

第二章 关系模型介绍

一、关系数据库的一些概念

- 关系数据库由**表(table)**的集合构成，每个表有唯一的名字
- 在关系模型的术语中，
 - **关系(relation)**：用来指代表
 - **元组(tuple)**：用于指代行
 - **属性(attribute)**：用于指代列
- **关系实例(relation instance)**：表示一个关系的特定实例，也就是所包含的一组特定的行
- **该属性的域**：对于关系的每个属性，都存在一个允许取值的集合
- **域是否是原子的**：看域中的元素是否被看作是**不可再分的单元**

- 数据库模式与数据库实例：
 - **数据库模式**：数据库的逻辑设计
 - **数据库实例**：给定时刻数据库中数据的一个快照
- **码(key)**
 - 一个关系中没有两个元组在所有属性上的取值都相同
 - **超码(super key)**：是一个或多个属性的集合，这些属性的组合可以使我们在一个关系中唯一地标识一个元组
 - **候选码(candidate key)**：最小超码
 - **主码(primary key)**：从候选码中挑选，主要用来在一个关系中区分不同元组的候选码；主码应该选择那些值从不或极少变化的属性
 - **外码**：一个关系模式r1可能在它的属性中包括另一个关系模式r2的主码。这个属性在r1上称作参照r2的外码。作用是对数据进行约束
 - **参照关系**：关系r1称为外码依赖的参照关系
 - **被参照关系**：r2叫做外码的被参照关系

二、基本的关系运算(查询)

- 关系代数基本运算有：选择、投影、并、集合差、笛卡尔积、更名

1.选择 σ

$\sigma_{dept_name="Physics"}(instructor)$

选择关系instructor中属于“物理(Physics)”系的那些元组

$\sigma_{salary>90000}(instructor)$

选择关系instructor中工资额大于90000美元的所有元组

可以使用的谓词：=, \neq , <, \leq , >, \geq

可以使用的连词： \wedge

ID	name	dept_name	salary
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

图.关系instructor

2. 投影运算 π

关系 r :

A	B	C
α	10	1
α	20	1
β	30	1
β	40	2

$\pi_{A,C}(r)$:

A	C
α	1
α	1
β	1
β	2

 $=$

A	C
α	1
β	1
β	2

只要A,C列

去掉重复行

3. 并运算 \cup

- $r \cup s = \{t \mid t \in r \text{ or } t \in s\}$

- 关系 r, s :

A	B
α	1
α	2
β	1

r

A	B
α	2
β	3

s

$r \cup s$:

A	B
α	1
α	2
β	1
β	3

- 条件:

要使 $r \cup s$ 有意义, 我们要求以下两个条件同时成立:

- 关系 r 和 s 必须是同元的, 即它们的属性数目必须相同;
- 对所有的 i , r 的第 i 个属性的域必须和 s 的第 i 个属性的域相同, 即对应属性的允许取值的集合相同

(该属性的域: 对于关系的每个属性, 都存在一个允许取值的集合)

- 注意: r 和 s 可以是数据库关系或者作为关系代数表达式结果的临时关系

4.差运算-

- $A-B$: 在A关系中而不在B关系中
- 关系 r,s :

A	B
α	1
α	2
β	1

r

A	B
α	2
β	3

s

$r-s$:

A	B
α	1
β	1

5.笛卡尔积 \times （当关系中有同名属性时，应使用更名运算进行换名）

关系 r,s :

A	B
α	1
β	2

r

C	D	E
α	10	a
β	10	a
β	20	b
γ	10	b

s

$r \times s$:

A	B	C	D	E
α	1	α	10	a
α	1	β	10	a
α	1	β	20	b
α	1	γ	10	b
β	2	α	10	a
β	2	β	10	a
β	2	β	20	b
β	2	γ	10	b

6. 更名 ρ

- $\rho_x(E)$: 返回表达式E的结果并把名字x赋给了它。
- 关系r:

A	B
α	1
β	2

r

$r \times \rho_s(r)$:

<i>r.A</i>	<i>r.B</i>	<i>s.A</i>	<i>s.B</i>
α	1	α	1
α	1	β	2
β	2	α	1
β	2	β	2

三、附加关系运算

- 附加关系运算有：集合交、自然连接、商集、赋值
- 附加关系运算不能增加关系代数的表达能力，但是可以使复杂查询的表达变得简单

1.集合交 \cap

- $r \cap s = \{ t \mid t \in r \text{ and } t \in s \}$

- 条件:**

要使 $r \cap s$ 有意义, 我们要求以下两个条件同时成立:

关系 r 和 s 必须是**同元的**, 即它们的**属性数目**必须相同;

对所有的 i , r 的第 i 个属性的域必须和 s 的第 i 个**属性的域**相同, 即**对应属性的允许取值的集合相同**

(该属性的域: 对于关系的每个属性, 都存在一个允许取值的集合)

- 注意: r 和 s 可以是**数据库关系**或者作为**关系代数表达式结果的临时关系**

- 关系 r , s :

A	B
α	1
α	2
β	1

r

A	B
α	2
β	3

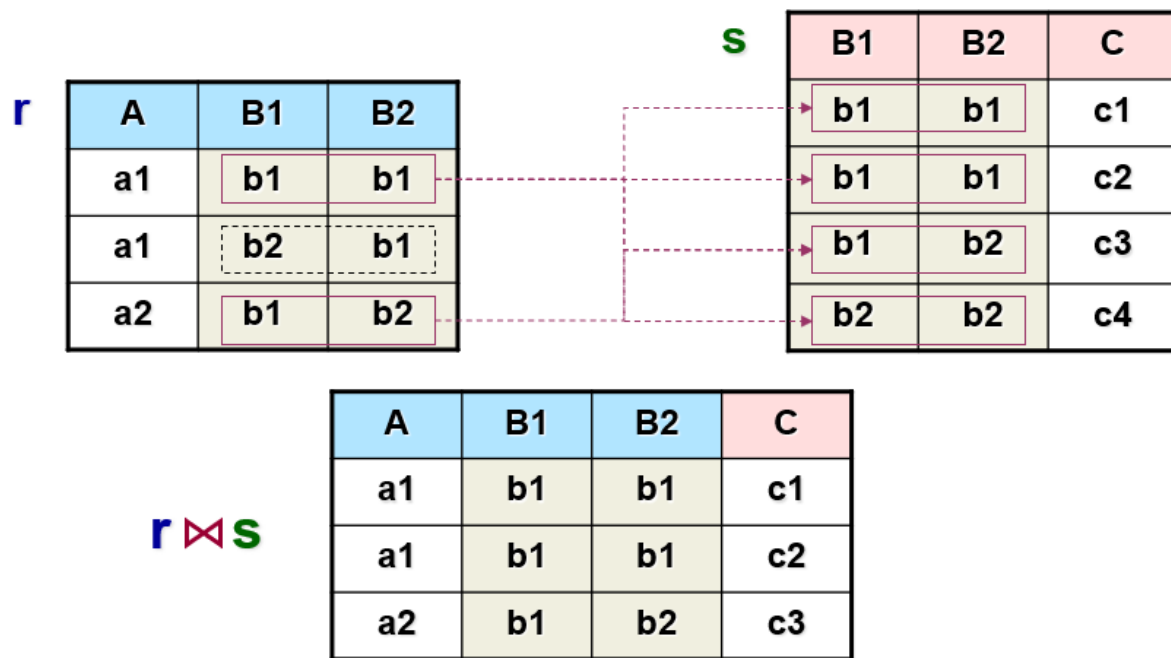
s

$r \cap s$:

A	B
α	2

- $r \cap s = r - (r - s)$

2.自然连接 ⋈



以左边为例,

$R=(A,B1,B2)$

$S=(B1,B2,C)$

结果的属性= $(A,B1,B2,C)$

$r \bowtie s = \pi_{r.A, r.B1, r.B2, s.C} (\sigma_{r.B1=s.B1 \wedge r.B2=s.B2} (r \times s))$

看关系r与s有什么相同的属性，然后用关系r里的值为蓝本在关系s中寻找是否有值全部相同的元组（可能有多），如果有，则将两个元组合并。

3. 商集 \div

关系 r 和 s :

A	B
α	1
α	2
α	3
β	1
γ	1
δ	1
δ	3
δ	4
ϵ	6
ϵ	1
β	2

r

B
1
2

s

S 是学生选课, B 是某老师开的课
 $S \div B$ 意为选了该老师所有课的学生

$r \div s$ 结果属性的集合 = $R - S$
 看 r 中完全包含关系 s 取值的项

$r \div s$:

A
α
β

4.赋值←

- 对关系代数查询而言，赋值必须是赋给一个临时关系变量。
- 对永久关系的赋值形成了对数据库的修改。

四、扩展的关系代数运算

- 扩展的关系代数运算有： 广义投影(`generalized-projection`)、 外连接运算(`outer-join`)

1. 广义投影 π

- $\pi_{F_1, F_2, \dots, F_n}(E)$
- 其中 F_1, F_2, \dots, F_n 可以是：一个属性或常量、可以使用 $+ - * /$ 等代数运算、还允许其他数据类型上的运算，比如对字符串的连接

2.外连接运算

- 类似于自然连接运算，不同之处在于它在结果中创建带空值的元组，以此来保留在连接中丢失的那些元组。
- 分类：
 - 左外连接 $\sqcup \bowtie$ ：取出左侧关系中所有于右侧关系的任一元组都不匹配的元组，用空值填充所有来自右侧关系的属性，再把产生的元组加到自然连接的结果中。（以左边的关系为蓝本，在右边找，如果有相同的直接连，如果没有相同的就添null再连接）
 - 右外连接 $\bowtie \sqcup$ ：与左外连接对称
 - 全外连接 $\sqcup \bowtie \sqcup$ ：既做左外连接又做右外连接

五、修改数据库

- 删除(deletion)、插入(insertion)、更新/修改(updating)
- 对于数据库中的修改操作，一次只能针对一个关系进行修改

1.删除(deletion)

- $r \leftarrow r - E$

2. 插入(insertion)

- $r \leftarrow r \cup E$

3.更新(updating)

- $r \leftarrow \Pi_{F_1, F_2, \dots, F_L}(r)$
- 针对一个关系中的某一个属性或某几个属性的值进行变动