

Các ngôn ngữ dữ liệu đối với mô hình quan hệ

Ngôn ngữ Đại số quan hệ

Nguyễn Thị Oanh
oanhnt@soict.hust.edu.vn

Bộ môn Hệ thống thông tin (<http://is.hust.edu.vn/>)
Viện CNTT&TT
Trường Đại học Bách Khoa Hà Nội

Hệ

Ứng dụng

- Định nghĩa DL
- Thao tác DL

Hệ QTCSDL

CSDL

CSDL

Nội dung

- Đặt vấn đề
 - Giới thiệu một số ngôn ngữ và phân loại
 - So sánh và đánh giá
- Ngôn ngữ dữ liệu mức thấp:
 - **Đại số quan hệ**
 - Tính toán vị từ
- Một số ngôn ngữ dữ liệu mức cao
 - QBE (Query By Example)
 - **SQL** (Structured Query Language)
- Kết luận



Đặt vấn đề

- Mục đích của ngôn ngữ dữ liệu
- Tại sao có nhiều ngôn ngữ dữ liệu?
- Ngôn ngữ cấp thấp vs. Ngôn ngữ cấp cao?

Ví dụ

- Tìm tên của các sinh viên nào sống ở Bundoora
 - Tìm các bộ của bảng Student có Suburb = Bundoora
 - Đưa ra các giá trị của thuộc tính Name của các bộ này

Student

Id	Name	Suburb
1108	Robert	Kew
3936	Glen	Bundoora
8507	Norman	Bundoora
8452	Mary	Balwyn

Ví dụ (2)

- Tìm các sinh viên đăng ký khoá học có mã số 113
 - Tìm các giá trị SID trong bảng Enrol có trường course tương ứng là 113
 - Đưa các bộ của bảng Student có Id trong các giá trị tìm thấy ở trên

Student

Id	Name	Suburb
1108	Robert	Kew
3936	Glen	Bundoora
8507	Norman	Bundoora
8452	Mary	Balwyn

Enrol

SID	Course
3936	101
1108	113
8507	101

Course

No	Name	Dept
113	BCS	CSCE
101	MCS	CSCE

Phân loại ngôn ngữ truy vấn

- Ngôn ngữ cấp thấp
 - **Đại số quan hệ**
 - 1 câu hỏi = 1 tập các phép toán trên các quan hệ
 - Được biểu diễn bởi một biểu thức đại số (quan hệ)
 - **Tính toán vị từ**
 - 1 câu hỏi = 1 mô tả của các bộ mong muốn
 - Được đặc tả bởi một vị từ mà các bộ phải thoả mãn
 - Phân biệt 2 lớp:
 - ngôn ngữ tính toán vị từ biến bộ
 - ngôn ngữ tính toán vị từ biến miền
- Ngôn ngữ cấp cao
 - **QBE**
 - **SQL**



Ngôn ngữ đại số quan hệ

Tổng quan

- Gồm các phép toán tương ứng với các thao tác trên các quan hệ
- Mỗi phép toán
 - Đầu vào: một hay nhiều quan hệ
 - Đầu ra: một quan hệ
- Biểu thức đại số quan hệ = chuỗi các phép toán
- Kết quả thực hiện một biểu thức đại số là một quan hệ
- Được cài đặt trong phần lớn các hệ CSDL hiện nay

Phân loại các phép toán

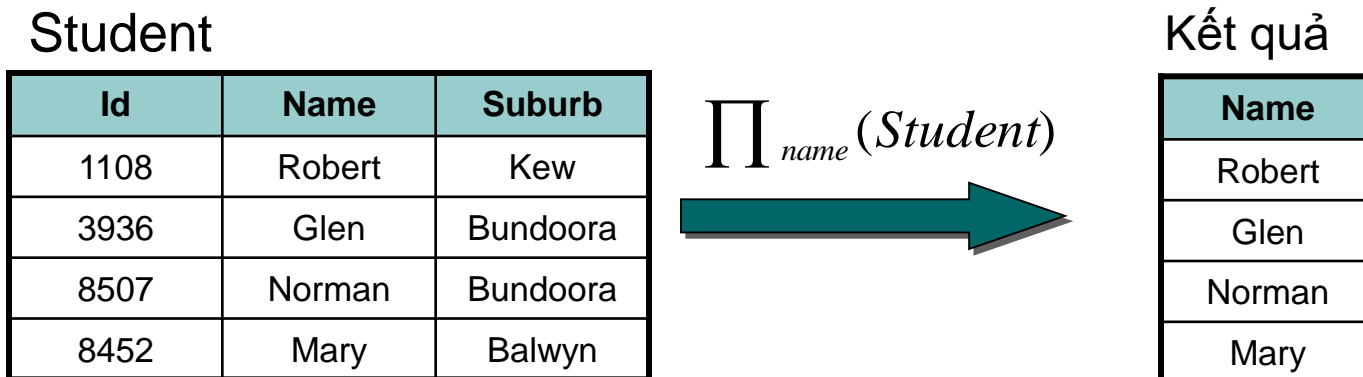
- Phép toán quan hệ
 - Phép chiếu (*projection*)
 - Phép chọn (*selection*)
 - Phép kết nối (*join*)
 - Phép chia (*division*)
- Phép toán tập hợp
 - Phép hợp (*union*)
 - Phép giao (*intersection*)
 - Phép trừ (*difference*)
 - Phép tích đề-các (*cartesian product*)

Phép chiếu

- Đ/n: Lựa chọn một số thuộc tính từ một quan hệ
- Cú pháp: $\Pi_{A1, A2, \dots}(R)$



- ❖ Ví dụ: đưa ra danh sách tên của tất cả các sinh viên



Phép chọn

- Đ/n: Lựa chọn các bộ trong một quan hệ thoả mãn điều kiện cho trước
- Cú pháp: $\sigma_{\langle condition \rangle}(R)$

R1
R2
R3
R4



R2
R3

- ❖ Ví dụ: đưa ra danh sách những sinh viên sống ở Bundoora

$$\sigma_{suburb="Bundoora"}(Student)$$

Student

Id	Name	Suburb
1108	Robert	Kew
3936	Glen	Bundoora
8507	Norman	Bundoora
8452	Mary	Balwyn



Kết quả

Id	Name	Suburb
3936	Glen	Bundoora
8507	Norman	Bundoora

Vi dụ - chọn và chiếu

- đưa ra tên của các sinh viên sống ở Bundoora

$$\Pi_{name}(\sigma_{suburb="Bundoora"}(Student))$$

Student

Id	Name	Suburb
1108	Robert	Kew
3936	Glen	Bundoora
8507	Norman	Bundoora
8452	Mary	Balwyn

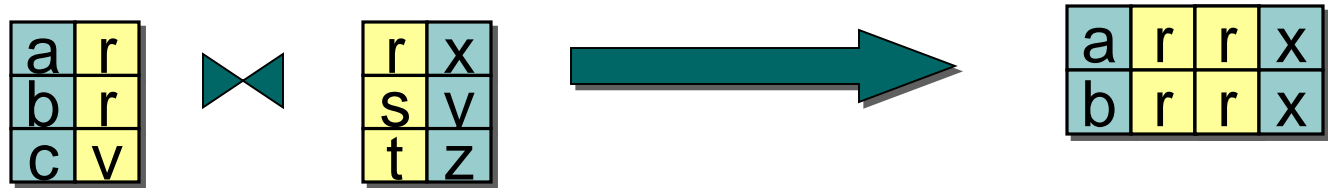


Kết quả

Name
Glen
Norman

Phép kết nối

- Đ/n: ghép các bộ từ 2 quan hệ thoả mãn điều kiện kết nối $R_1 \triangleright \triangleleft_{\langle join_condition \rangle} R_2$
- Cú pháp:



- ❖ Ví dụ: đưa ra danh sách các sinh viên và khoá học $Student \triangleright \triangleleft_{Id=SID} Enrol$

Student

Id	Name	Suburb
1108	Robert	Kew
3936	Glen	Bundoora
8507	Norman	Bundoora
8452	Mary	Balwyn



Enrol

SID	Course
3936	101
1108	113
8507	101



Kết quả

SID	Id	Name	Suburb	Course
1108	1108	Robert	Kew	113
3936	3936	Glen	Bundoora	101
8507	8507	Norman	Bundoora	101

Ví dụ - chọn, chiếu và kết nối

- đưa ra **tên** của các sinh viên sống ở Bundoora và **mã khoá học** mà sinh viên đó đăng ký

$$\Pi_{Name, Course} (\sigma_{Suburb="Bundoora"} (Student \bowtie_{Id=SID} Enrol))$$

Student

Id	Name	Suburb
1108	Robert	Kew
3936	Glen	Bundoora
8507	Norman	Bundoora
8452	Mary	Balwyn

Enrol

SID	Course
3936	101
1108	113
8507	101



Kết quả

Name	Course
Glen	101
Norman	101

Phép kết nối tự nhiên

- Đ/n: là phép kết nối với điều kiện bằng trên các thuộc tính trùng tên

❖ Ví dụ:

Takes

SID	SNO
1108	21
1108	23
8507	23
8507	29

Enrol

SID	Course
3936	101
1108	113
8507	101

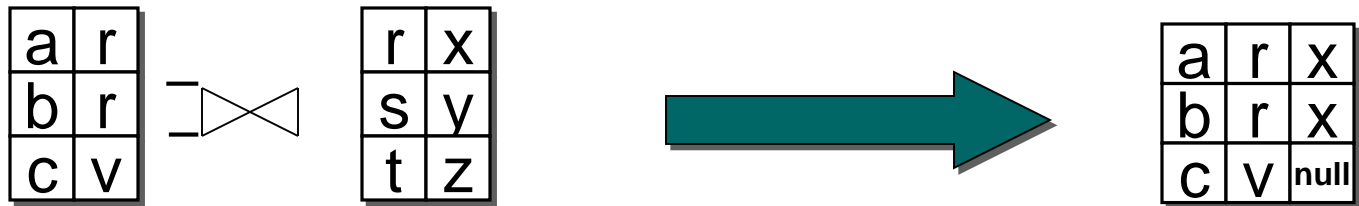
*



SID	SNO	Course
1108	21	113
1108	23	113
8507	23	101
8507	29	101

Phép kết nối ngoài

- Phép kết nối ngoài trái



- Phép kết nối ngoài phải




Ví dụ về phép kết nối ngoài

- Đưa ra danh sách mã số các sinh viên và mã khoá học mà sinh viên đó đăng ký nếu có

Student

ID	Name	Suburb
1108	Robert	Kew
3936	Glen	Bundoora
8507	Norman	Bundoora
8452	Mary	Balwyn


Id = SID

Enrol

SID	Course
3936	101
1108	113
8507	101

Kết quả

ID	Name	Suburb	Course
1108	Robert	Kew	113
3936	Glen	Bundoora	101
8507	Norman	Bundoora	101
8452	Mary	Balwyn	null

Phép tích đề-các

- Đ/n: là kết nối giữa từng bộ của quan hệ thứ nhất và mỗi bộ của quan hệ thứ hai
- Cú pháp: $R_1 \times R_2$

a
b
c
d

\times

x
y



a	x
a	y
b	x
b	y
c	x
c	y
d	x
d	y

Ví dụ phép tích đề-các

Student

Id	Name	Suburb
1108	Robert	Kew
3936	Glen	Bundoora
8507	Norman	Bundoora
8452	Mary	Balwyn

Sport

SportID	Sport
05	Swimming
09	Dancing

X

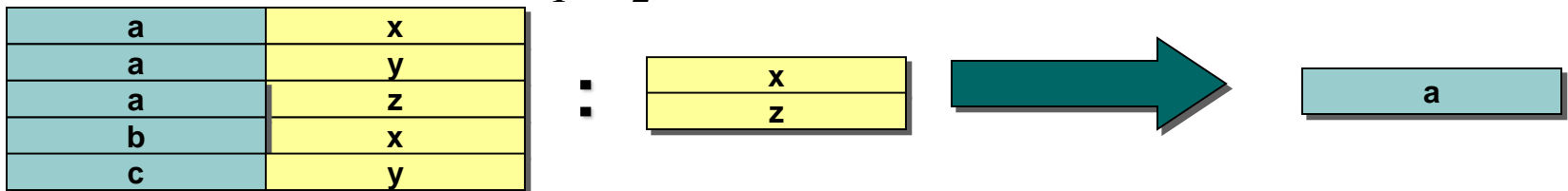
Student_Sport



Id	Name	Suburb	SportID	Sport
1108	Robert	Kew	05	Swimming
3936	Glen	Bundoora	05	Swimming
8507	Norman	Bundoora	05	Swimming
8452	Mary	Balwyn	05	Swimming
1108	Robert	Kew	09	Dancing
3936	Glen	Bundoora	09	Dancing
8507	Norman	Bundoora	09	Dancing
8452	Mary	Balwyn	09	Dancing

Phép chia

- Đ/n: cho R_1 và R_2 lần lượt là các quan hệ n và m ngôi. Kết quả của phép chia R_1 cho R_2 là một quan hệ (n-m) ngôi
- Cú pháp: $R_1 : R_2$



❖ Ví dụ:

Subject

Name	Course
Systems	BCS
Database	BCS
Database	MCS
Algebra	MCS

Course

Course
BCS
MCS

:



Kết quả

Name
Database

Phép hợp

- Đ/n: gồm các bộ thuộc ít nhất một trong hai quan hệ đầu vào
 - 2 quan hệ khả hợp được xác định trên cùng miền giá trị
- Cú pháp: $R_1 \cup R_2$



❖ Ví dụ:

Subject

Name	Course
Systems	BCS
Database	BCS
Database	MCS
Algebra	MCS

∪

Subject2

Name	Course
DataMining	MCS
Writing	BCS

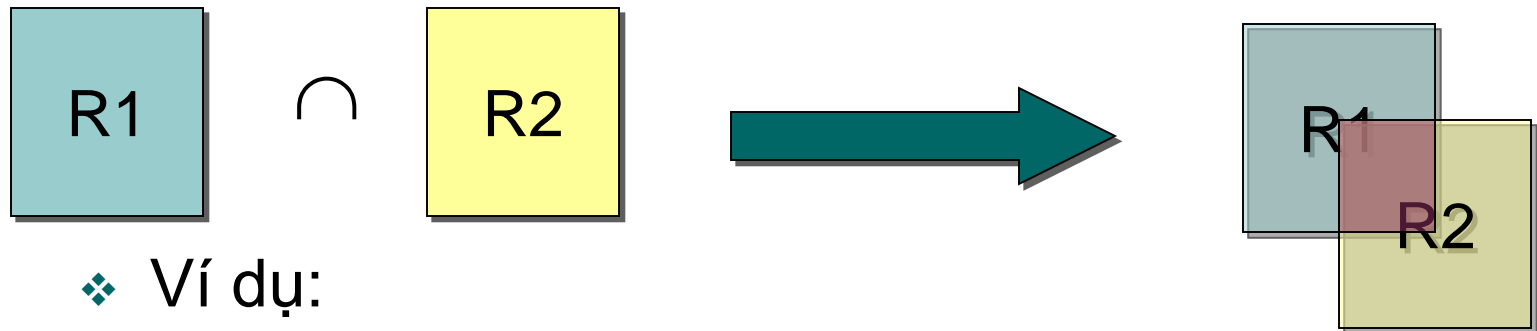
Kết quả



Name	Course
Systems	BCS
Database	BCS
Database	MCS
Algebra	MCS
DataMining	MCS
Writing	BCS

Phép giao

- Đ/n: gồm các bộ thuộc cả hai quan hệ đầu vào
- Cú pháp: $R_1 \cap R_2$



❖ Ví dụ:

Subject

Name	Course
Systems	BCS
Database	BCS
Database	MCS
Algebra	MCS

Subject2

Name	Course
DataMining	MCS
Database	MCS
Systems	BCS
Writing	BCS

Kết quả

Name	Course
Systems	BCS
Database	MCS

Phép trừ

- Đ/n: gồm các bộ thuộc quan hệ thứ nhất nhưng không thuộc quan hệ thứ hai
 - 2 quan hệ phải là khả hợp
- Cú pháp: $R_1 \setminus R_2$



❖ Ví dụ:

Subject

Name	Course
Systems	BCS
Database	BCS
Database	MCS
Algebra	MCS

Subject2

Name	Course
DataMining	MCS
Database	MCS
Systems	BCS
Writing	BCS

\



Kết quả

Name	Course
Database	BCS
Algebra	MCS

Ví dụ

- Phép giao và phép trừ

r	(A	B	C)
	a ₁	b ₁	c ₁
	a ₁	b ₁	c ₂
	a ₁	b ₂	c ₂
	a ₂	b ₂	c ₂
	a ₃	b ₂	c ₂

s	(A	B	C)
	a ₁	b ₁	c ₁
	a ₁	b ₂	c ₁
	a ₁	b ₂	c ₂

$r \cap s = g$	(A	B	C)
	a ₁	b ₁	c ₁
	a ₁	b ₂	c ₂

$r - s = t$	(A	B	C)
	a ₁	b ₁	c ₂
	a ₂	b ₂	c ₂
	a ₃	b ₂	c ₂

Ví dụ (...)

- Phép chiếu

$$\Pi_X(r) = \{ t[X] \mid t \in r \}$$

$$X = \{ A, B \} ; Y = \{ C \}$$

r	(A	B	C)
	a ₁	b ₁	c ₁
	a ₁	b ₁	c ₂
	a ₁	b ₂	c ₂
	a ₂	b ₂	c ₂
	a ₃	b ₂	c ₂

$$\Pi_X(r) = s_1 \left(\begin{array}{cc} A & B \\ a_1 & b_1 \\ a_1 & b_2 \\ a_2 & b_2 \\ a_3 & b_2 \end{array} \right)$$

$$\Pi_Y(r) = s_2 \left(\begin{array}{c} C \\ c_1 \\ c_2 \end{array} \right)$$

Ví dụ (...)

- Phép chọn:

$$\sigma_F(r) = \{ t \mid t \in r \wedge F(t) = \text{đúng} \}$$

r	(A	B	C)
	a ₁	b ₁	c ₁
	a ₁	b ₁	c ₂
	a ₁	b ₂	c ₂
	a ₂	b ₂	c ₂
	a ₃	b ₂	c ₂

$$\sigma_{A=a_1}(r) = r_1$$

(A	B	C)
a ₁	b ₁	c ₁
a ₁	b ₂	c ₂
a ₁	b ₁	c ₂

$$\sigma_{A=a_1 \wedge C=c_2}(r) = r_2$$

(A	B	C)
a ₁	b ₁	c ₂
a ₁	b ₂	c ₂



Bài tập

- Biểu diễn câu hỏi truy vấn bằng ngôn ngữ đại số quan hệ
- Tính kết quả của biểu thức

Bài tập 1

- Cho CSDL gồm 3 quan hệ sau: S(Các hãng cung ứng), P (các mặt hàng), SP(các sự cung ứng)

S (S# SNAME STATUS CITY)					SP (S# P# QTY)		
S1	Smith	20	London		S1	P1	300
S2	Jones	10	Paris		S1	P2	200
S3	Black	30	Paris		S1	P3	400
					S2	P1	300
					S2	P2	400
					S3	P2	200
P (P# PNAME COLOR WEIGHT CITY)							
P1	Nut	red	12	London			
P2	Bolt	green	17	Paris			
P3	Screw	blue	17	Rom			
P4	Screw	red	14	London			

Bài tập 1 – Yêu cầu

- Biểu diễn câu hỏi truy vấn bằng ngôn ngữ đại số quan hệ
 - Đưa ra danh sách các mặt hàng màu đỏ
 - Cho biết S# của các hãng cung ứng mặt hàng 'P1' hoặc 'P2'
 - Liệt kê S# của các hãng cung ứng cả hai mặt hàng 'P1' và 'P2'
 - Đưa ra S# của các hãng cung ứng ít nhất một mặt hàng màu đỏ
 - Đưa ra S# của các hãng cung ứng tất cả các mặt hàng.
- Tính kết quả của các câu truy vấn

Bài tập 1 – đáp án

- Đưa ra danh sách các mặt hàng màu đỏ:

$$\sigma_{\text{COLOR} = \text{'red'}}(P)$$

- Cho biết S# của các hãng cung ứng mặt hàng 'P1' hoặc 'P2':

$$\Pi_{S\#}(\sigma_{P\# = \text{'P1'}} \vee P\# = \text{'P2'}}(SP))$$

- Liệt kê S# của các hãng cung ứng cả hai mặt hàng 'P1' và 'P2':

$$\Pi_{S\#}(\sigma_{P\# = \text{'P1'}}(SP)) \cap \Pi_{S\#}(\sigma_{P\# = \text{'P2'}}(SP))$$

- Đưa ra S# của các hãng cung ứng ít nhất một mặt hàng màu đỏ:

$$\Pi_{S\#}(SP * \sigma_{\text{COLOR} = \text{'red'}}(P))$$

- Đưa ra S# của các hãng cung ứng tất cả các mặt hàng:

$$\Pi_{S\#,P\#}(SP) \div \Pi_{P\#}(P)$$

