

Chương 3: Ngôn ngữ đại số quan hệ

Nội dung

1. Giới thiệu
2. Biểu thức đại số quan hệ
3. Các phép toán
4. Biểu thức đại số quan hệ
5. Ví dụ

1. Giới thiệu

- ◆ Đại số quan hệ (ĐSQH) có nền tảng toán học (cụ thể là lý thuyết tập hợp) để mô hình hóa CSDL quan hệ. Đối tượng xử lý là các quan hệ trong cơ sở dữ liệu quan hệ.
- ◆ **Chức năng:**
 - Cho phép mô tả các phép toán rút trích dữ liệu từ các quan hệ trong cơ sở dữ liệu quan hệ.
 - Cho phép tối ưu quá trình rút trích bằng các phép toán có sẵn của lý thuyết tập hợp.

2. Biểu thức ĐSQH

- ◆ Biểu thức ĐSQH là một biểu thức gồm các phép toán ĐSQH.
- ◆ Biểu thức ĐSQH được xem như một quan hệ (không có tên).
- ◆ Có thể đặt tên cho quan hệ được tạo từ một biểu thức ĐSQH.
- ◆ Có thể đổi tên các thuộc tính của quan hệ được tạo từ một biểu thức ĐSQH.

3. Các phép toán

3.1 Giới thiệu

3.2 Phép chọn

3.3 Phép chiếu

3.4 Phép gán

3.5 Các phép toán trên tập hợp

3.6 Phép kết

3.7 Phép chia

3.8 Hàm tính toán và gom nhóm

3.1 Giới thiệu (1)

- ♦ Có năm phép toán cơ bản:
 - Chọn (σ) hoặc ($:$)
 - Chiếu (π) hoặc ($[]$)
 - Tích (\times)
 - Hiệu ($-$)
 - Hội (\cup)

3.1 Giới thiệu (2)

- ◆ Các phép toán khác không cơ bản nhưng hữu ích:
 - Giao (\cap)
 - Kết ($\triangleright\triangleleft$)
 - Chia (\div)
 - Phép bù (\neg)
 - Đổi tên (ρ)
 - Phép gán (\leftarrow)
- ◆ Kết quả sau khi thực hiện các phép toán là các quan hệ, do đó có thể kết hợp giữa các phép toán để tạo nên phép toán mới.

3.2 Phép chọn (Selection)

- ♦ Trích chọn các bộ (dòng) từ quan hệ R. Các bộ được trích chọn phải thỏa mãn điều kiện chọn p.
- ♦ **Ký hiệu:** $\sigma_p(R)$
- ♦ **Định nghĩa:** $\sigma_p(R) = \{t / t \in R, p(t)\}$ p(t):thỏa điều kiện p
- ♦ Kết quả trả về là một quan hệ, có cùng danh sách thuộc tính với quan hệ R. Không có kết quả trùng.
- ♦ Phép chọn có tính giao hoán

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

Lược đồ CSDL quản lý giáo vụ

HOCVIEN (MAHV, HO, TEN, NGSINH, GIOITINH, NOISINH, MALOP)

LOP (MALOP, TENLOP, TRGLOP, SISO, MAGVCN)

KHOA (MAKHOA, TENKHOA, NGTLAP, TRGKHOA)

MONHOC (MAMH, TENMH, TCLT, TCTH, MAKHOA)

DIEUKIEN (MAMH, MAMH_TRUOC)

GIAOVIEN(MAGV,HOTEN,HOCVI,HOCHAM,GIOITINH,NGSINH,NGVL,
HESO, MUCLUONG, MAKHOA)

GIANGDAY(MALOP,MAMH,MAGV,HOCKY, NAM,TUNGAY,DENNGAY)

KETQUATHI (MAHV, MAMH, LANTHI, NGTHI, DIEM, KQUA)

3.2 Ví dụ phép chọn

- Tìm những học viên có giới tính là nam và có nơi sinh ở TpHCM

$\sigma_{(\text{Gioitinh}=\text{'Nam'}) \wedge (\text{Noisinh}=\text{'TpHCM'})}(\text{HOCVIEN})$

HOCVIEN				
Mahv	HoTen	Gioitinh	Noisinh	Malop
K1103	Ha Duy Lap	Nam	Nghe An	K11
K1102	Tran Ngoc Han	Nu	Kien Giang	K11
K1104	Tran Ngoc Linh	Nu	Tay Ninh	K11
K1105	Tran Minh Long	Nam	TpHCM	K11
K1106	Le Nhat Minh	Nam	TpHCM	K11

3.3 Phép chiếu (Project)

- ♦ Sử dụng để trích chọn giá trị một vài thuộc tính của quan hệ
- ♦ **Ký hiệu:** $\pi_{A_1, A_2, \dots, A_k}(R)$

trong đó A_i là tên các thuộc tính được chiếu.

- ♦ Kết quả trả về một quan hệ có k thuộc tính theo thứ tự như liệt kê. Các dòng trùng nhau chỉ lấy một.
- ♦ Phép chiếu không có tính giao hoán

3.3 Ví dụ

- Tìm mã số, họ tên những học viên có giới tính là nam và có nơi sinh ở TpHCM

$\pi_{Mahv, Hoten} \sigma_{(Gioitinh='Nam') \wedge (Noisinh='TpHCM')}(HOCVIEN)$

HOCVIEN				
Mahv	HoTen	Gioitinh	Noisinh	Malop
K1103	Ha Duy Lap	Nam	Nghe An	K11
K1102	Tran Ngoc Han	Nu	Kien Giang	K11
K1104	Tran Ngoc Linh	Nu	Tay Ninh	K11
K1105	Tran Minh Long	Nam	TpHCM	K11
K1106	Le Nhat Minh	Nam	TpHCM	K11

3.4 Phép gán (Assignment)

- ♦ Dùng để diễn tả câu truy vấn phức tạp.
- ♦ **Ký hiệu:** $A \leftarrow B$
- ♦ **Ví dụ:**
 $R(\text{HO}, \text{TEN}, \text{LUONG}) \leftarrow \pi_{\text{HONV}, \text{TENNV}, \text{LUONG}}(\text{NHANVIEN})$
- ♦ Kết quả bên phải của phép gán được gán cho biến quan hệ nằm bên trái.

3.5 Các phép toán tập hợp

3.5.1 Giới thiệu

3.5.2 Phép hội

3.5.3 Phép trừ

3.5.4 Phép giao

3.5.5 Phép tích

3.5.1 Giới thiệu

- ◆ Các phép toán thực hiện trên 2 quan hệ xuất phát từ lý thuyết tập hợp của toán học: phép hội ($R \cup S$), phép giao ($R \cap S$), phép trừ ($R - S$), phép tích ($R \times S$).
- ◆ Đối với các phép hội, giao, trừ, các quan hệ R và S phải **khả hợp**:
 - Số lượng thuộc tính của R và S phải bằng nhau:
 $R(A_1, A_2, \dots, A_n)$ và $S(B_1, B_2, \dots, B_n)$
 - Miền giá trị của thuộc tính phải tương thích
 $\text{dom}(A_i) = \text{dom}(B_i)$
- ◆ Quan hệ kết quả của phép hội, giao, trừ có cùng tên thuộc tính với quan hệ đầu tiên.

3.5.2 Phép hội (Union)

- ♦ **Ký hiệu:** $R \cup S$
- ♦ **Định nghĩa:** $R \cup S = \{t | t \in R \vee t \in S\}$ trong đó R, S là hai quan hệ khả hợp.
- ♦ **Ví dụ:** Học viên được khen thưởng đợt 1 hoặc đợt 2

DOT1	
Mahv	Hoten
K1103	Le Van Tam
K1114	Tran Ngoc Han
K1203	Le Thanh Hau
K1308	Nguyen Gia

DOT2	
Mahv	Hoten
K1101	Le Kieu My
K1114	Tran Ngoc Han

Mahv	Hoten
K1101	Le Kieu My
K1103	Le Van Tam
K1114	Tran Ngoc Han
K1203	Le Thanh Hau
K1308	Nguyen Gia

DOT1 \cup DOT2

3.5.3 Phép trừ (Set Difference)

- ♦ **Ký hiệu:** $R-S$
- ♦ **Định nghĩa:** $R-S = \{t | t \in R \wedge t \notin S\}$ trong đó R, S là hai quan hệ khả hợp.
- ♦ **Ví dụ:** Học viên được khen thưởng đợt 1 nhưng không được khen thưởng đợt 2

DOT1	
Mahv	Hoten
K1103	Le Van Tam
K1114	Tran Ngoc Han
K1203	Le Thanh Hau
K1308	Nguyen Gia

DOT2	
Mahv	Hoten
K1101	Le Kieu My
K1114	Tran Ngoc Han

Mahv	Hoten
K1103	Le Van Tam
K1203	Le Thanh Hau
K1308	Nguyen Gia

DOT1- DOT2

3.5.4 Phép giao (Set-Intersection)

- ♦ **Ký hiệu:** $R \cap S$
- ♦ **Định nghĩa:** $R \cap S = \{t | t \in R \wedge t \in S\}$ trong đó R, S là hai quan hệ khả hợp. Hoặc $R \cap S = R - (R - S)$
- ♦ **Ví dụ:** Học viên được khen thưởng cả hai đợt 1 và 2

KT_D1	
Mahv	Hoten
K1103	Le Van Tam
K1114	Tran Ngoc Han
K1203	Le Thanh Hau
K1308	Nguyen Gia

KT_D2	
Mahv	Hoten
K1101	Le Kieu My
K1114	Tran Ngoc Han

Mahv	Hoten
K1114	Tran Ngoc Han

DOT1 \cap DOT2

3.5.5 Phép tích (1)

- ♦ **Ký hiệu:** $R \times S$
- ♦ **Định nghĩa:** $R \times S = \{t_r t_s / t_r \in R \wedge t_s \in S\}$
- ♦ Nếu R có n bộ và S có m bộ thì kết quả là $n * m$ bộ
 $KQ(A_1, A_2, \dots, A_m, B_1, B_2, \dots, B_n) \leftarrow R(A_1, A_2, \dots, A_m) \times S(B_1, B_2, \dots, B_n)$
- ♦ Phép tích thường dùng kết hợp với các phép chọn để kết hợp các bộ có liên quan từ hai quan hệ.
- ♦ **Ví dụ:** từ hai quan hệ HOCVIEN và MONHOC, có tất cả những trường hợp nào “*học viên đăng ký học môn học*”, giả sử không có bất kỳ điều kiện nào

3.5.5 Phép tích (2)

HOCVIEN	
Mahv	Hoten
K1103	Le Van Tam
K1114	Tran Ngoc Han
K1203	Le Thanh Hau

MONHOC
Mamh
CTRR
THDC
CTDL

Mahv	Hoten	Mamh
K1103	Le Van Tam	CTRR
K1114	Tran Ngoc Han	CTRR
K1203	Le Thanh Hau	CTRR
K1103	Le Van Tam	THDC
K1114	Tran Ngoc Han	THDC
K1203	Le Thanh Hau	THDC
K1103	Le Van Tam	CTDL
K1114	Tran Ngoc Han	CTDL
K1203	Le Thanh Hau	CTDL

HOCVIEN×MONHOC

3.6 Phép kết

3.6.1 Phép kết

3.6.2 Phép kết bằng, phép kết tự nhiên

3.6.3 Phép kết ngoài

3.6.1 Phép kết (Theta-Join) (1)

- ♦ Theta-join (θ): Tương tự như phép tích kết hợp với phép chọn. Điều kiện chọn gọi là *điều kiện kết*.
- ♦ **Ký hiệu:** $R \overset{p}{\bowtie} S$
trong đó R,S là các quan hệ, p là điều kiện kết
- ♦ Các bộ có giá trị NULL tại thuộc tính kết nối không xuất hiện trong kết quả của phép kết.
- ♦ Phép kết với điều kiện tổng quát gọi là θ -kết với θ là một trong những phép so sánh ($\neq, =, >, \geq, <, \leq$)

3.6.1 Phép kết (2)

$$A_1 > B_2$$

$$R \triangleright \triangleleft S$$

R	
A ₁	A ₂
1	2
1	8
0	0
8	4
0	3

S		
B ₁	B ₂	B ₃
0	2	8
7	8	7
8	0	4
1	0	7
2	1	5

A ₁	A ₂	B ₁	B ₂	B ₃
1	2	8	0	4
1	2	1	0	7
1	8	8	0	4
1	8	1	0	7
8	4	0	2	8
8	4	8	0	4
8	4	1	0	7
8	4	2	1	5

3.6.2 Phép kết bằng, kết tự nhiên

- ◆ Nếu θ là phép so sánh bằng ($=$), phép kết gọi là phép kết bằng (equi-join).

Ký hiệu: $HOCVIEN \overset{Mahv=Trglop}{\triangleright \triangleleft} LOP$

- ◆ Nếu điều kiện của equi-join là các thuộc tính giống nhau thì gọi là phép kết tự nhiên (natural-join). Khi đó kết quả của phép kết loại bỏ bớt 1 cột (bỏ 1 trong 2 cột giống nhau)

Ký hiệu: $HOCVIEN \overset{Mahv}{\triangleright \triangleleft} KETQUATHI$ hoặc

$HOCVIEN * KETQUATHI$

3.6.3 Phép kết ngoài (outer join)

- ◆ Mở rộng phép kết để tránh mất thông tin
- ◆ Thực hiện phép kết và sau đó thêm vào kết quả của phép kết các bộ của quan hệ mà không phù hợp với các bộ trong quan hệ kia.
- ◆ Có 3 loại:
 - Left outer join $R \bowtie_{\text{left}} S$
 - Right outer join $R \bowtie_{\text{right}} S$
 - Full outer join $R \bowtie_{\text{full}} S$
- ◆ **Ví dụ:** In ra danh sách tất cả các học viên và điểm số của các môn học mà học viên đó thi (nếu có)

Left outer join

- ◆ HOCVIEN  KETQUATHI

Mahv	Hoten	Mahv	Mamh	Diem
HV01	Nguyen Van Lan	HV01	CSDL	7.0
HV01	Nguyen Van Lan	HV01	CTRR	8.5
HV02	Tran Hong Son	HV02	CSDL	8.5
HV03	Nguyen Le	HV03	CTRR	9.0
HV04	Le Minh	Null	Null	Null

HOCVIEN	
Mahv	Hoten
HV01	Nguyen Van Lan
HV02	Tran Hong Son
HV03	Nguyen Le
HV04	Le Minh

KETQUATHI		
Mahv	Mamh	Diem
HV01	CSDL	7.0
HV02	CSDL	8.5
HV01	CTRR	8.5
HV03	CTRR	9.0

Right outer join

♦ HOCVIEN \bowtie KETQUATHI

Mahv	Hoten	Mahv	Mamh	Diem
HV01	Nguyen Van Lan	HV01	CSDL	7.0
HV01	Nguyen Van Lan	HV01	CTRR	8.5
HV02	Tran Hong Son	HV02	CSDL	8.5
HV03	Nguyen Le	HV03	CTRR	9.0
Null	Null	HV05	CTRR	8.0

HOCVIEN	
Mahv	Hoten
HV01	Nguyen Van Lan
HV02	Tran Hong Son
HV03	Nguyen Le
HV04	Le Minh

KETQUATHI		
Mahv	Mamh	Diem
HV01	CSDL	7.0
HV02	CSDL	8.5
HV01	CTRR	8.5
HV03	CTRR	9.0
HV05	CTRR	8.0

Full outer join

- Mahv
 HOCVIEN  KETQUATHI

Mahv	Hoten	Mahv	Mamh	Diem
HV01	Nguyen Van Lan	HV01	CSDL	7.0
HV01	Nguyen Van Lan	HV01	CTRR	8.5
HV02	Tran Hong Son	HV02	CSDL	8.5
HV03	Nguyen Le	HV03	CTRR	9.0
HV04	Le Minh	Null	Null	Null
Null	Null	HV05	CTRR	8.0

HOCVIEN	
Mahv	Hoten
HV01	Nguyen Van Lan
HV02	Tran Hong Son
HV03	Nguyen Le
HV04	Le Minh

KETQUATHI		
Mahv	Mamh	Diem
HV01	CSDL	7.0
HV02	CSDL	8.5
HV01	CTRR	8.5
HV03	CTRR	9.0
HV05	CTRR	8.0

3.7 Phép chia (Division)

- ♦ Được dùng để lấy ra một số bộ trong quan hệ R sao cho thỏa với tất cả các bộ trong quan hệ S
- ♦ Ký hiệu $R \div S$
 - $R(Z)$ và $S(X)$
 - Z là tập thuộc tính của R, X là tập thuộc tính của S
 - $X \subseteq Z$
- ♦ Kết quả của phép chia là một quan hệ $T(Y)$
 - Với $Y=Z-X$
 - Có t là một bộ của T nếu với mọi bộ $t_s \in S$, tồn tại bộ $t_R \in R$ thỏa 2 điều kiện
 - $t_R(Y) = t$
 - $t_R(X) = t_s(X)$

R(Z)	
X	Y

S(X)
X

T(Y)
Y

3.7 Phép chia (Division)

- ♦ **Định nghĩa:** $Q = R \div S = \{t / \forall s \in S, (t, s) \in R\}$
- ♦ R và S là hai quan hệ, R^+ và S^+ lần lượt là tập thuộc tính của R và S. Điều kiện $S^+ \neq \emptyset$ là **tập con không bằng** của R^+ . Q là kết quả phép chia giữa R và S, $Q^+ = R^+ - S^+$
- ♦ Có thể diễn đạt bằng phép toán đại số như sau:

$$T_1 \leftarrow \pi_{R^+ - S^+}(R)$$

$$T_2 \leftarrow \pi_{R^+ - S^+}((S \times T_1) - R)$$

$$T \leftarrow T_1 - T_2$$

3.7 Phép chia (2)

♦ Ví dụ

$R \div S$

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

3.7 Phép chia (tt)

Câu hỏi: Cho biết Mahv nào đã thi hết các môn học?

KETQUATHI		
Mahv	Mamh	Diem
HV01	CSDL	7.0
HV02	CSDL	8.5
HV01	CTRR	8.5
HV03	CTRR	9.0
HV01	THDC	7.0
HV02	THDC	5.0
HV03	THDC	7.5
HV03	CSDL	6.0

KETQUA (R)

MONHOC	
Mamh	Tenmh
CSDL	Cơ sở dữ liệu
CTRR	Cấu trúc rời rạc
THDC	Tin học đại cương

MONHOC (S)



Mahv
HV01
HV03

KETQUA ÷ MONHOC

$KETQUA \leftarrow KETQUATHI[Mahv, Mamh]$

$MONHOC \leftarrow MONHOC[Mamh]$

3.7 Phép chia (tt)

$$T_1 \leftarrow \pi_{R^+ - S^+}(R)$$

T1
Mahv
HV01
HV02
HV03

3.7 Phép chia (tt)

$$T_2 \leftarrow \pi_{R^+ - S^+}((S \times T_1) - R)$$

T1
Mahv
HV01
HV02
HV03

X

MONHOC
Mamh
CSDL
CTRR
THDC

T1-MONHOC	
Mahv	Mamh
HV01	CSDL
HV01	CTRR
HV01	THDC
HV02	CSDL
HV02	CTRR
HV02	THDC
HV03	CSDL
HV03	CTRR
HV03	THDC

—

KETQUA (R)	
Mahv	Mamh
HV01	CSDL
HV02	CSDL
HV01	CTRR
HV03	CTRR
HV01	THDC
HV02	THDC
HV03	THDC
HV03	CSDL

T2	
Mahv	Mamh
HV02	CTRR

3.7 Phép chia (tt)

$$T \leftarrow T_1 - T_2$$

T1
Mahv
HV01
HV02
HV03

—

T2
Mahv
HV02

T
Mahv
HV01
HV03

3.7 Phép chia (4)

- ♦ Biểu diễn phép chia thông qua tập đầy đủ các phép toán ĐSQH

$$Q1 \leftarrow \pi_Y(R)$$

$$Q2 \leftarrow Q1 \times S$$

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

$$T \leftarrow Q1 - Q3$$

3.8 Hàm tính toán và gom nhóm (1)

- ◆ Hàm tính toán gồm các hàm: avg(giatri), min(giatri), max(giatri), sum(giatri), count(giatri).
- ◆ Phép toán gom nhóm:

$$G_1, G_2, \dots, G_n \mathfrak{S} F_1(A_1), F_2(A_2), \dots, F_n(A_n) (E)$$

- E là biểu thức đại số quan hệ
- G_i là thuộc tính gom nhóm (rỗng, nếu không gom nhóm)
- F_i là hàm tính toán
- A_i là tên thuộc tính

3.8 Hàm tính toán và gom nhóm (2)

- ♦ Điểm thi cao nhất, thấp nhất, trung bình của môn CSDL ?

$\mathfrak{J}_{\max(Diem), \min(Diem), agv(Diem)} \sigma_{Mamh='CSDL'} (KETQUATHI)$

KETQUATHI		
Mahv	Mamh	Diem
HV01	CSDL	7.0
HV02	CSDL	8.5
HV01	CTRR	8.5
HV03	CTRR	9.0
HV05	CTRR	8.0

3.8 Hàm tính toán và gom nhóm (3)

- ◆ Điểm thi cao nhất, thấp nhất, trung bình của từng môn ?

Mamh $\mathfrak{J}_{\max(Diem), \min(Diem), \text{avg}(Diem)}$ (*KETQUATHI*)

KETQUATHI		
Mahv	Mamh	Diem
HV01	CSDL	7.0
HV02	CSDL	8.5
HV01	CTRR	8.5
HV03	CTRR	9.0
HV05	CTRR	8.0

BÀI TẬP

Lược đồ CSDL quản lý bán hàng gồm có các quan hệ sau:

KHACHHANG (MAKH, HOTEN, DCHI, SODT, NGSINH,
DOANHISO, NGDK)

NHANVIEN (MANV,HOTEN, NGVL, SODT)

SANPHAM (MASP,TENSP, DVT, NUOCSX, GIA)

HOADON (SOHD, NGHD, MAKH, MANV, TRIGIA)

CTHD (SOHD,MASP,SL)

Mô tả các câu truy vấn sau bằng ĐSQH

1. In ra danh sách các sản phẩm (MASP,TENSP) do “Việt Nam” sản xuất có giá từ 30.000 đến 40.000
2. In ra danh sách các khách hàng (MAKH, HOTEN) đã mua hàng trong ngày 1/1/2007.
3. In ra danh sách các sản phẩm (MASP,TENSP) do “Việt Nam” sản xuất hoặc các sản phẩm được bán ra trong ngày 1/1/2007.
4. Tìm các số hóa đơn mua cùng lúc 2 sản phẩm có mã số “BB01” và “BB02”.
5. In ra danh sách các sản phẩm (MASP,TENSP) do “Việt Nam” sản xuất không bán được trong năm 2006.
6. Tìm số hóa đơn đã mua tất cả các sản phẩm do Singapore sản xuất