CSC10006 – Introduction to Database

Chapter 5
Relational Algebra

Instructor: **Vũ Thị Mỹ Hằng, Ph.D.** (**vtmhang**@fit.hcmus.edu.vn)

Lab Assistants: **Tiết Gia Hồng, M.Sc.** (**tghong**@fit.hcmus.edu.vn)

Lương Hán Cơ, M.Sc. (**lhco**@fit.hcmus.edu.vn)

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

fit@hcmus

cdio 4.0

fit@hcmus

Content

1. Introduction
2. Relational Algebra Operations From Set Theory
3. Selection Operation
4. Projection Operation
5. CARTESIAN Operations
6. Join Operation
7. Division Operation
8. Additional Relational Operations
9. Update Operations

cdio 4.0

fit@hcmus

Content

1. Introduction
2. Relational Algebra Operations From Set Theory
3. Selection Operation
4. Projection Operation
5. CARTESIAN Operations
6. Join Operation
7. Division Operation
8. Additional Relational Operations
9. Update Operations

1. Introduction

Update on the relation "KHOA"

- Add a new Khoa 'Hóa học' into the relation
- Move "CNTT" to the room "B12"

MÃKHOA	TÊNKHOA	NĂMTL	PHÒNG	ĐIỆNTHOAI	TRƯỞNGKHOA	NGÀYNHẬNCHỨC
CNTT	Công nghệ thông tin	1995	B12	0838123456	002	20/02/2005
VL	Vật lý	1976	B21	0838223223	005	18/09/2003
SH	Sinh học	1980	B31	0838454545	004	11/10/2000
HH	Hóa học	1980	B41	NULL	007	15/10/2001

- Query: name of the faculty found from the year 1980

TÊNKHOA
Sinh học
Hóa học

1. Introduction (cont.)

- Two types of operations:
  - Make change the **data (update): add new, delete and modify**
  - Don't make change the data **(extract data): query**
- Query Language – QL:
  - Allow to extract or update data stored in a database schema
- Types of Relational Query:
  - Relational Algebra (Procedural Language)
    - A sequence of relational algebra operations forms a **relational algebra expression**, whose result will also be a relation that represents the result of a database query (or retrieval request).
  - Relational Calculus (Nonprocedural Language)
    - This differs from relational algebra, where we must write a sequence of operations to specify a retrieval request; hence relational algebra can be considered as a procedural way of stating a query
  - SQL (*Structured Query Language*)

1. Introduction (cont.)

- Algebra
  - Operator (toán tử)
  - Operand (toán hạng)

- Arithmetic (số học)
  - Operators: +, -, *, /
  - Operands or variables: x, y, z
  - Constant
  - Expression
    - (x+7) / (y-3)
    - (x+y)*z and/or (x+7) / (y-3)

1. Introduction (cont.)

Relational Algebra

- Operands – variables are relations
  - Or set (tập hợp)
- Operators are relational operations
  - Relational Algebra Operations From Set Theory
    - Union ∪ (hội)
    - Intersect ∩ (giao)
    - Difference set/ Minus – (trừ)
  - Unary Relational Operations
    - Selection σ (chọn)
    - Projection π (chiếu)
  - Additional relational operations
    - Cartesian product × (Tích Cartesian)
    - Join ⋈ (kết)

1. Introduction (cont.)

- In Algebra relation, a query is expressed through a **relational algebra expression**:
  - The result of a retrieval is a new relation, which may have been formed from one or more relations.
  - The algebra operations thus produce new relations, which can be further manipulated using operations of the same algebra
  - It is formed by a sequence of relational algebra operations
  - the result of the previous operation is input to the next operation.

1. Introduction (cont.)

	Algebra	**Relational algebra**
Operand	- Variable : x, y, z, … - Constant : 150, …	- Schema : NhanVien, … - Relation/Instances/ tuples : t, v, …
Operator	- Operations between operands to form a new value: +, -, *, /, …	- Operations between relations to form a new relation: *selection* σ, *union* $\cup$, …
Expression	- Sequence of algebra operators - Return a new value *(x+7) / (y-3)* *(x+y)*z and/or (x+7) / (y-3)*	- Sequence of relational algebra (query) - Return a new relation $\pi_{MANV}(NHANVIEN)$

cdio 4.0

fit@hcmus

Content

1. Introduction
2. Relational Algebra Operations From Set Theory
3. Selection Operation
4. Projection Operation
5. CARTESIAN Operations
6. Join Operation
7. Division Operation
8. Additional Relational Operations
9. Update Operations

2. Relational Algebra Operations From Set Theory

- Operations
  - Union $r \cup s$
  - Intersect $r \cap s$
  - Difference set/ Minus $r - s$

- Union Compatibility (khả hợp)
  - The operand relations $R(A_1, A_2, \ldots, A_n)$ and $S(B_1, B_2, \ldots, B_n)$:
  - Two are compatible if
    - R and S have the same number of attributes (e.g. n)
    - and $DOM(A_i)=DOM(B_i)$, $1 \leq i \leq n$

- The resulting relation for $R \cup S$, $R \cap S$, or R-S has the same attribute names as the first operand relation R (by convention).

2. Union

- Example

SINHVIEN	TENSV	NGSINH	PHAI
	Tung	12/08/1955	Nam
	Hang	07/19/1968	Nu
	Nhu	06/20/1951	Nu
	Hung	09/15/1962	Nam

GIAOVIEN	TENGV	NG_SINH	GIOITINH
	Trinh	04/05/1986	Nu
	Khang	10/25/1983	Nam
	Phuong	05/03/1958	Nu
	Minh	02/28/1942	Nam
	Chau	12/30/1988	Nu

Degree n=3
DOM(TENSV) = DOM(TENGV)
DOM(NGSINH) = DOM(NG_SINH)
DOM(PHAI) = DOM(GIOITINH)

2. Union

- Given two compatible relations **r** and **s**
- Union of r and s
  - Notation: $r \cup s$
  - The result of the union operation includes tuples in r or in s, or in both relations (duplicated tuples are eliminated)

$$r \cup s = \{ t / t \in r \vee t \in s \}$$

- Example

r	A	B
	α	1
	α	2
	β	1

s	A	B
	α	2
	β	3

r ∪ s	A	B
	α	1
	α	2
	β	1
	α	2
	β	3

2. Union - example

SinhVien	
HOTEN	DIACHI
Đinh Bá Tiến	119 Cống Quỳnh, Tp HCM
Nguyễn Thanh Tùng	222 Nguyễn Văn Cừ, Tp HCM
Lê Quỳnh Như	291 Hồ Văn Huê, Tp HCM

GiaoVien	
HOTEN	DIACHI
Đinh Bá Tiến	119 Cống Quỳnh, Tp HCM
Trần Thanh Tâm	553 Mai Thị Lựu, Tp HCM

SinhVien ∪ GiaoVien	
HOTEN	DIACHI
Đinh Bá Tiến	119 Cống Quỳnh, Tp HCM
Nguyễn Thanh Tùng	222 Nguyễn Văn Cừ, Tp HCM
Trần Thanh Tâm	553 Mai Thị Lựu, Tp HCM

2. Intersect

- Given two compatible relations **r** and **s**
- Intersect of r and s
  - Notation: $r \cap s$
  - The result relation of the intersect operation includes tuples that are both in r and s

$$r \cap s = \{ t / t \in r \wedge t \in s \}$$

r	A	B
	α	1
	α	2
	β	1

s	A	B
	α	2
	β	3

r ∩ s	A	B
	α	2

2. Intersect (cont.) - Example

SinhVien	
HOTEN	DIACHI
Đinh Bá Tiến	119 Cống Quỳnh, Tp HCM
Nguyễn Thanh Tùng	222 Nguyễn Văn Cừ, Tp HCM
Lê Quỳnh Như	291 Hồ Văn Huê, Tp HCM

GiaoVien	
HOTEN	DIACHI
Đinh Bá Tiến	119 Cống Quỳnh, Tp HCM
Trần Thanh Tâm	553 Mai Thị Lựu, Tp HCM

SinhVien ∩ GiaoVien	
HOTEN	DIACHI
Đinh Bá Tiến	119 Cống Quỳnh, Tp HCM

2. Minus

- Given two compatible relations **r** and **s**
- Minus of r and s
  - Notation: r – s
  - The result of the minus operation includes tuples **in r** and **not in s**

$$r - s = \{ t / t \in r \wedge t \notin s \}$$

r	A	B
	α	1
	α	2
	β	1

s	A	B
	α	2
	β	3

r – s	A	B
	α	1
	β	1

2. Minus - example

SinhVien	
HOTEN	DIACHI
Đinh Bá Tiến	119 Cống Quỳnh, Tp HCM
Nguyễn Thanh Tùng	222 Nguyễn Văn Cừ, Tp HCM
Lê Quỳnh Như	291 Hồ Văn Huê, Tp HCM

GiaoVien	
HOTEN	DIACHI
Đinh Bá Tiến	119 Cống Quỳnh, Tp HCM
Trần Thanh Tâm	553 Mai Thị Lựu, Tp HCM

SinhVien – GiaoVien	
HOTEN	DIACHI
Nguyễn Thanh Tùng	222 Nguyễn Văn Cừ, Tp HCM
Lê Quỳnh Như	291 Hồ Văn Huê, Tp HCM

2. Characteristics

- Notice that both union and intersection are *commutative operations (tính giao hoán)*; that is

$$r \cup s = s \cup r$$

$$r \cap s = s \cap r$$

- Both union and intersection can be treated as n-ary operations applicable to any number of relations as both are *associative operations (tính kết hợp)*; that is

$$r \cup (s \cup t) = (r \cup s) \cup t$$

$$r \cap (s \cap t) = (r \cap s) \cap t$$

- The minus operation is *not commutative*; that is, in general

$$R - S \neq S - R$$

cdio 4.0

fit@hcmus

Content

1. Introduction
2. Relational Algebra Operations From Set Theory
3. Selection Operation
4. Projection Operation
5. CARTESIAN Operations
6. Join Operation
7. Division Operation
8. Additional Relational Operations
9. Update Operations

3. Selection Operation

- Is used to select a *subset* of the tuples from a relation that satisfy a selection condition P. It is a filter that keeps only those tuples that satisfy a qualifying condition – those satisfying the condition are selected while others are discarded
- Denoted by

$$\sigma_P(r)$$

σ: sigma

- P is an expression consisting of clauses as:
  - <attribute> <comparison> <constant>
  - <attribute> <comparison> <comparision>
    - <comparison>: < , > , ≤ , ≥ , ≠ , =
    - Clauses are connected by logical connections ∧ , ∨ , ¬

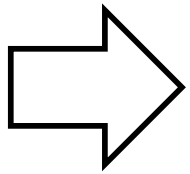
3. Selection Operation (cont.)

- The result relation:
  - Attributes: the same attributes of the relation **r**
  - Tuples: less or equal of **r**

$$\sigma_{(A=B)\wedge(D>5)}(r)$$

- Example

r	A	B	C	D
	α	α	1	7
	α	β	5	7
	β	β	12	3
	β	β	23	10



A	B	C	D
α	α	1	7
β	β	23	10

3. Selection Operation (cont.)

- The select is *commutative operations*; that is

$$\sigma_{p1}(\sigma_{p2}(r)) = \sigma_{p2}(\sigma_{p1}(r)) = \sigma_{p1 \wedge p2}(r)$$

Content

1. Introduction
2. Relational Algebra Operations From Set Theory
3. Selection Operation
4. Projection Operation
5. CARTESIAN Operations
6. Join Operation
7. Division Operation
8. Additional Relational Operations
9. Update Operations

4. Projection Operation

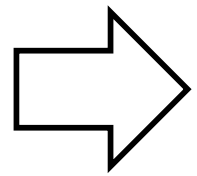
- This operation selects certain columns from a relation **r** and discards the other columns
- Notation

$$\pi_{A1, A2, \ldots, Ak}(r)$$

- The result is a relation
  - having k attributes
  - Number of tuples are less than r
- Example

$$\pi_{A,C}(r)$$

r	A	B	C
	α	10	1
	α	20	1
	β	30	1
	β	40	2



$\pi_{A,C}(r)$	A	C
	α	1
	β	1
	β	2

4. Projection Operation (cont.)

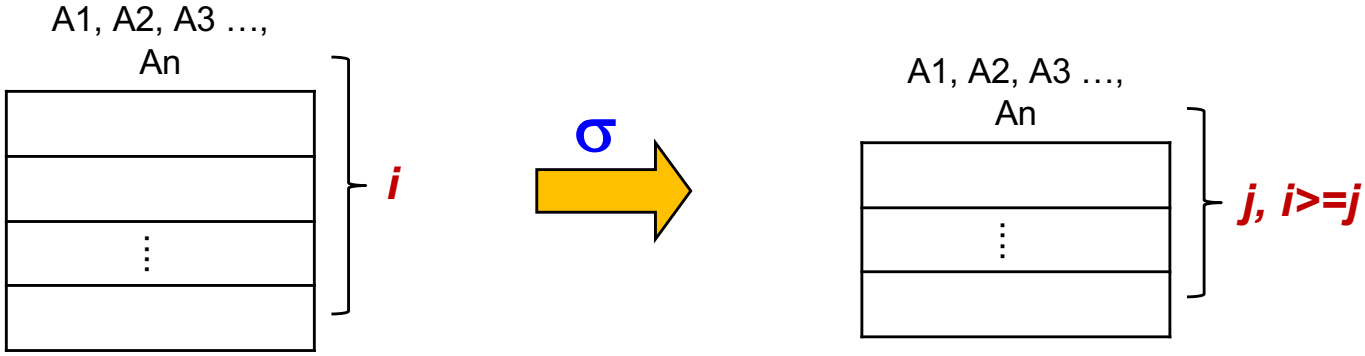
- The projection is *not commutative operations*

$$\pi_{X,Y}(r) = \pi_X(\pi_Y(r))$$

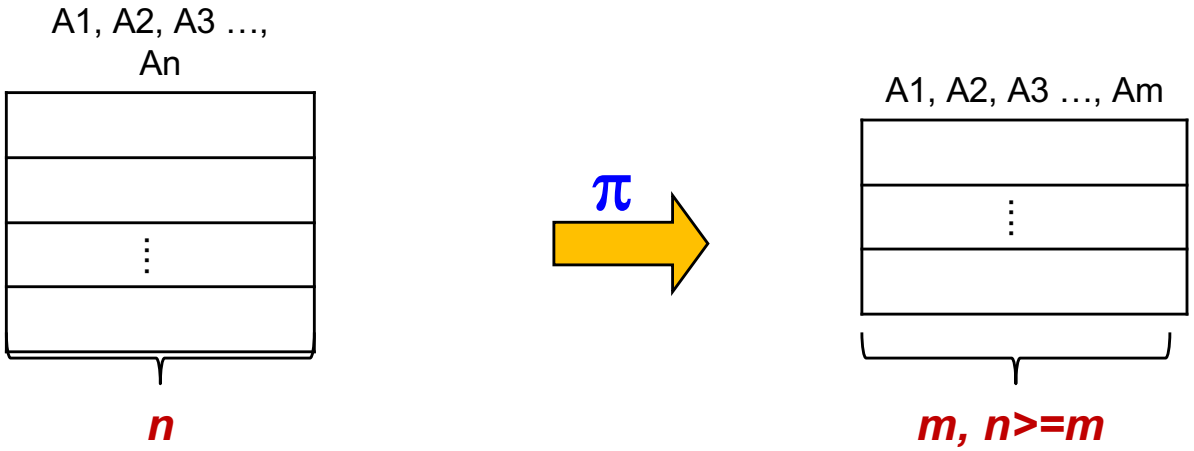
$$\pi_{A1, A2, \ldots, An}(\pi_{A1, A2, \ldots, Am}(r)) = \pi_{A1, A2, \ldots, An}(r) \text{ , với } n \leq m$$

4. Selection vs Projection

Selection



Projection



4. SQL – Rel. Algebra Mapping

List all columns

The entire tuple is produced

```
SELECT *
FROM KHOA
WHERE PHONG='I53'
AND NAMTL = '1995'
```

MaKhoa	TenKhoa	Phong	NamTL	DienThoai	TruongKhoa	NgayNhanChuc
CNTT	Công nghệ thông tin	I53	1995	08313964145	GV130	01/01/2007
~~SH~~	~~Sinh học~~	~~B32~~	~~1975~~	~~08313123545~~	~~GV250~~	~~01/01/1990~~

$$\sigma_{PHG='I53' \wedge NamTL='1995'}(KHOA)$$

4. SQL – Rel. Algebra Mapping

List some specific columns

```
SELECT MAKHOA, TENKHOA, PHONG
FROM KHOA
WHERE PHONG='I53'
AND NAMTL = '1995'
```

MaKhoa	TenKhoa	Phong
CNTT	Công nghệ thông tin	I53

$$\pi_{MAKHOA,\ TENKHOA,\ PHONG}(\sigma_{PHG='I53' \wedge NamTL='1995'}(KHOA))$$

4. Example 1

- Retrieve the name and salary of female teachers

$$\pi_{\text{HOTEN, LUONG}}(\sigma_{\text{PHAI='Nữ'}}(\text{GIAOVIEN}))$$

4. Example 2

- List out the codes/IDs of teachers who belong to the dept. “HTTT” or participate in the topic code “001”

$$\pi_{MAGV}(\sigma_{MABM='HTTT'}(GIAOVIEN)) \cup \pi_{MAGV}(\sigma_{MAĐT='001'}(TG\_ĐETAI))$$

4. Example 3

- Retrieve the code/ID of the deans who are also the project managers

$$\pi_{TRUONGKHOA}(KHOA) \cap \pi_{GVCNĐT}(ĐETAI)$$

4. Example 4

- Retrieve the names of the jobs that started between 01/01/2007 and 01/08/2007

$$\sigma_{(NGAYBĐ>='1/1/2007' \wedge NGAYBĐ<='1/8/2007')}(CONGVIEC)$$

Quiz #1

- Given the schema **PHANCONG**(MÃNV, MÃĐA, THOIGIAN) and the query "*Cho biết danh sách mã nhân viên vừa có tham gia đề án số 1 vừa có tham gia đề án số 2*". Choose correct answers.

A. $\Pi_{MÃNV}(\sigma_{MÃĐA =1 \text{ AND } MÃĐA =2}(PHÂNCÔNG))$

B. $\Pi_{MÃNV}((\sigma_{MÃĐA =1}(PHÂNCÔNG) \cap (\sigma_{MÃĐA =2}(PHÂNCÔNG))$

C. $\Pi_{MÃNV}(\sigma_{MÃĐA =1}(PHÂNCÔNG)) \cap \Pi_{MÃNV}(\sigma_{MÃĐA =2}(PHÂNCÔNG))$

4. Example 5

- Example
  - Retrieve the name and salary of teachers after increasing 10%

$$\pi_{\text{HOTEN, LUONG*1.1}}\,(\text{GIAOVIEN})$$

4. Sequence of Operations

- Combining algebra operations
  - Combining expressions

$$\pi_{A1, A2, \ldots, Ak}(\sigma_P(r)) \qquad \sigma_P(\pi_{A1, A2, \ldots, Ak}(r))$$

  - Execute step by step
    - B1

$$\sigma_P(r)$$

    - B2

$$\pi_{A1, A2, \ldots, Ak}(\text{result of B1})$$

Rename the resulting relation

4. Assignment

- Is used to retrieve a result relation from an algebra operation

- Notation: $\leftarrow$
- Example:
  - B1
  - B2 $s \leftarrow \sigma_P(r)$

$KQ \leftarrow \pi_{A1, A2, \ldots, Ak}(s)$

4. Rename Operation

- We may want to apply several relational algebra operations one after the other. Either we can write the operations as a single **relational algebra expression** by nesting the operations, or we can apply one operation at a time and create **intermediate result relations**. In the latter case, we must give names to the relations that hold the intermediate results.
  - Given relation r(B, C, D)

$$\rho_s(r) \ : \ \text{rename the relation } \mathbf{r} \text{ to } \mathbf{s}$$

  - Rename attribute

$$\rho_{X,\ C,\ D}(r) \ : \ \text{rename the attribute B to X}$$

  - Rename relation and attribute:

$$\rho_{s(X,C,D)}(r) : \text{rename the relation r to s, and the attribute B to X}$$

4. Example 5

- Retrieve the code/ID and name of teachers who work in the dept. “HTTT”.

$$\pi_{MAGV,\ HOTEN}(\sigma_{MABM='HTTT'}(GIAOVIEN))$$

- C1:

$$GV\_HTTT \leftarrow \sigma_{MABM='HTTT'}(GIAOVIEN)$$

$$KQ \leftarrow \pi_{MAGV,\ HOTEN}(GV\_HTTT)$$

- C2:

$$KQ(MA,\ TEN) \leftarrow \pi_{MAGV,\ HOTEN}(GV\_HTTT)$$

$$\rho_{KQ(MA,\ TEN)}(\pi_{MAGV,\ HOTEN}(GV\_HTTT))$$

cdio 4.0

fit@hcmus

Content

1. Introduction
2. Relational Algebra Operations From Set Theory
3. Select Operation
4. Projection Operation
5. CARTESIAN Operations
6. Join Operation
7. Division Operation
8. Additional Relational Operations
9. Update Operations

5. CARTESIAN Operations

- This operation is used to combine tuples from two relations in a combinatorial fashion.
- Notation

$$r \times s$$

- The result relation is a relation q with:
  - Each tuple of q is a combination between 1 tuple in **r** and 1 tuple in **s**
  - If the relation **r** has *u* tuples and **s** has *v* tuples, then **q** has $u \times v$ tuples
  - If **r** has *n* attributes and **s** has *m* attributes then q has $n + m$ attributes ($R^+ \cap S^+ = \varnothing$)

5. CARTESIAN Operations - Example

r	A	B
	α	1
	β	2

S	X	C	D
	α	10	+
	β	10	+
	β	20	-
	γ	10	-

unambiguous

r × s	A	R.B	X	C	D
	α	1	α	10	+
	α	1	β	10	+
	α	1	β	20	-
	α	1	γ	10	-
	β	2	α	10	+
	β	2	β	10	+
	β	2	β	20	-
	β	2	γ	10	-

$\rho_{(X,C,D)}(s)$

5. CARTESIAN Operations

- Usually, the Cartesian operation is followed by the select operation

$r \times s$

A	R.B	S.B	C	D
α	1	α	10	+
α	1	β	10	+
α	1	β	20	-
α	1	γ	10	-
β	2	α	10	+
β	2	β	10	+
β	2	β	20	-
β	2	γ	10	-

$\sigma_{A=S.B}(r \times s)$

A	R.B	S.B	C	D
α	1	α	10	+
β	2	β	10	+
β	2	β	20	-

5. CARTESIAN Operations – Example 6

- Retrieve the information of the dept. and the information of the lecturer who is the head of that department

TENBM	MABM	TRUONGBM	NGAYNHANCHUC	…
Hệ thống thông tin	HTTT	002	20/09/2004	
Công nghệ tri thức	CNTT			
Mạng máy tính	MMT	001	15/05/2005	

MAGV	HOTEN	NGSINH	MABM	PHAI	LUONG	….
001	Nguyễn Hoài An	15/02/1973	MMT	Nam	2000	
002	Trần Trà Dương	20/06/1960	HTTT	Nu	2500	
003	Nguyễn Ngọc Anh	11/05/1975	HTTT	Nu	2200	
004	Trương Nam Sơn	20/06/1959	VS	Nam	2300	

5. CARTESIAN Operations – Example 6

TENBM	MABM	TRUONGBM	NGAYNHANCHUC	GV	HOTEN	…
Hệ thống thông tin	HTTT	**002**	20/09/2004	**002**	Trần Trà Dương	
Mạng máy tính	MMT	**001**	15/05/2005	**001**	Trương Nam Sơn	

5. CARTESIAN Operations – Example 6

- B1: Cartesian operation BOMON and GIAOVIEN

  BM_GV ← (GIÁOVIÊN × BỘMÔN)

- B2: select tuples that satisfy TRUONGBM = MAGV

  $KQ \leftarrow \sigma_{TRUONGBM=MAGV}(BM\_GV)$

GIÁOVIÊN

MÃGV	HỌTÊN	….	NGÀYSINH	SỐNHÀ	….
001	Nguyễn Hoài An	….	15/02/1973	25/3	….
002	Trần Trà Hương	….	20/06/1960	125	….
003	Nguyễn Ngọc Ánh	….	11/05/1975	12/21	….
….	….	….	….	….	….

BỘMÔN

MÃBM	TÊNBM	PHÒNG	….	TRƯỞNGBM	….
HTTT	Hệ thống thông tin	B13	….	002	….
CNTT	Công nghệ tri thức	B15	….		….
MMT	Mạng máy tính	B16	….	001	….
….	….	….	….	….	….

$\sigma_{TRUONGBM=MAGV}(BM\_GV)$

MÃGV	HỌTÊN	….	NGÀYSINH	….	MÃBM	TÊNBM	PHÒNG	….	TRƯỞNGBM	….
001	Nguyễn Hoài An		15/02/1973	….	HTTT	Hệ thống thông tin	B13		002	
001	Nguyễn Hoài An		15/02/1973	….	CNTT	Công nghệ tri thức	B15			
001	Nguyễn Hoài An	….	15/02/1973	….	MMT	Mạng máy tính	B16	….	001	….
….	….	….	….	….	….	….	….	….	….	….

© Bộ môn HTTT - Khoa CNTT – Trường ĐH KHTN

5. CARTESIAN Operations – Example 7

- Retrieve the highest salary of the lecturers?
  - GIANGVIEN(MSGV, HOTEN, LUONG, …)

HOTEN	…	LUONG	…	…	LUONG	…
Nguyễn Hoài An	…	2000	…	…	2000	…
Trần Trà Hương	…	2500	…	…	2500	…
Nguyễn Ngọc Anh	…	2200	…	…	2200	…
	…		…	…		…

5. CARTESIAN Operations – Example 7

- B1: Retrieve the salaries that are not highest (R3)

$$r1 \leftarrow (\pi_{LUONG}(GIAOVIEN))$$

$$r2 \leftarrow \sigma_{GIAOVIEN.LUONG < R1.LUONG}(GIAOVIEN \times r1)$$

$$r3 \leftarrow \pi_{R2.LUONG}(r2)$$

- B2: Full salary set minuses the salary set in R3

$$KQ \leftarrow \pi_{LUONG}(GIAOVIEN) - r3$$

5. CARTESIAN Operations – Example 8

- Retrieve the name of lecturers working the same department with ‘Trần Trà Hương’
  - Relation: GIAOVIEN
  - Attributes: HOTEN, MABM
  - Condition: HOTEN = ‘Trần Trà Hương’

What dept. “Trần Trà Hương” belongs to?

MABM	HOTEN
MMT	Nguyễn Hoài An
HTTT	Trần Trà Hương
HTTT	Nguyễn Ngọc Anh
VS	Trương Nam Sơn
…	…

What lecturers belong to that dept.?

MABM	HOTEN
MMT	Nguyễn Hoài An
HTTT	Trần Trà Hương
HTTT	Nguyễn Ngọc Anh
VS	Trương Nam Sơn
…	…

5. CARTESIAN Operations – Example 8

- B1: Find the dept. of 'Trần Trà Hương'

$$r1 \leftarrow \pi_{MABM,MAGV}(\sigma_{HOTEN='Trần\ Trà\ Hương'}(GIAOVIEN))$$

- B2: Retrieve the name of lecturers of that dept.

$$r2 \leftarrow \sigma_{HOTEN <>'Trần\ Trà\ Hương'}(GIAOVIEN)$$

$$r3 \leftarrow \sigma_{R1.MABM=R2.MABM}(r1 \times r2)$$

$$KQ \leftarrow \pi_{HOTEN}(r3)$$

Content

1. Introduction
2. Relational Algebra Operations From Set Theory
3. Select Operation
4. Projection Operation
5. CARTESIAN Operations
6. **Join Operation**
    1. Natural join (kết tự nhiên)
    2. Theta join (kết có điều kiện tổng quát)
    3. Equi join (kết bằng)
7. Division Operation
8. Additional Relational Operations
9. Update Operations

6. Join Operation

- To identify and select related tuples from two relations
- Denoted by: r ⋈ s
  - R(A_1, A_2, …, A_n) and S(B_1, B_2, …, B_m)
- The result relation is a relation q
  - Haivng *n* + *m* attributes: Q(A_1, A_2, …, A_n, B_1, B_2, …, B_m)
  - Each tuple of q is a combination of 1 in r and 1 in s that satisfy join condition:
    - Join condition: A_i θ B_j
    - A_i attribute of R, B_j attribute of S
    - A_i and B_j have the same value domain
    - θ is a comparision: ≠, =, <, >, ≤, ≥

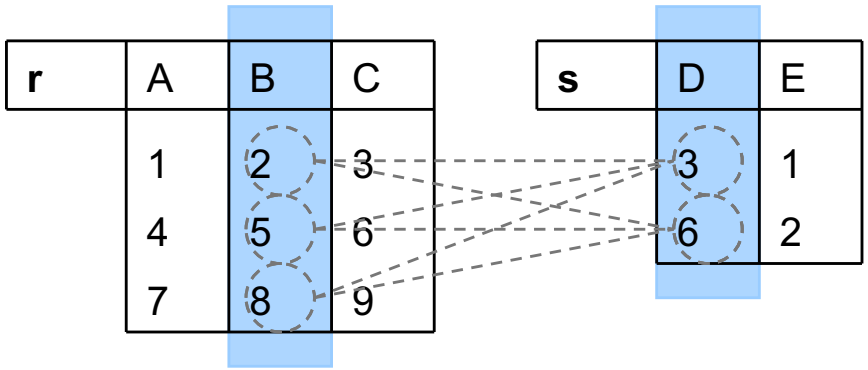
6. Join Operation

- **Theta join**
  - Notation $r \bowtie_C s$
  - C join condition
- **Equi join**: when the condition C is a equal join

- **Natural join**: equi join on two attribute that have the same name:
  - Denoted by: $r \bowtie s$ or $r * s$
  - $R^+ \cap S^+ \neq \varnothing$
  - The attributes of the result relation are in **r** or **s**, minus the attributes in the join condition

6. Join Operation

Theta join - Example



r	A	B	C
	1	2	3
	4	5	6
	7	8	9

s	D	E
	3	1
	6	2

$r \bowtie_{B<D} s$

A	B	C	D	E
1	2	3	3	1
1	2	3	6	2
4	5	6	6	2

$$r \bowtie_C s = \sigma_C(r \times s)$$

6. Join Operation

Equi join - Example

r	A	B	C
	1	2	3
	4	5	6
	7	8	9

s	D	E
	3	1
	6	2

$r \bowtie_{C=D} s$

A	B	C	D	E
1	2	3	3	1
4	5	6	6	2

r	A	B	C
	1	2	3
	4	5	6
	7	8	9

s	SCC	D
	3	1
	6	2

$\rho_{(S.C,D)}\ s$

$r \bowtie_{C=s.C} s$

A	B	C	s.C	D
1	2	3	3	1
4	5	6	6	2

6. Join Operation

Natural join - Example

r	A	B	C
	1	2	3
	4	5	6
	7	8	9

s	C	D
	3	1
	6	2

$r \bowtie s$

A	B	C	C	D
1	2	3	3	1
4	5	6	6	2

6. Join Operation – Example 10

- Retrieve the lecturers who have the salary greater than the lecturer 'Nguyễn Hoài An'
  - Relation: **GIAOVIEN**
  - Attribute: **LUONG**

GIAOVIEN(MAGV, HOTEN, **LUONG**, PHAI, NGAYSINH,…)

$R1(LG) \leftarrow \pi_{LUONG}(\sigma_{HOTEN='Nguyễn\ Hoài\ An'}(GIAOVIEN))$

$KQ \leftarrow GIAOVIEN \bowtie_{LUONG>LG} R1$

KQ(MAGV, HOTEN, **LUONG**, PHAI, NGAYSINH,…, **LG**))

6. Join Operation – Example 11

- For each lecturer, list out the information of the department that they are working for
  - Relation: **GIAOVIEN**, **BOMON**

GIAOVIEN(MAGV, HOTEN, LUONG, PHAI, …, **MABM**, …)

BOMON(**MABM**, TENBM, PHONG, DIENTHOAI, …)

KQ ← GIAOVIEN ⋈ BOMON

KQ(MAGV, HOTEN, …, **MABM**, TENBM, PHONG, …))

6. Join Operation – Example 12

- For each project, list out the information of the lecturers who manage that project
  - Relation: **ĐETAI**, **GIAOVIEN**

ĐETAI(MAĐT, TENĐT, KINHPHI, …, **GVCNĐT**)

GIAOVIEN(**MAGV**, HOTEN, LUONG, PHAI, …)

$KQ \leftarrow ĐETAI \bowtie_{GVCNĐT = MAGV} GIAOVIEN$

KQ(MAĐT, TENĐT, KINHPHI, …, **GVCNĐT, MAGV,** HOTEN, …)

6. Join Operation – Example 13

- Retrieve the highest salary of the department ‘HTTT’?

Quiz

- Given the DB schema below:
  - PHÒNG(Mã_Phòng, Diện_tích, Giá_tiền, Loại)
  - THUÊ(Mã_Phòng, Mã_KH, Ngày_nhận , Số_ngày)
  - KHÁCH_HÀNG(Mã_KH, Họ_tên, Điện_thoại, Giới_tính)
- Find out the name of customers who rent the room 'R101':

A) $\pi_{\text{Họ\_Tên}}(\sigma_{\text{Mã\_Phòng}='\text{R101}'}(\text{KHÁCH\_HÀNG} * \text{THUÊ}))$

B) $\pi_{\text{Họ\_Tên}}(\sigma_{\text{Mã\_Phòng}='\text{R101}'}(\text{KHÁCH\_HÀNG} \bowtie_{\text{KHÁCH\_HÀNG.Mã\_KH} <> \text{THUÊ.Mã\_KH}} \text{THUÊ}))$

C) $\pi_{\text{Họ\_Tên}}(\text{KHÁCH\_HÀNG} \bowtie_{\text{KHÁCH\_HÀNG.Mã\_KH} = \text{THUÊ.Mã\_KH AND Mã\_Phòng} <> '\text{R101}'} \text{THUÊ}))$

D) Chỉ có A và B đúng.

E) Tất cả A, B và C đúng.

Complete Set of Relational Operations

- The set of operations including **select σ, project π , union ∪, set difference -** , and **cartesian product X** is called a complete set because any other relational algebra expression can be expressed by a combination of these five operations.

  - Example
    - $r \cap s = r \cup s - ((r-s) \cup (s-r))$
    - $r \bowtie_c s = \sigma_c(r \times s)$

cdio 4.0

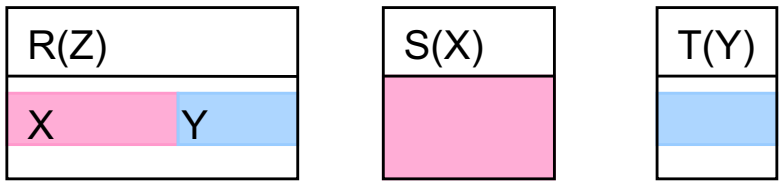
fit@hcmus

Content

1. Introduction
2. Relational Algebra Operations From Set Theory
3. Select Operation
4. Projection Operation
5. CARTESIAN Operations
6. Join Operation
7. Division Operation
8. Additional Relational Operations
9. Update Operations

7. Division Operation

- To extract tuples in the relation **r** that satisfy <u>all tuples</u> in the relation **s**
- Denoted by r ÷ s
  - r(Z) và x(X)
    - Z are attrbites of **r**, X are attributes of **s**
    - $X \subseteq Z$
- The result relation of the division **t(Y)**
  - With Y=Z-X
  - t_0 is a tuple of **t** if all tuples $t_S \in$ **s**, there exists a tuple $t_R \in$ **r** that satisfies 2 two below conditions:
    - $t_R(Y) = t_0$
    - $t_R(X) = t_S(X)$



7. Division Operation - Example

r ÷ s

r	A	B	C	D	E
	α	a	α	a	1
	α	a	γ	a	1
	α	a	γ	b	1
	β	a	γ	a	1
	β	a	γ	b	3
	γ	a	γ	a	1
	γ	a	γ	b	1
	γ	a	β	b	1

s	D	E
	a	1
	b	1

A	B	C
α	a	γ
γ	a	γ

7. Division Operation - Example

DANGKY	MASV	MANH
	SV01	MH01
	SV02	MH01
	SV03	MH01
	SV01	MH02
	SV02	MH02
	SV03	MH03
	SV01	MH03
	SV04	MH01

MONHOC	MANH
	MH01
	MH02
	MH03

DANGKY ÷ MONHOC	MASV
	SV01

7. Division Operation – Example 17

- Retrieve the codes/IDs of lecturers who participate all works of the project “001”

7. Division Operation – Example 18

- Retrieve the name of the projects that all lecturers of the dept. “Hệ thống thông tin” are working on.

7. Division Operation

- **Division operation** can be expressed through a complete set of relational algebra operations

$$Q1 \leftarrow \pi_Y(r)$$

$$Q2 \leftarrow Q1 \times s$$

$$Q3 \leftarrow \pi_Y(Q2 - r)$$

$$KQ \leftarrow Q1 - Q3$$

Recap of Relational Algebra Operations

TABLE 6.1 OPERATIONS OF RELATIONAL ALGEBRA

Operation	Purpose	Notation
SELECT	Selects all tuples that satisfy the selection condition from a relation R.	$\sigma_{\text{<SELECTION CONDITION>}}(R)$
PROJECT	Produces a new relation with only some of the attributes of R, and removes duplicate tuples.	$\pi_{\text{<ATTRIBUTE LIST>}}(R)$
THETA JOIN	Produces all combinations of tuples from R_1 and R_2 that satisfy the join condition.	$R_1 \bowtie_{\text{<JOIN CONDITION>}} R_2$
EQUIJOIN	Produces all the combinations of tuples from R_1 and R_2 that satisfy a join condition with only equality comparisons.	$R_1 \bowtie_{\text{<JOIN CONDITION>}} R_2$, OR $R_1 \bowtie (\text{<JOIN ATTRIBUTES 1>}), (\text{<JOIN ATTRIBUTES 2>})\ R_2$
NATURAL JOIN	Same as EQUIJOIN except that the join attributes of R_2 are not included in the resulting relation; if the join attributes have the same names, they do not have to be specified at all.	$R_1 *_{\text{<JOIN CONDITION>}} R_2$, OR $R_1 * (\text{<JOIN ATTRIBUTES 1>}), (\text{<JOIN ATTRIBUTES 2>}) R_2$ OR $R_1 * R_2$
UNION	Produces a relation that includes all the tuples in R_1 or R_2 or both R_1 and R_2; R_1 and R_2 must be union compatible.	$R_1 \cup R_2$
INTERSECTION	Produces a relation that includes all the tuples in both R_1 and R_2; R_1 and R_2 must be union compatible.	$R_1 \cap R_2$
DIFFERENCE	Produces a relation that includes all the tuples in R_1 that are not in R_2; R_1 and R_2 must be union compatible.	$R_1 - R_2$
CARTESIAN PRODUCT	Produces a relation that has the attributes of R_1 and R_2 and includes as tuples all possible combinations of tuples from R_1 and R_2.	$R_1 \times R_2$
DIVISION	Produces a relation $R(X)$ that includes all tuples $t[X]$ in $R_1(Z)$ that appear in R_1 in combination with every tuple from $R_2(Y)$, where $Z = X \cup Y$.	$R_1(Z) \div R_2(Y)$

Content

1. Introduction
2. Relational Algebra Operations From Set Theory
3. Select Operation
4. Projection Operation
5. CARTESIAN Operations
6. Join Operation
7. Division Operation
8. Additional Relational Operations
   - Aggregate function (Hàm kết hợp)
   - Grouping (gom nhóm)
   - Outer join (Kết ngoài)
9. Update Operations

8. Aggregate functions

- A type of request that cannot be expressed in the basic relational algebra is to specify mathematical **aggregate functions** on collections of values from the database.

- Examples of such functions include retrieving the average or total salary of all employees or the total number of employee tuples. These functions are used in simple statistical queries that summarize information from the database tuples.

- Common functions applied to collections of numeric values include **SUM, AVERAGE, MAXIMUM,** and **MINIMUM**. The **COUNT** function is used for counting tuples or values.

8. Aggregate functions

- Input: set of values or relation
- Output: single value

  - AVG
  - MIN
  - MAX
  - SUM
  - COUNT

r	A	B
	1	2
	3	4
	1	2
	1	2

SUM(B) = 10
AVG(A) = 1.5
MIN(A) = 1
MAX(B) = 4
COUNT(A) = 4

8. Grouping

- Grouping the relation into groups based on a grouping condition
- Notation

$$G1, G2, \ldots, Gn\,\mathfrak{S}_{F1(A1), F2(A2), \ldots, Fn(An)}(R)$$

  - R: Relation/Relational algebra expression
  - G1, G2, …, Gn: Grouping attributes.
  - F1, F2, …, Fn: Aggregate functions.
  - A1, A2, …, An: R's Attributes.

8. Grouping - Example

r	A	B	C
	α	2	7
	α	4	7
	β	2	3
	γ	2	10

$\Im_{SUM(C)}(r)$

SUM_C
27

$_{A}\Im_{SUM(C)}(r)$

A	SUM_C
α	14
β	3
γ	10

8. Grouping – Example 19

- Retrieve the number of lecturers and their summary of salary

8. Grouping – Example 19

- Retrieve the number of lecturers and average of salary of each department

8. Grouping – Example 20

- What is/are the name(s) of the faculty with the most teachers?
  - Calculate the number of lecturers according to faculties

    $r1 \leftarrow {}_{MaKhoa,\ TênKhoa}\Im_{COUNT(HoTen)}(GIAOVIEN * BOMON * KHOA)$

  - Get the highest number of lecturers

    $r2(MAX\_COUNT) \leftarrow \Im_{MAX(COUNT\_HoTen)}(r1)$

  - Faculty with the most lecturers

    $KQ \leftarrow \pi_{MaKhoa,\ TenKhoa}(r1 \bowtie_{COUNT\_HoTen = MAX\_COUNT} r2)$

8. Grouping – Example 21

- What are the subject and the number of projects of that subject?

Quiz

- DB schema:
  - PHÒNG(Mã_Phòng, Diện_tích, Giá_tiền, Loại)
  - THUÊ(Mã_Phòng, Mã_KH, Ngày_nhận , Số_ngày)
  - KHÁCH_HÀNG(Mã_KH, Họ_tên, Điện_thoại, Giới_tính)
- Retrieve the types of rooms that have never been rented :

A. $\pi_{Loại}(\sigma_{COUNT\_Mã\_Phòng=0}(_{Loại}\Im_{COUNT(Mã\_Phòng)}(THUÊ * PHÒNG)))$

B. $\pi_{Mã\_Phòng, Loại}(KHÁCH\_HÀNG * THUÊ * PHÒNG) \div \pi_{Mã\_Phòng}(PHÒNG)$

C. $\pi_{Mã\_Phòng, Loại}(PHÒNG) \div \pi_{Mã\_Phòng}(PHÒNG)$

D. $\pi_{Loại}(KHÁCH\_HÀNG * THUÊ * PHÒNG) - \pi_{Loại}(PHÒNG)$

E. All wrong

8. Outer Join

- Extend the JOIN operation to avoid information loss
  - Perform the Join operation
  - Get more tuples that do not satisfy the join condition
- 3 types:
  - Left outer join ⟕
  - Right outer join ⟖
  - Full outer join ⟗

8. Outer Join – Example 22

- Retrieve the name of the lecturer and the name of the department he/she is heading, if any

$$R1 \leftarrow GIAOVIEN \ ⟕_{MAGV=TRUONGBM} \ BOMON$$

$$KQ \leftarrow \pi_{HOTEN,\ TENBM}(R1)$$

HOTEN	TENBM
Nguyễn Hoài An	Mạng máy tính
Trần Trà Hương	Hệ thống thông tin
Nguyễn Ngọc Ánh	null
…	…

8. Outer Join – Example 23

- Retrieve the dept. name and dept. head's name if any

Content

1. Introduction
2. Relational Algebra Operations From Set Theory
3. Select Operation
4. Projection Operation
5. CARTESIAN Operations
6. Join Operation
7. Division Operation
8. Additional Relational Operations
9. Update Operations

9. Update Operations

- Date can be updated by:
  - Insertion
  - Deletion
  - Update/Modify
- Update Operations is expressed through the assignment:

$$r_{new} \leftarrow \text{relational operations } r_{old}$$

9. Insertion

- Denoted by

$$r_{new} \leftarrow r_{old} \cup E$$

  - r relation
  - E relational algebra expression
- Example
  - Assign lecturer code “001” to participate in work 4 of the project “001” with an allowance level 2.

$$THAMGIAĐT \leftarrow THAMGIAĐT \cup (‘001’, ‘001’, 4, 2)$$

9. Deletion

- Denoted by

$$r_{new} \leftarrow r_{old} - E$$

  - r relation
  - E relational algebra expression

- Example
  - Remove works of the lecturer “001”

$$THAMGIAĐT \leftarrow THAMGIAĐT - \sigma_{MAGV='001'}(THAMGIAĐT)$$

9. Update/ Modify

- Denoted by

$$r_{new} \leftarrow \pi_{F1, F2, \ldots, Fn} (r_{old})$$

  - r relation
  - Fi expression that calculates the new value for the updating attribute

- Example
  - Increase the allowance for all works up to 1.5 times

$$THAMGIAĐT \leftarrow \pi_{MAGV, MAĐT, STT, PHUCAP*1.5}(THAMGIAĐT)$$