

## 3.3 Mô hình hóa dữ liệu

### (Phân tích hệ thống về dữ liệu)

---

Viện Công nghệ thông tin & truyền thông  
Đại học Bách Khoa Hà Nội

# Một số định nghĩa

---

## □ **Mô hình dữ liệu**

- Hình thức hóa biểu diễn dữ liệu được tạo và sử dụng bởi hệ thống
- Gồm dữ liệu và các ràng buộc, liên hệ

## □ **Mô hình dữ liệu logic**

- Mức tổ chức dữ liệu

## □ **Mô hình dữ liệu vật lý**

- Mức lưu trữ dữ liệu

## □ **Chuẩn hóa**

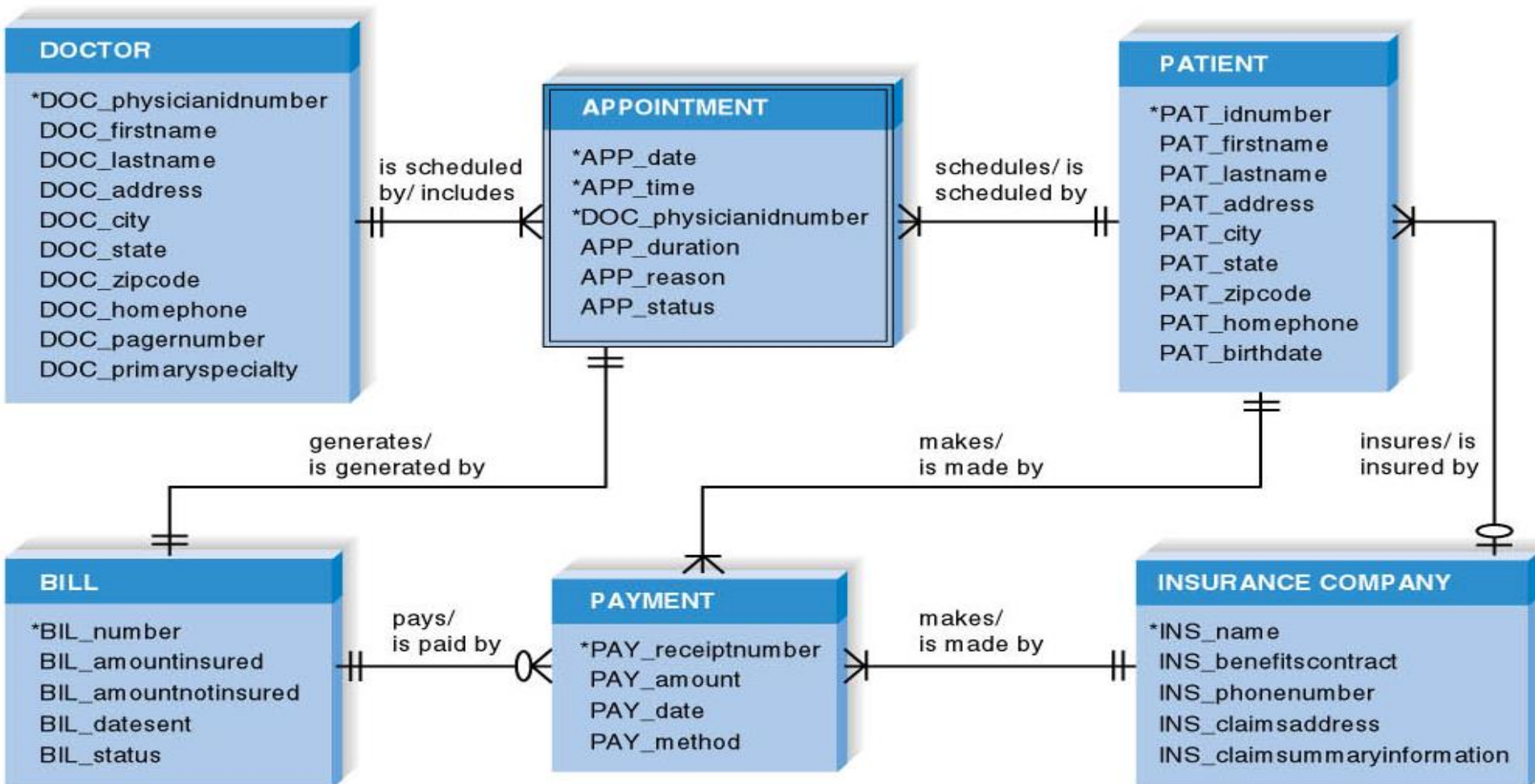
- Quá trình phân tích dữ liệu để xây dựng được cấu trúc phù hợp

## ❖ Cân bằng giữa mô hình dữ liệu và mô hình nghiệp vụ




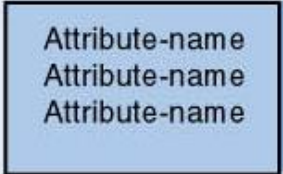

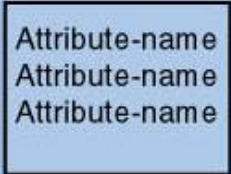
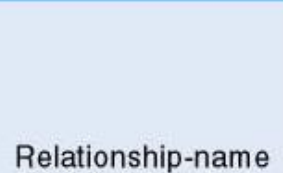
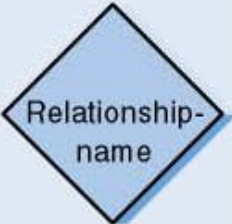

# Mô hình dữ liệu thực thể - liên kết

- Biểu diễn: bằng sơ đồ thực thể - liên kết
- Các khái niệm cơ bản
  - **Thực thể**: một đối tượng trong thế giới thực, tồn tại độc lập và phân biệt được với các đối tượng khác
  - **Tập thực thể**: gồm các thực thể có tính chất giống nhau
  - **Thuộc tính**: một đặc tính của một tập thực thể
  - **Khoá**:
  - **Liên kết**: một mối liên hệ có nghĩa giữa nhiều thực thể
    - Mỗi liên kết có thể có các thuộc tính
    - 1-1, 1-n, n-m, đệ quy
  - **Tập liên kết**: một tập hợp các liên kết cùng kiểu

# Sơ đồ thực thể liên kết

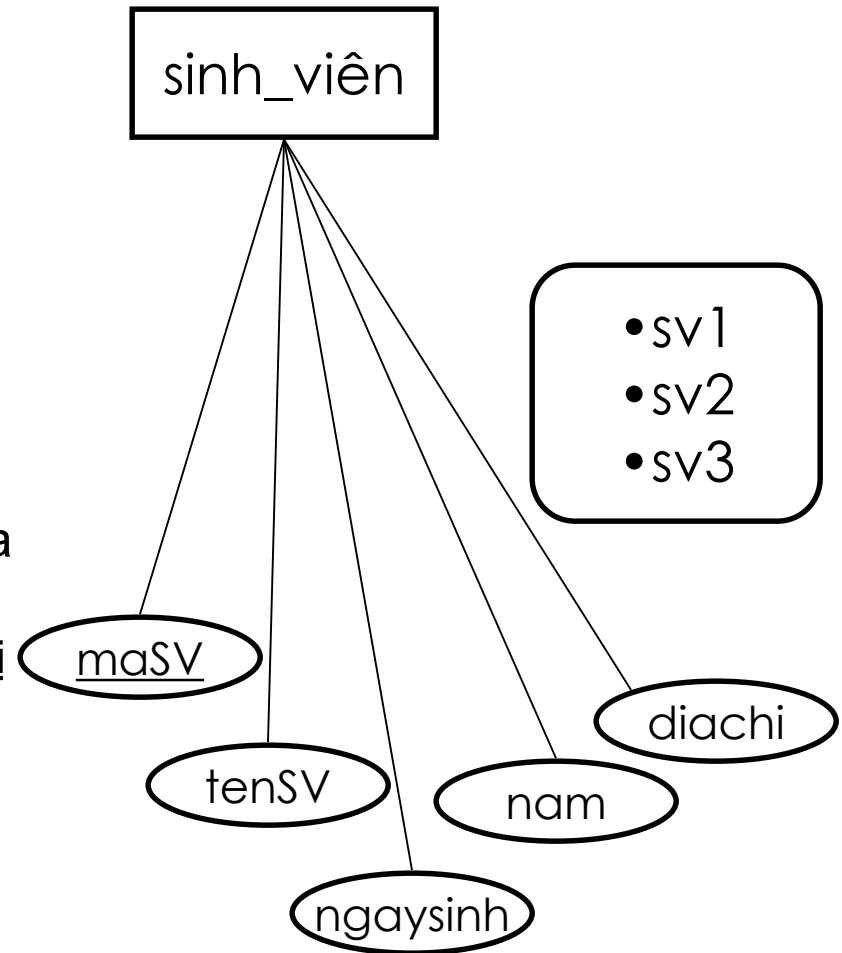


# Quy ước

	IDEF1X	Chen	Crow's Foot
<b>An ENTITY:</b> <ul style="list-style-type: none"> <li>✓ Is a person, place, or thing</li> <li>✓ Has a singular name spelled in all capital letters</li> <li>✓ Has an identifier</li> <li>✓ Should contain more than one instance of data</li> </ul>	ENTITY-NAME 	ENTITY-NAME 	ENTITY-NAME 
<b>An ATTRIBUTE:</b> <ul style="list-style-type: none"> <li>✓ Is a property of an entity</li> <li>✓ Should be used by at least one business process</li> <li>✓ Is broken down to its most useful level of detail</li> </ul>	ENTITY-NAME 		ENTITY-NAME 
<b>A RELATIONSHIP:</b> <ul style="list-style-type: none"> <li>✓ Shows the association between two entities</li> <li>✓ Has a parent entity and a child entity</li> <li>✓ Is described with a verb phrase</li> <li>✓ Has cardinality (1 : 1, 1 : N, or M : N)</li> <li>✓ Has modality (null, not null)</li> <li>✓ Is dependent or independent</li> </ul>	<u>Relationship-name</u> 		<u>Relationship-name</u> 

# Thực thể và thuộc tính

- **Thực thể**: một đối tượng trong thế giới thực
- **Tập thực thể**: gồm các thực thể có tính chất giống nhau
- **Thuộc tính**: một đặc tính của một tập thực thể
  - **Miền giá trị** ~ tập các giá trị có thể
  - **Khoá** ~ xác định sự duy nhất của 1 thực thể

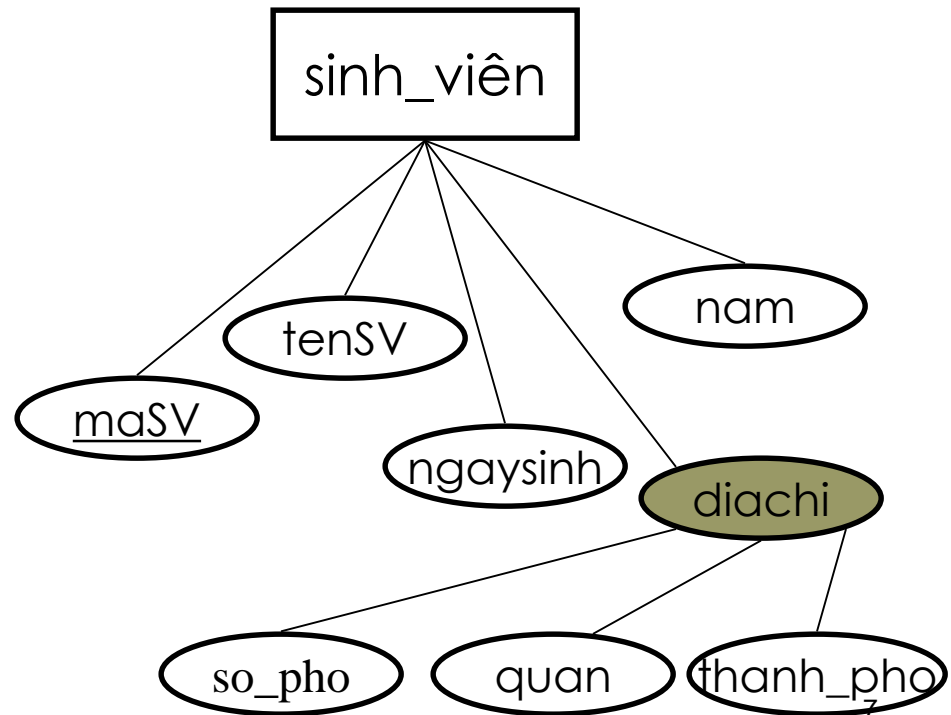


# Kiểu thuộc tính

- Thuộc tính đơn giản (thuộc tính nguyên tố)
  - có kiểu dữ liệu nguyên tố

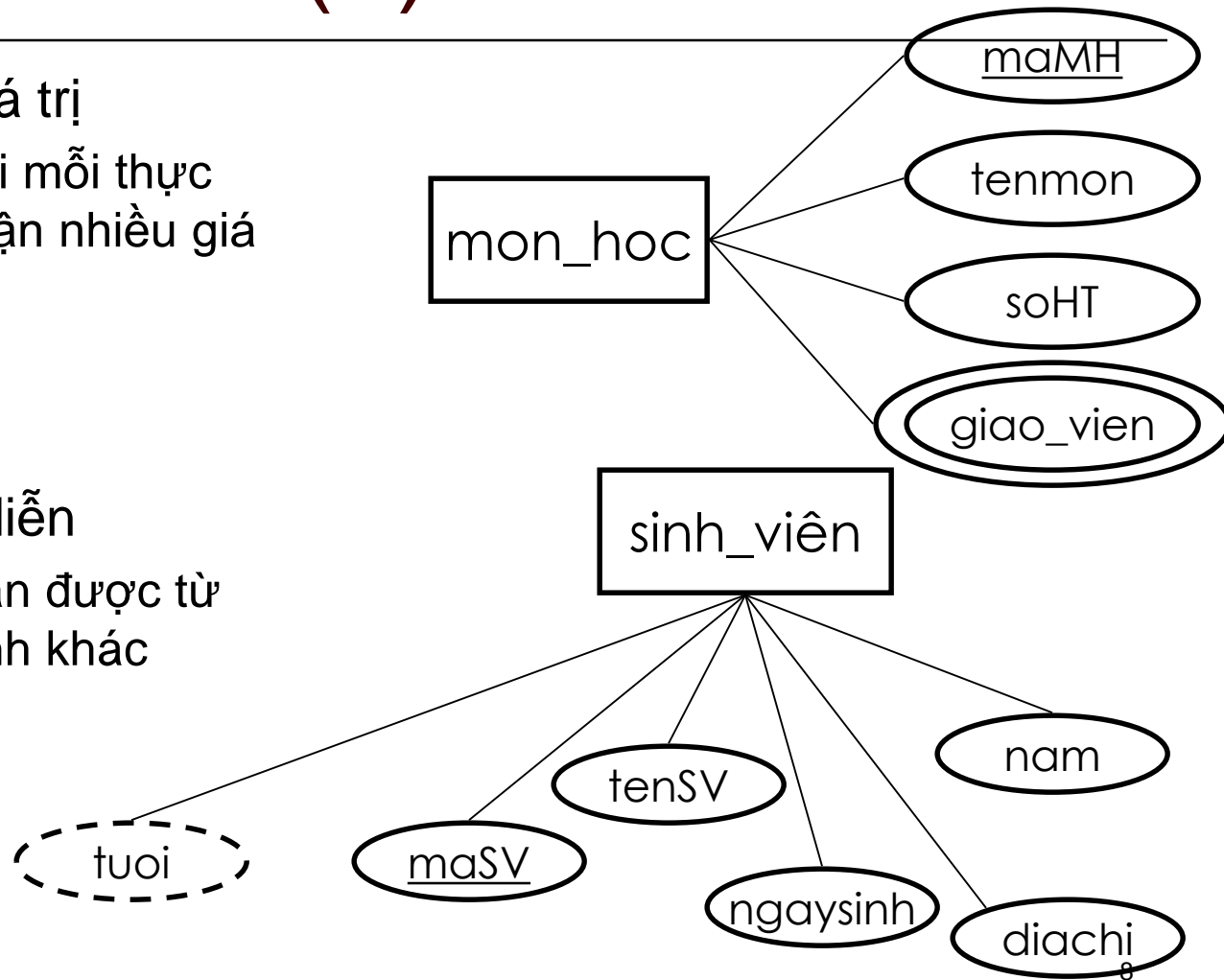
tenSV = “Trần T. Bình”  
= “Ng. Đ. Trung”

- Thuộc tính phức
  - có kiểu phức, định nghĩa bởi các thuộc tính khác



# Kiểu thuộc tính (2)

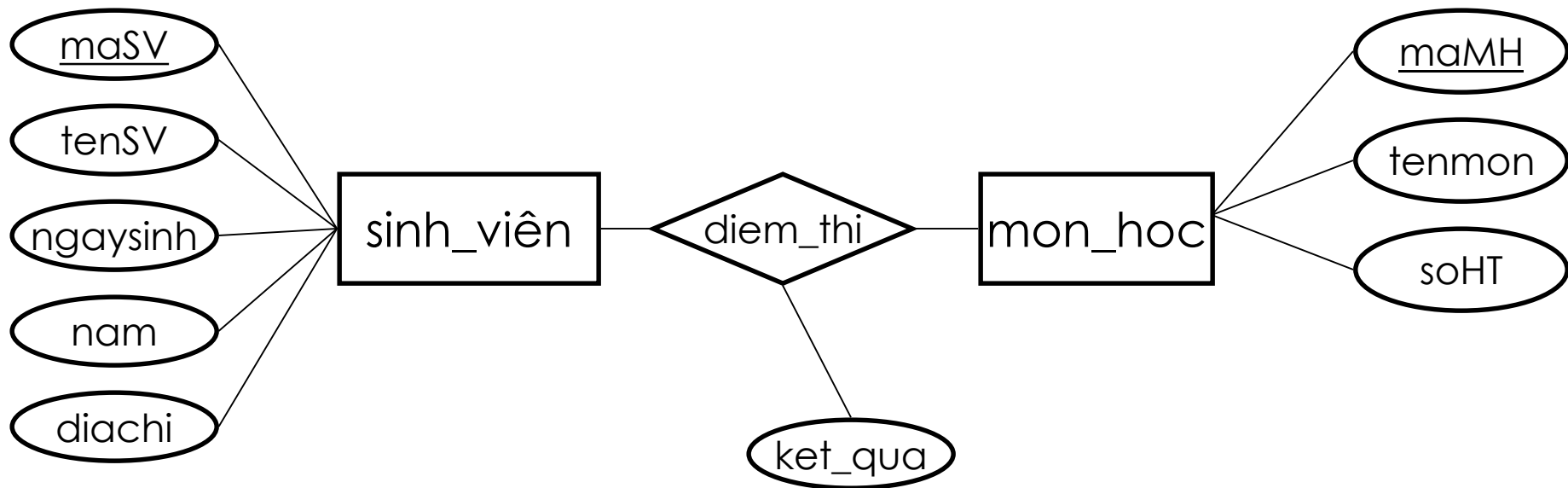
- Thuộc tính đa giá trị
  - tương ứng với mỗi thực thể, có thể nhận nhiều giá trị
- Thuộc tính suy diễn
  - có thể tính toán được từ (các) thuộc tính khác





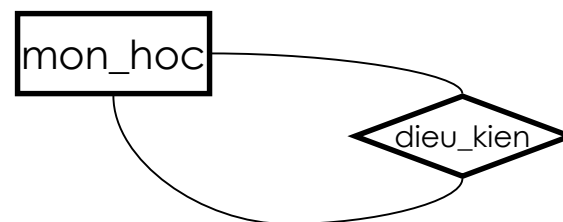
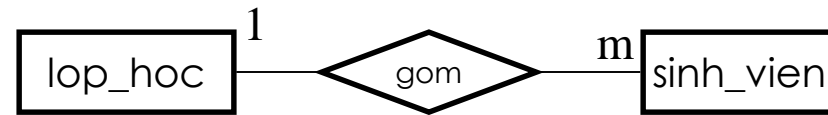
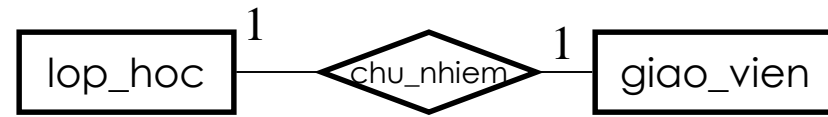
# Liên kết

- ❑ Đ/n: là sự kết hợp giữa một số thực thể
- ❑ Thuộc tính



# Ràng buộc của kết nối

- **1-1**: Liên kết 1 thực thể của một tập thực thể với nhiều nhất 1 thực thể của tập thực thể khác
- **1-n**: Liên kết 1 thực thể của một tập thực thể với nhiều thực thể của tập thực thể khác
- **n-m**: Liên kết 1 thực thể của một tập thực thể với nhiều thực thể của tập thực thể khác và ngược lại
- **đệ quy**: Liên kết giữa các thực thể cùng kiểu



# Từ điển dữ liệu và siêu dữ liệu

---

## □ Siêu dữ liệu

- thông tin mô tả về các thành phần trong sơ đồ dữ liệu
- Lưu trữ trong từ điển dữ liệu, được chia sẻ giữa các (nhóm) người phát triển hệ thống

## □ Từ điển dữ liệu

- Giúp cải thiện chất lượng hệ thống đang xây dựng

# Một số lưu ý khi thiết kế sơ đồ TTLK

---

- ❑ Tập thực thể chỉ gồm 1 thực thể  
→ không cần biểu diễn
- ❑ Tên tập thực thể nên trùng với tên đối tượng nó biểu diễn
- ❑ Nếu 1 (số) thuộc tính của tập thực thể là không quan trọng cho ứng dụng → nên loại bỏ
- ❑ Tránh đặt tên trùng nhau cho các thuộc tính của các tập thực thể khác nhau → có thể đặt EntName\_attrName
- ❑ Nên phân chia các thuộc tính riêng rẽ. Ví dụ: Họ, đệm, tên nếu cần thao tác đến các thuộc tính riêng rẽ này
- ❑ Xác nhận, kiểm tra các ràng buộc, liên kết

# Sơ đồ thực thể - liên kết

## → Sơ đồ quan hệ

---

- Biến đổi tập các thực thể
- Biến đổi các liên kết
  - Liên kết 1-1
  - Liên kết 1-n
  - Liên kết n-m
- Các khoá của các sơ đồ quan hệ
- Các sơ đồ quan hệ với khoá chung

# Quá trình chuẩn hóa

---

## □ Vấn đề đặt ra

- Có cần phải tinh chỉnh thiết kế nữa hay không?
- Thiết kế đã là tốt hay chưa?
- Định nghĩa về các dạng chuẩn.

## □ Mục đích:

Mỗi dạng chuẩn đảm bảo ngăn ngừa (giảm thiểu) một số các dạng dư thừa hay dị thường dữ liệu

## □ Các dạng chuẩn hay sử dụng

- Dạng chuẩn 1 (1NF)
- Dạng chuẩn 2 (2NF)
- Dạng chuẩn 3 (3NF)
- Dạng chuẩn Boye-Code (BCNF)
- Dạng chuẩn 4 (4NF)

