



# Thiết kế mô hình ERD

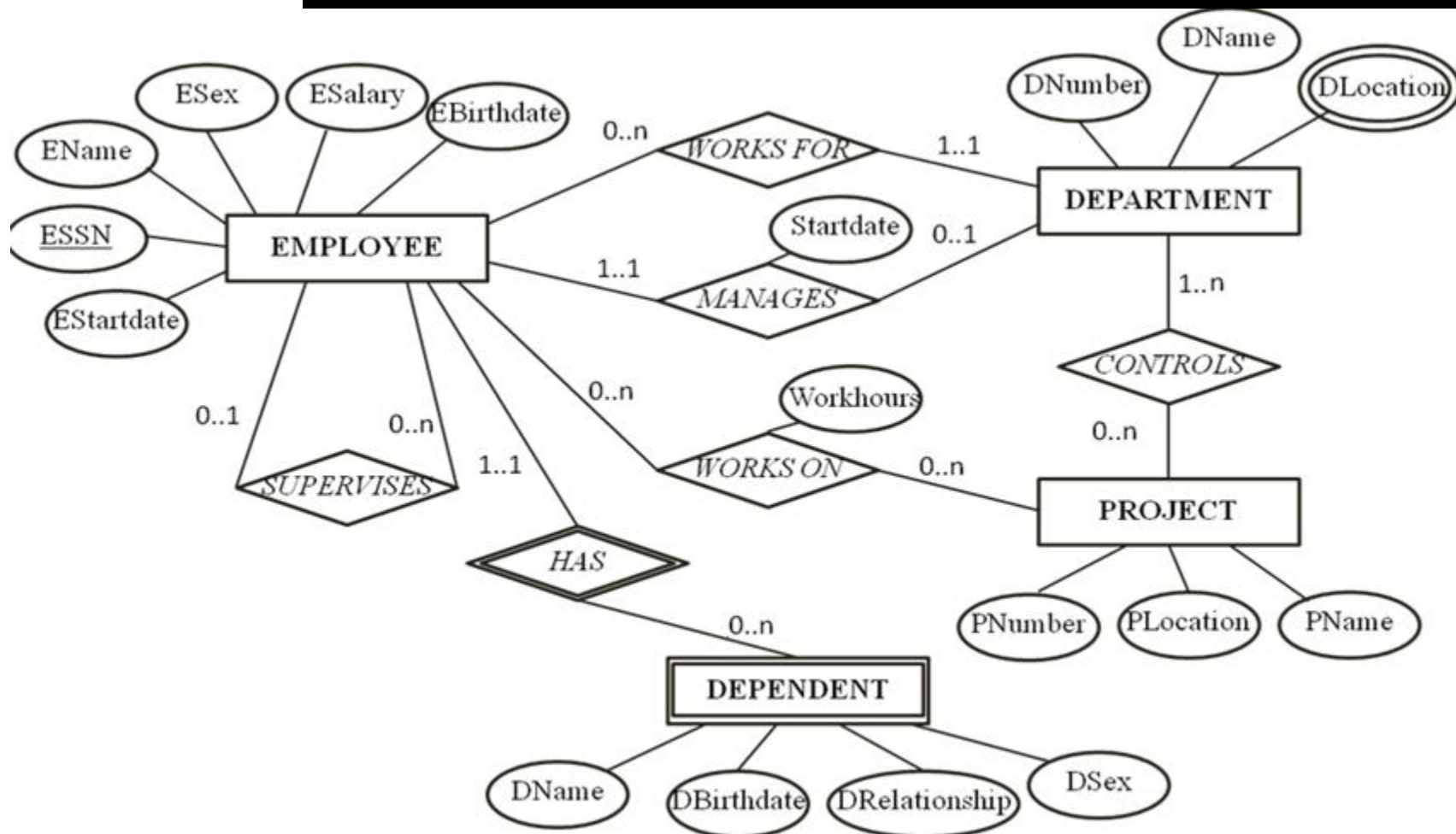
- Cho một cơ sở dữ liệu phục vụ cho một hội nghị khoa học.
- Trong đó, các tác giả sẽ gửi những bài báo của mình cho hội nghị. Thông tin tác giả bao gồm địa chỉ email (định danh), họ và tên, bộ môn, khoa, trường công tác, học vị, chức vị.
- Mỗi bài báo được gửi lên hội nghị sẽ được gán một mã số, và có những thông tin như tựa đề, abstract, tên file lưu trữ. Một bài báo có thể có nhiều tác giả nhưng chỉ có một tác giả có vai trò là tác giả chính.
- Hội nghị sẽ phân công một số nhà khoa học tham gia đánh giá các bài báo. Một bài báo sẽ được nhiều người đánh giá. Thông tin người đánh giá bao gồm địa chỉ email (định danh), họ, tên, số điện thoại liên lạc, học vị, chức vị, và những hướng nghiên cứu.
- Khi đánh giá bài báo, hệ thống cần lưu lại: câu hỏi mà người đánh giá đặt ra cho bài báo, phần tóm tắt bài báo, nhận xét chung, điểm cho chất lượng của bài báo.



- Công ty được tổ chức thành nhiều đơn vị với các thông tin bao gồm: tên đơn vị, số hiệu đơn vị. Mỗi đơn vị có thể có nhiều văn phòng đại diện tại những địa điểm khác nhau. Đơn vị được quản lý bởi một nhân viên giữ chức vụ trưởng đơn vị. Hệ thống cũng yêu cầu phải duy trì thông tin về ngày ký quyết định bổ nhiệm cho từng trưởng đơn vị.
- Mỗi đơn vị điều hành nhiều dự án. Thông tin về dự án bao gồm: tên dự án, số hiệu, và địa điểm thực hiện dự án.
- Công ty duy trì thông tin về nhân viên bao gồm: họ và tên, mã số, mức lương, giới tính, ngày sinh và ngày ký hợp đồng làm việc với công ty. Mỗi nhân viên thuộc biên chế một đơn vị nhưng có thể tham gia nhiều dự án khác nhau. Các dự án này không nhất thiết phải do cùng một đơn vị quản lý. Công ty cũng theo dõi thời gian (tính bằng số giờ làm việc trong tuần) tham gia từng dự án của mỗi nhân viên. Ngoài ra, mỗi nhân viên còn chịu sự giám sát của một nhân viên khác.
- Cuối cùng, công ty có chế độ bảo hiểm dành cho người thân của nhân viên. Thông tin về người thân bao gồm tên, giới tính, ngày sinh và mối liên hệ với nhân viên.



# VÍ DỤ MẪU



Biểu diễn cơ sở dữ liệu **COMPANY** bằng mô hình thực thể kết hợp.



# CƠ SỞ DỮ LIỆU

## CHƯƠNG 4: MÔ HÌNH DỮ LIỆU QUAN HỆ

**Tham khảo tài liệu [A]:  
Chapter 3: The Relational Model**





# MÔ HÌNH DỮ LIỆU QUAN HỆ

1

- Tổ chức dữ liệu

2

- Ràng buộc trên dữ liệu

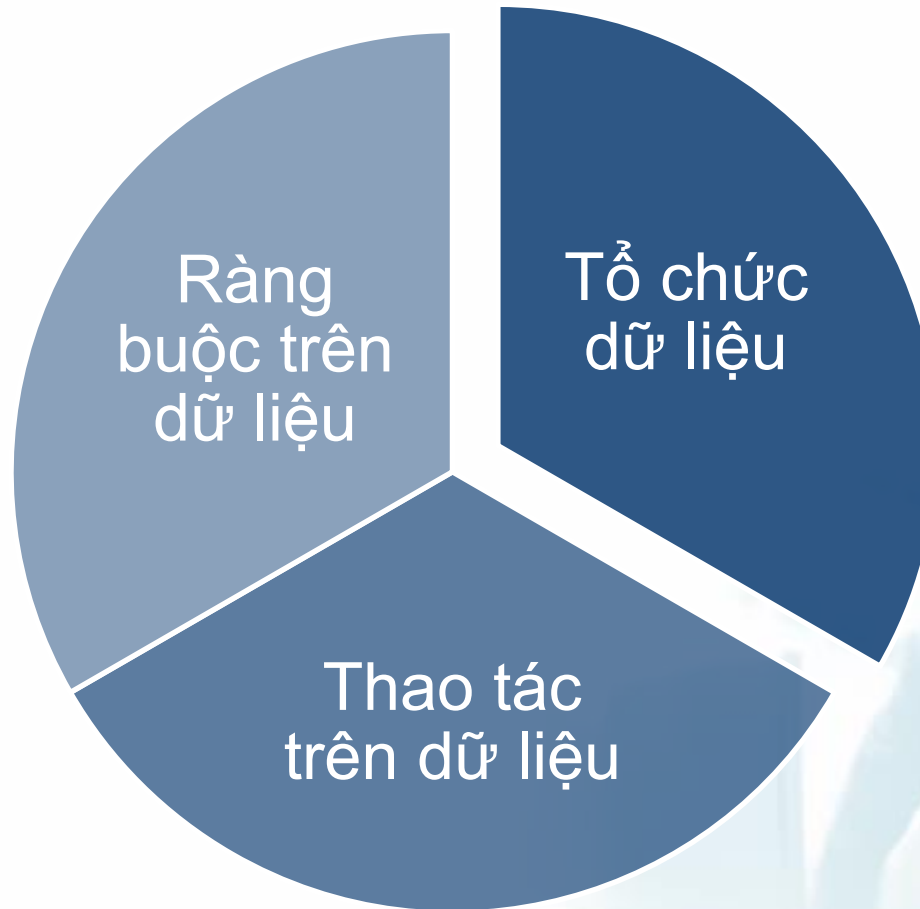
3

- Thao tác trên dữ liệu





# BA THÀNH PHẦN CỦA MÔ HÌNH DỮ LIỆU QUAN HỆ







# CÁC KHÁI NIỆM CƠ BẢN CỦA MÔ HÌNH DỮ LIỆU QUAN HỆ

- Lược đồ quan hệ
- Bậc của quan hệ
- Thuộc tính
- Miền giá trị
- Bộ dữ liệu
- Thể hiện quan hệ
- Lược đồ cơ sở dữ liệu quan hệ



# VÍ DỤ MẪU

Cơ sở dữ liệu - Mô hình dữ liệu quan hệ

Thuộc tính

ESSN	ENAME	ESALARY	ESSEX	EBIRTHDATE	ESTARTDATE	DNUM	SUPERVISORSSN
30121050004	Mai Duy An	30000	F	1968-02-17...	2000-01...	1	30121050037
30121050015	Huỳnh Mai Anh	58000	F	1963-05-04...	2005-01...	1	30121050004
30121050027	Nguyễn Thúy ...	91000	F	1977-04-10...	2005-02...	1	30121050004
30121050035	Tống Thị Lan A...	78000	F	1978-10-28...	2004-04...	1	30121050037
30121050037	Võ Việt Anh	110000	M	1974-11-11...	2000-05...	1	NULL

Bộ dữ liệu

Giá trị thành phần

Hình 4.1. Một thể hiện của quan hệ **EMPLOYEE**



## WORKSON

🔑	ESSN
🔑	PNum
	workHours

## PROJECT

🔑	PNumber
	PName
	PLocation
	DNum

## EMPLOYEE

🔑	ESSN
	EName
	ESalary
	ESex
	EBirthdate
	EStartdate
	DNum
	supervisorSSN

## DEPENDENT

🔑	DName
🔑	ESSN
	DSex
	DBirthdate
	DRelationship

## DEPARTMENT

🔑	DNumber
	DName
	mgrSSN
	mgrStartdate

## DEPLOYMENT

🔑	DNum
🔑	DLocation

Hình 4.2. Lược đồ cơ sở dữ liệu **COMPANY**



# MỘT SỐ ĐẶC TRƯNG CỦA MÔ HÌNH DỮ LIỆU QUAN HỆ

## ■ Thứ tự của các bộ trong quan hệ

ESSN	ENAME	ESALARY	ESSEX	EBIRTHDATE	ESTARTDATE	DNUM	SUPERVISORSSN
30121050004	Mai Duy An	30000	F	1968-02-17...	2000-01...	1	30121050037
30121050015	Huỳnh Mai Anh	58000	F	1963-05-04...	2005-01...	1	30121050004
30121050027	Nguyễn Thúy Qu...	91000	F	1977-04-10...	2005-02...	1	30121050004
30121050035	Tống Thị Lan Anh	78000	F	1978-10-28...	2004-04...	1	30121050037
30121050037	Võ Việt Anh	110000	M	1974-11-11...	2000-05...	1	NULL

ESSN	ENAME	ESALARY	ESSEX	EBIRTHDATE	ESTARTDATE	DNUM	SUPERVISORSSN
30121050004	Mai Duy An	30000	F	1968-02-17...	2000-01...	1	30121050037
30121050015	Huỳnh Mai Anh	58000	F	1963-05-04...	2005-01...	1	30121050004
30121050035	Tống Thị Lan Anh	78000	F	1978-10-28...	2004-04...	1	30121050037
30121050027	Nguyễn Thúy Qu...	91000	F	1977-04-10...	2005-02...	1	30121050004
30121050037	Võ Việt Anh	110000	M	1974-11-11...	2000-05...	1	NULL

Hình 4.3. Hai thể hiện của quan hệ **EMPLOYEE**

**TƯƠNG ĐƯƠNG**



# MỘT SỐ ĐẶC TRƯNG CỦA MÔ HÌNH DỮ LIỆU QUAN HỆ

## ■ Thứ tự của các giá trị bên trong một bộ

ESSN	ENAME	ESALARY	ESSEX	EBIRTHDATE	ESTARTDATE	DNUM	SUPERVISORSSN
30121050004	Mai Duy An	30000	F	1968-02-17...	2000-01-...	1	30121050037
30121050015	Huỳnh Mai Anh	58000	F	1963-05-04...	2005-01-...	1	30121050004
30121050027	Nguyễn Thúy Q...	91000	F	1977-04-10...	2005-02-...	1	30121050004
30121050035	Tống Thị Lan Anh	78000	F	1978-10-28...	2004-04-...	1	30121050037
30121050037	Võ Việt Anh	110000	M	1974-11-11...	2000-05-...	1	NULL

ENAME	ESSEX	ESALARY	EBIRTHDATE	ESTARTDATE	DNUM	SUPERVISORSSN	ESSN
Mai Duy An	F	30000	1968-02-17...	2000-01-...	1	30121050037	30121050004
Huỳnh Mai Anh	F	58000	1963-05-04...	2005-01-...	1	30121050004	30121050015
Tống Thị Lan Anh	F	78000	1978-10-28...	2004-04-...	1	30121050037	30121050035
Nguyễn Thúy Q...	F	91000	1977-04-10...	2005-02-...	1	30121050004	30121050027
Võ Việt Anh	M	110000	1974-11-11...	2000-05-...	1	NULL	30121050037

Hình 4.3. Hai thể hiện của quan hệ **EMPLOYEE**

**TƯƠNG ĐƯƠNG**



# MỘT SỐ ĐẶC TRƯNG CỦA MÔ HÌNH DỮ LIỆU QUAN HỆ

## ■ Các giá trị thuộc tính trong bộ

Giá trị thuộc tính là giá trị đơn

ESSN	ENAME	ESALARY	ESSEX	EBIRTHDATE	ESTARTDATE	DNUM	SUPERVISORSSN
30121050004	Mai Duy An	30000	F	1968-02-17...	2000-01...	1	30121050037
30121050015	Huỳnh Mai Anh	58000	F	1963-05-04...	2005-01...	1	30121050004
30121050027	Nguyễn Thúy ...	91000	F	1977-04-10...	2005-02...	1	30121050004
30121050035	Tống Thị Lan A...	78000	F	1978-10-28...	2004-04...	1	30121050037
30121050037	Võ Việt Anh	110000	M	1974-11-11...	2000-05...	1	NULL

Sử dụng NULL để biểu diễn giá trị không biết



# MỘT SỐ ĐẶC TRƯNG CỦA MÔ HÌNH DỮ LIỆU QUAN HỆ

- Quan hệ biểu diễn thực thể và quan hệ biểu diễn liên kết

Quan hệ biểu diễn thực thể

**EMPLOYEE** (ESSN, EName, ESalary, ..., DNum, supervisorSSN)

**PROJECT** (PNumber, PName, PLocation, DNum)

**WORKSON** (ESSN, PNum, workHours)

Quan hệ biểu diễn liên kết





# RÀNG BUỘC TOÀN VỆN DỮ LIỆU

## Ràng buộc miền giá trị

- Giá trị ứng với thuộc tính  $A$  phải là đơn trị và phải thuộc miền giá trị  $\text{dom}(A)$

## Ví dụ

Mọi bộ trong quan hệ **EMPLOYEE** có giá trị tại  $ESex$  là 'M', 'F', hoặc *null*



# RÀNG BUỘC TOÀN VỆN DỮ LIỆU

## ■ Ràng buộc khóa

- Hai bộ bất kỳ của quan hệ không có giá trị bằng nhau tại **thuộc tính khóa**

## ■ Định nghĩa

- $SK$  là siêu khóa của quan hệ  $R$  khi hai bộ bất kỳ của  $R$  không có giá trị bằng nhau tại  $SK$
- $K$  là khóa của quan hệ  $R$  khi
  - $K$  là siêu khóa
  - $K$  là siêu khóa nhỏ nhất





ESSN	ENAME	ESALARY	ESSEX	EBIRTHDATE	ESTARTDATE	DNUM	SUPERVISORSSN
30121050004	Mai Duy An	30000	F	1968-02-17...	2000-01...	1	30121050037
30121050015	Huỳnh Mai Anh	58000	F	1963-05-04...	2005-01...	1	30121050004
30121050027	Nguyễn Thúy ...	91000	F	1977-04-10...	2005-02...	1	30121050004
30121050035	Tống Thị Lan A...	78000	F	1978-10-28...	2004-04...	1	30121050037
30121050037	Võ Việt Anh	110000	M	1974-11-11...	2000-05...	1	NULL

- Khóa của **EMPLOYEE** là gì?
  - {ESSN}?
  - {ENAME, ESSEX}?

1

- Xác định tất cả khóa của các quan hệ trong ví dụ mẫu

2

- Khóa nào được chọn làm khóa chính trong số các khóa của từng quan hệ, tại sao?



# RÀNG BUỘC TOÀN VẬN DỮ LIỆU

## Ràng buộc giá trị *NULL*

- Khi thuộc tính *A* được thiết lập là *NOT NULL* thì các bộ dữ liệu không được nhận giá trị *null* tại thuộc tính này

## Ví dụ

Mọi bộ trong quan hệ **EMPLOYEE** không thể nhận giá trị *null* tại thuộc tính *ENAME*



# RÀNG BUỘC TOÀN VẬN DỮ LIỆU

## Ràng buộc thực thể

- Nếu  $K$  là khóa chính của quan hệ  $R$  thì tồn tại đồng thời ràng buộc *NOT NULL* trên các thuộc tính của  $K$

## Ví dụ

Nếu  $\{ESSM\}$  là khóa chính của **EMPLOYEE** thì các bộ dữ liệu không được có giá trị *null* tại thuộc tính này



# RÀNG BUỘC TOÀN VỆN DỮ LIỆU

## Ràng buộc toàn vẹn tham chiếu

Một giá trị xuất hiện trong cột  $A$  của quan hệ  $R$  thì phải xuất hiện trong cột  $B$  của quan hệ  $S$


Để tồn tại một ràng buộc toàn vẹn tham chiếu từ  $A$  của  $R$  tới  $B$  của  $S$ , thì


$$\text{dom}(A) = \text{dom}(B)$$

$B$  phải là khóa (chính hoặc phụ) của  $S$

$$\pi_A(R) \subseteq \pi_B(S)$$


## ■ Xét hai lược đồ quan hệ


DEPARTMENT	
	DNumber
	DName
	mgrSSN
	mgrStartdate

PROJECT	
	PNumber
	PName
	PLocation
	DNum

## ■ Ràng buộc toàn vẹn tham chiếu?

## ■ Xét hai lược đồ quan hệ

EMPLOYEE	
	ESSN
	EName
	ESalary
	ESex
	EBirthdate
	EStartdate
	DNum
	supervisorSSN

DEPARTMENT	
	DNumber
	DName
	mgrSSN
	mgrStartdate

## ■ Ràng buộc toàn vẹn tham chiếu?





# BA TOÁN TỬ CẬP NHẬT TRÊN DỮ LIỆU QUAN HỆ

## Thêm mới

- Thêm mới một (số) bộ dữ liệu vào quan hệ

## Chỉnh sửa

- Cập nhật giá trị một (số) bộ dữ liệu hiện có trong quan hệ

## Xóa bỏ

- Xóa bỏ một (số) bộ dữ liệu hiện có ra khỏi quan hệ



# HOẠT ĐỘNG 2

1

- Toán tử thêm mới có khả năng gây ra vi phạm ràng buộc toàn vẹn nào?


2

- Toán tử chỉnh sửa có khả năng gây ra vi phạm ràng buộc toàn vẹn nào?

3

- Toán tử xóa bỏ có khả năng gây ra vi phạm ràng buộc toàn vẹn nào?

## EMPLOYEE

	ESSN
	ENAME
	ESALARY
	ESSEX
	EBIRTHDATE
	ESTARTDATE
	DNUM
	SUPERVISORSSN

- Khi nào RBTV trên **EMPLOYEE** bị vi phạm?



# BA CƠ CHẾ XỬ LÝ VI PHẠM RÀNG BUỘC TOÀN VỆN

1

- Cơ chế mặc định

Tự động từ chối mọi hành vi vi phạm ràng buộc toàn vẹn

2

- Cơ chế thay thế

Tự động cập nhật giá trị khóa ngoại khi giá trị khóa chính tương ứng bị thay đổi

3

- Cơ chế sử dụng giá trị null

Mọi giá trị vi phạm ràng buộc toàn vẹn được thiết lập giá trị là null



# Bài tập

Việc tổ chức kỳ thi tốt nghiệp của một khoa như sau:

- Mỗi thí sinh có một Mã số sinh viên duy nhất (MASV), mỗi MASV xác định được các thông tin: họ và tên (HOTEN), ngày sinh (NGAYSINH), nơi sinh, phái, dân tộc. Mỗi thí sinh thuộc một lớp duy nhất.
- Mỗi lớp có một mã lớp (MALOP) duy nhất, mỗi mã lớp xác định các thông tin: tên lớp (TENLOP), mỗi lớp chỉ thuộc sự quản lý của một khoa nào đó. Mỗi khoa có một mã khoa duy nhất (MAKHOA), mỗi mã khoa xác định tên khoa (TENKHOA).
- Mỗi thí sinh đều phải dự thi tốt nghiệp ba môn. Mỗi môn thi có một mã môn thi (MAMT) duy nhất, mỗi mã môn thi xác định các thông tin: tên môn thi (TENMT), thời gian làm bài – được tính bằng phút (PHUT), ngày thi (NGAYTHI), buổi thi (BUOITHI), môn thi này là môn lý thuyết hay thực hành (LYTHUYET). Chú ý rằng, nếu một môn học được cho thi ở nhiều hệ thì được đặt MAMT khác nhau (chẳng hạn cả trung cấp và cao đẳng ngành công nghệ thông tin đều thi môn Cơ Sở Dữ Liệu), để diễn tả điều này, mỗi mã môn học cần phải được ghi chú (GHICHU) để cho biết môn thi đó dành cho khối nào trung cấp, hay cao đẳng). Mỗi thí sinh ứng với một môn thi có một điểm thi (DIEMTHI) duy nhất, điểm thi được chấm theo thang điểm 10 và có lấy điểm lẻ đến 0.5. Một thí sinh được coi là đậu tốt nghiệp nếu điểm thi của tất cả các môn của thí sinh đó đều lớn hơn hoặc bằng 5.



# Bài tập

**Hãy xác định khoá cho mỗi lược đồ quan hệ trong cơ sở dữ liệu bên dưới. Và phát biểu các ràng buộc toàn vẹn có trong cơ sở dữ liệu đó.**

Việc tổ chức kỳ thi tốt nghiệp của một khoa như sau:

- Mỗi thí sinh có một Mã số sinh viên duy nhất (MASV), mỗi MASV xác định được các thông tin: họ tên (HOTEN), ngày sinh (NGAYSINH), nơi sinh, phái, dân tộc. **Mỗi sinh viên thuộc 1 lớp**
- Mỗi lớp có một mã lớp (MALOP) duy nhất, mỗi mã lớp xác định các thông tin: tên lớp (TENLOP), mỗi lớp chỉ thuộc sự quản lý của một khoa nào đó. Mỗi khoa có một mã khoa duy nhất (MAKHOA), mỗi mã khoa xác định tên khoa (TENKHOA).
- Mỗi thí sinh đều phải dự thi tốt nghiệp ba môn. Mỗi môn thi có một mã môn thi (MAMT) duy nhất, mỗi mã môn thi xác định các thông tin: tên môn thi (TENMT), thời gian làm bài – được tính bằng phút (PHUT), ngày thi (NGAYTHI), buổi thi (BUOITHI), môn thi này là môn lý thuyết hay thực hành (LYTHUYET). Chú ý rằng, nếu một môn học được cho thi ở nhiều hệ thì được đặt MAMT khác nhau (chẳng hạn cả trung cấp và cao đẳng ngành công nghệ thông tin đều thi môn Cơ Sở Dữ Liệu), để diễn tả điều này, mỗi mã môn học cần phải được ghi chú (GHICHU) để cho biết môn thi đó dành cho khối nào trung cấp, hay cao đẳng). Mỗi thí sinh ứng với một môn thi có một điểm thi (DIEMTHI) duy nhất, điểm thi được chấm theo thang điểm 10 và có lấy điểm lẻ đến 0.5. Một thí sinh được coi là đậu tốt nghiệp nếu điểm thi của tất cả các môn của thí sinh đó đều lớn hơn hoặc bằng 5.





Một thư viện tổ chức việc cho mượn sách như sau:

- Mỗi quyền sách được đánh một mã sách (MASH) dùng để phân biệt với các quyền sách khác (giả sử nếu một tác phẩm có nhiều bản giống nhau hoặc có nhiều tập thì cũng xem là có mã sách khác nhau), mỗi mã sách xác định các thông tin khác như: tên sách (TENSACH), tên tác giả (TACGIA), nhà xuất bản (NHAXB), năm xuất bản (NAMXB).
- Mỗi độc giả được thư viện cấp cho một thẻ thư viện, trong đó có ghi rõ mã độc giả (MAĐG), cùng với các thông tin khác như: họ tên (HOTEN), ngày sinh (NGAYSINH), địa chỉ (ĐIACHI), nghề nghiệp (NGHENGHIEP).
- Cứ mỗi lượt mượn sách, độc giả phải đăng ký các quyền sách cần mượn vào một phiếu mượn, mỗi phiếu mượn có một số phiếu mượn (SOPM) khác nhau, mỗi phiếu mượn xác định các thông tin như: ngày mượn sách (NGAYMUON), mã độc giả. Các quyền sách trong cùng một phiếu mượn không nhất thiết phải trả trong một lần. Mỗi quyền sách có thể thuộc nhiều phiếu mượn khác nhau (tất nhiên là tại các thời điểm khác nhau).





# Bài Tập

Cho một cơ sở dữ liệu phục vụ cho một hội nghị khoa học.

- Trong đó, các tác giả sẽ gửi những bài báo của mình cho hội nghị. Thông tin tác giả bao gồm địa chỉ email (định danh), họ và tên, bộ môn, khoa, trường công tác, học vị, chức vị.
- Mỗi bài báo được gửi lên hội nghị sẽ được gán một mã số, và có những thông tin như tựa đề, abstract, tên file lưu trữ. Một bài báo có thể có nhiều tác giả nhưng chỉ có một tác giả có vai trò là tác giả chính.
- Hội nghị sẽ phân công một số nhà khoa học tham gia đánh giá các bài báo. Một bài báo sẽ được nhiều người đánh giá. Thông tin người đánh giá bao gồm địa chỉ email (định danh), họ, tên, số điện thoại liên lạc, học vị, chức vị, và những hướng nghiên cứu.
- Khi đánh giá bài báo, hệ thống cần lưu lại những câu hỏi mà người đánh giá đặt ra cho bài báo, phần tóm tắt bài báo, nhận xét chung, điểm cho chất lượng, tính độc đáo, tính tương thích, tính trình bày, khả năng tiến cử của bài báo.



# Khung nhìn dữ liệu (Views) (1)

## ■ Quan hệ nền (Base Relation)

- Là quan hệ tương ứng với một thực thể trong lược đồ mức ý niệm
- Các bộ của quan hệ nền được lưu thật sự (mức vật lý) trong cơ sở dữ liệu

## ■ Khung nhìn dữ liệu (View)

- Là các kết quả có được nhờ sử dụng các toán tử quan hệ áp dụng trên các quan hệ nền
- Có vai trò như là các quan hệ mới



## Khung nhìn dữ liệu (Views) (2)

- Một view là một quan hệ ảo (virtual relation), quan hệ ảo đó không cần phải lưu vào cơ sở dữ liệu và chỉ được tạo ra khi có nhu cầu.
- Nội dung của một view được định nghĩa như là một truy vấn (query) trên một hay nhiều quan hệ nền
- View có tính chất động, nghĩa là các thay đổi trên các quan hệ nền sẽ ảnh hưởng lập tức lên các view có liên quan



# Mục đích của View

- Cung cấp một cơ chế dùng dữ liệu an toàn, mềm dẻo và mạnh mẽ bằng cách dấu bớt các phần không cần thiết đối với các người dùng khác nhau
- Cho phép user truy xuất dữ liệu theo cách thức có thể tùy biến được, nhờ vậy cùng một dữ liệu có thể được xem theo các cách khác nhau, bởi các user khác nhau, vào cùng một thời điểm
- Giúp đơn giản hóa việc thực hiện các phép toán trên các quan hệ nền



# View và việc cập nhật dữ liệu (1)

- Việc cập nhật dữ liệu trên các quan hệ nền sẽ ảnh hưởng tức thời đến các view có tham khảo đến các quan hệ nền đó
- Nếu view cập nhật được dữ liệu thì các quan hệ nền phía dưới cũng sẽ chịu ảnh hưởng theo
- Phân view làm 3 lớp:
  - Không cập nhật dữ liệu được (not updateable)
  - Cập nhật dữ liệu được (updateable)
  - Cập nhật dữ liệu một phần (partially updateable)



## View và việc cập nhật dữ liệu (2)

- Các giới hạn dùng khi cập nhật dữ liệu với view:
  - Cho phép cập nhật nếu việc truy vấn chỉ trên 1 quan hệ nền và đã bao gồm khóa ứng viên của quan hệ nền.
  - Không cho phép cập nhật nếu việc truy vấn liên quan đến nhiều quan hệ nền
  - Không cho phép cập nhật nếu có dùng các phép toán thống kê hay gom nhóm dữ liệu