text{GiangDay})$
- $R2 \leftarrow R1 \bowtie_{\text{MaKhoaHoc}} \text{KetQua}$
- $KQ \leftarrow \Pi_{\text{MSSV}} (\delta_{8 \leq \text{Diem} \leq 10} (R2))$

```
Create View v12CSDL AS
SELECT      *
FROM        GiangDay
WHERE       MaMH='CSDL'
```

```
SELECT      MSSV
FROM        KetQua kq, v12CSDL v
WHERE       kq.MaKhoaHoc=v.MaKhoaHoc and
            (Diem between 8 and 10)
```

## 12. Cho biết các sv học môn CSDL có điểm từ 8 đến 10

- $R1 \leftarrow \delta_{\text{MaMH}='CSDL'} (\text{GiangDay})$
- $R2 \leftarrow R1 \bowtie_{\text{MaKhoaHoc}} \text{KetQua}$
- $KQ \leftarrow \Pi_{\text{MSSV}} (\delta_{8 \leq \text{Diem} \leq 10} (R2))$

```
SELECT      MSSV
FROM        KetQua kq, GiangDay gd
WHERE       kq.MaKhoaHoc=gd.MaKhoaHoc and
            MaMH='CSDL' and
            (Diem between 8 and 10)
```

## 14. Cho biết tên khoa, tên môn học mà các sv trong khoa đã học

- $R1 \leftarrow \Pi_{\text{TenKhoa}, \text{MSSV}} (\text{Khoa} \bowtie_{\text{MaKhoa}} \text{SinhVien})$
- $R2 \leftarrow \Pi_{\text{TenKhoa}, \text{MaKhoaHoc}} (R1 \bowtie_{\text{MSSV}} \text{KetQua})$
- $R3 \leftarrow \Pi_{\text{TenKhoa}, \text{MaMH}} (R2 \bowtie_{\text{MaKhoaHoc}} \text{GiangDay})$
- $KQ \leftarrow \Pi_{\text{TenKhoa}, \text{TenMH}} (R3 \bowtie_{\text{MaMH}} \text{MonHoc})$

```
SELECT      TenKhoa, TenMH
FROM        Khoa k, SinhVien sv, KetQua kq,
            GiangDay gd, MonHoc mh
WHERE       k.MaKhoa=sv.MaKhoa and
            sv.MSSV=kq.MSSV and
            kq.MaKhoaHoc=gd.MaKhoaHoc and
            gd.MaMH=mh.MaMH
```



## 19. Cho biết mã, tên các sv có kết quả một môn học nào đó trên 8

---

- $R1 \leftarrow \Pi_{MSSV} (\delta_{Diem > 8} (KetQua))$
- $KQ \leftarrow \Pi_{MSSV, Ten} (R1 \bowtie_{MSSV} SinhVien)$

```
SELECT    sv.MSSV, Ten
FROM      KetQua kq, SinhVien sv
WHERE     kq.MSSV=sv.MSSV and
          Diem > 8
```



## 20. Cho biết mã, tên các sv có kết quả các môn học đều trên 8

---

- $R1 \leftarrow \Pi_{MSSV} (\delta_{Diem \leq 8} (KetQua))$
- $R2 \leftarrow (\Pi_{MSSV} (SinhVien)) - R1$
- $KQ \leftarrow \Pi_{MSSV, Ten} (R2 \bowtie_{MSSV} SinhVien)$

```
SELECT      MSSV, Ten
FROM        SinhVien
WHERE       MSSV NOT IN (
            SELECT MSSV
            FROM      KetQua
            WHERE     Diem <= 8
            )
```



## 21. Cho biết có bao nhiêu sv

---

- $KQ(SLSV) \leftarrow \mathcal{F}_{COUNT(MSSV)}(SinhVien)$

```
SELECT 'SLSV'=COUNT(MSSV)
FROM   SinhVien
```





## 23. Cho biết có bao nhiêu sv nữ khoa CNTT

---

- $R1 \leftarrow \delta_{\text{PhaiNu}='Yes' \wedge \text{MaKhoa}='CNTT'} (\text{SinhVien})$
- $KQ(\text{SLSV}) \leftarrow \mathcal{F}_{\text{COUNT}(\text{MSSV})} (R1)$

```
SELECT    'SLSV'=COUNT(MSSV)
FROM      SinhVien
WHERE     PhaiNu='Yes' and
          MaKhoa='CNTT'
```



## 28. Cho biết mã, tên, địa chỉ và điểm trung bình của từng sv

---

- $R1(MSSV, DTB) \leftarrow MSSV \mathcal{F}_{AVG(Diem)} (KetQua)$
- $KQ \leftarrow \Pi_{MSSV, Ten, DiaChi, DTB} (R1 \bowtie_{MSSV} SinhVien)$

```
SELECT    sv.MSSV, Ten, DiaChi,
          `DTB`=AVG(Diem)
FROM      KetQua kq, SinhVien sv
WHERE     kq.MSSV=sv.MSSV
GROUP BY  sv.MSSV, Ten, DiaChi
```

## 29. Cho biết số lượng điểm $\geq 8$ của từng sv

- $R1 \leftarrow \delta_{\text{Diem} \geq 8} (\text{KetQua})$
- $R2(\text{MSSV}, \text{SLDT8}) \leftarrow \text{MSSV} \overset{\curvearrowright}{\text{COUNT}}(\text{Diem}) (R1)$
- $KQ \leftarrow \Pi_{\text{MSSV}, \text{SLDT8}} (\text{SinhVien} \bowtie_{\text{MSSV}} R2)$

Create View v29tam AS

```
SELECT      MSSV, 'SLDT8'=COUNT(Diem)
FROM        KetQua
WHERE       Diem >= 8
GROUP BY    MSSV
```

```
SELECT SinhVien.MSSV, Ten, SLDT8
FROM   SinhVien LEFT JOIN v29tam on
       SinhVien.MSSV=v29tam.MSSV
```

### 33. Cho biết số lượng tín chỉ của từng sv

- $R1 \leftarrow \Pi_{MSSV, MaMH} (KetQua \bowtie_{MaKhoaHoc} GiangDay)$
- $R2 \leftarrow \Pi_{MSSV, SoTC} (R1 \bowtie_{MaMH} MonHoc)$
- $R3(MSSV, TSTC) \leftarrow MSSV \mathfrak{S}_{SUM(SoTC)}(R2)$
- $KQ \leftarrow \Pi_{MSSV, Ten, TSTC} (R3 \bowtie_{MSSV} SinhVien)$

```
SELECT      sv.MSSV, Ten, 'TSTC'=SUM(SoTC)
FROM        SinhVien sv, KetQua kq, GiangDay gd,
           MonHoc mh
WHERE       sv.MSSV=kq.MSSV and
           kq.MaKhoaHoc=gd.MaKhoaHoc and
           gd.MaMH=mh.MaMH
GROUP BY   sv.MSSV, Ten
```

## 34. Cho biết tên những sv chỉ mới thi đúng 1 môn học

- $R1 \leftarrow \Pi_{MSSV, MaMH} (KetQua \bowtie_{MaKhoaHoc} GiangDay)$
- $R2(MSSV, SLMH) \leftarrow MSSV \mathcal{J}_{COUNT(MaMH)}(R1)$
- $KQ \leftarrow \Pi_{Ten} (\delta_{SLMH=1} (R2 \bowtie_{MSSV} SinhVien))$

```
SELECT    Ten
FROM      SinhVien sv, v34tam v
WHERE     sv.MSSV=v.MSSV
GROUP BY  Ten
HAVING    COUNT(MaMH)=1
```

Create View v34tam AS

```
SELECT Distinct MSSV, MaMH
FROM    KetQua kq, GiangDay gd
WHERE   kq.MaKhoaHoc = gd.MaKhoaHoc
```

## 40. Cho biết tên môn học có số tín chỉ nhiều nhất

- $R1(\text{SoTCNN}) \leftarrow \mathcal{F}_{\text{MAX}(\text{SoTC})}(\text{MonHoc})$
- $KQ \leftarrow \Pi_{\text{TenMH}}(R1 \bowtie_{\text{SoTCNN}=\text{SoTC}} \text{MonHoc})$

```
SELECT    MaMH, TenMH, SoTC
FROM      MonHoc
WHERE     SoTC = ( SELECT    MAX(SoTC)
                  FROM      MonHoc
                )
```

## 43. Cho biết mã, tên, địa chỉ của các sv có điểm CSDL lớn nhất

- $R1 \leftarrow \delta_{MaMH='CSDL'} (KetQua \bowtie_{MaKhoaHoc} GiangDay)$
- $R2 \leftarrow \Pi_{MSSV, Diem} (R1)$
- $R3(DCSDLLN) \leftarrow \mathcal{J}_{MAX(Diem)} (R2)$
- $R4 \leftarrow \Pi_{MSSV} (R3 \bowtie_{DCSDLLN=Diem} R2)$

■  $KQ \leftarrow \Pi_{MSSV, Ten, DiaChi} (R4 \bowtie_{MSSV} SinhVien)$

SELECT

sv.MSSV, Ten, DiaChi

FROM

KetQua kq, GiangDay gd, SinhVien sv

WHERE

kq.MaKhoaHoc=gd.MaKhoaHoc and

kq.MSSV=sv.MSSV and gd.MaMH='CSDL' and

Diem = (SELECT MAX(Diem)

FROM

KetQua kq, GiangDay gd

WHERE

kq.MaKhoaHoc=gd.MaKhoaHoc and

gd.MaMH='CSDL' )

## 45. Với mỗi môn học, cho biết tên, điểm của các sv có điểm lớn nhất

- $R1 \leftarrow \Pi_{\text{MaMH}, \text{Diem}, \text{MSSV}} (\text{KetQua} \bowtie_{\text{MaKhoaHoc}} \text{GiangDay})$
- $R2(\text{MaMH}, \text{DLN}) \leftarrow \text{MaMH} \mathfrak{S}_{\text{MAX}(\text{Diem})} (R1)$
- $R3 \leftarrow \Pi_{\text{MaMH}, \text{MSSV}, \text{Diem}} (R2 \bowtie_{\text{DLN}=\text{Diem} \wedge R1.\text{MaMH}=R2.\text{MaMH}} R1)$
- $KQ \leftarrow \Pi_{\text{MaMH}, \text{Ten}, \text{Diem}} (R3 \bowtie_{\text{MSSV}} \text{SinhVien})$

Create View v45tam as

```
SELECT MaMH, 'DLN'=MAX(Diem)
FROM KetQua kq, GiangDay gd
WHERE kq.MaKhoaHoc=gd.MaKhoaHoc
GROUP BY MaMH
```

```
SELECT v.MaMH, Ten, Diem
FROM KetQua kq, GiangDay gd,
v45tam v, SinhVien sv
WHERE kq.MaKhoaHoc=gd.MaKhoaHoc and
sv.MSSV=kq.MSSV and
gd.MaMH=v.MaMH and
kq.Diem=v.DLN
```



## 45. Với mỗi môn học, cho biết tên, điểm của các sv có điểm lớn nhất

- $R1 \leftarrow \Pi_{\text{MaMH}, \text{Diem}, \text{MSSV}} (\text{KetQua} \bowtie_{\text{MaKhoaHoc}} \text{GiangDay})$
- $R2(\text{MaMH}, \text{DLN}) \leftarrow \text{MaMH} \mathcal{S}_{\text{MAX}(\text{Diem})} (R1)$
- $R3 \leftarrow \Pi_{\text{MaMH}, \text{MSSV}, \text{Diem}} (R2 \bowtie_{\text{DLN}=\text{Diem} \wedge R1.\text{MaMH}=R2.\text{MaMH}} R1)$
- $KQ \leftarrow \Pi_{\text{MaMH}, \text{Ten}, \text{Diem}} (R3 \bowtie_{\text{MSSV}} \text{SinhVien})$

```
SELECT  gd.MaMH, Ten, Diem
FROM    KetQua kq, GiangDay gd, SinhVien sv
WHERE   kq.MaKhoaHoc=gd.MaKhoaHoc and
        sv.MSSV=kq.MSSV and
        Diem = (
            SELECT  MAX(Diem)
            FROM    KetQua kq, GiangDay gd1
            WHERE   kq.MaKhoaHoc=gd1.MaKhoaHoc and
                    gd1.MaMH=gd.MaMH
        )
```

# 51. Tên các giáo viên không tham gia giảng dạy năm 2001

- $R1 \leftarrow \Pi_{\text{MaGV}} (\delta_{\text{Nam}=2001} \text{GiangDay})$
- $R2 \leftarrow (\Pi_{\text{MaGV}} (\text{GiaoVien})) - R1$
- $KQ \leftarrow \Pi_{\text{TenGV}} (R2 \bowtie_{\text{MaGV}} \text{GiaoVien})$

```
SELECT      TenGV
FROM        GiaoVien
WHERE       MaGV NOT IN (
                SELECT      MaGV
                FROM        GiangDay
                WHERE       Nam=2001
            )
```

## 57. Tên sv, tên môn học mà sv chưa học

- $R1 \leftarrow \Pi_{MSSV, MaMH} (\text{SinhVien X MonHoc})$
- $R2 \leftarrow \Pi_{MSSV, MaMH} (\text{KetQua} \bowtie_{MaKhoaHoc} \text{GiangDay})$
- $R3 \leftarrow (R1 - R2)$
- $KQ \leftarrow \Pi_{Ten, TenGV} (R3 \bowtie_{MSSV} \text{SinhVien}) \bowtie_{MaGV} \text{GiaoVien}$   

```
SELECT Ten, TenMH
FROM SinhVien sv, MonHoc
WHERE MaMH NOT IN (
    SELECT MaMH
    FROM KetQua kq, GiangDay gd
    WHERE kq.MaKhoaHoc=gd.MaKhoaHoc
    and kq.MSSV=sv.MSSV
)
```

# 61. Tên các giáo viên tham gia dạy tất cả môn học

- $R \leftarrow \Pi_{\text{MaGV}, \text{MaMH}} (\text{GiangDay})$
- $S \leftarrow \Pi_{\text{MaMH}} (\text{MonHoc})$
- $R3 \leftarrow R \div S$
- $KQ \leftarrow \Pi_{\text{TenGV}} (R3 \bowtie_{\text{MaGV}} \text{GiaoVien})$

```
SELECT TenGV
FROM GiaoVien gv
WHERE NOT EXISTS (
  SELECT *
  FROM MonHoc mh
  WHERE NOT EXISTS (
    SELECT *
    FROM GiangDay gd
    WHERE gd.MaMH=mh.MaMH and gd.MaGV=gv.MaGV ) )
```

## 66. Tên các sv đã học tất cả môn học

■  $R \leftarrow \Pi_{\text{MSSV}, \text{MaMH}} (\text{GiangDay} \bowtie_{\text{MaKhoaHoc}} \text{KetQua})$

■  $S \leftarrow \Pi_{\text{MaMH}} (\text{MonHoc})$

■  $R3 \leftarrow R \div S$

■  $\text{KQ} \leftarrow \Pi_{\text{Ten}} (R3 \bowtie_{\text{MSSV}} \text{SinhVien})$

SELECT Ten

FROM SinhVien sv

WHERE NOT EXISTS (

SELECT \*

FROM MonHoc mh

WHERE NOT EXISTS (

SELECT \*

FROM KetQua kq, GiangDay gd

WHERE kq.MaKhoaHoc=gd.MaKhoaHoc and

gd.MaMH=mh.MaMH and kq.MSSV=sv.MSSV ) )

# 68. Tên giáo viên đã dạy tất cả môn do giáo viên GV03 dạy

- $R \leftarrow \Pi_{\text{MaGV}, \text{MaMH}} (\text{GiangDay})$
- $S \leftarrow \Pi_{\text{MaMH}} (\delta_{\text{MaGV}='GV03'} \text{GiangDay})$
- $R3 \leftarrow R \div S$

- $KQ \leftarrow \Pi_{\text{TenGV}} (R3 \bowtie_{\text{MaGV}} \text{GiaoVien})$

```
SELECT Ten
FROM SinhVien sv
WHERE NOT EXISTS (
  SELECT *
  FROM GiangDay gd
  WHERE MaGV='GV03' and NOT EXISTS (
    SELECT *
    FROM KetQua kq, GiangDay gd1
    WHERE kq.MaKhoaHoc=gd1.MaKhoaHoc and
          gd1.MaMH=gd.MaMH and kq.MSSV=sv.MSSV ) )
```

# CƠ SỞ DỮ LIỆU

ThS.Lê Văn Hạnh

# RÀNG BUỘC TOÀN VỊ

(Integrity Constraint)



- **RÀNG BUỘC TOÀN VỆN (RBTV) - CÁC YẾU TỐ CỦA RBTV**
- **CÁC LOẠI RÀNG BUỘC TOÀN VỆN**
  - A)-BỐI CẢNH MỘT QUAN HỆ
  - B)-BỐI CẢNH NHIỀU QUAN HỆ
- **BIỂU DIỄN RBTV BẰNG PHỤ THUỘC HÀM**



## \$5.1. RBTV & Các yếu tố

---

- RBTV là những điều kiện bất biến luôn phải được đảm bảo giữa tất cả các bộ của tất cả các quan hệ
- RBTV dùng mô tả các mối liên hệ, ràng buộc giữa các bộ, các thuộc tính
  - R1: Mỗi sinh viên có một mã số riêng biệt, không trùng với bất cứ sinh viên nào khác.
  - R2: Mỗi sinh viên phải đăng ký vào 1 Khoa của Trường.
  - R3: Mỗi sinh viên chỉ được thi tối đa 2 lần cho một môn học.



# \$5.1. RBTV & Các yếu tố

---

- Mỗi RBTV gồm 3 yếu tố:
  - Điều kiện
    - Ngôn ngữ tự nhiên
    - Mã giả
    - Đại số tập hợp, đại số quan hệ
  - Bối cảnh
    - Là những quan hệ được RBTV đề cập đến
  - Tâm ảnh hưởng
    - Những thao tác trên bối cảnh có thể gây vi phạm RBTV

# \$5.1. RBTV & Các yếu tố

## ■ Yếu tố tầm ảnh hưởng của RBTV

Các thao tác Các quan hệ	Các thao tác		
	Thêm	Sửa	Xóa
Q1	+/-/(T)	+/-/(T)	+/-/(T)
Q2	+/-/(T)	+/-/(T)	+/-/(T)
...	...	...	...
Qn	+/-/(T)	+/-/(T)	+/-/(T)

# \$5.1. RBTV & Các yếu tố

- Yếu tố tầm ảnh hưởng của RBTV
  - R1: Mỗi sinh viên có một mã số riêng biệt, không trùng với bất cứ sinh viên nào khác

- $\forall sv1 \in T_{\text{SinhVien}}, \forall sv2 \in T_{\text{SinhVien}}$ 
  - $sv1 \neq sv2 \Rightarrow sv1.MaSV \neq sv2.MaSV$

- Cuối  $\forall$

	Thêm	Sửa	Xóa
$T_{\text{SinhVien}}$	+	-(MaSV)	-

- R1 có bối cảnh một quan hệ

# \$5.1. RBTV & Các yếu tố

- Yếu tố tầm ảnh hưởng của RBTV
  - R2: Mỗi sinh viên phải đăng ký vào 1 Khoa của Trường
    - $\forall sv \in T_{SinhVien}, \exists k \in T_{Khoa}$ 
      - $k.MaKhoa = sv.MaKhoa$

■ Cuối  $\forall$

	Thêm	Sửa	Xóa
$T_{SinhVien}$	+	-(MaKhoa)	-
$T_{Khoa}$	-	-(MaKhoa)	+

- R2 có bối cảnh hai quan hệ

# \$5.1. RBTV & Các yếu tố

- Yếu tố tầm ảnh hưởng của RBTV
  - R3: Mỗi sinh viên chỉ được thi tối đa 2 lần cho một môn học

- $SV \in T_{KetQua}$

- $\exists sv: MaSV, MaMH \exists_{COUNT(Diem)} (KetQua) > 2$

- Cuối  $\forall$

	Thêm	Sửa	Xóa
$T_{KetQua}$	+	-(MaMH, MaSV)	-

- R3 có bối cảnh một quan hệ

- RÀNG BUỘC TOÀN VỆN (RBTV) - CÁC YẾU TỐ CỦA RBTV
- **CÁC LOẠI RÀNG BUỘC TOÀN VỆN**
  - A)-BỐI CẢNH MỘT QUAN HỆ
  - B)-BỐI CẢNH NHIỀU QUAN HỆ
- BIỂU DIỄN RBTV BẰNG PHỤ THUỘC HÀM





## \$5.2. Các loại RBTV

---

- Bối cảnh một quan hệ
  - Miền giá trị
  - Liên thuộc tính
  - Liên bộ
- Bối cảnh nhiều quan hệ
  - Phụ thuộc tồn tại
  - Liên bộ - liên quan hệ
  - Liên thuộc tính – liên quan hệ
  - Thuộc tính tổng hợp
  - Chu trình
- Các bước xác định RBTV

# Bối cảnh một quan hệ

- Miền giá trị:
  - Liên tục (trong 1 hoặc nhiều khoảng)
    - Điểm của sinh viên phải trong  $[0..10]$ 
      - $\forall kq \in T_{KetQua}$ 
        - $kq.Diem$  BETWEEN 0 AND 10
      - Cuối  $\forall$
    - Rời rạc
      - Quan hệ thân nhân phải là vợ, chồng, con
        - $\forall tn \in T_{ThanNhan}$ 
          - $tn.QuanHe$  IN ('Vo','Chong','Con')
        - Cuối  $\forall$
  - Dạng thức định trước

# Bối cảnh một quan hệ

- Liên thuộc tính:
  - Điều kiện ràng buộc giữa các thuộc tính trong cùng một quan hệ
  - Xét lược đồ quan hệ HangHoa(MaHH, TenHH, GiaNhap, GiaBan), ta có RBTV:
    - Giá bán phải lớn hơn giá nhập
      - $\forall hh \in T_{\text{HangHoa}}$ 
        - $hh.GiaBan \geq hh.GiaNhap$
    - Cuối  $\forall$

# Bối cảnh một quan hệ

- Liên bộ:

- Điều kiện ràng buộc giữa các bộ trong cùng một quan hệ, rất phổ biến
- Thông thường loại ràng buộc này là các RBTV được biểu diễn bằng phụ thuộc hàm
- Sinh viên chỉ được thi lần 2 khi đã thi lần 1
  - $\forall kq1, kq2 \in T_{KetQua} \wedge kq1 \neq kq2 \wedge kq1.MaSV = kq2.MaSV \wedge kq1.MaMH = kq2.MaMH$ 
    - $\exists kq2.LanThi = 2 \Rightarrow \exists kq1.LanThi = 1$
  - Cuối  $\forall$

# Bối cảnh nhiều quan hệ

- Phụ thuộc tồn tại
  - Còn được gọi là RBTV khóa ngoại, rất phổ biến
  - Gọi :  $K_Q$  là tập thuộc tính của lược đồ quan hệ  $Q$ .
  - Và  $K_R$  là tập thuộc tính làm khóa nội của lược đồ quan hệ  $R$ .
  - $K_Q$  được xem là khóa ngoại - hay,  $K_q$  phụ thuộc tồn tại vào  $K_R$  - khi :
    1.  $K_Q \not\subseteq K_R$  nhưng  $K_Q \subseteq R^+$ .
    2.  $K_Q \subseteq K_R$ .

# Bối cảnh nhiều quan hệ

- Liên bộ - liên quan hệ
  - Ràng buộc quy định trên nhiều bộ giữa nhiều quan hệ với nhau
  - Ví dụ: mỗi khoa phải có ít nhất một sinh viên
    - $\forall k \in T_{\text{Khoa}}, \forall sv \in T_{\text{SinhVien}}$ 
      - $\exists sv.MaKhoa : sv.MaKhoa = k.MaKhoa$
    - Cuối  $\forall$
  - Khác biệt với phụ thuộc tồn tại ở điểm: bắt buộc phải có ít nhất một

# Bối cảnh nhiều quan hệ

- Liên thuộc tính - liên quan hệ
  - Cho 2 lược đồ quan hệ
    - **HANGHOA**(MaHH, TenHH, SLTon, DGVon)
    - **CT\_HOADON**(SoHD, MaHH, SLBan, DGBan)
  - Ta có,  $SLBan \leq SLTon$ , vậy
    - $\forall cthd \in T_{CT\_HOADON}, \forall hh \in T_{HangHoa} \wedge$   
 $cthd.MaHH = hh.MaHH$ 
      - $cthd.SLBan \leq hh.SLTon$
    - Cuối  $\forall$

# Bối cảnh nhiều quan hệ

- Thuộc tính tổng hợp
  - Giá trị thuộc tính A của một LĐQH Q được tính toán từ một hay nhiều thuộc tính của một hay nhiều LĐQH khác
  - Cho 3 LĐQH
    - KHACHHANG(MaKH, TenCT, TienNo)
    - HOADON(SoHD, MaKH, NgayHD, TriGiaHD)
    - PHIEUTHU(SoPT, MaKH, NgayPT, SoTien)

$\forall kh \in KhachHang, \forall hd \in HoaDon, \forall pt \in PhieuThu$

$$kh.TienNo = \sum_{hd.MaKH=kh.MaKH} TriGiaHD - \sum_{pt.MaKH=kh.MaKH} SoTien$$



# Bối cảnh nhiều quan hệ

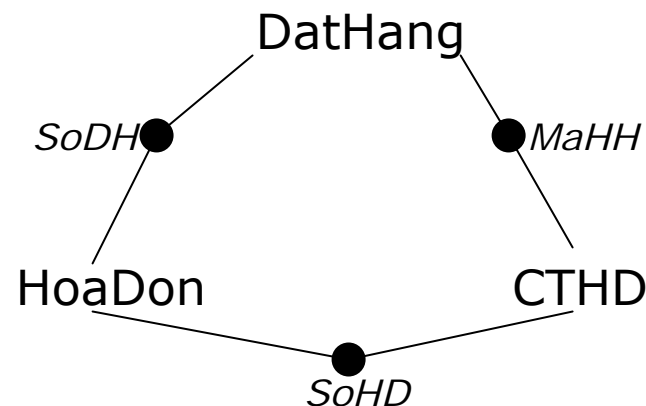
## ■ Chu trình

### ■ Cho 3 LĐQH

- $\text{DatHang} (\underline{\text{SoDH}}, \text{MaHH})$
- $\text{HoaDon} (\underline{\text{SoHD}}, \text{SoDH})$
- $\text{CTHD} (\underline{\text{SoHD}}, \text{MaHH})$

### ■ Có RBTV chu trình thể hiện sự liên quan giữa:

- $A = T_{\text{DatHang}}[\text{SoDH}, \text{MaHH}]$
- $B = T_{\text{HoaDon}} \bowtie T_{\text{CTHD}}[\text{SoHD}, \text{MaHH}]$

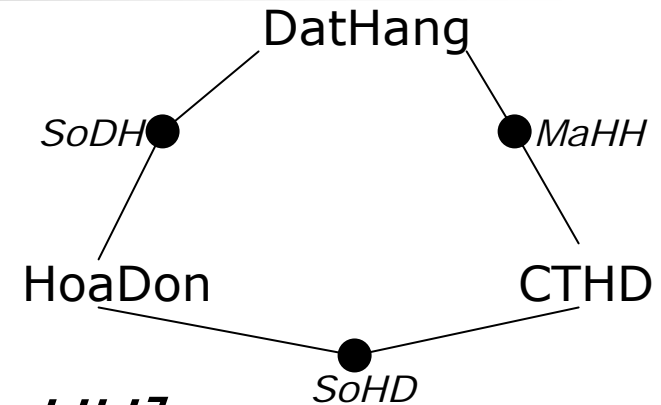


# Bối cảnh nhiều quan hệ

## ■ Chu trình

■  $A = T_{DatHang}[SoDH, MaHH]$

■  $B = T_{HoaDon} \bowtie T_{CTHD}[SoHD, MaHH]$



- $A = B$ : một hóa đơn được lập cho đúng và đủ số lượng hàng hóa của 1 đơn đặt hàng
- $B \subseteq A$ : một hóa đơn có thể giao đủ hoặc ít hơn số lượng hàng được đặt
- $A \not\subseteq B$  và  $B \not\subseteq A$ : hóa đơn giao hàng tùy ý không cần đúng những mặt hàng hoặc số lượng đã đặt

# Các bước xác định RBTV

- Bước 1: Lập bảng mô tả các RBTV*

		Gồm các table có trong LD CSDL			
<i>Bối cảnh</i>	<i>Loại RBTV</i>	SV	Khoa	GD	Table <i>n</i>
Một quan hệ	Miền giá trị	R2			
	Liên thuộc tính				
	Liên bộ	R1	R3		
Nhiều quan hệ	Phụ thuộc tồn tại	R11	R11		
	Liên bộ – Liên quan hệ				
	Liên thuộc tính – Liên quan hệ				
	Thuộc tính tổng hợp				
	Chu trình				251

# Các bước xác định RBTV

## ■ Bước 2: Lập bảng tầm ảnh hưởng

① RBTV về <u>khóa nội</u> trong quan hệ Q		<b>Thêm</b>	<b>Sửa</b>	<b>Xóa</b>
	Q	+	-	-

② RBTV về <u>miền giá trị</u> trong quan hệ Q		<b>Thêm</b>	<b>Sửa</b>	<b>Xóa</b>
	Q	+	+	-

③ RBTV về <u>liên thuộc tính</u> trong quan hệ Q		<b>Thêm</b>	<b>Sửa</b>	<b>Xóa</b>
	Q	+	+	-

④ RBTV về <u>khóa ngoại</u> : $R[a] \subseteq Q[a]$		<b>Thêm</b>	<b>Sửa</b>	<b>Xóa</b>
	Q	-	+	+
		+	+	-

# Các bước xác định RBTV

- *Bước 3: Lập bảng tầm ảnh hưởng tổng hợp*

**BẢNG TẦM ẢNH HƯỞNG TỔNG HỢP CÁC RBTV**

	$T_{\text{SinhVien}}$			$T_{\text{Khoa}}$			$T_{\text{KetQua}}$		
	<b><i>T</i></b>	<b><i>S</i></b>	<b><i>X</i></b>	<b><i>T</i></b>	<b><i>S</i></b>	<b><i>X</i></b>	<b><i>T</i></b>	<b><i>S</i></b>	<b><i>X</i></b>
<b><i>R1</i></b>	+	-	-						
<b><i>R2</i></b>	+	+	-	-	-	+			
<b><i>R3</i></b>							+	-	-



# Nội Dung

---

- RÀNG BUỘC TOÀN VỆN (RBTV) - CÁC YẾU TỐ CỦA RBTV
- CÁC LOẠI RÀNG BUỘC TOÀN VỆN
  - A)-BỐI CẢNH MỘT QUAN HỆ
  - B)-BỐI CẢNH NHIỀU QUAN HỆ
- **BIỂU DIỄN RBTV BẰNG PHỤ THUỘC HÀM**



## \$5.3. Biểu diễn RBTV bằng Phụ thuộc hàm

---

- **Khái niệm phụ thuộc hàm**
- Hệ tiên đề Armstrong
- Một số bài toán
- Xác định khóa

# Khái niệm phụ thuộc hàm

- Phụ thuộc hàm (functional dependency) là một ***công cụ dùng để biểu diễn*** một cách hình thức các ràng buộc toàn vẹn (vấn tắt: ràng buộc). Phương pháp biểu diễn này có rất nhiều ưu điểm, và đây là một công cụ cực kỳ quan trọng, gắn chặt với lý thuyết thiết kế cơ sở dữ liệu.
- Phụ thuộc hàm được ứng dụng trong việc ***giải quyết các bài toán tìm khóa, tìm phủ tối thiểu và chuẩn hóa cơ sở dữ liệu.***



# Khái niệm phụ thuộc hàm

- Quan hệ PhiCong diễn tả phi công nào lái máy bay nào và máy bay khởi hành vào thời gian nào tuân theo những ràng buộc sau:
  - Mỗi máy bay có một giờ khởi hành duy nhất.
  - Nếu biết phi công, biết ngày giờ khởi hành thì biết được máy bay do phi công ấy lái.
  - Nếu biết máy bay, biết ngày khởi hành thì biết phi công lái chuyến bay ấy.

<b>PHICONG</b>	<b>MBAY</b>	<b>NGKH</b>	<b>GKH</b>
Cushing	83	9/8	10:15a
Cushing	116	10/8	1:25p
Clark	281	8/8	5:50a
Clark	301	12/8	6:35p
Clark	83	11/8	10:15a
Chin	83	13/8	10:15a
Chin	116	12/8	1:25p
Copely	281	9/8	5:50a
Copely	281	13/8	5:50a
Copely	412	15/8	1:25p

# Khái niệm phụ thuộc hàm

- Các ràng buộc này là các ví dụ về phụ thuộc hàm và được phát biểu lại như sau:
  - MAYBAY ***xác định*** GIOKH
  - {PHICONG,NGAYKH,GIOKH} ***xác định*** MABAY
  - {MAYBAY,NGAYKH} ***xác định*** PHICONG
- hay
  - GIOKH ***phụ thuộc hàm*** vào MAYBAY
  - MABAY ***phụ thuộc hàm*** vào {PHICONG,NGAYKH,GIOKH}
  - PHICONG ***phụ thuộc hàm*** vào {MAYBAY,NGAYKH}

# Khái niệm phụ thuộc hàm

- Định nghĩa: Cho một quan hệ  $R$ .  
 **$R.Y$  là phụ thuộc xác định vào  $R.X$**  nếu mỗi giá trị- $X$  được kết hợp với chính xác một giá trị- $Y$  (vào bất kỳ thời điểm nào).
- $X, Y$  có thể là tập các thuộc tính.
- Ký hiệu:  $X \rightarrow Y$ .
- Đọc là:  $X$  xác định (hàm)  $Y$ .

# Khái niệm phụ thuộc hàm

- Cho lược đồ quan hệ  $Q$  với tập thuộc tính  $Q^+$ .  $X, Y, Z$  là tập hợp con của  $Q^+$  thoả 2 điều kiện

$$\left\{ \begin{array}{l} X, Y \neq \emptyset \\ \hline Q^+ = X \cup Y \cup Z \end{array} \right.$$

- $X \rightarrow Y$  được gọi là 1 phụ thuộc hàm định nghĩa trên  $Q$  nếu:
  - $\forall q_1, q_2 \in \mathcal{T}Q : q_1.X = q_2.X \Rightarrow q_1.Y = q_2.Y$
  - Với  $\mathcal{T}Q$  là 1 quan hệ bất kỳ định nghĩa trên  $Q$ .

# Khái niệm phụ thuộc hàm

- Cho lược đồ quan hệ:
  - **SinhVien(MaSV, HoTenSV, MaKhoa)**
- Ta thấy:
  - **MaSV  $\rightarrow$  HoTenSV** là phụ thuộc hàm đúng.
  - **HoTenSV  $\rightarrow$  MaSV** là phụ thuộc hàm sai.

# Khái niệm phụ thuộc hàm

- Phụ thuộc hàm đầy đủ:
  - Cho tập phụ thuộc hàm  $F$  định nghĩa trên lược đồ quan hệ  $R$ .  $X \rightarrow Y$  là một phụ thuộc hàm trong  $F$ .
  - $X \rightarrow Y$  được gọi là phụ thuộc hàm đầy đủ nếu  $Y$  không phụ thuộc hàm vào tập con khác rỗng  $Z$  nào của  $X$ , nghĩa là  $\forall Z \subset X, Z \neq \emptyset, Z \not\rightarrow Y$
- Cho  $F = \{AB \rightarrow C, A \rightarrow C\}$ . Ta có,  $AB \rightarrow C$  không phải là phụ thuộc hàm đầy đủ.  $A \rightarrow C$  là phụ thuộc hàm đầy đủ.

# Khái niệm phụ thuộc hàm

- Phụ thuộc hàm hiển nhiên (phụ thuộc hàm tầm thường):
  - Cho lược đồ quan hệ  $R(XY)$
  - $X \rightarrow X$  là phụ thuộc hàm hiển nhiên
- VD:  $MaSV \rightarrow MaSV$
- Cho  $F$  là tập phụ thuộc hàm định nghĩa trên  $R$ , không cần mô tả những phụ thuộc hàm hiển nhiên trong  $F$ .

# Khái niệm phụ thuộc hàm

- Tính chất:
  - Cho 2 lược đồ quan hệ  $R(W)$  và  $Q(V)$ ,  
 $V \subseteq W, V \cap W \neq \emptyset$
  - $f: X \rightarrow Y$  là một PTH định nghĩa trên  $R$ ,  
 $X \subseteq V$
  - Ta nói  $Q$  thỏa  $X \rightarrow (Y \cap V)$
- Ví dụ:
  - $R(ABCD), F: AB \rightarrow C, C \rightarrow D; Q(ABC)$
  - Ta nói  $Q$  cũng thỏa  $AB \rightarrow C$



# Khái niệm phụ thuộc hàm

- Tính phản chiếu:
  - Cho 2 lược đồ quan hệ  $R(W)$  và  $Q(V)$ ,  
 $W \subseteq V, V \cap W \neq \emptyset$
  - $f: X \rightarrow Y$  là một PTH định nghĩa trên  $R$
  - Ta nói  $Q$  thỏa  $X \rightarrow Y$
- Ví dụ:
  - $R(ABC), F: AB \rightarrow C; Q(ABCD)$
  - Ta nói  $Q$  cũng thỏa  $AB \rightarrow C$

# Khái niệm phụ thuộc hàm

- Phụ thuộc hàm suy diễn:
  - Cho lược đồ quan hệ  $R$ , tập phụ thuộc hàm  $F$  định nghĩa trên  $R$
  - Phụ thuộc hàm  $X \rightarrow Y$  là được suy diễn từ  $F$  nếu mọi quan hệ  $r(R)$  thỏa  $F$  thì thỏa  $X \rightarrow Y$
  - Ký hiệu:  $F \models X \rightarrow Y$
- Ví dụ:
  - $SV(\text{MaSV}, \text{HoTen}, \text{SoCMND})$
  - $F = \{\text{MaSV} \rightarrow \text{SoCMND}, \text{SoCMND} \rightarrow \text{HoTen}\}$
  - Nếu mọi  $sv(SV)$  thỏa  $F$  cũng thỏa  $\text{MaSV} \rightarrow \text{HoTen}$
  - Vậy  $\text{MaSV} \rightarrow \text{HoTen}$  là PTH suy diễn từ  $F$



# § 5.5. Điều kiện KĐTV bằng Phụ thuộc hàm

---

- Khái niệm phụ thuộc hàm
- **Hệ tiên đề Armstrong**
- Một số bài toán
- Xác định khóa

# Hệ tiên đề Armstrong

- Do Armstrong đưa ra lần đầu vào năm 1974.
- Là một tập các quy tắc suy diễn áp dụng cho phụ thuộc hàm.
- Cho lược đồ quan hệ  $R$  với tập thuộc tính  $U$ , ký hiệu  $R(U)$ , và tập phụ thuộc hàm  $F$  được định nghĩa trên  $R$ .
- Các quy tắc suy diễn trong hệ tiên đề Armstrong được phát biểu như sau:

# Hệ tiên đề Armstrong

① Luật phản xạ	$Y \subseteq X$		$\Rightarrow$	$X \rightarrow Y$
② Luật tăng trưởng	$X \rightarrow Y$	}	$\Rightarrow$	$XZ \rightarrow YZ$
	$Z \subset U$			
③ Luật bắc cầu	$X \rightarrow Y$	}	$\Rightarrow$	$X \rightarrow Z$
	$Y \rightarrow Z$			
④ Luật hợp	$X \rightarrow Y$	}	$\Rightarrow$	$X \rightarrow YZ$
	$X \rightarrow Z$			
⑤ Luật phân rã	$X \rightarrow Y$	}	$\Rightarrow$	$X \rightarrow Z$
	$Z \subseteq Y$			
⑥ Luật bắc cầu giả	$X \rightarrow Y$	}	$\Rightarrow$	$XW \rightarrow Z$
	$YW \rightarrow Z$			

# Hệ tiên đề Armstrong

## ■ **Chứng minh:**

- ① Hiển nhiên là đúng. Vì *không thể tồn tại hai bộ*  $t_1, t_2$  bất kỳ thuộc một quan hệ  $R$  *giống nhau trên*  $X$  nhưng lại *khác nhau ở một tập con* nào đó của  $X$ .
- ② Đặt  $R(U)$  là một *quan hệ thỏa*  $\{X \rightarrow Y\}$ ;  $t_1, t_2$  là hai bộ thuộc  $R$  và  $Z \subset U$ . *Ta có: nếu*  $t_1.X = t_2.X$  *thì*  $t_1.Y = t_2.Y$ .
  - *Giả sử*  $R$  không thỏa  $XZ \rightarrow YZ$ , nghĩa là  $t_1.XZ = t_2.XZ$  và  $t_1.YZ \neq t_2.YZ$ . Nhưng *chắc chắn*  $t_1$  và  $t_2$  *không thể khác nhau ở*  $Z$  nên chúng phải khác nhau ở  $Y$ . *Điều này dẫn đến*  $R$  *không thỏa*  $X \rightarrow Y \rightarrow$  *trái với giả thiết.*
  - Vậy  $R$  phải thỏa  $XZ \rightarrow YZ$ , tức nếu  $X \rightarrow Y$  đúng thì  $XZ \rightarrow YZ$  đúng

# Hệ tiên đề Armstrong

## ■ **Chứng minh:**

- ③ Đặt  $R(U)$  là một quan hệ thỏa  $\{X \rightarrow Y, Y \rightarrow Z\}$ ,  $t_1, t_2$  là hai bộ thuộc  $R$ .
  - Nếu  $t_1.Y \neq t_2.Y$  thì rõ ràng là  $R$  vi phạm  $X \rightarrow Y$ , điều này trái với đề bài ban đầu. Nếu  $t_1.Y = t_2.Y$  thì  $X \rightarrow Z$  sai là do  $t_1.Z \neq t_2.Z$ , điều này dẫn đến  $R$  vi phạm  $Y \rightarrow Z$ . Vậy  $R$  phải thỏa  $X \rightarrow Z$ .
- ④ Tự chứng minh
- ⑤ Tự chứng minh
- ⑥ Tự chứng minh



## § 5.5. Điều kiện KĐTV bằng Phụ thuộc hàm

---

- Khái niệm phụ thuộc hàm
- Hệ tiên đề Armstrong
- **Một số bài toán**
- Xác định khóa



# Một số bài toán

- **Bao đóng tập phụ thuộc hàm –  $F^+$  – Closure(F)**
- Bao đóng tập thuộc tính –  $X^+$  – Closure(X)
- Bài toán thành viên – Member
- Tập phụ thuộc hàm tương đương
- Phủ tối thiểu

## Bao đóng tập PTH – $F^+$

- Cho lược đồ quan hệ  $R(U)$ ,  $F$  là tập phụ thuộc hàm được định nghĩa trên  $R$ .
- ***Bao đóng của  $F$*** , ký hiệu là  $F^+$ , là tập hợp tất cả các phụ thuộc hàm được suy diễn (khẳng định logic) từ  $F$  dựa vào hệ tiên đề Armstrong.
- Vậy  $F$  chính là một tập con của  $F^+$ .

## Bao đóng tập PTH – $F^+$

- Xét tập phụ thuộc hàm  $F = \{A \rightarrow B_1, A \rightarrow B_2, \dots, A \rightarrow B_n\}$ .
- Ta có:  $F^+ = \{A \rightarrow Y_i \mid Y_i \subset \{B_1, B_2, \dots, B_n\}\}$
- **$\rightarrow F^+$  có  $2^n$  phụ thuộc hàm**
  
- Thông thường, chúng ta chỉ muốn biết một phụ thuộc hàm  $X \rightarrow Y$  có thuộc  $F^+$  hay không? Nghĩa là  $X \rightarrow Y$  có được suy diễn từ  $F$  hay không? Điều này có thể được giải quyết dễ dàng bằng bài toán thành viên.

# Một số bài toán

- Bao đóng tập phụ thuộc hàm –  $F^+$  – Closure(F)
- **Bao đóng tập thuộc tính –  $X^+$  – Closure(X)**
- Bài toán thành viên – Member
- Tập phụ thuộc hàm tương đương
- Phủ tối thiểu

## Bao đóng tập thuộc tính

- Cho lược đồ quan hệ  $R(U)$ , tập phụ thuộc hàm  $F$ .
- $X$  là tập con của  $U$ ,  $X \subset U$
- Bao đóng của tập thuộc tính  $X$  trên  $F$ , ký hiệu  $X_F^+$ , hay  $\text{Closure}(X, F)$
- $\text{Closure}(X, F) = \{A \subset U \mid X \rightarrow A \in F^+\}$

# Bao đóng tập thuộc tính

- Thuật toán tìm Closure( $X, F$ )
  - Vào: Tập thuộc tính  $X$  và tập phụ thuộc hàm  $F$
  - Ra: closure( $X, F$ )
  - Begin
    - TậpCũ =  $\emptyset$
    - TậpMới =  $X$
    - While TậpMới  $\neq$  TậpCũ do Begin
      - TậpCũ = TậpMới
      - For each pth  $W \rightarrow Z \in F$  do
        - If  $W \subseteq$  TậpMới then TậpMới = TậpMới  $\cup$   $Z$
        - End { if }
      - End { for }
    - End { while }
    - Return ( TậpMới )
  - End

## Bao đóng tập thuộc tính

- Cho lược đồ quan hệ  $Q(ABCDEFGH)$  và tập phụ thuộc hàm:
- $F = \{ f1: B \rightarrow A, f2: DA \rightarrow CE, f3: D \rightarrow H, f4: GH \rightarrow C, f5: AC \rightarrow D \}$
- Cho  $X = \{BDC\}$
- Tìm  $\text{Closure}(X, F)$

## Bao đóng tập thuộc tính

- TậpCũ= $\emptyset$ , TậpMới= $\{BDC\}$
- $f1: B \rightarrow A$ ,  
 $B \subset \text{TậpMới} \rightarrow \text{TậpMới} = \{BDC\} \cup \{A\}$
- $f2: DA \rightarrow CE$ ,  
 $DA \subset \text{TậpMới} \rightarrow \text{TậpMới} = \{BDCA\} \cup \{E\}$
- $f3: D \rightarrow H$ ,  
 $D \subset \text{TậpMới} \rightarrow \text{TậpMới} = \{BDCAE\} \cup \{H\}$
- $f4: GH \rightarrow C, GH \not\subset \text{TậpMới}$
- $f5: AC \rightarrow D, AC \subset \text{TậpMới}, D \subset \text{TậpMới}$
- **$Closure(X, F) = \{BDCAEH\}$**



# Một số bài toán

- Bao đóng tập phụ thuộc hàm –  $F^+$  – Closure(F)
- Bao đóng tập thuộc tính –  $X^+$  – Closure(X)
- **Bài toán thành viên – Member**
- Tập phụ thuộc hàm tương đương
- Phủ tối thiểu

## Bài toán thành viên

- Cho lược đồ quan hệ  $R(U)$ , tập phụ thuộc hàm  $F$  được định nghĩa trên  $R$ .
- Bài toán thành viên là bài toán kiểm tra một phụ thuộc hàm  $f$  có thuộc bao đóng tập phụ thuộc hàm  $F^+$  hay không?
- Hay, xác định PTH  $f$  có được suy diễn từ tập PTH  $F$  hay không?
- Ký hiệu:  $F \models f$

## Bài toán thành viên

- Thuật toán kiểm tra thành viên
- Vào:  $F$  và  $f: X \rightarrow Y$
- Ra: Đúng nếu  $F \models X \rightarrow Y$  và sai nếu ngược lại.
- MEMBER (  $F, X \rightarrow Y$  )
- Begin
- If  $Y \subset \text{Closure}(X, F)$  then
- Return ( true );
- Return ( false );
- End.

## Bài toán thành viên

- Cho lược đồ quan hệ  $R(A, B, C, D, E, G, H)$  và tập các phụ thuộc hàm  $F = \{f1: AB \rightarrow C, f2: B \rightarrow D, f3: CD \rightarrow E, f4: CE \rightarrow GH, f5: G \rightarrow A\}$ . Áp dụng hệ tiên đề Armstrong, xác định  $F \models AB \rightarrow E$ ?
- Giải ?

# Bài toán thành viên

## ■ Giải

- 1.  $AB \rightarrow C$  (cho trước - phụ thuộc hàm f1)
- 2.  $AB \rightarrow AB$  (luật phản xạ)
- 3.  $AB \rightarrow B$  (luật phân rã)
- 4.  $B \rightarrow D$  (cho trước - phụ thuộc hàm f2)
- 5.  $AB \rightarrow D$  (bắc cầu 3 & 4)
- 6.  $AB \rightarrow CD$  (hợp 1 & 5)
- 7.  $CD \rightarrow E$  (cho trước - phụ thuộc hàm f3)
- 8.  $AB \rightarrow E$  (bắc cầu 6 & 7).
- 9. Kết thúc.

## Bài toán thành viên

- Cho lược đồ quan hệ R  
(A,B,C,D,E,G,H,I,J) và tập các phụ thuộc  
hàm  $F = \{f1:AB \rightarrow E, f2:AG \rightarrow J, f3:BE \rightarrow I,$   
 $f4:E \rightarrow G, f5:GI \rightarrow H \}$ . Xác định  $F \models$   
 $AB \rightarrow GH?$
- Giải ?

## Bài toán thành viên

### ■ Giải

- 1.  $AB \rightarrow E$  (cho trước - phụ thuộc hàm f1)
- 2.  $AB \rightarrow AB$  (phản xạ)
- 3.  $AB \rightarrow B$  (luật phân rã )
- 4.  $AB \rightarrow BE$  (hợp của 1 & 3)
- 5.  $BE \rightarrow I$  (cho trước - phụ thuộc hàm f3)
- 6.  $AB \rightarrow I$  (bắc cầu 4 & 5)
- 7.  $E \rightarrow G$  (cho trước - phụ thuộc hàm f4)
- 8.  $AB \rightarrow G$  (bắc cầu 1 & 7)
- 9.  $AB \rightarrow GI$  (hợp 6 & 8)
- 10.  $GI \rightarrow H$  (cho trước - phụ thuộc hàm f5)
- 11.  $AB \rightarrow H$  (bắc cầu 9 & 10)
- 12.  $AB \rightarrow GH$  (hợp 8 & 11)

# Một số bài toán

- Bao đóng tập phụ thuộc hàm –  $F^+$  – Closure(F)
- Bao đóng tập thuộc tính –  $X^+$  – Closure(X)
- Bài toán thành viên – Member
- **Tập phụ thuộc hàm tương đương**
- Phủ tối thiểu



## Tập PTH tương đương

- Cho lược đồ quan hệ  $R(U)$ , tập PTH  $F, G$  được định nghĩa trên  $R$
- Hai tập phụ thuộc hàm  $F$  và  $G$  được cho là tương đương, ký hiệu  $F \equiv G$ , nếu  $F^+ = G^+$
- Nghĩa là,  $F \equiv G \Leftrightarrow F \vDash G$  và  $G \vDash F$ 
  - $F \vDash G \Leftrightarrow \forall g \in G, F \vDash g$
  - $G \vDash F \Leftrightarrow \forall f \in F, G \vDash f$

# Tập PTH tương đương

- Thuật toán xác định  $F \models G$
- Vào: hai tập phụ thuộc hàm  $F$  và  $G$
- Ra: Đúng nếu  $F \models G$  và sai nếu ngược lại.
- DERIVES (F,G)
- Begin
  - $V := \text{true}$
  - For each  $X \rightarrow Y \in G$  do
    - $V := V \text{ AND MEMBER}(F, X \rightarrow Y)$
  - Return (V);
- End.

# Tập PTH tương đương

- Thuật toán xác định  $F \equiv G$
- Vào: hai tập phụ thuộc hàm  $F$  và  $G$
- Ra: Đúng nếu  $F$  tương đương  $G$ , sai nếu không
- $EQUI ( F, G)$
- Begin
- $V := DERIVES(F,G) \text{ AND } DERIVES(G,F)$
- Return ( $V$ )
- End.

## Tập PTH tương đương

- Cho  $G = \{AB \rightarrow C, A \rightarrow B, B \rightarrow C, A \rightarrow C\}$
- $F = \{AB \rightarrow C, A \rightarrow B, B \rightarrow C\}$
- $F \equiv G$  ?
- Giải
- $G \vDash F$ ? Đương nhiên.
- $F \vDash G$ ? Xét  $F \vDash A \rightarrow C$ ?
  - $A \rightarrow B, B \rightarrow C$ , luật bắc cầu  $\rightarrow A \rightarrow C$
- Vậy  $F \equiv G$

# Một số bài toán

- Bao đóng tập phụ thuộc hàm –  $F^+$  – Closure(F)
- Bao đóng tập thuộc tính –  $X^+$  – Closure(X)
- Bài toán thành viên – Member
- Tập phụ thuộc hàm tương đương
- **Phủ tối thiểu**

## Phủ tối thiểu

- Đây là bài toán quan trọng thường cần phải giải quyết khi chuẩn hóa CSDL.
- Cho lược đồ quan hệ  $R(U)$  và tập phụ thuộc hàm  $F$  được định nghĩa trên  $R$
- Tập phụ thuộc hàm  $G$  được gọi là phủ tối thiểu của  $F$  nếu thỏa các tính chất sau:

## Phủ tối thiểu

1.  $F \equiv G$
2. Các phụ thuộc hàm trong  $G$  có vẻ phải chỉ gồm 1 thuộc tính
3. Với mọi phụ thuộc hàm  $X \rightarrow A \in G$ , không tồn tại một tập con thực sự  $Z$  của  $X$  sao cho  $G \setminus \{X \rightarrow A\} \cup \{Z \rightarrow A\} \equiv G$
4. Không tồn tại bất kỳ một phụ thuộc hàm  $X \rightarrow A$  nào trong  $G$  sao cho  $G \setminus \{X \rightarrow A\} \equiv G$

# Phủ tối thiểu

- Thuật toán tìm phủ tối thiểu
- Vào: tập phụ thuộc hàm  $F$  - Ra:  $G$  là phủ tối thiểu của  $F$
- MINIMALCOVER( $F, G$ )
  - 1.  $G := F$
  - 2. Dùng luật phân rã để đảm bảo vế phải các phụ thuộc hàm chỉ gồm 1 thuộc tính.
  - 3. For each  $X \rightarrow A$  trong  $G$ 
    - For each  $B \in X$
    - If  $(G - \{X \rightarrow A\}) \cup (X - \{B\} \rightarrow A) \equiv G$  Then
    - Thay  $X \rightarrow A$  bằng  $(X - \{B\}) \rightarrow A$  trong  $G$
  - 4. For each  $X \rightarrow A$  trong  $G$ 
    - If  $(G - \{X \rightarrow A\}) \equiv G$  then
    - Loại  $X \rightarrow A$  khỏi  $G$
  - 5. Return ( $G$ )



## Phủ tối thiểu

- Cho lược đồ quan hệ  $Q(ABCDEFGHIG)$ ,  
 $F = \{AB \rightarrow C, ABC \rightarrow D, C \rightarrow B, \mathbf{D \rightarrow EG}, H \rightarrow D, B \rightarrow D\}$
- 1.  $G = F$
- 2.  $G = \{AB \rightarrow C, \mathbf{ABC \rightarrow D}, C \rightarrow B, \mathbf{D \rightarrow E}, \mathbf{D \rightarrow G}, H \rightarrow D, B \rightarrow D\}$
- 3.  $G = \{AB \rightarrow C, C \rightarrow B, D \rightarrow E, D \rightarrow G, H \rightarrow D, B \rightarrow D\}$
- 4.  $G = \{AB \rightarrow C, C \rightarrow B, D \rightarrow E, D \rightarrow G, H \rightarrow D, B \rightarrow D\}$



## \$5.3. Biểu diễn RBTV bằng Phụ thuộc hàm

---

- Khái niệm phụ thuộc hàm
- Hệ tiên đề Armstrong
- Một số bài toán
- **Xác định khóa**

# Xác định khóa

## ■ Thuật toán 1

1. Xây dựng  $2^n$  tập con của  $Q^+$  bằng phương pháp đường chạy nhị phân.
2. Xây dựng tập  $S$  chứa tất cả các siêu khóa của  $Q$ ,  $S = \{S_i \mid (S_i)^+ = Q^+\}$ .
3. Xây dựng tập  $K$  chứa tất cả các khóa của  $Q$  từ tập siêu khóa  $S$  bằng cách loại bỏ khỏi  $S$  những siêu khóa  $S_i$  không tối tiểu.

# Xác định khóa

- Cho lược đồ  $Q(ABC)$  và  $F = \{AB \rightarrow C, C \rightarrow A\}$

A	B	C	$X_i$	$X_i, F^+$	
0	0	0			
0	0	1	C	AC	
0	1	0	B	B	
0	1	1	BC	ABC	= $Q^+$
1	0	0	A	A	
1	0	1	AC	AC	
1	1	0	AB	ABC	= $Q^+$
1	1	1	ABC	ABC	= $Q^+$

- Khóa  $K = \{BC, AB\}$

# Xác định khóa

- Thuật toán 2
- Vào:  $R(U)$  ,  $F$
- Ra :  $K$  ( khóa )
- Bước 1: Gán  $K = U$
- Bước 2: Lặp :
  - Loại khỏi  $K$  phần tử  $A$  mà  
 $\text{Closure}( K - A, F ) = U$
- Thuật toán trên chỉ tìm được một khóa.
- Nếu cần tìm nhiều khóa, ta thay đổi trật tự loại bỏ các phần tử của  $K$ .

# Xác định khóa

- Cho  $R = \{A, B, C, D, E, G, H, I\}$ ,  $F = \{AC \rightarrow B, BI \rightarrow ACD, ABC \rightarrow D, H \rightarrow I, ACE \rightarrow BCG, CG \rightarrow AE\}$   
Tìm K ?
  - **Bước 1:** Gán  $K = R = \{A, B, C, D, E, G, H, I\}$
  - **Bước 2:** Lần lượt loại bớt các thuộc tính của K
  - **Loại phân tử A:** ta có  $\{B, C, D, E, G, H, I\}^+ = R$  vì pth  $CG \rightarrow AE$  khiến A thuộc về  $\{B, C, D, E, G, H, I\}^+$  nên  $K = \{B, C, D, E, G, H, I\}$ .
  - **Loại phân tử B,** ta có  $\{C, D, E, G, H, I\}^+ = R$  vì pth  $CG \rightarrow AE$  khiến A thuộc về  $\{C, D, E, G, H, I\}^+$  và pth  $AC \rightarrow B$  nên  $K = \{C, D, E, G, H, I\}$ .
  - **Loại phân tử C,** ta có  $\{D, E, G, H, I\}^+ \neq R$  nên K vẫn là  $\{C, D, E, G, H, I\}$

# Xác định khóa

- **Loại phần tử D**, ta có:  $\{C, E, G, H, I\}^+ = R$  vì pth  $CG \rightarrow AE$  khiến A thuộc về  $\{C, E, G, H, I\}^+$  và pth  $AC \rightarrow B$  nên  $K = \{C, E, G, H, I\}$ .
- **Loại phần tử E**, ta có:  $\{C, G, H, I\}^+ = R$  vì pth  $CG \rightarrow AE, AC \rightarrow B, ABC \rightarrow D$  nên  $K = \{C, G, H, I\}$ .
- **Loại phần tử G**, ta có:  $\{C, H, I\}^+ \neq R$  nên K vẫn là  $\{C, G, H, I\}$ .
- **Loại phần tử H**, ta có:  $\{C, G, I\}^+ \neq R$  nên K vẫn là  $\{C, G, H, I\}$ .
- **Loại phần tử I**, ta có:  $\{C, G, H\}^+ = R$  vì  $CG \rightarrow AE, AC \rightarrow B, ABC \rightarrow D$  nên  $K = \{C, G, H\}$ .
- **Vậy  $K = \{C, G, H\}$  là một khóa của R**

# Xác định khóa

- Thuật toán 3 (cải tiến dựa trên ý tưởng phân loại thuộc tính).  $\forall A \in U$ , phân A vào:
- **Nguồn (ký hiệu N)**: nếu A **không thuộc về phải** của bất kỳ PTH không hiển nhiên nào của F. Những thuộc tính không xuất hiện trong tất cả những PTH đều phải đưa vào đây.
- **Đích (ký hiệu D)**: nếu A **không thuộc Nguồn** và A **không thuộc về trái** của bất kỳ PTH không hiển nhiên nào của F.
- **Trung Gian (ký hiệu L)**: nếu A **không thuộc Nguồn** và cũng **không thuộc Đích**.



# Xác định khóa

- Nếu  $Z$  là khóa của  $Q$  thì  $Z$  chứa tất cả các thuộc tính Nguồn và không chứa bất kỳ thuộc tính Đích nào.
- **Bước 1:** Xác định tập  $N$
- **Bước 2:** Nếu  $(N)^+ = Q^+ \rightarrow Q$  có một khóa duy nhất  $\rightarrow$  dừng.
- **Bước 3:** Xác định  $L$ . Thực hiện các bước như thuật toán 1 nhưng chỉ xét đến các thuộc tính thuộc tập  $N$  và  $L$ .

# Xác định khóa

- $Q(ABCDE), F = \{CE \rightarrow AB, B \rightarrow DE\}$
- Bước 1:  $N = \{C\}$
- Bước 2:  $(N)^+ = (C)^+ = \{C\}$
- Bước 3:  $L = \{B, E\}, D = \{A, D\}$

L		N	$X_i$	$(X_i)^+_F$					Super Key
B	E	C		A	B	C	D	E	
0	0	1	C			C			
0	1	1	CE	A	B	C	D	E	= Q+
1	0	1	CB	A	B	C	D	E	= Q+
1	1	1	CBE	A	B	C	D	E	= Q+



# Nhập Môn Cơ Sở Dữ Liệu

---

Minh họa bài tập phụ thuộc hàm



# Chương 5 – Bài B1

---

- Cho lược đồ quan hệ R (ABCDEFGH) , quan hệ r xác định trên R và tập phụ thuộc hàm:
  - $F = \{AB \longrightarrow C, B \longrightarrow D, CD \longrightarrow E, CE \longrightarrow GH, G \longrightarrow A\}$
- Sử dụng các luật dẫn trong hệ tiên đề Amstrong, chứng minh rằng nếu r thỏa F thì cũng thỏa:
  - ***B.1.1.***-  $AB \longrightarrow E$
  - ***B.1.2.***-  $AB \longrightarrow G.$

$F = \{AB \rightarrow C, B \rightarrow D, CD \rightarrow E, CE \rightarrow GH, G \rightarrow A\} \models AB \rightarrow E ?$

1.  $AB \rightarrow C$  (gt)
2.  $B \rightarrow D$  (gt)
3.  $BC \rightarrow CD$  (tăng trưởng)
4.  $AB \rightarrow CD$  (bắc cầu giả 1&3)
5.  $CD \rightarrow E$  (gt)
6.  $AB \rightarrow E$  (bắc cầu 4&5)
7. đpcm

$F = \{AB \rightarrow C, B \rightarrow D, CD \rightarrow E, CE \rightarrow GH, G \rightarrow A\}$

$\vdash AB \rightarrow G ?$

1.  $AB \rightarrow C$  (gt)
2.  $B \rightarrow D$  (gt)
3.  $BC \rightarrow CD$  (tăng trưởng)
4.  $AB \rightarrow CD$  (bắc cầu giả 1&3)
5.  $CD \rightarrow E$  (gt)
6.  $CD \rightarrow CE$  (tăng trưởng)
7.  $CE \rightarrow GH$  (gt)
8.  $AB \rightarrow GH$  (bắc cầu 4&6&7)
9.  $AB \rightarrow G$  (phân rã)
10. đpcm

# Chương 5 – Bài B2

- Cho lược đồ quan hệ R (ABCDE) và tập phụ thuộc hàm :
  - $F = \{A \rightarrow B, C \rightarrow D, BD \rightarrow E\}$
- Sử dụng các luật dẫn trong hệ tiên đề Amstrong, CMR  **$AC \rightarrow E$** .
  1.  $A \rightarrow B$  (gt)
  2.  $BD \rightarrow E$  (gt)
  3.  $AD \rightarrow E$  (bắc cầu giả 1&2)
  4.  $C \rightarrow D$  (gt)
  5.  $AC \rightarrow E$  (bắc cầu giả 3&4) đpcm

# Chương 5 – Bài B3

- 3.1. Cho lược đồ quan hệ R (ABCD):
  - Với  $\mathbf{C} \subset \mathbf{B}$  và tập phụ thuộc hàm  $F = \{A \rightarrow B, C \rightarrow D\}$ . CMR:  $\mathbf{A} \rightarrow \mathbf{D}$ .
    1.  $C \subset B$  (gt)
    2.  $B \rightarrow C$  (phản xạ)
    3.  $A \rightarrow B$  (gt)
    4.  $A \rightarrow C$  (bắc cầu 3&2)
    5.  $C \rightarrow D$  (gt)
    6.  $A \rightarrow D$  (bắc cầu 4&5) đpcm.



# Chương 5 – Bài B3

- 3.2. Cho lược đồ quan hệ R (ABCD):
  - $F = \{AB \rightarrow C, D \rightarrow B, C \rightarrow ABD\}$
  - $AB \rightarrow D$  thuộc  $F^+$ ?
    1.  $AB \rightarrow C$  (gt)
    2.  $C \rightarrow ABD$  (gt)
    3.  $C \rightarrow D$  (phân rã)
    4.  $AB \rightarrow D$  (bắc cầu 1&3)
    5. Vậy  $AB \rightarrow D$  thuộc  $F^+$ . Đpcm.

# Chương 5 – Bài B4

- Cho lược đồ quan hệ  $Q$  (ABCDEG), và  $F = \{AB \rightarrow E, AG \rightarrow D, BE \rightarrow D, E \rightarrow G, GD \rightarrow C\}$ . CMR:  **$AB \rightarrow GC$**

1.  $AB \rightarrow E$  (gt)
2.  $E \rightarrow G$  (gt)
3.  $AB \rightarrow G$  (bắc cầu 1&2)
4.  $BE \rightarrow D$  (gt)
5.  $AB \rightarrow D$  (bắc cầu giả 1&4)
6.  $AB \rightarrow GD$  (hợp 3&5)
7.  $GD \rightarrow C$  (gt)
8.  $AB \rightarrow C$  (bắc cầu 6&7)
9.  $AB \rightarrow GC$  (hợp 3&8) đpcm.

# Chương 5 – Bài B5.2

- $F = \{A \rightarrow D, AB \rightarrow DE, CE \rightarrow G, E \rightarrow H\}$   
Tìm  $(AB)_F^+$ .
  1.  $(AB)_F^+ = AB$
  2.  $(AB)_F^+ = ABDE$  (do PTH  $AB \rightarrow DE$ )
  3.  $(AB)_F^+ = ABDEH$  (do PTH  $E \rightarrow H$ )
  4. Không tìm thêm được thuộc tính nào  
 $\rightarrow (AB)_F^+ = ABDEH$

# Chương 5 – Bài B5.6.1

- $F = \{AB \rightarrow C, B \rightarrow D, CD \rightarrow E, CE \rightarrow GH, G \rightarrow A\}$ . Tìm  $(AB)_F^+$ .

1.  $(AB)_F^+ = AB$

2.  $(AB)_F^+ = ABC$  (do PTH  $AB \rightarrow C$ )

3.  $(AB)_F^+ = ABCD$  (do PTH  $B \rightarrow D$ )

4.  $(AB)_F^+ = ABCDE$  (do PTH  $CD \rightarrow E$ )

5.  $(AB)_F^+ = ABCDEGH$  (do PTH  $CE \rightarrow GH$ )

# Chương 5 – Bài B6.1

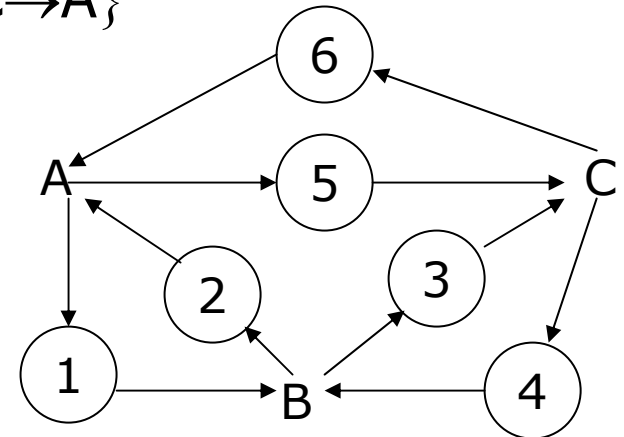
- Chứng minh tính đúng đắn của luật sau: **Nếu  $X \rightarrow Y$ ,  $Z \rightarrow W$  thì  $XZ \rightarrow YW$ .**
  1.  $X \rightarrow Y$  (gt)
  2.  $XZ \rightarrow YZ$  (tăng trưởng)
  3.  $Z \rightarrow W$  (gt)
  4.  $YZ \rightarrow YW$  (tăng trưởng)
  5.  $XZ \rightarrow YW$  (bắc cầu 2&4) đpcm.

# Chương 5 – Xác định PTH thừa

- B1: Phân rã để  $F$  chỉ có các phụ thuộc hàm mà vế phải gồm 1 thuộc tính.
- B2: Tìm và loại bỏ các phụ thuộc không đầy đủ.
- B3: Lần lượt xét từng phụ thuộc hàm có trong  $F$ .  
Đối với mỗi phụ thuộc hàm ( $F_i$ ):
  - Không dùng đến phụ thuộc hàm đang xét, tính bao đóng của tập thuộc tính ở vế trái của  $F_i$ .
    - Nếu bao đóng của tập thuộc tính trong vế trái của  $F_i$ :
      - Có chứa thuộc tính ở vế phải của  $F_i$  thì  $F_i$  thừa.
      - Không chứa các thuộc tính ở vế phải của  $F_i$  thì  $F_i$  không thừa.

# Chương 5 – Bài D1 – xác định PTH thừa

- D1.1.  $Q(ABC)$ ,  $F = \{A \rightarrow B, B \rightarrow A, B \rightarrow C, C \rightarrow B, A \rightarrow C, C \rightarrow A\}$ 
  1.  $F = \{A \rightarrow B, B \rightarrow A, B \rightarrow C, C \rightarrow B, A \rightarrow C, C \rightarrow A\}$
  2.  $F = \{A \rightarrow B, B \rightarrow A, B \rightarrow C, C \rightarrow B, A \rightarrow C, C \rightarrow A\}$
  3.  $A \rightarrow B$ ,  $(A)_F^+ = ACB \rightarrow$  thừa
    - $F = \{B \rightarrow A, B \rightarrow C, C \rightarrow B, A \rightarrow C, C \rightarrow A\}$
  4.  $B \rightarrow A$ ,  $(B)_F^+ = BCA \rightarrow$  thừa
    - $F = \{B \rightarrow C, C \rightarrow B, A \rightarrow C, C \rightarrow A\}$
  5.  $B \rightarrow C$ ,  $(B)_F^+ = B \rightarrow$  không thừa
  6.  $C \rightarrow B$ ,  $(C)_F^+ = CA \rightarrow$  không thừa
  7.  $A \rightarrow C$ ,  $(A)_F^+ = A \rightarrow$  không thừa
  8.  $C \rightarrow A$ ,  $(C)_F^+ = CB \rightarrow$  không thừa
  9.  $F = \{B \rightarrow C, C \rightarrow B, A \rightarrow C, C \rightarrow A\}$



# Chương 5 – Bài D1 – xác định PTH thừa

- $Q(A\ B\ C\ D)$  với

$$F = \{A \rightarrow BD, A \rightarrow C, C \rightarrow B, AB \rightarrow C\}$$

1.  $F = \{A \rightarrow B, A \rightarrow D, A \rightarrow C, C \rightarrow B, AB \rightarrow C\}$ , bỏ  $AB \rightarrow C$

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

3.  $A \rightarrow B$ ,  $(A)_F^+ = ADCB \rightarrow$  thừa

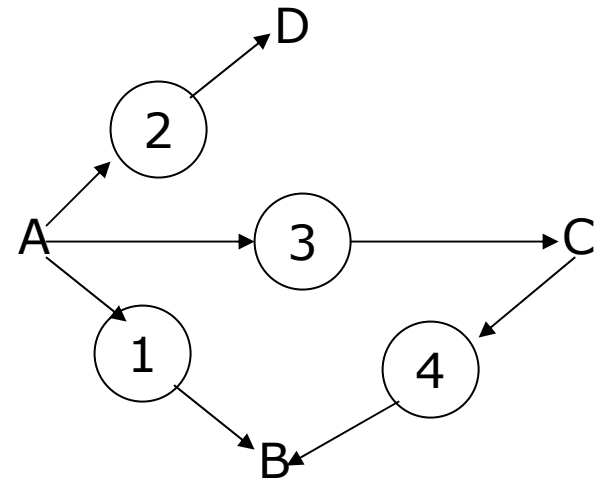
- $F = \{A \rightarrow D, A \rightarrow C, C \rightarrow B\}$

4.  $A \rightarrow D$ ,  $(A)_F^+ = ACB \rightarrow$  không thừa

5.  $A \rightarrow C$ ,  $(A)_F^+ = AD \rightarrow$  không thừa

6.  $C \rightarrow B$ ,  $(C)_F^+ = C \rightarrow$  không thừa

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





# Chương 5 – Bài D1 – xác định PTH thừa

- $Q(A B C D)$  với  
 $F = \{A \rightarrow C, C \rightarrow A, AD \rightarrow B, BC \rightarrow D, AB \rightarrow D, CD \rightarrow B\}$ 
  1.  $F = \{A \rightarrow C, C \rightarrow A, AD \rightarrow B, BC \rightarrow D, AB \rightarrow D, CD \rightarrow B\}$
  2.  $F = \{A \rightarrow C, C \rightarrow A, AD \rightarrow B, BC \rightarrow D, AB \rightarrow D, CD \rightarrow B\}$
  3.  $A \rightarrow C, (A)_F^+ = A \rightarrow$  không thừa
  4.  $C \rightarrow A, (C)_F^+ = C \rightarrow$  không thừa
  5.  $AD \rightarrow B, (AD)_F^+ = ADCB \rightarrow$  thừa
    - $F = \{A \rightarrow C, C \rightarrow A, BC \rightarrow D, AB \rightarrow D, CD \rightarrow B\}$
  6.  $BC \rightarrow D, (BC)_F^+ = BCAD \rightarrow$  thừa
    - $F = \{A \rightarrow C, C \rightarrow A, AB \rightarrow D, CD \rightarrow B\}$
  7.  $AB \rightarrow D, (AB)_F^+ = ABC \rightarrow$  không thừa
  8.  $CD \rightarrow B, (CD)_F^+ = CDA \rightarrow$  không thừa
  9.  $F = \{A \rightarrow C, C \rightarrow A, AB \rightarrow D, CD \rightarrow B\}$

# Chương 5 – Tập PTH tương đương

- Q (A B C D),  $F \equiv G$  ?
  - $F = \{A \rightarrow B, A \rightarrow C, B \rightarrow A, B \rightarrow C, C \rightarrow A\}$
  - $G = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$
  - 1. Rã vể phải của F sao cho chỉ có 1 th/tính
  - 2.  $F' = F \setminus G = \{A \rightarrow C, B \rightarrow A\}$ , CM  $G \not\models F'$ .
  - 3.  $f'1: A \rightarrow C, (A)_G^+ = ABC \rightarrow G \models f'1$
  - 4.  $f'2: B \rightarrow A, (B)_G^+ = BCA \rightarrow G \models f'2$
  - 5.  $\rightarrow G \models F'$
  - 6. Vậy,  $F \equiv G$ . đpcm



# Chương 5 – Phủ tối thiểu

- Cho tập phụ thuộc hàm  $F$  định nghĩa trên lược đồ quan hệ  $R$ .
- Phủ tối thiểu của  $F$  là một tập phụ thuộc hàm  $G$  thỏa 3 tính chất sau :
  1.  $F \equiv G$
  2.  $G$  chỉ có các phụ thuộc hàm mà vế phải gồm 1 thuộc tính. Ví dụ:
    - $AB \longrightarrow C$  được
    - $A \longrightarrow BC$  không, vì vế phải có 2 th/tính.
  3.  $G$  không có phụ thuộc hàm thừa.
- Áp dụng các bước loại PTH thừa và tìm tập PTH tương đương để tìm phủ tối thiểu.

# Chương 5– Tìm khóa

- ***Thuộc tính nguồn*** : (ký hiệu là ***N***) nếu A không xuất hiện ở vế phải của bất kỳ phụ thuộc hàm không hiển nhiên nào của ***F***. Những thuộc tính không xuất hiện trong tất cả các phụ thuộc hàm phải được đưa vào tập này.
- ***Thuộc tính đích*** : (ký hiệu là ***D***) nếu A không phải thuộc tính nguồn và A không xuất hiện ở vế trái của bất kỳ phụ thuộc hàm không hiển nhiên nào của ***F***.
- ***Thuộc tính không phải nguồn và cũng không phải đích***: (ký hiệu là ***L***) nếu A không phải thuộc tính nguồn và cũng không phải thuộc tính đích.
- Thuật toán
  - Bước 1 : Xác định tập ***N*** và tập ***L***.
  - Bước 2 : Thực hiện các bước như trong thuật toán nguyên thủy nhưng chỉ xét đến các thuộc tính có trong ***N*** với các thuộc tính có trong ***L***.

# Chương 5 – Tìm khóa

- E3.1.  $Q(ABCDEFG)$ ,  $F = \{AE \rightarrow C, CG \rightarrow A, BD \rightarrow G, GA \rightarrow E\}$ 
  - $N = \{BD\}$ ,  $L = \{A, C, E, G\}$
  - $(BD)_F^+ \neq Q^+$
  - $(BDA)_F^+ = Q^+ \rightarrow$  siêu khóa
  - $(BDC)_F^+ = Q^+ \rightarrow$  siêu khóa
  - $(BDE)_F^+ \neq Q^+$
  - $(BDG)_F^+ \neq Q^+$
  - $(BDEG)_F^+ \neq Q^+$
  - Vậy có 2 khóa:  $\{ABD, BCD\}$

# Chương 5 – Chuẩn hóa

- **C.1.1.** Q(A B C) với tập F như sau:
- $F = \{A \rightarrow B, C \rightarrow B, B \rightarrow C\}$ 
  1.  $Q^+ = \{A, B, C\}$ ,  $F = \{A \rightarrow B ; C \rightarrow B ; B \rightarrow C\}$ ,  $Keys = \{A\}$
  2. Lấy phủ tối tiểu:  $F = \{A \rightarrow B ; C \rightarrow B ; B \rightarrow C\}$
  3. Chọn PTH  $C \rightarrow B$  và  $B \rightarrow C$  để tạo quan hệ mới:  
 $Q1^+ = \{C, B\}$ ,  $F = \{C \rightarrow B ; B \rightarrow C\}$ ,  $Keys = \{C ; B\}$
  4. Chọn PTH  $A \rightarrow B$  để tạo quan hệ mới:  
 $Q2^+ = \{A, B\}$ ,  $F = \{A \rightarrow B\}$ ,  $Keys = \{A\}$
  5. Vậy, từ lược đồ ban đầu, ta phân rã thành các lược đồ quan hệ sau:
  6.  $Q1(CB)$ ,  $F = \{C \rightarrow B, B \rightarrow C\}$ ,  $Keys = \{C, B\}$ , 3NF
  7.  $Q2(AB)$ ,  $F = \{A \rightarrow B\}$ ,  $Keys = \{A\}$ , 3NF

# Chương 5 – Chuẩn hóa

- **C.1.5.**  $Q(ABCDEFGH)$  với tập  $F$  như sau:
- $F = \{BG \rightarrow D, D \rightarrow H, GH \rightarrow C, B \rightarrow A, BC \rightarrow E\}$ 
  - $Q_+ = \{ABCDEFGH\}$ ,  $F = \{BG \rightarrow D, D \rightarrow H, GH \rightarrow C, B \rightarrow A, BC \rightarrow E\}$ ,  $Keys = \{BG\}$
  - Lấy phủ tối thiểu:  $F = \{BG \rightarrow D, D \rightarrow H, GH \rightarrow C, B \rightarrow A, BC \rightarrow E\}$
  - Chọn PTH  $B, G \rightarrow D$  để tạo quan hệ mới:  $Q_1(BGD)$ ,  $F = \{BG \rightarrow D\}$ ,  $Keys = \{BG\}$
  - Chọn PTH  $D \rightarrow H$  để tạo quan hệ mới:  $Q_2(DH)$ ,  $F = \{D \rightarrow H\}$ ,  $Keys = \{D\}$
  - Chọn PTH  $G, H \rightarrow C$  để tạo quan hệ mới:  $Q_3(GHC)$ ,  $F = \{GH \rightarrow C\}$ ,  $Keys = \{GH\}$
  - Chọn PTH  $B \rightarrow A$  để tạo quan hệ mới:  $Q_4(BA)$ ,  $F = \{B \rightarrow A\}$ ,  $Keys = \{B\}$
  - Chọn PTH  $B, C \rightarrow E$  để tạo quan hệ mới:  $Q_5(BCE)$ ,  $F = \{BC \rightarrow E\}$ ,  $Keys = \{BC\}$
  - Vậy, từ lược đồ ban đầu, ta phân rã thành các lược đồ quan hệ sau:
    - $Q_1(B, G, D)$ ,  $F = \{BG \rightarrow D\}$ ,  $Keys = \{BG\}$
    - $Q_2(D, H)$ ,  $F = \{D \rightarrow H\}$ ,  $Keys = \{D\}$
    - $Q_3(G, H, C)$ ,  $F = \{GH \rightarrow C\}$ ,  $Keys = \{GH\}$
    - $Q_4(B, A)$ ,  $F = \{B \rightarrow A\}$ ,  $Keys = \{B\}$
    - $Q_5(B, C, E)$ ,  $F = \{BC \rightarrow E\}$ ,  $Keys = \{BC\}$

# CƠ SỞ DỮ LIỆU

ThS.Lê Văn Hạnh



# CHUẨN HÓA LƯỢNG ĐỒ CSDL



# Nội dung

---

- **Giới thiệu**
- Các dạng chuẩn
  - 1NF
  - 2NF
  - 3NF
  - BCNF
- Phương pháp phân rã
- Phương pháp kiểm tra bảo toàn thông tin



## \$6.1. Giới thiệu

---

- Chất lượng thiết kế của 1 lược đồ CSDL có thể được đánh giá dựa trên:
  - Sự trùng lặp thông tin
  - Chi phí kiểm tra các RBTV
- Một lược đồ CSDL-Quan Hệ ban đầu có thể được phân tích thành nhiều lược đồ quan hệ con với chất lượng tốt hơn

# \$6.1. Giới thiệu

- DATHANG(SoDH, NgayDH, MaKH, MaHH, SLDat)
- $F = \{ f1: \text{SoDH} \longrightarrow \text{NgayDH, MaKH}$   
 $f2: \text{SoDH, MaHH} \longrightarrow \text{SLDat} \}$

<i>SoDH</i>	<i>NgayDH</i>	<i>MaK</i> <i>H</i>	<i>MaH</i> <i>H</i>	<i>SLDa</i> <i>t</i>
DH00	5/12/200	K000	S001	5
DH00	5/12/200	K000	S003	2
DH00	5/12/200	K000	T003	5
DH00	8/12/200	K000	S003	1
DH00	8/12/200	K000	T004	1

2

0 3





# Nội dung

---

- Giới thiệu
- **Các dạng chuẩn**
  - 1NF
  - 2NF
  - 3NF
  - BCNF
- Phương pháp phân rã
- Phương pháp kiểm tra bảo toàn thông tin



## \$6.2. Các dạng chuẩn

---

- **Một số khái niệm**
- 1NF
- 2NF
- 3NF
- BCNF

# Một số khái niệm

- Thuộc tính khóa/ thuộc tính không khóa
- Thuộc tính phụ thuộc đầy đủ :
  - Một thuộc tính A được gọi là phụ thuộc đầy đủ vào tập thuộc tính X nếu  $X \rightarrow A$  là phụ thuộc hàm đầy đủ
  - $X \rightarrow Y$  được gọi là phụ thuộc hàm đầy đủ nếu về phải Y không phụ thuộc hàm vào tập con khác rỗng Z nào của X
- Thuộc tính đơn / kép
  - Thuộc tính đơn là thuộc tính chỉ chứa các giá trị nguyên tố.
    - Không bị phân chia bởi các phép toán ĐSQH
    - Không bị tích hợp từ nhiều giá trị có cùng ý nghĩa





## \$6.2. Các dạng chuẩn

---

- Một số khái niệm
- **1NF**
- 2NF
- 3NF
- BCNF

# Dạng chuẩn 1 – 1NF

- Một lược đồ quan hệ  $Q$  được gọi là ở *1NF* nếu mọi thuộc tính của  $Q$  đều là thuộc tính đơn.
- Một lược đồ CSDL  $C$  được gọi là ở *1NF* nếu mọi lược đồ quan hệ của  $C$  đều ở *1NF*.
- Các lược đồ quan hệ đạt 1NF vẫn bị gặp phải 4 vấn đề: dư thừa, mâu thuẫn, bất thường khi thêm, bất thường khi xóa.

# Dạng chuẩn 1 – 1NF

- Q(SoHD, MaKH, HoTenKH, MaHH, TenHH, DonGia, SL)
- F = {MAKH → HOTENKH; MAHH → TENHH;  
MAHH → DONGIA; SOHD → MAKH;  
SOHD, MAHH → SL }
- Khóa = ?

SOHD	MAKH	HOTENKH	MAHH	TENHH	DONGIA	SL
HD01	KH03	Đông Tà	HH02	Sáo trúc	2000	5
HD01	KH03	Đông Tà	HH03	Cửu Âm	4000	10
HD02	KH04	Bắc Cái	HH01	Kentucky	10000	15
HD03	KH01	Tây Độc	HH03	Cửu Âm	4000	30
HD03	KH01	Tây Độc	HH07	Axit	500	100

# Dạng chuẩn 1 – 1NF

- Nếu công ty muốn thêm mặt hàng mới?
- Khóa chính không thể chứa giá trị NULL

SOHD	MAKH	HOTENKH	MAHH	TENHH	DONGIA	SL
HD01	KH03	Đông Tà	HH02	Sáo trúc	2000	5
HD01	KH03	Đông Tà	HH03	Cửu Âm	4000	10
?	?	?	<b>HH08</b>	<b>Rượu</b>	<b>7000</b>	<b>?</b>
HD02	KH04	Bắc Cái	HH01	Kentucky	10000	15
HD03	KH01	Tây Độc	HH03	Cửu Âm	4000	30
HD03	KH01	Tây Độc	HH07	Axit	500	100

- Nếu công ty muốn loại bỏ mặt hàng Kentucky? → Bắc Cái qua đời

# Dạng chuẩn 1 – 1NF

- Xét tập phụ thuộc hàm được định nghĩa trên lược đồ quan hệ:
- $F = \{ \text{MAKH} \rightarrow \text{HOTENKH}; \text{MAHH} \rightarrow \text{TENHH}; \text{MAHH} \rightarrow \text{DONGIA}; \text{SOHD} \rightarrow \text{MAKH}; \text{SOHD, MAHH} \rightarrow \text{SL} \}$
- Khóa?  $K = \{ \text{SOHD, MAHH} \}$
- **→ Tồn tại nhiều thuộc tính không phụ thuộc đầy đủ vào khóa.**



## \$6.2. Các dạng chuẩn

---

- Một số khái niệm
- 1NF
- **2NF**
- 3NF
- BCNF

# Dạng chuẩn 2 – 2NF

- Từ đây, chúng ta giả sử rằng tập phụ thuộc hàm  $F$  chỉ chứa những phụ thuộc hàm mà vế phải của nó chỉ có 1 thuộc tính.
- Một lược đồ quan hệ  $Q$  được gọi là ở  $2NF$  nếu :
  - $Q$  ở  $1NF$ .
  - Mọi thuộc tính không khóa của  $Q$  đều phụ thuộc đầy đủ vào các khóa của  $Q$ .
- Một lược đồ CSDL  $C$  được gọi là ở  $2NF$  nếu mọi lược đồ quan hệ con  $Q_i$  của  $C$  đều ở  $2NF$ .

# Dạng chuẩn 2 – 2NF

SOHD	MAKH	HOTENKH
HD01	KH03	Đông Tà
HD02	KH04	Bắc Cái
HD03	KH01	Tây Độc

SOHD	MAHH	SL
HD01	HH02	5
HD01	HH03	10
HD02	HH01	15
HD03	HH03	30
HD03	HH07	100

$F = \{SOHD \rightarrow MAKH; MAKH \rightarrow HOTENKH\}$

$F = \{SOHD, MAHH \rightarrow SL\}$

MAHH	TENHH	DONGIA
HH02	Sáo trúc	2000
HH03	Cửu Âm	4000
HH01	Kentucky	10000
HH03	Cửu Âm	4000
HH07	Axit	500

$F = \{MAHH \rightarrow TENHH; MAHH \rightarrow DONGIA\}$

**Các lược đồ quan hệ trên đạt 2NF**



# Dạng chuẩn 2 – 2NF

- Xét quan hệ  
 $Q(\text{SOHD}, \text{MAKH}, \text{HOTENKH})$
- $F = \{\text{SOHD} \rightarrow \text{MAKH};$   
 $\text{MAKH} \rightarrow \text{HOTENKH}\}$

SOHD	MAKH	HOTENKH
HD01	KH03	Đông Tà
HD02	KH04	Bắc Cái
HD03	KH01	Tây Độc

- **Tồn tại thuộc tính không khóa phụ thuộc bắc cầu vào khóa**
- Các lược đồ quan hệ đạt 2NF vẫn có thể bị gặp phải 4 vấn đề: dư thừa, mâu thuẫn, bất thường khi thêm, bất thường khi xóa.

# Dạng chuẩn 2 – 2NF

- Hoàn toàn không thể thêm 1 khách hàng mới khi khách hàng này chưa mua hàng.
- Khi hủy 1 hóa đơn HD02 nghĩa là đồng thời hủy luôn thông tin về khách hàng KH04.
- Khi thêm một hóa đơn mới dễ xảy ra hiện tượng nhập sai dữ liệu → mâu thuẫn.

SOHD	MAKH	HOTENKH
HD01	KH03	Đông Tà
HD02	KH04	Bắc Cái
?	<b>KH08</b>	<b>Nam Đế</b>
HD03	KH01	Tây Độc
HD04	KH03	Đông Tửu



## \$6.2. Các dạng chuẩn

---

- Một số khái niệm
- 1NF
- 2NF
- **3NF**
- BCNF

# Dạng chuẩn 3 – 3NF

- **Định nghĩa:** Một lược đồ quan hệ Q được cho là đạt 3NF nếu:
  - Q đạt 2NF
  - Mọi thuộc tính **không khóa** của Q đều **không** phụ thuộc bắc cầu vào khóa của Q.
- Một lược đồ CSDL C được gọi là đạt 3NF nếu mọi lược đồ quan hệ Q của nó đều đạt 3NF.

# Dạng chuẩn 3 – 3NF

SOHD	MAKH
HD01	KH03
HD02	KH04
HD03	KH01

$F = \{SOHD \rightarrow MAKH\}$

$F = \{SOHD, MAHH \rightarrow SL\}$

SOHD	MAHH	SL
HD01	HH02	5
HD01	HH03	10
HD02	HH01	15
HD03	HH03	30
HD03	HH07	100

MAHH	TENHH	DONGIA
HH02	Sáo trúc	2000
HH03	Cửu Âm	4000
HH01	Kentucky	10000
HH03	Cửu Âm	4000
HH07	Axit	500

MAKH	HOTENKH
KH03	Đông Tà
KH04	Bắc Cái
KH01	Tây Độc

$F = \{MAHH \rightarrow TENHH; MAHH \rightarrow DONGIA\}$   $F = \{MAKH \rightarrow HOTENKH\}$

**4 lược đồ quan hệ trên đều đạt 3NF.**

# Dạng chuẩn 3 – 3NF

- Ngoài ra, ta còn có **định nghĩa khác** về 3NF:
- Cho lược đồ quan hệ  $Q$  và tập phụ thuộc hàm  $F$ .  $Q$  được cho là đạt 3NF khi **1 trong 3** điều kiện sau thỏa:
  - $\forall f : X \rightarrow A \in F$ ,  $f$  là phụ thuộc hàm hiển nhiên.
  - $\forall f : X \rightarrow A \in F$ ,  $X$  là siêu khóa của  $Q$ .
  - Vế phải của phụ thuộc hàm chứa thuộc tính khóa.

# Dạng chuẩn 3 – 3NF

- Các quan hệ đạt 3NF vẫn có thể gây ra dữ liệu xấu trong các trường hợp sau:
  - Khi thuộc tính không khóa xác định thuộc tính khóa.
  - Khi lược đồ quan hệ có 2 tập khóa gối nhau.
    - **DMLop**(Tinh, MonHoc, HocPhan, GioHoc, TrTam)
    - **F** = {Tinh → MonHoc, HocPhan, GioHoc, TrTam, TrTam → Tinh}

# Dạng chuẩn 3 – 3NF

Tỉnh	MonHo	HocPhan	GioHoc	TrTam
TpHCM	SQL	Căn bản	8:00-10:00	LTV
AnGiang	SQL	Căn bản	10:00-12:00	Van Kiep
DongNai	SQL	Nâng cao	8:00-10:00	Van Hoa

- Quan hệ này đạt 3NF nhưng sẽ gây dữ liệu xấu nếu ta xóa bất kỳ bộ nào. Điều này sẽ làm ta mất luôn thông tin về địa điểm của trung tâm.





## \$6.2. Các dạng chuẩn

---

- Một số khái niệm
- 1NF
- 2NF
- 3NF
- **BCNF**

# Dạng chuẩn BC – BCNF

- Định nghĩa 1 : lược đồ quan hệ Q ở dạng chuẩn BC nếu :
  - Q ở dạng chuẩn 2.
  - Mọi thuộc tính của Q không phụ thuộc bắc cầu vào khóa của Q
    - khác dạng chuẩn 3 ở chỗ mọi thuộc tính, trong khi dạng chuẩn 3 chỉ là những thuộc tính không khóa.

# Dạng chuẩn BC – BCNF

- Định nghĩa 2 : lược đồ quan hệ Q ở dạng chuẩn BC nếu :
  - Q ở dạng chuẩn 2.
  - Mọi phụ thuộc hàm không hiển nhiên đều có vế trái chứa khóa.
- Nhận xét:
  - Q có 1 PTH có  $(\text{vế trái})^+ = Q^+ \rightarrow \text{BCNF}$
  - Q đạt BCNF  $\rightarrow$  Q đạt 3NF

# Dạng chuẩn BC – BCNF

- **DMLop**(Tinh, MonHoc, HocPhan, GioHoc, TrTam)
- **F** = {Tinh → MonHoc, HocPhan, GioHoc, TrTam, TrTam → Tinh}

TrTam	Tinh
LTV	TpHCM
Van Kiep	AnGiang
Van Hoa	DongNai

Q1(TrTam, Tinh)

F1 = {TrTam → Tinh}

Tinh	MonHo	HocPhan	GioHoc
TpHCM	SQL	Căn bản	8:00-10:00
AnGiang	SQL	Căn bản	10:00-12:00
DongNai	SQL	Nâng cao	8:00-10:00

Q2(Tinh, MonHoc, HocPhan, GioHoc)

F1 = {Tinh → MonHoc, HocPhan, GioHoc}

# Dạng chuẩn BC – BCNF

SOHD	MAKH
HD01	KH03
HD02	KH04
HD03	KH01

$F = \{SOHD \rightarrow MAKH\}$

$F = \{SOHD, MAHH \rightarrow SL\}$

SOHD	MAHH	SL
HD01	HH02	5
HD01	HH03	10
HD02	HH01	15
HD03	HH03	30
HD03	HH07	100

MAHH	TENHH	DONGIA
HH02	Sáo trúc	2000
HH03	Cửu Âm	4000
HH01	Kentucky	10000
HH03	Cửu Âm	4000
HH07	Axit	500

MAKH	HOTENKH
KH03	Đông Tà
KH04	Bắc Cái
KH01	Tây Độc

$F = \{MAHH \rightarrow TENHH; MAHH \rightarrow DONGIA\}$   $F = \{MAKH \rightarrow HOTENKH\}$

**4 lược đồ quan hệ trên đều đạt BCNF?**



# Nội dung

---

- Giới thiệu
- Các dạng chuẩn
  - 1NF
  - 2NF
  - 3NF
  - BCNF
- **Phương pháp phân rã**
- Phương pháp kiểm tra bảo toàn thông tin

## \$6.3. Phương pháp phân rã

### ■ **Định lý Delobel (1973):**

- Giả sử có một lược đồ quan hệ  $Q(XYZ)$  và tập phụ thuộc hàm  $F$ .
- Nếu  $X \longrightarrow Y \in F^+$  thì phép phân rã  $Q$  thành 2 lược đồ quan hệ con  $Q_1(XY)$  và  $Q_2(XZ)$  là bảo toàn thông tin khi:

- $\forall TQ: \quad TQ = TQ [XY] \bowtie TQ [XZ]$

### ■ **Lược đồ quan hệ phổ quát:** là lược đồ chứa tất cả các thuộc tính của ứng dụng

- Vậy ta sẽ có 1 quan hệ phổ quát, nó chính là 1 thể hiện của lược đồ quan hệ phổ quát

- ***Mô tả:***

- *Input* : lược đồ quan hệ phổ quát  $R$  và tập phụ thuộc hàm  $F$ .
- *Output* : phân rã  $C = (R_1, R_2, \dots, R_m)$  thỏa:
  - Mọi  $R_i$  đều đạt BCNF.
  - Phân rã bảo toàn thông tin.



# Phương pháp phân rã nổi không mất thành BCNF

- Phương pháp:
  - Tìm những phụ thuộc hàm làm cho lược đồ quan hệ vi phạm dạng chuẩn 3.
  - Trong những phụ thuộc hàm vừa có xác định phụ thuộc hàm **X** thỏa yêu cầu: nếu bỏ phụ thuộc hàm này đi thì số lượng phụ thuộc hàm bị mất là ít nhất.
  - Dựa trên phụ thuộc hàm **X**, tạo lược đồ quan hệ mới với:
    - Các thuộc tính: là những thuộc tính có trong phụ thuộc hàm **X**.
    - Tập phụ thuộc hàm: chỉ gồm 1 phụ thuộc hàm **X**.
  - Trong lược đồ quan hệ cũ trước đó:
    - Tập thuộc tính: loại bỏ những thuộc tính có trong vế phải của phụ thuộc hàm **X**.
    - Tập phụ thuộc hàm: loại bỏ phụ thuộc hàm **X**.

■ Ví dụ:

- Cho lược đồ quan hệ  $Q$  ( $ABCDEG$ ) và tập phụ thuộc hàm :

$$F = \{A \rightarrow C, BC \rightarrow D, C \rightarrow G, D \rightarrow E\}$$

- Cho lược đồ quan hệ  $Q$  ( $ABCDEG$ ) và tập phụ thuộc hàm  $F$ :

$$F = \{AE \rightarrow C, CG \rightarrow A, BD \rightarrow G, BDA \rightarrow E\}$$

#### ■ **Mô tả:**

- *Input* : lược đồ quan hệ phổ quát  $R$  và tập phụ thuộc hàm  $F$ .
- *Output* : phân rã  $C = (R_1, R_2, \dots, R_m)$  thỏa:
  - $R_i$  đạt tối thiểu dạng chuẩn 3.
  - Phân rã bảo toàn thông tin.
  - Phân rã bảo toàn phụ thuộc hàm.

- Bước 1 : tìm 1 phủ tối tiểu ( có thể có nhiều phủ tối tiểu).
- Bước 2 : Chia phủ tối tiểu thành những nhóm **Si** sao cho trong mỗi nhóm chỉ chứa các phụ thuộc hàm có cùng vế trái. Mỗi vế trái ký hiệu là **Ki**, đó là siêu khóa của những quan hệ con tương lai.
- Bước 3 :
  - 3.1 : Gộp các **Si** có các siêu khóa tương đương lại thành cùng 1 nhóm. Bước này nên tính bằng cách lập các bao đóng từ các vế trái.
  - 3.2 : Gộp các **Si** có các phụ thuộc hàm dạng **Ki**→**Km** và **Km**→**Ki**.

#### ■ Bước 4 :

- 4.1 : Anh xạ mỗi ***Si*** thành một lược đồ quan hệ.
- 4.2 : *Nếu trong tập các lược đồ quan hệ (lđqh) vừa tạo không có lđqh chứa đủ các thuộc tính khóa của lđqh Q ban đầu, ta tạo thêm 1 lđqh nữa. Lđqh này sẽ bao gồm tập các thuộc tính khóa của lđqh Q (ban đầu).*

#### ■ Ví dụ:

- Cho  $R(ABCD)$ ;  $F = \{A \rightarrow B, A \rightarrow C, C \rightarrow D, B \rightarrow C, B \rightarrow A\}$ , Khóa =  $\{A\}$ 
  1. Phủ tối thiểu  $G = \{A \rightarrow B, C \rightarrow D, B \rightarrow C, B \rightarrow A\}$
  2. Chia  $G$  thành 4 nhóm
    - $S1: \{A \rightarrow B\}$
    - $S2: \{C \rightarrow D\}$
    - $S3: \{B \rightarrow C\}$
    - $S4: \{B \rightarrow A\}$

■ Ví dụ:

3. Gộp  $S1$  và  $S4$  vì hai nhóm này có phụ thuộc hàm dạng  $Ki \rightarrow Km$  và  $Km \rightarrow Ki$ :

- $S1: \{ A \rightarrow B, B \rightarrow A \}$
- $S2: \{ C \rightarrow D \}$
- $S3: \{ B \rightarrow C \}$

4. Ánh xạ mỗi nhóm  $S_i$  thành một lược đồ quan hệ con như sau:

- $R1(AB); F1 = \{ A \rightarrow B, B \rightarrow A \}$  đạt BCNF
- $R2(CD); F2 = \{ C \rightarrow D \}$  đạt BCNF
- $R3(BC); F3 = \{ B \rightarrow C \}$  đạt BCNF

## Các yếu tố cần bảo đảm khi thực hiện phân rã

- Đạt được dạng chuẩn mong muốn : trong thực tế, người ta chấp nhận 3NF hoặc BCNF.
- Bảo toàn thuộc tính :  $R^+ = R_1^+ \cup R_2^+ \dots \cup R_m^+$ .
- Bảo toàn phụ thuộc hàm:  $F^+ = (F_1 \cup F_2 \dots \cup F_n)^+$ .
- Bảo toàn thông tin : là dựa trên các quan hệ mới, ta có thể phục hồi lại được quan hệ cũ bằng cách dùng phép kết tự nhiên trên các quan hệ mới.



## Phép kiểm tra nổi bảo toàn thông tin

- ***B1: Xây dựng một bảng*** gồm  **$n+1$  cột** ( $n$  là số lượng thuộc tính của quan hệ gốc) và  **$m+1$  dòng** ( $m$  là số lượng quan hệ được phân rã ra từ quan hệ gốc).
  - Cột đầu tiên chứa các quan hệ của phép phân rã. Các cột còn lại được đặt tên là  $A_1, A_2, \dots, A_n$  tương ứng các thuộc tính
  - Dòng đầu tiên chứa tên các thuộc tính, các dòng còn lại được đặt là  $R_1, R_2, \dots, R_m$  tương ứng các quan hệ đã phân rã

## Phép kiểm tra nổi bảo toàn thông tin

- ***B2: Điền giá trị vào bảng*** theo cách sau:
  - Đối với từng thuộc tính  $A_i$ , kiểm tra xem thuộc tính này có phải là thuộc tính của  $R_J$  hay không?
    - Nếu thuộc tính  $A_i$  có trong quan hệ  $R_J$  thì ta điền giá trị  $a_i$  vào ô  $(A_i, R_J)$  của bảng
    - Ngược lại, nếu không có thì ta điền giá trị  $b_i$  vào đó.

## Phép kiểm tra nổi bảo toàn thông tin

- ***B3: Thay đổi các giá trị có trong bảng:***
  - Khi bắt đầu bước này, đặt giá trị của biến *CoThayDoi* = *False*.
  - Điều kiện dừng của ***B3***:
    - Hoặc có một dòng toàn giá trị a.
    - Hoặc sau khi xét tất cả các phụ thuộc hàm mà giá trị của biến *CoThayDoi* = *False*.
  - 3.1. Nếu có hơn 1 dòng có cùng giá trị tại X, giá trị tại Y của các dòng này sẽ là  $a_j$  nếu  $a_j$  có xuất hiện tại Y, ngược lại lấy b nhỏ nhất. Gán *CoThayDoi*=*True*.
  - 3.2. Nếu thỏa điều kiện dừng, qua B4. Ngược lại, xét PTH kế tiếp, quay lại B3.1. Nếu không còn PTH nào, quay lại B3.

## Phép kiểm tra nối bảo toàn thông tin

- **B4:** Kiểm tra tất cả các dòng của bảng.
  - Nếu có bất kỳ dòng nào chỉ chứa các giá trị  $a$  (tức là đủ các giá trị từ  $a_1$  đến  $a_n$ ) thì phép phân rã là bảo toàn thông tin.
  - Ngược lại, phép phân rã là không bảo toàn thông tin
- Xét quan hệ  $R(V, X, Y, Z, W)$ , có tập  $F = \{X \rightarrow Z, Y \rightarrow Z, Z \rightarrow W, WV \rightarrow Z, ZV \rightarrow X\}$ 
  - $R$  được phân rã thành  $R1(X, W), R2(X, Y), R3(Y, V), R4(Z, W, V), R5(X, V)$ .
  - Thực hiện kiểm tra bằng phép nối bảo toàn