

SQL VÀ ĐẠI SỐ QUAN HỆ

Nguyễn Đình Hóa

dinhhoa@gmail.com

094-280-7711

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, phép kết nối tự nhiên, phép chia, phép bán kết nối, và phép kết nối ngoài.

PHÉP CHỌN

Loại: Một ngôi

Ký hiệu: Sigma, σ

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

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

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: Π , π

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 (j) của s phải giống nhau, cho mọi giá trị của i, j .

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

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 (ρ).

❖ *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*.

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.
- ❖ Bài này 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:

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

- 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: $r \rhd \triangleleft s$

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

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

Định nghĩa:

$r \rhd \triangleleft s \equiv$ 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).

$r \triangleright \triangleleft s \equiv$ 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.

$r \rhd \triangleleft \triangleright \triangleleft s \equiv$ 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 $s(A,C) = \{(a, d), (s, t), (b, d)\}$

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

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

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

VÍ DỤ PHÉP KẾT NỐI NGOÀI

R

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

$r = R \bowtie S$

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

$r = R \bowtie S$

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

$r = R \bowtie S$

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:

$$r \triangleright_{(r.B > s.M)} s$$

- 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 BÀN 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 $r(\{X\})$ và $s(\{Y\})$

Lược đồ quan hệ kết quả: Z với $Z = X - Y$

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

Ví dụ:

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

và $s(C) = \{(c), (d)\}$

thì: $r \div s = t(A,B) = \{(a,b)\}$

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