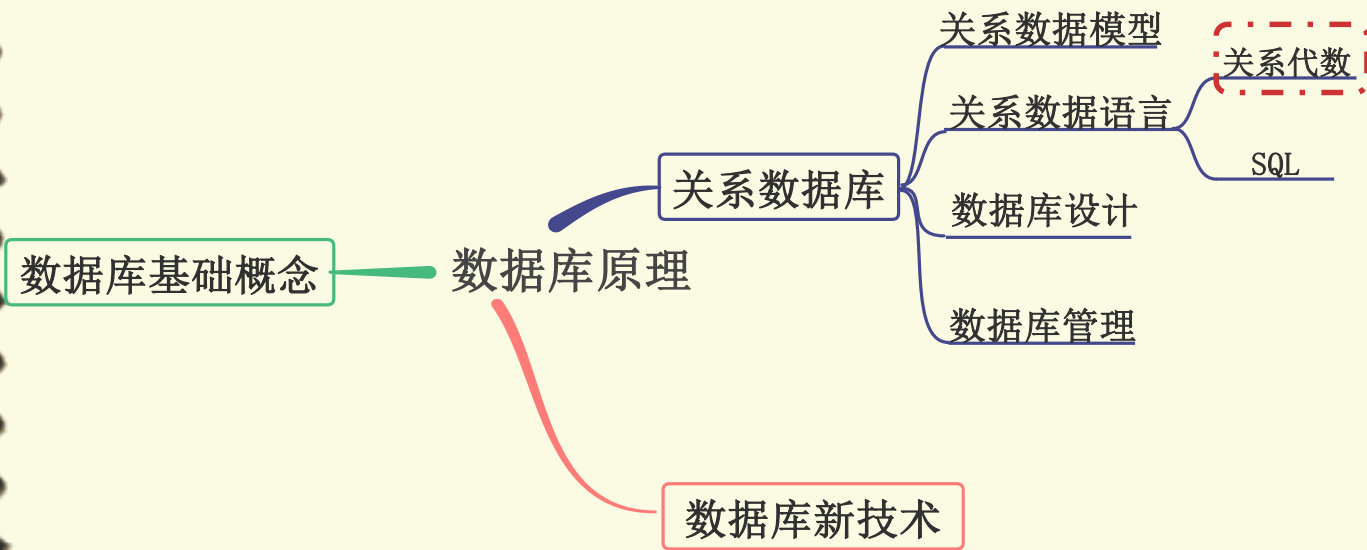


数据库原理

第3章 关系代数

辽东学院 鲁 琴

本章要点



关系数据查询语言

关系代数, 关系演算 等

- Formal
- 不用于现在商用的DBMS中

SQL

- Actual
- 当DBMS处理SQL查询时，首先是将其转换成关系代数或相似的内部表现形式

关系代数

- ◆ 并、差、交
- ◆ 投影、选择
- ◆ 笛卡尔积、连接
- ◆ 重命名

什么是关系代数

- ◆ 是一种抽象的数据查询语言
- ◆ 用对关系的运算来表达查询

- 运算对象：关系
- 运算符：4类
- 运算结果：关系

$$2 + 3 = 5$$

$$R \cup S = ?$$

$$R \bowtie S = ?$$

$$\sigma_c(R) = ?$$

关系运算符分类

◆ 传统的集合运算符

- \cup $-$ \cap
- 将关系看成元组的集合
- 所有的运算对象必须具有相同的结构

◆ 专门的关系运算符

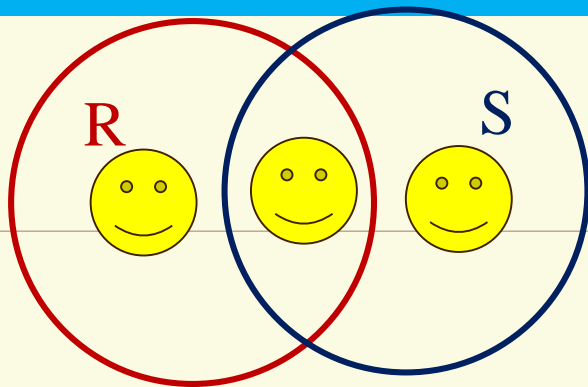
- 选择和投影 σ π
- 笛卡尔积, 连接 \times \bowtie
- 重命名 ρ

辅助专门的关系运算符

比较运算符 $<$ $<=$ $>$ $>=$ $=$ $<>$

逻辑运算符 \neg \wedge \vee

① 传统的集合运算



◆ $R \cup S$

- 并(**union**)
- 结果由属于R或属于S的所有元组组成

◆ $R \cap S$

- 交(**intersection**), 结果由既属于R又属于S的元组组成

◆ $R - S$

- 差(**difference**), 结果由属于R而不属于S的所有元组组成

◆ 先决条件

- R 和S 的属性个数必须相同
- 对应属性必须取自同一个域

实例：并运算

R

<i>A</i>	<i>B</i>	<i>C</i>
<i>a1</i>	<i>b1</i>	<i>c1</i>
<i>a1</i>	<i>b2</i>	<i>c2</i>
<i>a2</i>	<i>b2</i>	<i>c1</i>

S

<i>A</i>	<i>B</i>	<i>C</i>
<i>a1</i>	<i>b2</i>	<i>c2</i>
<i>a1</i>	<i>b3</i>	<i>c2</i>
<i>a2</i>	<i>b2</i>	<i>c1</i>

R ∪ S

<i>A</i>	<i>B</i>	<i>C</i>
<i>a1</i>	<i>b1</i>	<i>c1</i>
<i>a1</i>	<i>b2</i>	<i>c2</i>
<i>a2</i>	<i>b2</i>	<i>c1</i>
<i>a1</i>	<i>b3</i>	<i>c2</i>

实例：差运算

R

<i>A</i>	<i>B</i>	<i>C</i>
<i>a1</i>	<i>b1</i>	<i>c1</i>
<i>a1</i>	<i>b2</i>	<i>c2</i>
<i>a2</i>	<i>b2</i>	<i>c1</i>

S

<i>A</i>	<i>B</i>	<i>C</i>
<i>a1</i>	<i>b2</i>	<i>c2</i>
<i>a1</i>	<i>b3</i>	<i>c2</i>
<i>a2</i>	<i>b2</i>	<i>c1</i>

R-S

<i>A</i>	<i>B</i>	<i>C</i>
<i>a1</i>	<i>b1</i>	<i>c1</i>

实例：交运算

R

A	B	C
$a1$	$b1$	$c1$
$a1$	$b2$	$c2$
$a2$	$b2$	$c1$

S

A	B	C
$a1$	$b2$	$c2$
$a1$	$b3$	$c2$
$a2$	$b2$	$c1$

$R \cap S$

A	B	C
$a1$	$b2$	$c2$
$a2$	$b2$	$c1$

讨论

<i>name</i>	<i>address</i>	<i>gender</i>	<i>birthdate</i>
Carrie Fisher	123 Maple St., Hollywood	F	9/9/99
Mark Hamill	456 Oak Rd., Brentwood	M	8/8/88

Relation *R*

<i>name</i>	<i>address</i>	<i>gender</i>	<i>birthdate</i>
Carrie Fisher	123 Maple St., Hollywood	F	9/9/99
Harrison Ford	789 Palm Dr., Beverly Hills	M	7/7/77

Relation *S*

$R \cup S$

<i>name</i>	<i>address</i>	<i>gender</i>	<i>birthdate</i>
Carrie Fisher	123 Maple St., Hollywood	F	9/9/99
Mark Hamill	456 Oak Rd., Brentwood	M	8/8/88
Harrison Ford	789 Palm Dr., Beverly Hills	M	7/7/77

讨论

<i>name</i>	<i>address</i>	<i>gender</i>	<i>birthdate</i>
Carrie Fisher	123 Maple St., Hollywood	F	9/9/99
Mark Hamill	456 Oak Rd., Brentwood	M	8/8/88

Relation *R*

<i>name</i>	<i>address</i>	<i>gender</i>	<i>birthdate</i>
Carrie Fisher	123 Maple St., Hollywood	F	9/9/99
Harrison Ford	789 Palm Dr., Beverly Hills	M	7/7/77

Relation *S*

$$R \cap S$$

<i>name</i>	<i>address</i>	<i>gender</i>	<i>birthdate</i>
Carrie Fisher	123 Maple St., Hollywood	F	9/9/99

讨论

<i>name</i>	<i>address</i>	<i>gender</i>	<i>birthdate</i>
Carrie Fisher	123 Maple St., Hollywood	F	9/9/99
Mark Hamill	456 Oak Rd., Brentwood	M	8/8/88

Relation *R*

<i>name</i>	<i>address</i>	<i>gender</i>	<i>birthdate</i>
Carrie Fisher	123 Maple St., Hollywood	F	9/9/99
Harrison Ford	789 Palm Dr., Beverly Hills	M	7/7/77

Relation *S*

R - S

<i>name</i>	<i>address</i>	<i>gender</i>	<i>birthdate</i>
Mark Hamill	456 Oak Rd., Brentwood	M	8/8/88

②投影和选择

◆ 投影 (Projection)

◆ 选择 (Selection)

投影 (Projection)

- ◆ 用于从 R 中选择出若干属性列组成新的关系

$$\pi_L(R)$$

- ◆ L 为 R 中的属性列表
- ◆ 结果为只包含 R 中某些列的新的关系
- ◆ 结果要去掉重复元组

实例

Student

Sno	Sname	Ssex	Sage	Sdept
95001	李勇	男	20	CS
95002	刘晨	女	19	IS
95003	王敏	女	18	MA
95004	张立	男	19	IS

查询所有学生的姓名和所在系

- 即求Student关系上学生姓名和所在系两个属性上的投影

$\pi_{\text{Sname}, \text{Sdept}}(\text{Student})$

结果

或 $\pi_{2,5}(\text{Student})$

Sname	Sdept
李勇	CS
刘晨	IS
王敏	MA
张立	IS

实例


Student

Sno	Sname	Ssex	Sage	Sdept
95001	李勇	男	20	CS
95002	刘晨	女	19	IS
95003	王敏	女	18	MA
95004	张立	男	19	IS

查询学生关系Student中都有哪些系

$\pi_{Sdept}(Student)$

结果

Sdept
CS
IS
MA


选择 (Selection)

- ◆ 用于在关系**R**中选择满足给定条件的各个元组

$\sigma_c(R)$

- **C**: 选择条件, 是一个逻辑表达式
- 结果为只包含**R**中某些元组的新的关系

实例

Student

Sno	Sname	Ssex	Sage	Sdept
95001	李勇	男	20	CS
95002	刘晨	女	19	IS
95003	王敏	女	18	MA
95004	张立	男	19	IS

查询信息系（IS系）全体学生信息

$\sigma_{Sdept = 'IS'} (Student)$

或 $\sigma_5 = 'IS' (Student)$

结果

Sno	Sname	Ssex	Sage	Sdept
95002	刘晨	女	19	IS
95004	张立	男	19	IS

实例

Student

Sno	Sname	Ssex	Sage	Sdept
95001	李勇	男	20	CS
95002	刘晨	女	19	IS
95003	王敏	女	18	MA
95004	张立	男	19	IS

查询年龄小于20岁的学生信息

$\sigma_{Sage < 20}(Student)$

结果

$\sigma_{4 < 20}(Student)$

Sno	Sname	Ssex	Sage	Sdept
95002	刘晨	女	19	IS
95003	王敏	女	18	MA
95004	张立	男	19	IS

Movies

讨论

title	YEAR	LENGTH	genre	studioName
Galaxy Quest	1999	104	comdedy	DreamWorks
Pretty Woman	1990	119	comedy	Disney
Star Trek	1966	98	sciFi	Paramount
Star Wars	1977	124	sciFi	Fox
Superman	1978	143	sciFi	Warner Bros

查询时长至少为100分钟且为‘Fox’公司出品的影片信息

其它
方法?

$$\sigma_{length \geq 100 \wedge studioName = 'Fox'}(Movies)$$

$$\sigma_{length \geq 100}(Movies) \cap \sigma_{studioName = 'Fox'}(Movies)$$

结果

title	YEAR	LENGTH	genre	studioName
Star Wars	1977	124	sciFi	Fox

Movies

讨论

title	YEAR	LENGTH	genre	studioName
Galaxy Quest	1999	104	comdedy	DreamWorks
Pretty Woman	1990	119	comedy	Disney
Star Trek	1966	98	sciFi	Paramount
Star Wars	1977	124	sciFi	Fox
Superman	1978	143	sciFi	Warner Bros

查询时长至少为100分钟且为‘Fox’公司出品的影片的名字和发行年份

$\Pi_{\text{title,year}}(\sigma_{\text{length} \geq 100 \wedge \text{studioName} = \text{'Fox'}}(\text{Movies}))$

$\Pi_{\text{title,year}}(\sigma_{\text{length} \geq 100}(\text{Movies})) \cap \sigma_{\text{studioName} = \text{'Fox'}}(\text{Movies})$

$\Pi_{\text{title,year}}($

③笛卡尔积，连接

◆ 笛卡尔积 (Product)

◆ 连接 (Join)

- θ 连接
- 等值连接
- 自然连接

笛卡尔积 (Cartesian Product)

◆ R关系：n个属性，k1个元组

◆ S关系：m个属性，k2个元组

◆ $R \times S$

– 将R中的每个元组t1和S中的每个元组t2配对连接

– 列数：n+m

• 前n列是关系R的一个元组t1

• 后m列是关系S的一个元组t2

– 行数：k1 × k2

– 当R和S中有重名属性A时，则采用R.A和S.A分别命名对应的属性列

实例

R	A	B	C
	$a1$	$b1$	$c1$
	$a1$	$b2$	$c2$
	$a2$	$b2$	$c1$

S	A	B	C
	$a1$	$b2$	$c2$
	$a1$	$b3$	$c2$
	$a2$	$b2$	$c1$

$R \times S$

$R.A$	$R.B$	$R.C$	$S.A$	$S.B$	$S.C$
$a1$	$b1$	$c1$	$a1$	$b2$	$c2$
$a1$	$b1$	$c1$	$a1$	$b3$	$c2$
$a1$	$b1$	$c1$	$a2$	$b2$	$c1$
$a1$	$b2$	$c2$	$a1$	$b2$	$c2$
$a1$	$b2$	$c2$	$a1$	$b3$	$c2$
$a1$	$b2$	$c2$	$a2$	$b2$	$c1$
$a2$	$b2$	$c1$	$a1$	$b2$	$c2$
$a2$	$b2$	$c1$	$a1$	$b3$	$c2$
$a2$	$b2$	$c1$	$a2$	$b2$	$c1$

连接 (Join)

◆ 连接也称为 θ 连接

$$R \bowtie_{A\theta B} S$$

- A和B：分别为R和S上度数相等且可比的属性组
 - θ ：比较运算符
- 连接运算从R和S的笛卡尔积 $R \times S$ 中选取（R关系）在A属性组上的值与（S关系）在B属性组上值满足比较条件的元组

例

R

<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> ₁	<i>b</i> ₁	5
<i>a</i> ₁	<i>b</i> ₂	6
<i>a</i> ₂	<i>b</i> ₃	8
<i>a</i> ₂	<i>b</i> ₄	12

S

<i>B</i>	<i>E</i>
<i>b</i> ₁	3
<i>b</i> ₂	7
<i>b</i> ₃	10
<i>b</i> ₃	2
<i>b</i> ₅	2

R ⋈ *S*
C < *E*

<i>A</i>	<i>R.B</i>	<i>C</i>	<i>S.B</i>	<i>E</i>
<i>a</i> ₁	<i>b</i> ₁	5	<i>b</i> ₂	7
<i>a</i> ₁	<i>b</i> ₁	5	<i>b</i> ₃	10
<i>a</i> ₁	<i>b</i> ₂	6	<i>b</i> ₂	7
<i>a</i> ₁	<i>b</i> ₂	6	<i>b</i> ₃	10
<i>a</i> ₂	<i>b</i> ₃	8	<i>b</i> ₃	10

两类常用连接运算 —— 等值连接

◆ 什么是等值连接

- θ 为 “=” 的连接运算称为等值连接

◆ $R \bowtie_{A=B} S$

◆ 等值连接的含义

- 从关系R与S的笛卡尔积中选取A、B属性值相等的那些元组

实例

R

<i>A</i>	<i>B</i>	<i>C</i>
<i>a</i> ₁	<i>b</i> ₁	5
<i>a</i> ₁	<i>b</i> ₂	6
<i>a</i> ₂	<i>b</i> ₃	8
<i>a</i> ₂	<i>b</i> ₄	12

求等值连接 $R \bowtie S$

R.B=S.B

<i>A</i>	<i>R.B</i>	<i>C</i>	<i>S.B</i>	<i>E</i>
<i>a</i> ₁	<i>b</i> ₁	5	<i>b</i> ₁	3
<i>a</i> ₁	<i>b</i> ₂	6	<i>b</i> ₂	7
<i>a</i> ₂	<i>b</i> ₃	8	<i>b</i> ₃	10
<i>a</i> ₂	<i>b</i> ₃	8	<i>b</i> ₃	2

S

<i>B</i>	<i>E</i>
<i>b</i> ₁	3
<i>b</i> ₂	7
<i>b</i> ₃	10
<i>b</i> ₃	2
<i>b</i> ₅	2

两类常用连接运算 —— 自然连接

◆ 自然连接是在公共属性（组）上进行的等值连接

- 两个关系中必须具有公共属性（组）
- 在结果中把重复的属性列去掉

$R \bowtie S$

例

R

A	B	C
a_1	b_1	5
a_1	b_2	6
a_2	b_3	8
a_2	b_4	12

求自然连接 $R \bowtie S$

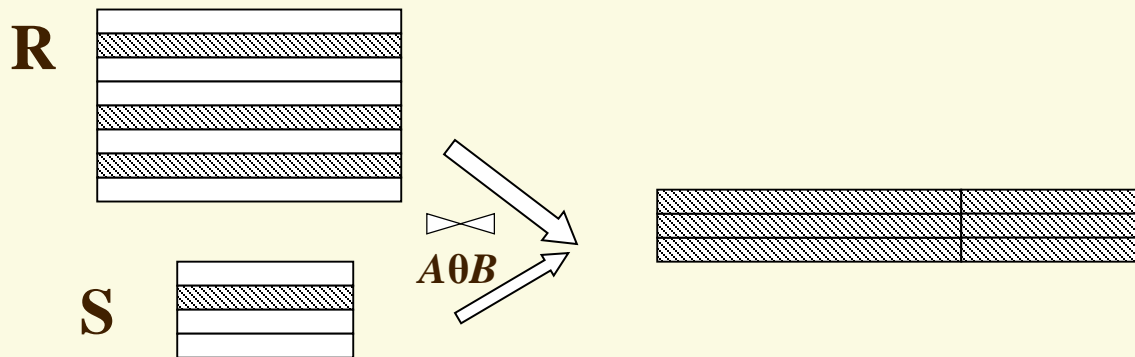
A	B	C	E
a_1	b_1	5	3
a_1	b_2	6	7
a_2	b_3	8	10
a_2	b_3	8	2

S

B	E
b_1	3
b_2	7
b_3	10
b_3	2
b_5	2

连接

- 一般的连接操作是从行的角度进行运算



- 自然连接还需要取消重复列，所以是同时从行和列的角度进行运算

实例: 关系R与S, 求 $R \bowtie_{C>D} S$, 以及R和S的等值连接 $R \bowtie_{C=D} S$

R

A	B	C
a1	b1	2
a1	b2	4
a2	b3	6
a2	b4	8

S

B	D
b1	5
b2	6
b3	7
b3	8

大于连接 ($C>D$)

A	R.B	C	S.B	D
a2	b3	6	b1	5
a2	b4	8	b1	5
a2	b4	8	b2	6
a2	b4	8	b3	7

等值连接 ($C=D$)

A	R.B	C	S.B	D
a2	b3	6	b2	6
a2	b4	8	b3	8

例：R与S，求 $R \bowtie S$ ，以及R和S的等值连接 $R \bowtie_{R.B=S.B} S$

R

A	B	C
a1	b1	2
a1	b2	4
a2	b3	6
a2	b4	8

S

B	D
b1	5
b2	6
b3	7
b3	8

等值连接 ($R.B=S.B$)

自然连接

A	R.B	C	S.B	D
a1	b1	2	b1	5
a1	b2	4	b2	6
a2	b3	6	b3	7
a2	b3	6	b3	8

例：R与S，求R \bowtie S, 以及R和S的等值连接R $\bowtie_{R.B=S.B}$ S

R			S	
A	B	C	B	C
a1	b1	2	b1	5
a1	b2	4	b2	4
a2	b3	6	b3	7
a2	b4	8	b3	8

等值连接(R.B=S.B)

自然连接

A	R.B	R.C	S.B	S.C				
a1	b1	2	b1	5				
a1	b2	4	b2	4				
a2	b3	6	b3	7				
a2	b3	6	b3	8				

关系代数

教学数据库中

用关系代数表达式表达:

$T(\underline{TID}, TNAME, TITLE)$

$C(\underline{CID}, \underline{CNAME}, \underline{TID})$

$S(\underline{SID}, SNAME, AGE, SEX)$

$SC(\underline{SID}, \underline{CID}, SCORE)$

1. 检索学习课程号为C2课程的学生学号与成绩

$\pi_{SID, SCORE}(\sigma_{CID='C2'}(SC)) \cap \pi_{1,3}(\sigma_{2='C2'}(SC))$

2. 检索学习课程号为C2课程的学生学号与姓名

$\pi_{SID, SNAME}(\sigma_{CID='C2'}(SC) \bowtie S)$

3. 检索至少选修LIU老师所授课程中一门课程的学生学号与姓名

$\pi_{SID, SNAME}(\sigma_{TNAME='LIU'}(T) \bowtie C \bowtie SC \bowtie S)$

关系代数运算的应用

T(TID,TNAME,TITLE)

C(CID,CNAME,TID)

S(SID,SNAME,AGE,SEX)

SC(SID,CID,SCORE)

4.检索选修课程号为C2或C4课程的学生学号

$\pi_{SID}(\sigma_{CID='C2' \vee CID='C4'}(SC))$

5. 检索选修课程号为C2和C4课程的学生学号

$\pi_1(\sigma_{2='C2'}(SC)) \cap \pi_1(\sigma_{2='C4'}(SC))$

6. 检索不学C2课程的学生姓名与年龄

$\pi_{SNAME,AGE}(\sigma_{CID \neq 'C2'}(SC) \bowtie S)$

选修关系SC

检索选修课程号为C2和C4课程的学生学号

SID	CID	SCORE
95001	C1	92
95001	C2	85
95001	C3	88
95001	C4	99
95002	C2	90
95002	C3	80

说明:

SID,CID,SCORE

属性分别表示

学号、课程号

和课程成绩

$\pi_{SID}(\sigma_{CID='C2' \wedge CID='C4'}(SC))$ 

④重命名 ρ

$\rho_{S(A_1, \dots, A_n)}(R)$

- 将关系R重命名为S
- 将关系S中的各属性命名为A1,A2...An

实例

Bars(

name,	addr
Joe's	Maple St.
Sue's	River Rd.

$\rho_{R(\text{bar}, \text{addr})}(\text{Bars})$

R(

bar,	addr
Joe's	Maple St.
Sue's	River Rd.

等价运算

$$R \cap S = R - (R - S)$$

$$R \bowtie_C S = \sigma_C (R \times S)$$

$$R \bowtie S = \pi_L (\sigma_C (R \times S))$$

关系运算符的优先级

1. 单目运算符优先级最高--- select, project
2. 笛卡尔积和连接运算符
3. 交
4. 并和差

括号的优先级最高

关系代数小结

— 关系代数运算

- 并、差、交
- 投影、选择
- 笛卡尔积、连接
- 重命名