

Chương 3

**Đại số quan hệ**

(Relational algebra)

Khoa HTTT - ĐHCT

# Nội dung

- Giới thiệu
- Một số khái niệm
- Các phép toán cơ bản
- Các phép toán khác

# Nội dung

- Giới thiệu
- Một số khái niệm
- Các phép toán cơ bản
- Các phép toán khác

# Giới thiệu

Relational Algebra



- Là ngôn ngữ hình thức cho mô hình quan hệ
- Được phát triển trước ngôn ngữ SQL
- Tập các thao tác trên mô hình quan hệ chính là ĐSQH
- Ý nghĩa và tầm quan trọng của ĐSQH:
  - Cung cấp một nền tảng cho các thao tác trên mô hình quan hệ.
  - Nó được sử dụng như là cơ sở cho việc cài đặt và tối ưu hóa các câu truy vấn
  - Một số khái niệm của nó được tích hợp vào ngôn ngữ truy vấn chuẩn SQL.

# Nội dung

- Giới thiệu
- Một số khái niệm
- Các phép toán cơ bản
- Các phép toán khác

# Ký hiệu

- Tập các thuộc tính  $U = \{A_1, A_2, \dots, A_n\}$
- $D_1, D_2, \dots, D_n$  tương ứng là miền giá trị của các thuộc tính  $A_1, A_2, \dots, A_n$
- $R, S$  là các quan hệ
- Quan hệ  $R$  gồm tập các thuộc tính  $U$ , ký hiệu  $R(U)$

# Định nghĩa quan hệ

- **Định nghĩa**
  - Quan hệ R trên tập thuộc tính  $U=\{A_1, A_2, \dots, A_n\}$  là *tập con* của tích Descartes  $D_1 \times D_2 \times \dots \times D_n$
  - Mỗi phần tử  $t=(d_1, d_2, \dots, d_n) / d_i \in D_i, i = 1, 2, \dots, n$  là một bộ của quan hệ
- **Ví dụ:** xét tập các thuộc tính  $U = \{\text{ML, tenlop}\}$ 
  - $\text{Dom}(\text{ML})= \{\text{DI1295A1, DI1295A2, DI1395A3}\} = D_1$
  - $\text{Dom}(\text{tenlop})= \{\text{HTTT 1, HTTT 2, HTTT 3 }\} = D_2$

=> tích Descartes  $D_1 \times D_2$  có  $3 \times 3$  bộ  
⇒ Có thể có nhiều quan hệ được hình thành từ 9 bộ trên. Một trong số các quan hệ đó là

ML	tenlop
DI1295A1	HTTT 1
DI1295A2	HTTT 2
DI1395A3	HTTT 3

# Đại số quan hệ

- Định nghĩa:
  - ĐSQH ký hiệu là  $\alpha=(R,O)$ 
    - R: quan hệ
    - O: là tập các phép toán quan hệ
  - Các phép toán quan hệ:
    - Đầu vào: *một hoặc nhiều* quan hệ
    - Đầu ra hay kết quả: là *một quan hệ mới* hay quan hệ kết quả
- Biểu thức quan hệ:
  - Biểu thức gồm các quan hệ trong CSDL quan hệ và các phép toán quan hệ

# Nội dung

- Giới thiệu
- Một số khái niệm
- Các phép toán cơ bản
- Các phép toán khác

# Các phép toán cơ bản

- Các phép toán một ngôi
- Các phép toán tập hợp
- Các phép toán nối kết

# Phép toán một ngôi

- Là phép toán chỉ thao tác trên một quan hệ. Đầu vào của chúng là một quan hệ và trả về kết quả là một quan hệ mới.
- Gồm các phép toán:
  - Chọn,
  - Chiếu

# Phép chọn

- Phép chọn (Selection) được sử dụng để chọn ra một tập hợp con các bộ từ một quan hệ thỏa mãn một điều kiện chọn. Ký hiệu  $\sigma$

## Định nghĩa

- Cho quan hệ  $R(U)$
- Cho  $E$ : biểu thức logic gồm các phép toán số học và Logic :  
 $<$ ,  $\leq$ ,  $>$ ,  $\geq$ ,  $\neq$ ,  $\wedge$ ,  $\vee$ ,  $\neg$
- Phép chọn  $\sigma$  trên quan hệ  $R$  theo điều kiện  $E$ , ký hiệu  $\sigma_E(R)$  có kết quả là 1 tập con của  $R$ , gồm các bộ  $t \in R$  thỏa điều kiện  $E$

$$\sigma_E(R) = \{ t / t \in R \wedge t(E) \} \quad t(E): \text{bộ } t \text{ thỏa điều kiện } E$$

# Phép chọn

- Ví dụ 1 : Cho quan hệ R(A,B,C) với các thể hiện như sau:

R	A	B	C
1	a		2
5	x		4
2	b		2

$\sigma_{C=2}(R)$  ?

$\sigma_{A=C}(R)$  ?

- Ví dụ 2: Tìm các sinh viên sinh năm 1990

$\sigma_{namsinh=1990}(\text{SINHVIEN})$

- Ví dụ 3: Tìm các môn từ 3 tín chỉ trở lên

- Ví dụ 4: Tìm các sinh viên lớp DI1496A1 ở Cần Thơ

SELECT \* from SINHVIEN  
Where namsinh = 1990;

# CSDL Ví dụ

- SINHVIEN (MASV, hoten, namsinh, diachi, # *ML*)
- LOP (ML, tenlop)
- MONHOC (MM, tenmon, TC, LT, TH)
- DIEM (#MASV, #MM, hk, nk, diem)
- GIAOVIEN (MAGV, hotenGv, namsinhGv, diachiGv)
- DAY (#MAGV, #MM, hk, nk)

- Tìm các môn từ 3 tín chỉ trở lên  $\sigma_{TC \geq 3} (MONHOC)$
- Tìm các sinh viên lớp DI1496A1 ở Cần Thơ
- Tìm các SV có học môn CSDL đạt điểm giỏi  
trở lên trong HK 2, NK 17-18

$\sigma_{tenlop = 'DI1496A1' \wedge diachi = 'Cần Thơ'} (SINHVIEN \bowtie LOP)$  14

$\sigma_{tenmon = 'CSDL' \wedge diem \geq 8.5 \wedge hk = 2 \wedge nk = '17-18'}$

# Phép chiếu

- Phép Chiếu (Projection) được sử dụng để chọn một vài cột hay thuộc tính từ một bảng , ký hiệu  $\pi$

## Định nghĩa

- Cho quan hệ  $R(U)$
- $X$  là tập con của  $U$  :  $X \subset U$ ,  $X \neq \emptyset$
- Phép chiếu  $\pi$  trên quan hệ  $R$  theo tập thuộc tính  $X$ , ký hiệu  $\pi_X(R)$  có kết quả là tập hợp các bộ  $t \in R$  nhưng chỉ lấy giá trị trên  $X$

$$\pi_X(R) = \{ t / t \in R \wedge t[X] \} \quad t[X]: t \text{ nhận giá trị trên } X$$

- $\pi_X(R)$  chỉ gồm các bộ phân biệt (Các bộ trùng nhau trong kết quả chỉ giữ lại một)

# Phép chiếu

- **Ví dụ 1:** Cho quan hệ R(A,B,C) với các thể hiện như sau:

R	A	B	C
1	a	2	
5	x	4	
2	b	2	

$\pi_C(R)$  ?

$\pi_{A, C}(R)$  ?

- **Ví dụ 2 :** - Liệt kê mã sinh viên và tên sinh viên của tất cả các sinh viên.
  - Liệt kê mã sinh viên và tên sinh viên của tất cả các sinh viên sinh năm 1990.

$\pi_{MSSV, hoten}(SINHVIEN)$

$\pi_{MSSV, hoten}(\sigma_{namsinh=1990}(SINHVIEN))$

# CSDL Ví dụ

- SINHVIEN (MASV, hoten, namsinh, diachi, # *ML*)
- LOP (ML, tenlop)
- MONHOC (MM, tenmon, TC, LT, TH)
- DIEM (#MASV, #MM, hk, nk, diem)
- GIAOVIEN (MAGV, hotenGv, namsinhGv, diachiGv)
- DAY (#MAGV, #MM, hk, nk)

- **Liệt kê mã sinh viên và tên sinh viên của tất cả các sinh viên.**

J MASV, hoten (SINHVIEN)

- **Liệt kê mã sinh viên và tên sinh viên của tất cả các sinh viên sinh năm 1990.**

σ MASV, hoten (σ namsinh = 1990 SINHVIEN)

- **Tìm MaSV có học môn CSDL đạt điểm giỏi trở lên trong HK 2, NK 17-18**

I SINHVIEN.MSV (σ tenmon = 'CSDL', ∧ diem > 8.5 ∧ hk = 2, nk = '17-18')

# Các phép toán trên tập hợp

- **Hai quan hệ tương thích**
  - Hai quan hệ tương thích nếu
    - Có cùng bậc n
    - Thuộc tính thứ  $j = 1, 2, \dots, n$  của hai quan hệ có cùng miền giá trị
- Các phép toán trên tập hợp gồm:
  - Hợp
  - Giao
  - Trừ

ĐÚ MÃ MÃNG C GÌ CÒN  $\wedge$  TRÙA  
( monhoc.)

# Phép hợp

- Ký hiệu  $\cup$

## Định nghĩa

Hợp của hai quan hệ tương thích  $R$  và  $S$ , ký hiệu  $R \cup S$ , gồm các bộ thuộc ít nhất một trong 2 quan hệ đã cho.

$$R \cup S = \{ t / t \in R \vee t \in S \}$$

- Ví dụ 1:** Cho 2 quan hệ  $R$  và  $S$ . Tìm  $R \cup S$

**R**

A	B	C
1	a	2
5	x	4
2	b	2

**S**

A	B	C
1	a	2
6	y	4
2	b	2

- Ví dụ 2 :** Liệt kê danh sách tên sinh viên và tên giáo viên

SQL:  $\pi_{hoten}(\text{SV}) \cup \pi_{hotenGV}(\text{GV})$

# Phép giao

- Ký hiệu  $\cap$

## Định nghĩa

Giao của hai quan hệ tương thích  $R$  và  $S$ , ký hiệu  $R \cap S$ , gồm các bộ thuộc về cả 2 quan hệ đã cho.

$$R \cap S = \{ t / t \in R \wedge t \in S \}$$

- Ví dụ 1:** Cho 2 quan hệ  $R$  và  $S$ . Tìm  $R \cap S$

R	A	B	C
1	a	2	{ } }
5	x	4	
2	b	2	{ } }

S	A	B	C
1	a	2	
6	y	4	
2	b	2	

- Ví dụ 2 :** Tìm mã số sinh viên vừa học môn CT180 vừa học môn CT181

EX: •

$$\begin{array}{l} \cup_{\text{MSV}} (\sigma_{MM = 'CT188'} (\text{ĐIỂM})) \\ \cap \cup_{\text{MSV}} (\sigma_{MM = 'CT181'} (\text{ĐIỂM})) \end{array}$$

# Phép trừ

- Ký hiệu \

## Định nghĩa

Hiệu của hai quan hệ tương thích  $R$  và  $S$ , ký hiệu  $R|S$ , gồm các bộ thuộc  $R$  và không thuộc  $S$ .

$$R \setminus S = \{ t / t \in R \wedge t \notin S \}$$

- Ví dụ 1:** Cho 2 quan hệ  $R$  và  $S$ . Tìm  $R \setminus S$

R	A	B	C
1	a	2	
5	x	4	}
2	b	2	

S	A	B	C
1	a	2	
6	y	4	
2	b	2	

1 row.

- Ví dụ 2 :** Tìm MSSV của sinh viên học môn TH409 và không học môn TH104

đu MSSV. ( 6 MM = 'TH409' ) ( SINH VIEN DA DIEM DA HOM HOC ) ✓

đu MSSV. ( 6 MM = 'TH104' ) ( SINH VIEN DA DIEM DA HOM HOC )

SLECT MASV from SINHVIEN, S, DIEM  
join where \_\_\_\_\_ and SMM = 'TH'

# CSDL Ví dụ

6

- SINHVIEN (MASV, hoten, namsinh, diachi, #ML)
- LOP (ML, tenlop)
- MONHOC (MM, tenmon, TC, LT, TH)
- DIEM (#MASV, #MM, hk, nk, diem)
- GIAOVIEN (MAGV, hotenGv, namsinhGv, diachiGv)
- DAY (#MAGV, #MM, hk, nk)

Tìm họ tên các sinh viên học môn Cơ sở dữ liệu  
trong hk 1 năm 23-24

JU hoten ( 6 tenmon = "Cơ sở dữ liệu"  $\wedge$  hk = 1  $\wedge$  nk = '23 - 24' [SINHVIEN  $\bowtie$  DIEM  $\bowtie$  MONHOC])

# Các phép toán kết nối

- Là các phép toán cho phép kết hợp các bộ từ nhiều quan hệ.
- Các phép toán:
  - Phép kết nối Theta ( $\theta$ )
  - Kết nối tự nhiên
  - Kết nối mở rộng

# Phép kết nối tự nhiên

- Ký hiệu \*

## Định nghĩa

- Phép kết nối tự nhiên trên 2 quan hệ  $R(U)$  và  $S(V)$ , ký hiệu  $R*S$ , là kết nối dựa trên phép so sánh = giữa 2 thuộc tính cùng miền giá trị của 2 quan hệ  $R$  và  $S$ , một trong 2 thuộc tính cùng tên này sẽ bị loại bỏ qua phép chiếu.

# CSDL Ví dụ

- SINHVIEN (MASV, hoten, namsinh, diachi, # *ML*)
- LOP (ML, tenlop)
- MONHOC (MM, tenmon, TC, LT, TH)
- DIEM (#MASV, #MM, hk, nk, diem)
- GIAOVIEN (MAGV, hotenGv, namsinhGv, diachiGv)
- DAY (#MAGV, #MM, hk, nk)

- . **Ví dụ 2:** Cho biết MSSV và họ tên các sinh viên học môn TH104  
JL MASV, hoten ( SINHVIEN \* σ<sub>MM = 'TH104'</sub>, (DIEM) ) Khoa JL Masv, Hoten ( σ<sub>MM = 'TH104'</sub>  
( SINHVIEN \*  
DIEM )  
Khoa )
- . VD3: Cho biết mã số và họ tên của sinh viên lớp Khoa học máy tính 1 K39 đạt điểm giỏi trở lên môn Cơ sở dữ liệu trong học kỳ 2 niên khóa 2015-2016

# CSDL Ví dụ

$\wedge \text{hk} = 2 \wedge \text{nk} = '2015-2016' \wedge \text{diem} \geq 8 \}$  SINH VIEN & LOP & DIEM & MON HOC  
SQL: Select a. MaSV, hoten from SinhVien a, Lop b, MonHoc c, Diem d  
where a. MaSV = d. MaSV AND d. MM = c. MM AND a. ML = b. ML AND  
tenlop = 'KHOA HOC MAY TINH 1-K39' AND tenmon = 'Co so du lieu' AND  
diem >= 8 and  
hk = 2  
AND nk = '2015 - 2016';

- SINHVIEN (MASV, hoten, namsinh, diachi, # ML)
- LOP (ML, tenlop)
- MONHOC (MM, tenmon, TC, LT, TH)
- DIEM (#MASV, #MM, hk, nk, diem)
- GIAOVIEN (MAGV, hotenGv, namsinhGv, diachiGv)
- DAY (#MAGV, #MM, hk, nk)

- . **Ví dụ:** sinh viên Phùng Phung Phí học môn CSDL được mấy điểm?

SELECT a. MaSV

// JSI Masv hien 2 bang  
nen gian cu phai

# Phép kết nối tự nhiên

- **Ví dụ 1:** Cho 2 quan hệ R và S

R	A	B	C
1	a	2	
5	x	4	
2	b	2	

S	A	E	F
1	x	m	
6	w	n	
2	z	p	

=> R\*S

R * S	A	B	C	E	F
1	a	2		x	m
2	b	2		z	p

- **Ví dụ 2:** Cho biết MSSV và họ tên các sinh viên học môn TH104
- VD3: Cho biết mã số và họ tên của sinh viên lớp Khoa học máy tính 1 K39 đạt điểm giỏi trở lên môn Cơ sở dữ liệu trong học kỳ 2 niên khóa 2015-2016

# Phép kết nối Theta

- Ký hiệu  $\triangleright\triangleleft$

## Định nghĩa

- Cho quan hệ  $R(U)$  và  $S(V)$
- $\theta$  là một trong các phép toán so sánh :  $<$ ,  $\leq$ ,  $>$ ,  $\geq$ ,  $=$ ,  $\neq$
- *Phép kết nối theo điều kiện  $\theta$  giữa quan hệ  $R$  theo thuộc tính  $A \in U$  và quan hệ  $S$  theo thuộc tính  $B \in V$ , ký hiệu  $R \triangleright_\theta \triangleleft S$ , là một quan hệ gồm các bộ có dạng  $(u, v)$ , trong đó  $u \in R$ ,  $v \in S$ , giá trị của chúng trên  $A$  và  $B$  thỏa  $\theta$*

$$R \triangleright_\theta \triangleleft S = \{ (u, v) / u \in R \wedge v \in S \wedge (u[A] \theta v[B]) \}$$

**Chú ý:** Phép kết nối chỉ thực hiện được khi  $\theta$  thực hiện được giữa  $A$  và  $B$

# Phép kết nối Theta

- Ví dụ 1 : Cho hai quan hệ R và S

R	A	B	C
1	a	2	
5	x	4	
2	b	2	

S	E	F
x		22
y		14
x		12

=>  $R \triangleright\triangleleft S$   
 $B=E$   
=?

$R \triangleright\triangleleft S$	A	B	C	E	F
	5	x	4	x	22
	5	x	4	x	12

- Ví dụ 2: Tìm MSSV và họ tên các sinh viên lớn tuổi hơn một giáo viên nào đó.

# Phép kết nối mở rộng (outer join)

- Cho phép thực hiện nối kết tự nhiên trên các giá trị null của thuộc tính dùng nối kết :

## Định nghĩa

- Cho  $R(U)$ ,  $S(V)$
- Đặt  $N=R^*S$
- Đặt  $P$  bao gồm các bộ  $(u,v)$  sao cho  $u \in R$ ,  $v \in S$ ,  $u \notin N$ , các giá trị của các thuộc tính trong  $v$  trên  $V$  đều là null
- Đặt  $Q$  bao gồm các bộ  $(u,v)$  sao cho  $u \in R$ ,  $v \in S$ ,  $v \notin N$ , các giá trị của các thuộc tính trong  $u$  trên  $U$  đều là null
- Kết nối mở rộng trái:  $R > \triangleleft S = P \cup N$
- Kết nối mở rộng phải:  $R > \triangleright S = N \cup Q$
- Kết nối mở rộng hai bên  $R > \triangleleft S = P \cup N \cup Q$

# Phép kết nối mở rộng

- Ví dụ:
  - 1) Hãy cho biết mỗi sinh viên đã học môn nào, kể cả các sinh viên chưa học môn nào
  - 2) hãy cho biết mỗi môn do các học sinh nào học, kể cả các môn không có học sinh nào học
  - 3) Hãy cho biết mỗi sinh viên đã học môn nào, kể cả các sinh viên chưa học môn nào và các môn không có học sinh nào học

# Phép kết nối mở rộng

- Ví dụ

$\pi_{MSSV, hoten}(SINHVIEN)$

MSSV	Hoten
1940636	Phạm Thị Ngọc Diễm
1940647	Trần Thanh Điền
1940763	Nguyễn Văn Đặng
1940836	Lê Thị Thùy Linh
1940852	Đặng Thùy Lan

$\pi_{MSSV, MM}(DIEM)$

MSSV	MM
1940636	CT104
1940647	CT165
1940763	CT304
1940763	CT106
1940630	CT114

## Kết nối tự nhiên (nối kết nội – inner join)

$$\pi_{\text{MSSV}, \text{hoten}}(\text{SINHVIEN}) * \pi_{\text{MSSV}, \text{MM}}(\text{DIEM})$$

MSSV	Hoten
1940636	Phạm Thị Ngọc Diễm
1940647	Trần Thanh Điền
1940763	Nguyễn Văn Đặng
1940836	Lê Thị Thùy Linh
1940852	Đặng Thùy Lan

MSSV	MM
1940636	CT104
1940647	CT165
1940763	CT304
1940763	CT106
1940630	CT114

MSSV	Hoten	MM
1940636	Phạm Thị Ngọc Diễm	CT104
1940647	Trần Thanh Điền	CT165
1940763	Nguyễn Văn Đặng	CT304
1940763	Nguyễn Văn Đặng	CT106

$\pi_{\text{MSSV}, \text{hoten}}(\text{SINHVIEN})$  $\pi_{\text{MSSV}, \text{MM}}(\text{DIEM})$ 

MSSV	Hoten
1940636	Phạm Thị Ngọc Diễm
1940647	Trần Thanh Điền
1940763	Nguyễn Văn Đặng
1940836	Lê Thị Thùy Linh
1940852	Đặng Thùy Lan

MSSV	MM
1940636	CT104
1940647	CT165
1940763	CT304
1940763	CT106
1940630	<u>CT114</u>

 $\pi_{\text{MSSV}, \text{hoten}}(\text{SINHVIEN}) > \triangle \pi_{\text{MSSV}, \text{MM}}(\text{DIEM})$ 

## Phép kết nối mở rộng

Mở rộng bên trái

MSSV	Hoten	MM
1940636	Phạm Thị Ngọc Diễm	CT104
1940647	Trần Thanh Điền	CT165
1940763	Nguyễn Văn Đặng	CT304
1940763	Nguyễn Văn Đặng	CT106
1940836	Lê Thị Thùy Linh	null
1940852	Đặng Thùy Lan	null

# Phép kết nối mở rộng

Mở rộng bên  
phải:

$$\pi_{\text{MSSV}, \text{hoten}}(\text{SINHVIEN}) \triangleright < \pi_{\text{MSSV}, \text{MM}}(\text{DIEM})$$

$\pi_{\text{MSSV}, \text{hoten}}(\text{SINHVIEN})$

$\pi_{\text{MSSV}, \text{MM}}(\text{DIEM})$

MSSV	Hoten
1940636	Phạm Thị Ngọc Diễm
1940647	Trần Thanh Điền
1940763	Nguyễn Văn Đặng
1940836	Lê Thị Thùy Linh
1940852	Đặng Thùy Lan

MSSV	MM
1940636	CT104
1940647	CT165
1940763	CT304
1940763	CT106
1940630	CT114

MSSV	Hoten	MM
1940636	Phạm Thị Ngọc Diễm	CT104
1940647	Trần Thanh Điền	CT165
1940763	Nguyễn Văn Đặng	CT304
1940763	Nguyễn Văn Đặng	CT106
1940630	null	CT114

# Phép kết nối mở rộng

Mở rộng hai  
bên:

$$\pi_{\text{MSSV}, \text{hoten}}(\text{SINHVIEN}) \sqsubset \pi_{\text{MSSV}, \text{MM}}(\text{DIEM})$$

MSSV	Hoten	MM
1940636	Phạm Thị Ngọc Diễm	CT104
1940647	Trần Thanh Điền	CT165
1940763	Nguyễn Văn Đặng	CT304
1940763	Nguyễn Văn Đặng	CT106
1940836	Lê Thị Thùy Linh	null
1940852	Đặng Thùy Lan	null
1940630	null	CT114

# Nội dung

- Giới thiệu
- Một số khái niệm
- Các phép toán cơ bản
- Các phép toán khác

# Các phép toán khác

- Phép chia
- Tích Descartes
- Các hàm kết tập

# Phép chia (Division)

- Ký hiệu /

## Định nghĩa

- $R(U)$  là quan hệ bậc n
- $S(V)$  là quan hệ bậc m,  $V \neq \emptyset$ ,  $U \cap V \neq \emptyset$
- *Phép chia quan hệ R cho S, ký hiệu  $R/S$ , cho kết quả là một quan hệ gồm các bộ t có  $n-m$  thuộc tính nhận giá trị trên  $U|V$  thuộc tính sao cho mỗi bộ  $v \in S$  thì bộ  $(t, v) \in R$*

$$R / S = \{ t / t[U \setminus V] \wedge (\forall v \in S, (t, v) \in R)\}$$

# Phép chia (Division)

SELECT MaSV, MM FROM DIEM  
where  
MM = ]All [  
All [IN]  
SELECT MM FROM MONHOC;

- Ví dụ 1 : Cho hai quan hệ R và S

R	A	B
a	1	
b	2	
a	2	
c	1	
c	2	

S	B
	1
	2

R/S	A
	a
	c

- Ví dụ 2: Tìm MSSV học tất cả các môn học

# Phép tích Descartes

- Ký hiệu  $\times$ 
  - $R(U)$  là quan hệ bậc n
  - $S(V)$  là quan hệ bậc m,  $U \cap V = \emptyset$ , nếu  $U \cap V \neq \emptyset$  thì đặt tên lại cho các thuộc tính trùng tên
  - *Phép tích Descartes của quan hệ R và S, ký hiệu  $R \times S$ , là một quan hệ có  $|R| \cdot |S|$  bộ có dạng  $(u, v)$ , mỗi bộ có  $n+m$  thuộc tính,  $u \in R$  và  $v \in S$*

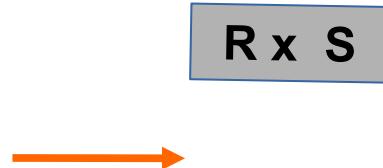
$$R \times S = \{ (u, v) / u \in R \wedge v \in S \}$$

# Phép tích Descartes

- Ví dụ: Cho 2 quan hệ R và S:

R	A	B
a	1	
b	2	
a	2	
c	1	
c	2	

S	C
	1
	2



R x S	A	B	C
a	1	1	
a	1	2	
b	2	1	
b	2	2	
a	2	1	
a	2	2	
c	1	1	
c	1	2	
c	2	1	
c	2	2	

# Các hàm kết tập

(Aggregate Functions)

- Đầu vào : một tập các giá trị của một thuộc tính
- Đầu ra : một giá trị duy nhất
- Cho quan hệ  $R(U)$  và thuộc tính  $A \in U$
- Các hàm :
  - $\text{AVG}_A(R)$  : tính giá trị trung bình của thuộc tính  $A$
  - $\text{MIN}_A(R)$  : tìm giá trị nhỏ nhất của thuộc tính  $A$
  - $\text{MAX}_A(R)$  : tìm giá trị lớn nhất của thuộc tính  $A$
  - $\text{SUM}_A(R)$  : tính tổng các giá trị của thuộc tính  $A$
  - $\text{COUNT}_A(R)$  : đếm số giá trị của thuộc tính  $A$

# CSDL Ví dụ

- SINHVIEN (MASV, hoten, namsinh, diachi, #ML)
- LOP (ML, tenlop)
- MONHOC (MM, tenmon, TC, LT, TH)
- DIEM (#MASV, #MM, hk, nk, diem)
- GIAOVIEN (MAGV, hotenGv, namsinhGv, diachiGv)
- DAY (#MAGV, #MM, hk, nk)

2) Hãy cho biết số môn mà sinh viên Nguyễn Văn An  
học

$\text{COUNT}_{\text{DISTINCT MM}} (\sigma_{\text{hoten} = 'Nguyễn Văn An'} (\text{SINHVIEN}) \times \text{DIEM.})$

3) Tổng số tín chỉ mà sinh viên Nguyễn Văn An học

1) Hãy cho biết điểm trung bình, điểm lớn nhất, điểm  
nhỏ nhất của môn CSDL

3.  $\text{sum}_{\text{TC}} (\bar{x}_{\text{MM}} (\sigma_{\text{hoten} = 'Nguyễn Văn An'} (\text{SINHVIEN}) \times \text{DIEM.}) \times \text{MONHOC})$

1.  $\pi_{\text{diem}} \text{AVG}(\text{diem}), \text{MAX}(\text{diem}), \text{MIN}(\text{diem})$  (  $\delta_{\text{tenmon}} = \text{'CSDL' ( MÔN HỌC ) } \times \text{ĐIỂM}$  )

# Các hàm kết tập (Aggregate Functions)

## • Ví dụ:

- 1) Hãy cho biết điểm trung bình, điểm lớn nhất, điểm nhỏ nhất của môn CSDL
- 2) Hãy cho biết số môn mà sinh viên Nguyễn Văn An học
- 3) Tổng số tín chỉ mà sinh viên Nguyễn Văn An học

1.  $\begin{array}{l} \text{SELECT AVG(d.diem), MAX(d.diem), MIN(d.diem)} \\ \text{FROM MONHOC m} \\ \text{JOIN Diem d ON m.MM = d.MM} \\ \text{Where m.tenmon = 'CSDL'}; \end{array}$

2.  $\begin{array}{l} \text{SELECT COUNT(DISTINCT d.MM)} \text{ FROM BINHVIEEN } \text{by} \\ \text{JOIN Diem d ON sv.NASV = d.MASV} \\ \text{WHERE sv.Hoten = 'Nguyễn Văn An';} \\ \text{DISTINCT.} \\ \text{3. SELECT SUM(m.TC) FROM SINHVIEN sv} \\ \text{JOIN Diem d ON sv.NASV = d.MASV.} \\ \text{JOIN MONHOC m ON d.MM = m.MM WHERE sv.hoten = "Nguyễn"} \\ \text{Văn An";} \end{array}$

# Các hàm kết tập trên nhóm

- Sử dụng các hàm kết tập trên **nhóm các tập giá trị, không phải trên tập tất cả các giá trị**

- Cho E là biểu thức quan hệ

- Hàm kết tập trên nhóm có dạng:

$$G_1, G_2, \dots, G_m \text{ } G F_1(A_1), F_2(A_2), \dots, F_p(A_p)(E)$$

- Trong đó:

- $G_i$  là một thuộc tính mà việc chia nhóm các bộ dựa trên đó
- $F_i$  là một hàm kết tập
- $A_i$  là một thuộc tính

- Kết quả sẽ gồm  $m+p$  thuộc tính (các thuộc tính phân nhóm và các giá trị của các hàm kết tập)

# Các hàm kết tập trên nhóm

- SINHVIEN (MASV, hoten, namsinh, diachi, #ML)
- LOP (ML, tenlop)
- MONHOC (MM, tenmon, TC, LT, TH)
- DIEM (#MASV, #MM, hk, nk, diem)
- GIAOVIEN (MAGV, hotenGv, namsinhGv, diachiGv)
- DAY (#MAGV, #MM, hk, nk)

1. Hãy cho biết điểm trung bình, điểm lớn nhất, điểm nhỏ nhất của từng sinh viên. MASV, hoten G AVG(diem), MAX(diem), MIN(diem) C SINHVIEN \* DIEM

2. Cho biết tổng số tiết dạy của từng giáo viên trong HK 2, NK 21-22 (biết rằng Số tiết = TC\*15) MAGV, hotenGv G SUM(TC \* 15) C GIAO VIEN  
X JOIN MM, TC (MONHOC) X G HK2 = 2 AND  
nK = '21-22' CDAY

3. Cho biết tổng số tín chỉ đã học của từng SV

MASV, hoten G SUM(TC) ( SINHVIEN \* DIEM \* (π MM, TC (MONHOC)))

SELECT MASV, hoten, SUM(TC);

4. Cho biết danh sách SV đã học trên 125 TC

σ tongTC > 125 ( MASV, hoten G SUM(TC) ( SINHVIEN \* DIEM \* (π MM, TC (MONHOC))) ) ) ) ) SELECT sv.MASV, sv.hoten, SUM(DISTINCT mh.TC) FROM SINHVIEN sv JOIN DIEM d ON sv.MASV = d.MASV JOIN MONHOC mh ON d.MM = mh.MM GROUP BY sv.MASV, sv.hoten HAVING SUM(DISTINCT mh.TC) > 125;