

# NGÔN NGỮ TRUY VẤN ĐẠI SỐ QUAN HỆ

# NGÔN NGỮ ĐẠI SỐ QUAN HỆ

- ❖ Là một ngôn ngữ truy vấn thủ tục, bao gồm các phép toán tập hợp một ngôi hoặc hai ngôi, nghĩa là các toán hạng của chúng là một quan hệ hoặc hai quan hệ. Kết quả đầu ra là một quan hệ.
- ❖ Năm phép toán cơ bản của đại số quan hệ là phép chọn, phép chiếu, phép hợp, phép trừ và phép tích Đề-các.
- ❖ Một số các phép toán mở rộng khác cũng được định nghĩa trong đại số quan hệ bao gồm: phép giao, kết nối tự nhiên, phép chia, phép kết nối nửa và kết nối ngoài.

# PHÉP CHỌN

Loại: Một ngôi

Ký hiệu: Sigma,  $\sigma$

Khuôn dạng chung:  $\sigma_{(\text{mệnh đề})}(\text{thể hiện của quan hệ})$

Lược đồ của quan hệ kết quả: tương tự như quan hệ toán hạng

Kích thước của quan hệ kết quả (số bộ):  $\leq | \text{quan hệ toán hạng} |$

*Ví dụ:*

$\sigma_{(\text{major} = \text{"CS"})}(\text{STUDENTS})$

$\sigma_{(\text{major} = \text{"CS"} \wedge \text{hair-color} = \text{"brown"})}(\text{STUDENTS})$

$\sigma_{(\text{hours-attempted} > \text{hours-earned})}(\text{STUDENTS})$

- Phép chọn lựa chọn ra các bộ từ thể hiện của quan hệ sao cho thỏa mãn mệnh đề điều kiện cụ thể nào đó.
- Mệnh đề có thể chứa các toán tử so sánh, như  $=$ ,  $\neq$ ,  $<$ ,  $\leq$ ,  $>$ ,  $\geq$ . Hoặc kết hợp với các phép toán liên kết và  $(\wedge)$ , hoặc  $(\vee)$ , và phủ định  $(\neg)$ .
- Phép chọn có thể được coi như một lát cắt ngang của quan hệ toán hạng.

# VÍ DỤ PHÉP CHỌN

R

A	B	C	D
a	a	yes	1
b	d	no	7
c	f	yes	34
a	d	no	6
a	c	no	7
b	b	no	69
c	a	yes	24
d	d	yes	47
h	d	yes	34
e	c	no	26
a	a	yes	5

$r = \sigma_{(A='a')}(R)$

A	B	C	D
a	a	yes	1
a	d	no	6
a	c	no	7
a	a	yes	5

$r = \sigma_{(A='a' \wedge C='yes')}(R)$

A	B	C	D
a	a	yes	1
a	a	yes	5

$r = \sigma_{(B='m')}(R)$

A	B	C	D
---	---	---	---

Một quan  
hệ rỗng

# PHÉP CHIẾU

Loại: Một ngôi

Ký hiệu:  $\Pi$ ,  $\pi$

Khuôn dạng chung:  $\pi_{(\text{danh sách các thuộc tính})}(\text{thể hiện của quan hệ})$

Lược đồ của quan hệ kết quả: được xác định bởi <danh sách các thuộc tính>

Kích thước của quan hệ kết quả (số bộ):  $\leq$  | quan hệ toán hạng |

*Ví dụ:*

$\pi_{(\text{student-id, name, major})}(\text{STUDENTS})$

$\pi_{(\text{name, advisor})}(\text{STUDENTS})$

$\pi_{(\text{name, gpa, hours-attempted})}(\text{STUDENTS})$

- Phép chiếu có thể được coi như một lát cắt dọc của quan hệ toán hạng.
- Nếu phép toán sinh ra các bộ giống hệt nhau, thì sẽ chỉ giữ lại một bộ và loại bỏ đi các bộ bị trùng.

# VÍ DỤ PHÉP CHIẾU

R

A	B	C	D
a	a	yes	1
b	d	no	7
c	f	yes	34
a	d	no	6
a	c	no	7
b	b	no	69
c	a	yes	24
d	d	yes	47
h	d	yes	34
e	c	no	26
a	a	yes	5

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

A	C
a	yes
b	no
c	yes
a	no
d	yes
h	yes
e	no

$r = \pi_{(A, D)}(R)$

A	D
a	1
b	7
c	34
a	6
a	7
b	69
c	24
d	47
h	34
e	26
a	5

$r = \pi_{(C)}(R)$

C
yes
no

# PHÉP HỢP

Loại: hai ngôi

Ký hiệu:  $\cup$

Khuôn dạng chung:  $r \cup s$ , với  $r$  và  $s$  là 2 quan hệ khả hợp

Lược đồ quan hệ kết quả: Lược đồ của các quan hệ toán hạng

Kích thước của quan hệ kết quả (số bộ):  $\leq \max\{|r| + |s|\}$

Ví dụ:

$$r \cup s$$

$$\pi_{(a, b)}(r) \cup \pi_{(a, b)}(s)$$

- Phép hợp cung cấp một phương tiện để trích lọc thông tin nằm trên hai quan hệ toán hạng khả hợp với nhau.
- **2 quan hệ  $r(R)$  và  $s(S)$  được gọi là khả hợp** khi thỏa mãn 2 điều kiện sau:
  - ✓ Chúng phải có cùng số bậc hay cùng số lượng thuộc tính.
  - ✓ Miền giá trị của thuộc tính thứ (i) của  $r$  và thuộc tính thứ (i) của  $s$  phải giống nhau, cho mọi giá trị của  $i$ .

# VÍ DỤ PHÉP HỢP

R

A	B	D
a	a	1
b	d	7
c	f	34
a	d	6
a	c	7

T

A	B
a	a
b	d
c	f
a	d
a	c

$$r = R \cup T$$

Không hợp lệ – R và T  
không phải là 2 quan hệ  
khả hợp

$$r = R \cup S$$

E	F	G
a	a	1
b	d	7
c	f	34
a	d	6
a	c	7
a	m	4
b	c	22
a	d	16

S

X	Y	Z
a	m	4
b	c	22
a	d	16
a	c	7

X

A	B
a	a
b	d
a	c

$$r = T \cup X$$

A	B
a	a
b	d
c	f
a	d
a	c



# PHÉP TRỪ

Loại: Hai ngôi

Ký hiệu:  $-$

Khuôn dạng chung:  $r - s$ , với  $r$  và  $s$  là hai quan hệ khả hợp

Lược đồ quan hệ kết quả: Lược đồ của quan hệ toán hạng

Kích thước của quan hệ kết quả (số bộ):  $\leq | \text{quan hệ } r |$

Ví dụ:  $r - s$

- Phép trừ cho phép trích lọc thông tin được chứa trong một quan hệ mà nó không được chứa trong quan hệ thứ hai.
- Tương tự như phép hợp, phép trừ yêu cầu 2 quan hệ toán hạng phải là khả hợp.

# VÍ DỤ PHÉP TRỪ

R

A	B	D
a	a	1
b	d	7
c	f	34
a	d	6
a	c	7

T

A	B
a	a
b	d
c	f
a	d
a	c

$$r = R - T$$

Không hợp lệ – R và T  
không phải là 2 quan hệ  
khả hợp

$$r = T - X$$

A	B
c	f
a	d

$$r = R - S$$

E	F	G
a	a	1
b	d	7
c	f	34
a	d	6

S

X	Y	Z
a	m	4
b	c	22
a	d	16
a	c	7

X

A	B
a	a
b	d
a	c

$$r = X - T$$

A	B
---	---

Quan hệ rỗng

$$r = S - R$$

E	F	G
a	m	4
b	c	22
a	d	16

# PHÉP TÍCH ĐỀ-CÁC

Loại: Hai ngôi

Ký hiệu:  $\times$

Khuôn dạng chung:  $r \times s$  (không giới hạn trên  $r$  và  $s$ )

Lược đồ quan hệ kết quả: lược đồ  $r \times$  lược đồ  $s$  với việc thay đổi tên gọi một số thuộc tính

Kích thước của quan hệ kết quả (số bộ):  $> | \text{quan hệ } r |$  và  $> | \text{quan hệ } s |$

Ví dụ:

$$r \times s$$

- Tích Đề-các cho phép kết nối 2 quan hệ bất kỳ thành một quan hệ đơn.
- Một quan hệ là một tập con của tích Đề-các tập các miền giá trị.

# VÍ DỤ PHÉP TÍCH ĐỀ-CÁC

T

A	B
a	a
b	d

X

A	B
a	a
b	d
a	c
c	a

 $r = T \times X$ 

T.A	T.B	X.A	X.B
a	a	a	a
a	a	b	d
a	a	a	c
a	a	c	a
b	d	a	a
b	d	b	d
b	d	a	c
b	d	c	a

# VÍ DỤ PHÉP TÍCH ĐỀ-CÁC

R

A	B	C	D
a	a	1	yes
b	d	7	yes
c	f	34	no

S

X	Y	Z
a	m	4
b	c	22
a	d	16
a	c	7

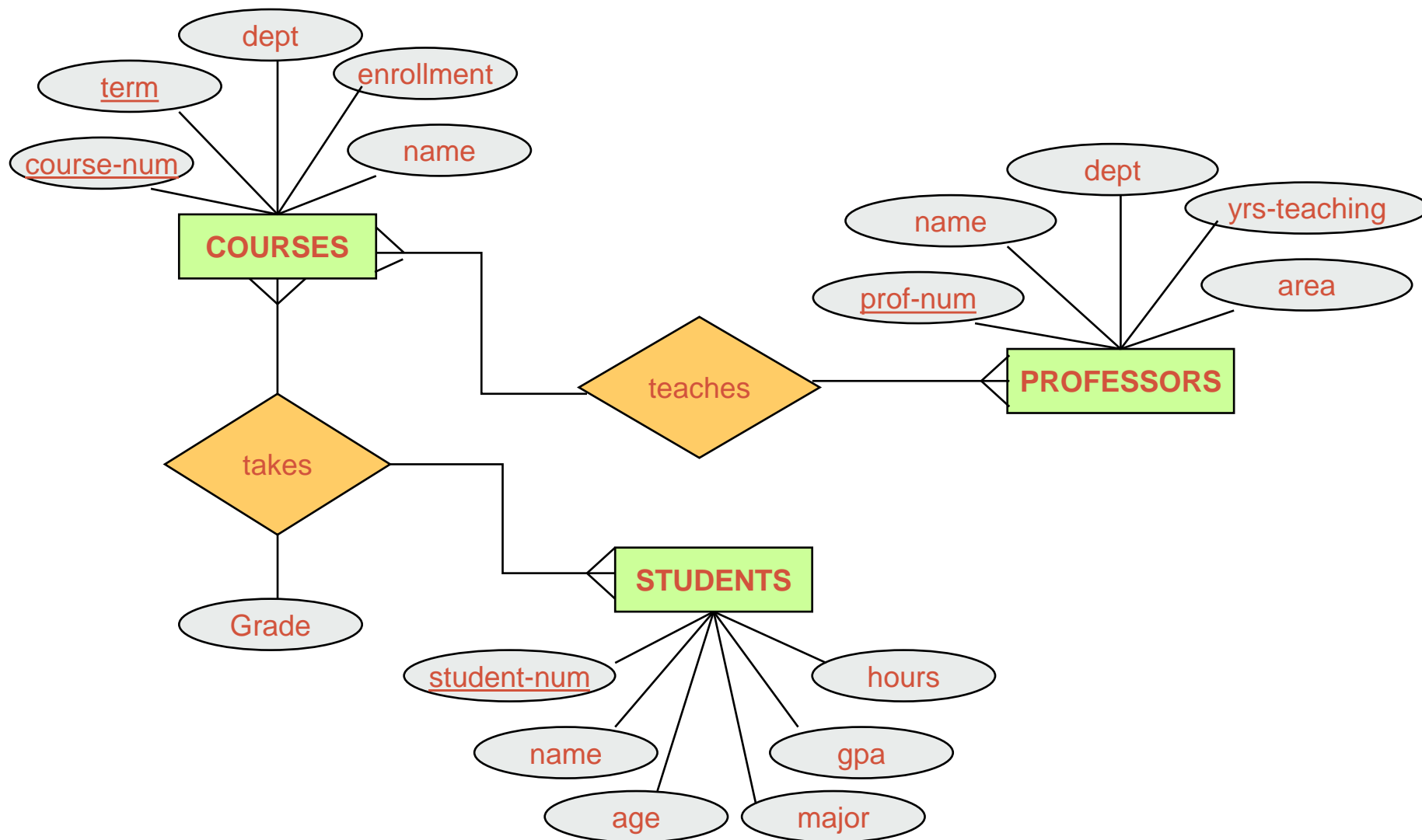
$r = R \times S$

A	B	C	D	X	Y	Z
a	a	1	yes	a	m	4
a	a	1	yes	b	c	22
a	a	1	yes	a	d	16
a	a	1	yes	a	c	7
b	d	7	yes	a	m	4
b	d	7	yes	b	c	22
b	d	7	yes	a	d	16
b	d	7	yes	a	c	7
c	f	34	no	a	m	4
c	f	34	no	b	c	22
c	f	34	no	a	d	16
c	f	34	no	a	c	7

## BIỂU DIỄN NGÔN NGỮ ĐẠI SỐ QUAN HỆ

- ❖ Mỗi toán tử đại số quan hệ cơ bản (trong số năm toán tử) có thể được sử dụng riêng rẽ để hình thành một truy vấn. Tuy nhiên, sức mạnh thực sự được cải thiện đáng kể khi chúng được kết hợp với nhau để hình thành các biểu thức truy vấn.
- ❖ Phần sau xem xét việc tạo ra các tổ hợp phức tạp hơn của 5 phép toán cơ bản (trước khi giới thiệu các phép toán mở rộng trong đại số quan hệ => Có được đánh giá cao hơn cho các phép toán mở rộng).

# VÍ DỤ: SƠ ĐỒ E-R



## VÍ DỤ: LƯỢC ĐỒ QUAN HỆ

- ❖ Dùng kỹ thuật để chuyển sơ đồ E-R trên thành lược đồ quan hệ như sau:

S = STUDENTS(s#, name, age, major, gpa, hours\_completed)

C = COURSES(c#, term, name, dept, enrollment)

P = PROFESSORS(p#, name, dept, yrs\_teaching, area)

TA = TAKES(s#, c#, term, grade)

TE = TEACH(p#, c#, term)



## VÍ DỤ: THỂ HIỆN QUAN HỆ S

<u>s#</u>	name	age	major	gpa	hrs_completed
S1	Michael Schumacher	19	Computer Science	4.00	45
S5	Jean Alesi	20	Physics	3.46	78
S3	Rubens Barrichello	21	Math	3.82	33
S2	Giancarlo Fisichella	18	Math	2.73	23
S4	Jarno Trulli	18	Computer Science	1.48	99
S7	Bernd Schneider	19	Computer Science	2.29	45
S6	Mika Hakkinen	20	English	2.37	33

## VÍ DỤ: THỂ HIỆN QUAN HỆ C

<u>c#</u>	<u>term</u>	name	dept	enrollment
C1	Fall 2002	CS1	CS	120
C1	Spring 2002	CS1	CS	100
C4	Fall 2002	Architecture	CS	97
C3	Fall 2002	Database	CS	86
C5	Spring 2001	Physics I	Physics	135
C5	Fall 2002	Physics I	Physics	125
C6	Summer 2002	Calculus III	Math	67

## VÍ DỤ: THỂ HIỆN QUAN HỆ TA

<u>s#</u>	<u>c#</u>	<u>term</u>	grade
S1	C3	Fall 03	A
S3	C4	Fall 02	B
S4	C6	Summer 03	C
S5	C5	Spring 01	D
S5	C1	Fall 02	A
S5	C3	Fall 03	C
S5	C6	Summer 02	C
S5	C4	Fall 03	A
S3	C5	Spring 01	C
S3	C1	Fall 03	A
S2	C4	Fall 03	D

## VÍ DỤ: THỂ HIỆN QUAN HỆ P

<u>p#</u>	name	dept	yrs_teaching
P1	Wilson	CS	5
P2	Davis	Math	32
P3	deMoser	CS	17
P4	Roberts	Physics	14

## VÍ DỤ: THỂ HIỆN QUAN HỆ TẾ

<u>p#</u>	<u>c#</u>	<u>term</u>
P1	C3	Fall 2002
P3	C4	Fall 2002
P4	C6	Summer 2002
P2	C5	Spring 2001
P2	C1	Spring 2002
P1	C4	Fall 2002
P3	C1	Fall 2002

# VÍ DỤ: TRUY VẤN 1

❖ *Tìm tên của tất cả các Sinh viên học ngành Công nghệ thông tin (Computer Science).*

❖ Cách tiếp cận:

- Đầu tiên, chọn ra tất cả những sinh viên học ngành CS:

$$r = \sigma_{(\text{major} = \text{"Computer Science"})}(S)$$

- Tiếp theo, chiếu thuộc tính tên trên kết quả vừa tìm được:

$$\text{result} = \pi_{(\text{name})}(r)$$

❖ Biểu thức truy vấn hoàn chỉnh:

$$\text{result} = \pi_{(\text{name})}(\sigma_{(\text{major} = \text{"Computer Science"})}(S))$$

## VÍ DỤ: TRUY VẤN 2

❖ *Tìm số hiệu sinh viên (s#) và tên của tất cả những sinh viên đã hoàn thành hơn 90 giờ học.*

❖ Cách tiếp cận:

- Đầu tiên, chọn ra tất cả những sinh viên đã hoàn thành hơn 90 giờ học:

$$r = \sigma_{(\text{hours\_completed} > 90)}(S)$$

- Tiếp theo, chiếu thuộc tính số hiệu sinh viên và tên trên kết quả vừa tìm được:

$$\text{result} = \pi_{(s\#, \text{name})}(r)$$

❖ Biểu thức truy vấn hoàn chỉnh:

$$\text{result} = \pi_{(s\#, \text{name})}(\sigma_{(\text{hours\_completed} > 90)}(S))$$

## VÍ DỤ: TRUY VẤN 3

❖ *Tìm tên của tất cả những sinh viên dưới 20 tuổi và đã hoàn thành hơn 80 giờ học.*

❖ Cách tiếp cận:

- Đầu tiên, chọn ra tất cả những sinh viên dưới 20 tuổi và đã hoàn thành hơn 80 giờ học:

$$r = \sigma_{((\text{hours\_completed} > 80) \wedge (\text{age} < 20))}(S)$$

- Tiếp theo, chiếu thuộc tính tên trên kết quả vừa tìm được:

$$\text{result} = \pi_{(\text{name})}(r)$$

❖ Biểu thức truy vấn hoàn chỉnh:

$$\text{result} = \pi_{(\text{name})}(\sigma_{((\text{hours\_completed} > 80) \wedge (\text{age} < 20))}(S))$$



## VÍ DỤ: TRUY VẤN 4

❖ *Tìm tên của tất cả các lớp thuộc Khoa Công nghệ thông tin (CS) hoặc Khoa Vật lý (Physics).*

❖ Cách tiếp cận:

- Đầu tiên, chọn ra tất cả các lớp hoặc thuộc Khoa CS hoặc Khoa Physics:

$$r = \sigma_{((\text{dept} = \text{Computer Science}) \vee (\text{dept} = \text{Physics}))}(C))$$

- Tiếp theo, chiếu thuộc tính tên trên kết quả vừa tìm được:

$$\text{result} = \pi_{(\text{name})}(r)$$

❖ Biểu thức truy vấn hoàn chỉnh:

$$\text{result} = \pi_{(\text{name})}(\sigma_{((\text{dept} = \text{Computer Science}) \vee (\text{dept} = \text{Physics}))}(C))$$

## VÍ DỤ: TRUY VẤN 5

- ❖ *Tìm tên của tất cả các giáo sư đã dạy một lớp trong kỳ mùa thu 2002 (Fall 2002).*
- ❖ Cách tiếp cận:
  - Đầu tiên, đặt toàn bộ thông tin về giáo sư vào cùng với thông tin về lớp học.
  - Tiếp theo, chỉ chọn các giáo sư và lớp học có liên quan đến nhau và thỏa mãn điều kiện (từ kết quả trên).
  - Cuối cùng, chiếu thuộc tính tên (giáo sư) trên kết quả vừa tìm được.
- ❖ Biểu thức truy vấn hoàn chỉnh:

$$\text{result} = \pi_{(P.\text{name})}(\sigma_{((TE.\text{term} = \text{Fall } 2002) \wedge (P.p\# = TE.p\#))}(P \times TE))$$

## VÍ DỤ: TRUY VẤN 6

- ❖ *Tìm tên của tất cả các sinh viên đã học một lớp trong kỳ mùa thu 2003 mà được dạy bởi một giáo sư có hơn 20 năm kinh nghiệm giảng dạy.*
- ❖ Cách tiếp cận:
  - Đầu tiên, đặt toàn bộ thông tin về giáo sư, thông tin về sinh viên, thông tin về việc dạy và thông tin về việc học vào cùng nhau.
  - Tiếp theo, chỉ chọn các sinh viên, giáo sư và lớp học có liên quan đến nhau và thỏa mãn điều kiện (từ kết quả trên).
  - Cuối cùng, chiếu thuộc tính tên (sinh viên) trên kết quả vừa tìm được.
- ❖ Biểu thức truy vấn hoàn chỉnh:

$$\text{result} = \pi_{(S.\text{name})}(\sigma_{((TA.\text{term} = \text{Fall } 2003) \wedge (P.\text{yrs\_teaching} > 20) \wedge (S.\text{s\#} = TA.\text{s\#}) \wedge (TA.\text{c\#} = TE.\text{c\#}) \wedge (TA.\text{term} = TE.\text{term}) \wedge (P.\text{p\#} = TE.\text{p\#}))}(S \times TA \times TE \times P))$$

## VÍ DỤ: TRUY VẤN 7

- ❖ *Tìm tên của tất cả các giáo sư hoặc dạy ở Khoa Công nghệ thông tin (CS) hoặc có hơn 20 năm kinh nghiệm giảng dạy.*
- ❖ Biểu thức truy vấn hoàn chỉnh:

$$\text{result} = [\pi_{(\text{name})}(\sigma_{(\text{dept} = \text{Computer Science})}(\text{P}))] \cup [\pi_{(\text{name})}(\sigma_{(\text{yrs\_teaching} > 20)}(\text{P}))]$$

hoặc:

$$\text{result} = \pi_{(\text{name})}(\sigma_{((\text{dept} = \text{Computer Science}) \vee (\text{yrs\_teaching} > 20))}(\text{P}))$$

## VÍ DỤ: TRUY VẤN 8

- ❖ *Tìm số hiệu của tất cả các sinh viên mà chỉ đăng ký học vào kỳ mùa xuân 2003.*
- ❖ Biểu thức truy vấn hoàn chỉnh:

$$\text{result} = \pi_{(TA.s\#)}[\sigma_{(TA.term = \text{Spring } 2003)}(TA)] - \pi_{(TA.s\#)}[\sigma_{(TA.term \neq \text{Spring } 2003)}(TA)]$$

Chú ý: Biểu thức truy vấn dưới đây sai, vì sao?

$$\text{result} = \pi_{(TA.s\#)}(\sigma_{(TA.term = \text{Spring } 2003)}(TA))$$

=> bởi vì ở đây chỉ liệt kê những sinh viên có đăng ký học vào kỳ mùa xuân 2003 và cũng có thể các sinh viên này có đăng ký các kỳ học khác nữa.

# CÁC TOÁN TỬ MỞ RỘNG

- ❖ Năm phép toán cơ bản (được trình bày ở phần trước) là đủ để thể hiện câu truy vấn trong đại số quan hệ (có thể chứng minh được).
- ❖ Trong các câu truy vấn phức tạp cần các biểu thức truy vấn khó và dài dòng. Để đơn giản hóa, trong đại số quan hệ, người ta đề xuất một số phép toán mở rộng để cung cấp sức mạnh thể hiện biểu thức truy vấn.
- ❖ Phần sau sẽ trình bày một số phép toán quan trọng và chung nhất.

# PHÉP GIAO

Loại: Hai ngôi

Ký hiệu:  $\cap$

Khuôn dạng chung:  $r \cap s$  với  $r$  và  $s$  là 2 quan hệ khả hợp

Lược đồ quan hệ kết quả: lược đồ của quan hệ toán hạng

Kích thước của quan hệ kết quả (số bộ):  $\leq \min \{ |r|, |s| \}$

Định nghĩa:  $r \cap s \equiv r - (r - s)$

Ví dụ:

$$(\pi_{(p\#)}(SPJ)) \cap (\pi_{(p\#)}(P))$$

- Phép giao tạo ra tập các bộ xuất hiện ở cả hai quan hệ toán hạng.

# VÍ DỤ PHÉP GIAO

R

A	B	C	D
a	a	yes	1
b	d	no	7
c	f	yes	34
a	d	no	6

$r = R \cap S$

A	B	C	D
a	a	yes	1
c	f	yes	34

$r = R \cap T$

A	B	C	D
---	---	---	---

S

E	F	G	H
a	a	yes	1
b	r	yes	3
c	f	yes	34
m	n	no	56

T

E	F	G	H
a	r	no	31
b	f	yes	30



# PHÉP KẾT NỐI

- ❖ Trong các biểu thức truy vấn có liên quan tới tích Đề-các, cần phải cung cấp thêm các phép chọn để loại bỏ đi những tổ hợp các bộ không liên quan tới nhau trong kết quả.  
  
=> Phép kết nối (join operation) là sự kết hợp tích Đề-các và các phép chọn.
- ❖ **Các loại phép kết nối:** kết nối theta, kết nối bằng, kết nối tự nhiên, kết nối ngoài và bán kết nối (Semi Join Operator).

# PHÉP KẾT NỐI THETA VÀ KẾT NỐI BẰNG

Loại: Hai ngôi

Ký hiệu/khuôn dạng chung:  $r \bowtie_{(\text{predicate})} s$

Lược đồ quan hệ kết quả: ghép nối các quan hệ toán hạng

Định nghĩa:  $r \bowtie_{(\text{predicate})} s \equiv \sigma_{(\text{predicate})}(r \times s)$

Ví dụ:

$$r \bowtie_{((color='blue') \wedge (size=3))} s$$

$$r \bowtie_{((color='blue') \wedge (size>3))} s$$

Kết nối bằng

Kết nối theta

- Phép kết nối theta là dạng rút gọn của tích Đề-các và sau đó là thực hiện phép chọn.
- Phép kết nối bằng là một trường hợp đặc biệt của kết nối theta mà trong đó tất cả các điều kiện trong mệnh đề đều là điều kiện bằng.
- Cả phép kết nối theta và kết nối bằng đều không loại bỏ các bộ dư thừa, vì vậy việc loại bỏ các bộ dư thừa cần phải được thực hiện một cách tường minh thông qua mệnh đề điều kiện.

# VÍ DỤ PHÉP KẾT NỐI THETA

R

A	B	C	D
a	a	yes	1
b	d	no	7
c	f	yes	34
a	d	no	6

S

E	F	G	H
a	a	yes	1
b	r	yes	3
c	f	yes	34
m	n	no	56

$r = R \bowtie_{(R.B < S.F)} S$

A	B	C	D	E	F	G	H
a	a	yes	1	b	r	yes	3
a	a	yes	1	c	f	yes	34
a	a	yes	1	m	n	no	56
b	d	no	7	b	r	yes	3
b	d	no	7	c	f	yes	34
b	d	no	7	m	n	no	56
c	f	yes	34	b	r	yes	3
c	f	yes	34	m	n	no	56
a	d	no	6	b	r	yes	3
a	d	no	6	c	f	yes	34
a	d	no	6	m	n	no	56

# PHÉP KẾT NỐI TỰ NHIÊN

Loại: Hai ngôi

Ký hiệu/khuôn dạng chung:  $r * s$

Lược đồ quan hệ kết quả: ghép nối các quan hệ toán hạng, với các thuộc tính được đặt tên chung và chỉ xuất hiện một lần.

Định nghĩa:  $r * s \equiv r \bowtie_{(r.commonattributes = s.commonattributes)} s$

Ví dụ:  $s * spj * p$

- Phép kết nối tự nhiên thực hiện kết nối bằng trên tất cả các thuộc tính có cùng tên của 2 quan hệ toán hạng.
- Bậc của quan hệ kết quả là tổng số bậc của 2 quan hệ toán hạng trừ đi số các thuộc tính chung của chúng.
- Phép kết nối tự nhiên là phổ biến nhất trong tất cả các phép kết nối. Nó rất có ích trong việc loại bỏ đi các bộ dư thừa. Các thuộc tính chung của 2 quan hệ toán hạng thường được gọi là **các thuộc tính kết nối**.

# VÍ DỤ PHÉP KẾT NỐI TỰ NHIÊN

**R**

A	B	C	D
a	a	yes	1
b	r	no	7
c	f	yes	34
a	m	no	6

**$r = R * S$**

A	B	C	D	M	G	H
a	a	yes	1	a	yes	1
a	a	yes	1	f	yes	34
a	m	no	6	n	no	56

**S**

B	M	G	H
a	a	yes	1
b	r	yes	3
a	f	yes	34
m	n	no	56

**$r = R * T$**

A	B	C	D	G	H
b	r	no	7	yes	30

**T**

A	B	G	H
a	f	no	31
b	r	yes	30

# PHÉP KẾT NỐI NGOÀI

Loại: Hai ngôi

Ký hiệu/khuôn dạng chung: Kết nối ngoài trái:  $r \supset \triangleleft s$

Kết nối ngoài phải:  $r \triangleright \subset s$       Kết nối ngoài đầy đủ:  $r \supset \triangleleft \triangleright \subset s$

Lược đồ quan hệ kết quả: ghép nối các quan hệ toán hạng

Định nghĩa:

$r \supset \triangleleft s \equiv$  kết nối tự nhiên  $r$  và  $s$  với các bộ của  $r$  không tương ứng trong  $s$  vẫn được giữ lại trong kết quả. Tất cả các giá trị còn khuyết ở thuộc tính của  $s$  đều được gán giá trị rỗng (null).

$r \triangleright \subset s \equiv$  kết nối tự nhiên  $r$  và  $s$  với các bộ của  $s$  không tương ứng trong  $r$  vẫn được giữ lại trong kết quả. Tất cả các giá trị còn khuyết ở thuộc tính của  $r$  đều được gán giá trị rỗng.

$r \supset \triangleleft \triangleright \subset s \equiv$  kết nối tự nhiên  $r$  và  $s$  với các bộ của  $r$  và  $s$  không tương ứng vẫn được giữ lại trong kết quả. Tất cả các giá trị còn khuyết sẽ được gán giá trị rỗng.

Ví dụ: Cho  $r(A,B) = \{(a, b), (c, d), (b,c)\}$  và  $s(A,C) = \{(a, d), (s, t), (b, d)\}$

$$r \supset \triangleleft s = (A,B,C) = \{(a,b,d), (b,c,d), (c,d,null)\}$$

$$r \triangleright \subset s = (A,B,C) = \{(a,b,d), (b,c,d), (s,null,t)\}$$

$$r \supset \triangleleft \triangleright \subset s = (A,B,C) = \{(a,b,d), (b,c,d), (s,null,t), (c,d,null)\}$$

# VÍ DỤ PHÉP KẾT NỐI NGOÀI

**R**

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

$r = R \bowtie_{\langle C \rangle} S$

A	B	C	D
1	2	3	10
1	2	3	11
4	5	6	null
7	8	9	null
null	6	7	12

**S**

B	C	D
2	3	10
2	3	11
6	7	12

$r = R \bowtie_{\langle A \rangle} S$

A	B	C	D
1	2	3	10
1	2	3	11
4	5	6	null
7	8	9	null

$r = R \bowtie_{\langle C \rangle} S$

B	C	D	A
2	3	10	1
2	3	11	1
6	7	12	null

# PHÉP BÁN KẾT NỐI

Loại: Hai ngôi

Ký hiệu/khuôn dạng chung:  $r \triangleright_{(\text{predicate})} s$

Lược đồ quan hệ kết quả: Lược đồ của  $r$

Định nghĩa:  $r \triangleright_{(\text{predicate})} s \equiv \pi_{(\text{thuộc tính của } r)}(r \triangleright_{\triangleleft(\text{predicate})} s)$

Ví dụ:

- Phép bán kết nối thực hiện một phép kết nối 2 quan hệ toán hạng, sau đó chiếu trên các thuộc tính của quan hệ toán hạng bên trái.
- Ưu điểm chính của phép bán kết nối là làm giảm số các bộ cần xử lý khi thực hiện phép kết nối  $\Rightarrow$  rất có ích trong môi trường phân tán.
- Theo một khuôn dạng chung như đã trình bày, phép bán kết nối chính là phép bán kết nối theta. Phép bán kết nối bằng và phép bán kết nối tự nhiên được định nghĩa theo cách tương tự.



# VÍ DỤ PHÉP BÁN KẾT NỐI

**R**

A	B	C	D
a	a	yes	1
b	r	no	7
c	f	yes	34
a	m	no	6

$$r = R \triangleright_{(R.B > S.M)} S$$

**S**

B	M	C
a	e	yes
b	r	yes
a	f	no
r	n	no

**T**

B	G	D
a	4	d
b	7	e
a	4	f
m	2	g

# PHÉP CHIA

Loại: Hai ngôi

Ký hiệu/khuôn dạng chung:  $r \div s$  với  $r(\{A\})$  và  $s(\{B\})$

Lược đồ quan hệ kết quả:  $C$  với  $C = A - B$

Định nghĩa:  $r \div s \equiv \pi_{(A-B)}(r) - (\pi_{(A-B)}((\pi_{(A-B)}(r) \times s) - r))$

Ví dụ:

Cho  $r(A,B,C) = \{(a,b,c), (a,d,c), (a,b,d), (a,c,c), (a,d,d)\}$

và  $s(C) = \{(c), (d)\}$

thì:  $r \div s = t(A,B) = \{(a,b), (a,d)\}$

## Các yêu cầu đối với phép chia:

1. Quan hệ  $r$  được định nghĩa trên tập thuộc tính  $A$  và quan hệ  $s$  được định nghĩa trên tập thuộc tính  $B$  sao cho  $B \subseteq A$ .
2. Cho  $C$  là tập các thuộc tính của  $A - B$ .  
 => Phép chia được định nghĩa như sau: một bộ  $t$  là thuộc  $r \div s$  nếu với mọi bộ  $t_s$  của  $s$  có một bộ  $t_r$  của  $r$  thỏa mãn cả 2 điều kiện:  

$$t_r[C] = t_s[C] \quad \text{và} \quad t_r[A-B] = t[A-B]$$

# VÍ DỤ PHÉP CHIA

**R**

A	B	C	D
a	f	yes	1
b	r	no	1
a	f	yes	34
e	g	yes	34
a	m	no	6
b	r	no	34

$r = R \div S$

A	B	C
a	f	yes
b	r	no

$r = R \div T$

A	B
a	f

$r = R \div U$

A	B
b	r

$r = R \div V$

A
a

$r = R \div W$

A
---

**S**

D
1
34

**T**

C	D
yes	1
yes	34

**U**

C	D
no	1
no	34

**V**

B	C	D
f	yes	1
f	yes	34
m	no	6

**W**

B	C	D
f	yes	1
g	yes	69

# TÍNH HỮU ÍCH CỦA CÁC TOÁN TỬ MỞ RỘNG

- ❖ Các toán tử đại số quan hệ mở rộng được định nghĩa dựa trên 5 phép toán cơ bản.
- ❖ Tính hữu ích của các phép toán này được thể hiện tốt nhất qua phép chia.
- ❖ Xem xét truy vấn sau dựa trên CSDL Nhà cung cấp-linh kiện-công việc-vận chuyển (**suppliers-parts-jobs-shipment**):

# LƯỢC ĐỒ QUAN HỆ

Câu truy vấn: Tìm số hiệu các nhà cung cấp vận chuyển mọi linh kiện?

SUPPLIERS (S)

<u>snum</u>	name	status	city
-------------	------	--------	------

PARTS (P)

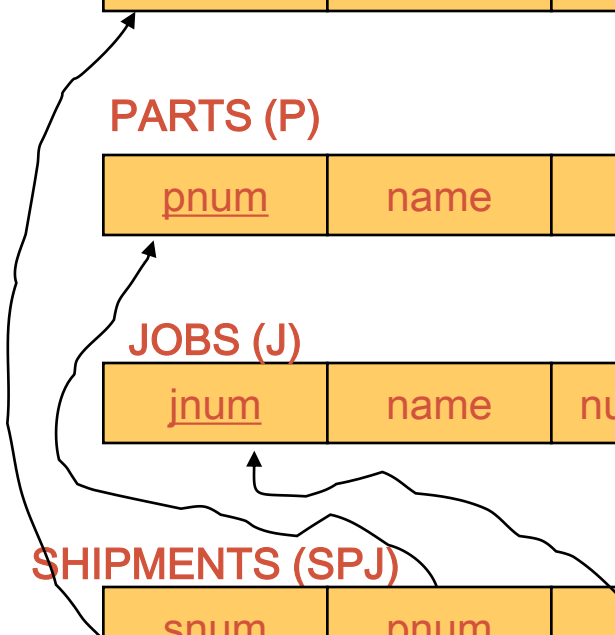
<u>pnum</u>	name	color	weight	city
-------------	------	-------	--------	------

JOBS (J)

<u>jnum</u>	name	numworkers	city
-------------	------	------------	------

SHIPMENTS (SPJ)

<u>snum</u>	<u>pnum</u>	<u>jnum</u>	qty	date
-------------	-------------	-------------	-----	------



# TÍNH HỮU ÍCH CỦA CÁC TOÁN TỬ MỞ RỘNG (Cont.)

## ❖ Giải pháp 1: Chỉ sử dụng 5 phép toán cơ bản:

$T1 = \pi_{(s\#, p\#)}(spj)$  //tất cả các cặp  $(s\#, p\#)$  cho các vận chuyển thực.

$T2 = \pi_{(p\#)}(p)$  //tất cả các linh kiện  $(p\#)$

$T3 = \pi_{(s\#)}(T1) \times T2$  //tất cả  $s\#$  trong  $T1$  được ghép với mỗi bộ trong  $T2$   
 $\{spj.s\#, p.p\#\}$

$T4 = T3 - T1$  //tất cả các bộ thuộc  $T3$  và không thuộc  $T1$  – các nhà cung cấp vận chuyển (một hoặc một số linh kiện).

$T5 = T1 - T4$  //tất cả các bộ thuộc  $T1$  mà không thuộc  $T4$  – các nhà cung cấp vận chuyển tất cả các linh kiện.

$T6 = \pi_{(s\#)}(T5)$  //lời giải

## Kết quả:

$$\pi_{(s\#)}[\pi_{(s\#, p\#)}(spj) - (\pi_{(s\#)}(spj) \times \pi_{(p\#)}(p) - \pi_{(s\#, p\#)}(spj))]$$

# TÍNH HỮU ÍCH CỦA CÁC TOÁN TỬ MỞ RỘNG (Cont.)

❖ *Giải pháp 2: Sử dụng các phép toán mở rộng:*

Kết quả:

$$(\pi_{(s\#, p\#)}(spj)) \div (\pi_{(p\#)}(p))$$

# VIẾT MỘT SỐ TRUY VẤN CHỈ SỬ DỤNG 5 TOÁN TỬ CƠ BẢN

1. *Tìm số hiệu của tất cả các nhà cung cấp hoặc ở tại Milan hoặc chuyển hàng tới bất kỳ công việc nào với số lượng lớn hơn 40.*

$$\pi_{(s\#)}(\sigma_{(city = Milan)}(S)) \cup [\pi_{(s\#)}(\sigma_{(qty > 40)}(SPJ))]$$

2. *Tìm tên của tất cả các nhà cung cấp mà chỉ vận chuyển các linh kiện màu đỏ.*

$$\pi_{(S.name)}[\pi_{(S\#,S.name)}(\sigma_{((SPJ.s\#=S.s\#)\wedge(SPJ.p\#=P.p\#)\wedge(color=red))}(SPJ \times S \times P)) \\ - \pi_{(S\#,S.name)}(\sigma_{((SPJ.s\#=S.s\#)\wedge(SPJ.p\#=P.p\#)\wedge(color \neq red))}(SPJ \times S \times P))]$$



# VIẾT MỘT SỐ TRUY VẤN CHỈ SỬ DỤNG 5 TOÁN TỬ CƠ BẢN (Cont.)

3. *Tìm tên của các nhà cung cấp mà ở cùng thành phố với công việc mà họ vận chuyển linh kiện đến cho.*

$$T1 = (S \times SPJ \times J)$$

$$T2 = \sigma_{(S.s\# = SPJ.s\#)}(T1) \quad // \text{chọn các bộ có cùng mã nhà cung cấp } s\#$$

$$T3 = \sigma_{(J.j\# = SPJ.j\#)}(T2) \quad // \text{chọn các bộ có cùng mã công việc } j\#$$

$$T4 = \sigma_{(J.city = S.city)}(T3) \quad // \text{chọn các bộ có cùng thành phố}$$

$$T5 = \pi_{(S.name)}(T4) \quad // \text{chiếu để lấy tập thuộc tính cuối cùng}$$

**Kết quả:**

$$\pi_{(S.name)}(\sigma_{(S.s\# = SPJ.s\#) \wedge (J.j\# = SPJ.j\#) \wedge (J.city = S.city)}(S \times SPJ \times J))$$

# VIẾT MỘT SỐ TRUY VẤN CHỈ SỬ DỤNG 5 TOÁN TỬ CƠ BẢN (Cont.)

4. *Tìm số hiệu của tất cả các linh kiện mà được vận chuyển bởi cả hai nhà cung cấp “S1” và “S2”.*

Lưu ý: Biểu thức sau đây không đúng, vì sao?

$$\pi_{(p\#)}(\sigma_{((s\# = "S1") \wedge (s\# = "S2"))}(\text{SPJ}))$$

→ Vì điều kiện lựa chọn luôn cho kết quả là sai.

**Kết quả đúng:**

$$[\pi_{(p\#)}(\sigma_{(s\#=S1)}(\text{SPJ}))] - ([\pi_{(p\#)}(\sigma_{(s\#=S1)}(\text{SPJ}))] - [\pi_{(p\#)}(\sigma_{(s\#=S2)}(\text{SPJ}))])$$

# VIẾT MỘT SỐ TRUY VẤN CHỈ SỬ DỤNG 5 TOÁN TỬ CƠ BẢN (Cont.)

5. *Tìm số hiệu của tất cả các nhà cung cấp mà vận chuyển cả các linh kiện màu đỏ (“red”) và màu xanh (“blue”).*

Lưu ý: Biểu thức sau đây không đúng, vì sao?

$$\pi_{(s\#)}(\sigma_{((color = blue) \wedge (SPJ.p\# = P.p\#) \wedge (color = red))}(P \times SPJ))$$

→ Vì điều kiện lựa chọn luôn cho kết quả là sai.

*Truy vấn đúng:*

$$T1 = \pi_{(s\#)}(\sigma_{((color = "blue") \wedge (SPJ.p\# = P.p\#))}(P \times SPJ))$$

$$T2 = \pi_{(s\#)}(\sigma_{((color = "red") \wedge (SPJ.p\# = P.p\#))}(P \times SPJ))$$

$$T3 = T2 - T1$$

$$T4 = T2 - T3$$

# VIẾT MỘT SỐ TRUY VẤN CHỈ SỬ DỤNG 5 TOÁN TỬ CƠ BẢN (Cont.)

6. *Tìm tất cả các cặp (s#, j#) của các nhà cung cấp và công việc ở cùng thành phố, nhưng các nhà cung cấp đó không vận chuyển bất kỳ linh kiện nào cho công việc này.*

Biểu thức truy vấn:

$$T1 = \pi_{(s\#, j\#)}(\sigma_{(S.city = J.city)}(S \times J)) \quad // \text{tất cả các cặp } (s\#, j\#) \text{ ở cùng thành phố.}$$

$$T2 = \pi_{(s\#, j\#)}(\sigma_{((S.city=J.city) \wedge (SPJ.j\#=J.j\#) \wedge (SPJ.s\#=S.s\#))}(S \times SPJ \times J))$$

//T2 chứa tất cả các cặp (s#,j#) trong đó nhà cung cấp vận chuyển các linh kiện tới công việc trong cùng thành phố.

$$T3 = T1 - T2$$

## TOÁN TỬ ĐẶT LẠI TÊN

- ❖ Không giống như các quan hệ trong CSDL, các quan hệ trung gian được sinh ra từ các kết quả của truy vấn không có tên gọi để tham chiếu đến.
- ❖ Nếu các quan hệ này không được lưu trữ một cách tường minh, nó sẽ bị mất sau khi truy vấn được thực thi.
- ❖ Tuy nhiên, trong một số trường hợp cần lưu trữ lại các quan hệ trung gian, ví dụ: sử dụng kết quả cho một truy vấn khác, ...

# TOÁN TỬ ĐẶT LẠI TÊN (Cont.)

❖ *Biểu diễn toán tử đặt lại tên:*

Chữ cái Hy Lạp thường rho ( $\rho$ ).

❖ *Khuôn dạng chung đầu tiên của toán tử đặt lại tên cho các quan hệ:*

$\rho_{\text{tên mới của quan hệ}}$  (quan hệ)

❖ *Ví dụ:*

$\rho_x(r)$ : đặt lại tên cho quan hệ  $r$  thành  $x$ .

## TOÁN TỬ ĐẶT LẠI TÊN (Cont.)

- ❖ Dạng thứ hai là *toán tử đặt lại tên cho cả quan hệ và thuộc tính*.
- ❖ Giả sử quan hệ toán hạng có cấp  $n$ , thì *dạng thức của toán tử đặt lại tên* là:

$\rho_{\text{tên mới của quan hệ } (A1, A2, \dots, AN)}(\text{quan hệ})$

- ❖ Ví dụ:

$\rho_{x(one, two, \dots, last)}(r)$

đặt lại tên quan hệ  $r$  thành  $x$  và  $n$  thuộc tính của quan hệ  $x$  có tên là *one, two, ..., last*.

# THỰC HÀNH TRUY VẤN VỚI TẤT CẢ CÁC TOÁN TỬ ĐẠI SỐ QUAN HỆ

1. *Liệt kê tất cả các cặp số hiệu nhà cung cấp ở cùng một thành phố.*

$$\pi_{(s.s\#,x.s\#)}(s \triangleright \triangleleft_{(s.city = x.city)}(\rho_x(s)))$$

2. *Liệt kê tất cả các vận chuyển có liên quan đến linh kiện màu xanh (“green”).*

$$spj \triangleright \triangleleft_{(spj.pnum = p.pnum)}(\sigma_{(color=green)}(p))$$



## THỰC HÀNH TRUY VẤN VỚI TẤT CẢ CÁC TOÁN TỬ ĐẠI SỐ QUAN HỆ (Cont.)

3. *Liệt kê tất cả số hiệu của các nhà cung cấp mà vận chuyển linh kiện được sản xuất ở cùng thành phố với nhà cung cấp đó.*

$$\pi_{(s.s\#)}(\sigma_{(s.city=p.city)}(s * p * spj))$$

4. *Liệt kê tên của tất cả các nhà cung cấp mà vận chuyển tất cả các linh kiện màu xanh (“blue”).*

$$\pi_{(s.name)}(s * (\pi_{(s\#,p\#)}(spj) \div \pi_{(p\#)}(\sigma_{(color=blue)}(p))))$$

# THỰC HÀNH TRUY VẤN VỚI TẤT CẢ CÁC TOÁN TỬ ĐẠI SỐ QUAN HỆ (Cont.)

5. *Liệt kê số hiệu của các nhà cung cấp mà chỉ vận chuyển các linh kiện màu xanh.*

$$(\pi_{(s\#)}(spj * (\sigma_{(color=blue)}(p)))) - (\pi_{(s\#)}(spj * (\sigma_{(color \neq blue)}(p))))$$