

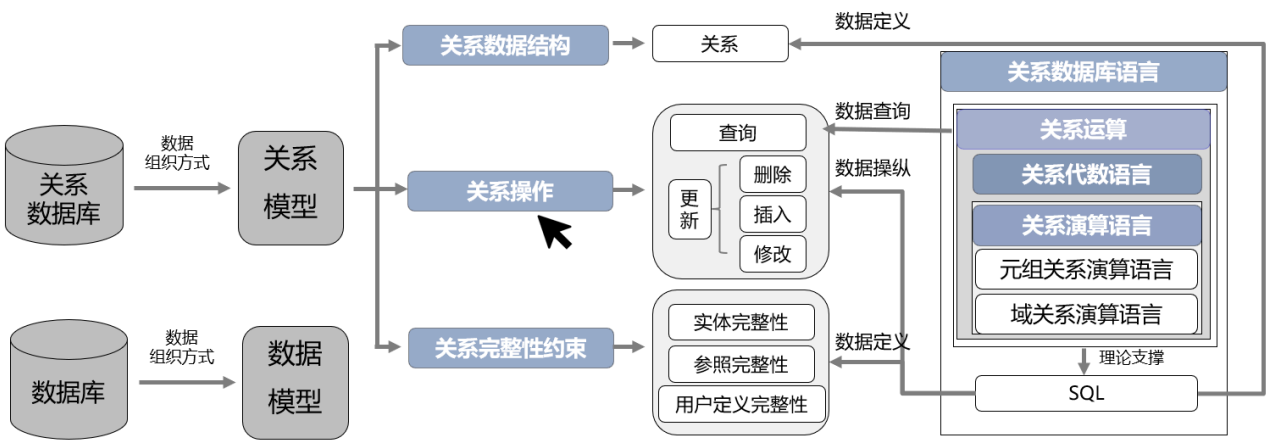
第四章 关系模型

串讲归纳

4.0 前言

有必要说明，这一章极为重要，一方面考点多，另一方面也是时至今日依旧最流行的数据库——关系型数据库。

其次，这一章主要介绍SQL查询的理论基础：关系代数运算。所以以题目、具体实例来深入学习练习最佳。



4.1 关系操作

关系数据结构主要是介绍关系中的各种名词概念，在第二章我们已经介绍过，如果仍不清楚，本章习题中依旧有解释。

如上图，常见的关系操作包括：查询和更新操作。更新则又包括删除、插入、修改。

关系数据库语言包括关系运算和SQL。关系运算包括关系代数【过程化查询语言】与关系演算【非过程化查询语言】，其中关系演算又分为元组关系演算和域关系演算两种。

而为了方便快捷地进行查询，我们使用结构化查询语言SQL进行数据操作处理，关系运算更多地是作为理论支撑。

4.2 关系完整性约束

实体完整性和参照完整性被称为关系的两个不变性，系统自动支持；用户定义的完整性当然结合用户所处领域自定义语义约束。

4.2.1 实体完整性约束

主要内容：每个元组都应唯一且可区分

如何确保？通过主键取值约束。

具体内容：若K为关系R的主键，则K不能去空值；若主键K为复合键，任一属性不能取空值；K不能重复。

4.2.2 参照完整性约束

主要内容：实体集之间关系的约束

如何确保？通过删除、插入规则。

具体解释：部门表与员工表两个关系，若删除一个部门，先要把员工表中该部门下所有员工删除；若插入一个新部门的员工，先要将新部门插入到部门表。

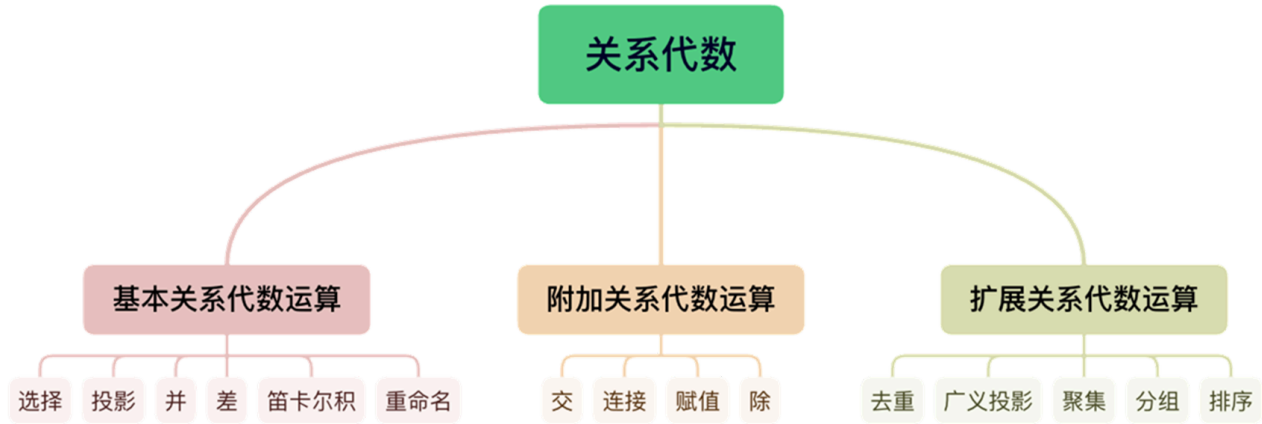
4.2.3 用户定义完整性约束

具体内容：用户对数据的特殊语义要求

通常通过其它约束实现：非空约束、唯一约束、自增长约束、默认值约束、检查约束。

其中，检查约束：即指定某一属性的取值范围，如年龄大于18。

4.3 关系运算



这部分就是关系代数运算部分，看着很难，但其实都是简单的集合运算（关系作为集合）。所以直接从例子入手，容易理解，但是要熟练掌握还需要不断练习做题、回顾。

- (1) (E_1) ; 关系

(2) $\sigma_p(E_1)$; 选择

(3) $\Pi_{A_1,A_2,\dots,A_k}(E_1)$; 投影

(4) $E_1 \cup E_2$; 交

(5) $E_1 - E_2$; 差

(6) $E_1 \times E_2$; 笛卡尔积

(7) $\rho_{S(A_1,A_2,\dots,A_n)}(E_1)$; 重命名
- (8) $E_1 \cap E_2$; 并

(9) $E_1 \bowtie_p E_2$; 连接

(10) $E_1 \div E_2$; 除

(11) $\delta(E_1)$; 去重

(12) $\Pi_{F_1,F_2,\dots,F_k}(E_1)$; 广义的投影

(13) $\mathcal{G}_{F_1(A_1),F_2(A_2),\dots,F_k(A_k)}(E_1)$; 聚集

(14) $\mathcal{G}_{G_1,G_2,\dots,G_l} \mathcal{G}_{F_1(A_1),F_2(A_2),\dots,F_k(A_k)}(E_1)$ 。分组聚集

一、基本关系代数运算

1. 选择 (Selection)

- 示例
 - 假设有一个“学生（学号，姓名，年龄，性别）”关系表：

学号	姓名	年龄	性别
1001	张三	20	男
1002	李四	22	女
1003	王五	20	男

- 选择年龄大于 20 岁的学生，关系代数表达式为
 $\sigma_{\text{年龄} > 20}(\text{学生})$

结果为：

学号	姓名	年龄	性别
1002	李四	22	女

2. 投影 (Projection)

- 示例
 - 对于上述“学生”关系表，若只想获取学生的学号和姓名，关系代数表达式为 $\pi_{\text{学号}, \text{姓名}}(\text{学生})$

结果为：

学号	姓名
1001	张三
1002	李四
1003	王五

3. 并 (Union)

- 示例
 - 假设有“参加社团 A 的学生（学号，姓名）”关系表 A 和“参加社团 B 的学生（学号，姓名）”关系表 B：

- 表 A:

学号	姓名
1001	张三
1003	王五

- 表 B:

学号	姓名
1001	张三
1004	孙六

- 计算 A 和 B 的并，关系代数表达式为 $A \cup B$

结果为：

学号	姓名
1001	张三
1003	王五
1004	孙六

4. 差 (Difference)

- 示例

- 使用上面的表 A 和表 B，计算 $A - B$ ，关系代数表达式为 $A - B$ 结果为：

学号	姓名
1003	王五

5. 笛卡尔积 (Cartesian Product)

- 示例
 - 假设有“课程（课程号，课程名）”关系表和“教师（教师号，教师姓名）”关系表：
 - 课程表：

课程号	课程名
C1	数据库
C2	操作系统

- 教师表：

教师号	教师姓名
T1	赵老师
T2	钱老师

- 计算这两个表的笛卡尔积，关系代数表达式为“课程 \times 教师”，结果为：

课程号	课程名	教师号	教师姓名
C1	数据库	T1	赵老师
C1	数据库	T2	钱老师
C2	操作系统	T1	赵老师
C2	操作系统	T2	钱老师

6. 重命名 (Rename)

- 示例
 - 假设有一个“员工（员工号，姓名，部门）”关系表，若要将其重命名为“职员”，可以用重命名操作，在不同的表示方法中有不同符号，假设用 $\rho_{\text{职员}}(\text{员工号}, \text{姓名}, \text{部门})(\text{员工})$ 表示职员的员工号、姓名、部门。

二、附加关系代数运算

1. 交 (Intersection)

- 示例
 - 使用前面的表 A 和表 B，计算 A 和 B 的交，关系代数表达式为

$A \cap B$

结果为：

学号	姓名
1001	张三

2. 连接 (Join)

- 等值连接示例
 - 假设有“学生选课（学号，课程号，成绩）”关系表和“课程（课程号，课程名，学分）”关系表：

- 学生选课表：

学号	课程号	成绩
1001	C1	80
1002	C2	70

- 课程表：

课程号	课程名	学分
C1	数据库	3
C2	操作系统	4

- 以课程号为连接条件进行等值连接，关系代数表达式为

$\sigma_{\text{学生选课.课程号}=\text{课程.课程号}}(\text{学生选课} \bowtie \text{课程})$

结果为：

学号	学生选课。课程号	成绩	课程名	学分
1001	C1	80	数据库	3
1002	C2	70	操作系统	4

- 自然连接示例
 - 对于上述“学生选课”和“课程”关系表进行自然连接，结果为：

学号	课程号	成绩	课程名	学分
1001	C1	80	数据库	3

学号	课程号	成绩	课程名	学分
1002	C2	70	操作系统	4

3. 赋值 (Assignment)

- 示例
 - 设关系 R 为“学生（学号，姓名，年龄）”，计算年龄大于 20 岁的学生，可先将选择操作的结果赋给一个临时关系变量 S，如 $S \leftarrow \sigma_{\text{年龄} > 20}(\text{学生})$ 。

4. 除 (Division)

- 示例
 - 假设有“学生选课（学生，课程）”关系表 R 和“必修课（课程）”关系表 S：

- R:

学生	课程
甲	语文
甲	数学
甲	英语
乙	语文
乙	数学

- S:

课程
语文
数学

- 计算 $R \div S$

结果为选了所有必修课的学生名单，即：

学生
甲
乙

三、扩展关系代数运算

1. 去重 (Distinct)

- 示例

- 假设有一个“成绩（学号，课程号，分数）”关系表：

学号	课程号	分数
1001	C1	80
1001	C1	80
1002	C2	70

- 使用去重操作，在 SQL 中用“SELECT DISTINCT”实现，结果为：

学号	课程号	分数
1001	C1	80
1002	C2	70

2. 广义投影 (Generalized Projection)

- 示例

- 对于上述“学生”关系表，若要计算学生年龄加 1 后的结果并投影出学号、姓名和新年龄，关系代数表达式为

$\pi_{\text{学号}, \text{姓名}, \text{年龄}+1}(\text{学生})$

结果（假设）为：

学号	姓名	新年龄
1001	张三	21
1002	李四	23
1003	王五	21

3. 聚集 (Aggregation)

- 示例

- 对于上述“成绩”关系表，计算所有学生成绩的平均分，

$\gamma_{\text{AVG}(\text{分数})}(\text{成绩})$

结果（假设平均分为 75）为：

AVG (分数)
75

4. 分组 (Grouping)

- 示例

- 假设有“销售（销售员，产品，销售额）”关系表：

销售员	产品	销售额
A	P1	1000
A	P2	1500
B	P1	800

- 按销售员分组计算销售额总和， $\gamma_{\text{销售员}} \text{SUM(销售额)}(\text{销售})$
结果为：

销售员	SUM (销售额)
A	2500
B	800

5. 排序 (Sorting)

- 示例

- 对于上述“学生”关系表，按年龄从小到大排序， $\tau_{\text{年龄}}(\text{学生})$
结果为：

学号	姓名	年龄	性别
1001	张三	20	男
1003	王五	20	男
1002	李四	22	女

Important

关系代数中，传统的集合运算【并，差，交和笛卡尔积】是从行的角度进行的，而专门的关系运算【选择、投影、连接、除运算】是从列的角度进行的。

关系代数的五个基本操作是【并、差、笛卡尔积、选择和投影】。

题库讲解

1 【单选题】根据参照完整性规则,对于基本关系R中的每个元组在外码上的值必须是()

A、基本关系S中某个元组的主码值 B、A或C C、空值 D、A和B

我的答案：B

💡 Tip

举个例子，有“学生（学号，姓名，班级编号）”和“班级（班级编号，班级名称）”两个关系。“学生”关系中的“班级编号”是外码，参照“班级”关系中的“班级编号”（主码）。如果一个新入学的学生还没有分配班级，那么该学生记录中的“班级编号”外码就可以为空值。当为学生分配了班级后，“学生”关系中的“班级编号”外码的值必须是“班级”关系中已经存在的某个“班级编号”的值。

2【单选题】建立学生表时,限定性别字段必须是男或女是实现数据的()。

A、以上都不是 B、域完整性 C、实体完整性 D、参照完整性

我的答案：B 域即范围

3【单选题】有两个关系R(A,B,C)和S(B,C,D),将R和S进行自然连接,得到的结果包含()列

A、3 B、6 C、5 D、*4

我的答案：D

4【单选题】设关系R、S、W各有10个元组，那么这三个关系的笛卡尔积的元组个数是()。

A、1000 B、30 C、不确定 D、10

我的答案：A

$$10 * 10 * 10$$

5【单选题】若一个关系有多个候选码,则选定其中一个为()

A、主属性 B、主码 C、全码 D、主候选码

我的答案：B

6【单选题】关系运算中花费时间可能最长的运算是()。

A、笛卡儿积 B、选择 C、除 D、投影

我的答案：A 遍历组合

7【单选题】有两个关系 R(A,B,C)和S(B,C,D),则R÷S结果的属性个数是()

A、2B、3C、1D、不一定

我的答案：C

8【单选题】从关系中挑选出指定的属性组成新关系的运算称为()

A、“连接”运算B、“投影”运算C、“选取”运算D、“交”运算

我的答案：B

9【单选题】已知关系课程表C，用关系代数查询“程军”老师所授课程的课程号(C#)和课程名(Cname)，正确的是__。

C#	Cname	Cteacher
A01	数据库	张帆
A02	英语	程军
A03	操作系统	王昂
A04	Python	郑磊

- A、 $\Pi_{C\# \text{ and } Cname}(\sigma_{Cteacher="程军"}(C))$
- B、 $\Pi_{C\#, Cname, Cteacher="程军"}(C)$
- C、 $\Pi_{C\# \text{ and } Cname \text{ and } Cteacher="程军"}(C)$
- D、 $\Pi_{C\#, Cname}(\sigma_{Cteacher="程军"}(C))$

我的答案：D 上面的符号应该是选择

10【单选题】在一个关系中不能相同的是

A、域B、分量C、属性D、记录

我的答案：D

11【单选题】若关系中的某一属性组的值能唯一的标识一个元组,而其子集不能,则称该属性组为()。

A、外码B、全码C、主码D、候选码

我的答案：D

Important

再强调一遍吧，必考点。

首先给出一个额外的定义辅助理解：超键是在关系模式中能够唯一标识元组的属性或属性组。

接下来给出选项的理解：一个关系中，可能有多个超键能够唯一标识元组，其中不含多余属性的超键就是候选码，从众多候选码中可以挑选一个作为主码，当只有所有属性组合在一起才能唯一标识元组时，这些属性叫全码。

12【单选题】对关系R进行投影运算后,得到关系S,则()

A、关系R的元组数大于关系S的元组数 B、关系R的元组数大于或等于关系S的元组数 C、关系R的元组数等于关系S的元组数 D、关系R的元组数小于或等于关系S的元组数

我的答案：B

13【单选题】从关系中挑选出指定的属性组成新关系的运算称为()。

A、“交”运算 B、“投影”运算 C、“连接”运算 D、“选取”运算

我的答案：B

14【单选题】设关系R与关系S具有相同的目数(或称度数)，且相对应属性的值取自同一个域，则 $R - (R \cap S)$ 等于_____。

A、 $S - B$ 、 $R \cap S$ C、 $S - R$ D、 $R \cup S$

我的答案：B

15【单选题】下面关于关系数据库特点的描述，错误的是()

- A、由于关系数据库将数据存储和数据表中，数据操作的瓶颈出现在多张数据表的操作中，而且数据表越多这个问题越严重
- B、为了规范化数据、减少重复数据以及充分利用好存储空间，把数据按照最小关系表的形式进行存储
- C、采用表格的储存方式，数据以行和列的方式进行存储，要读取和查询都十分方便
- D、关系数据库采用非结构化查询语言来对数据库进行查询

我的答案：D 结构化查询语言——SQL

16【单选题】在关系R(R#,RN,S#)和S(S#,SN,SD)中,R的主键是R#,S的主键是S#,则S#在R中称为

A、外键B、主键C、候选键D、以上都不是

我的答案：A

17【单选题】下列不属于传统集合运算的是()

A、交B、笛卡尔积C、并D、连接

我的答案：D

Important

按照关系代数运算划分：

扩展的有：去重、广义投影、聚集、分组、排序；

附加的有：交、连接、赋值、除；

基本的：选择、投影、并、差、笛卡尔积、重命名。

按照集合运算划分：传统的有交、并、差、笛卡尔积。

18【单选题】专门的关系运算包括选择,投影,连接,除运算等。其中从行的角度进行的运算是()
()

A、选择B、投影C、连接D、差

我的答案：A

Important

从运算角度来讲：

行的角度：选择、去重、笛卡尔积、交、并

列的角度：投影、连接、除运算等

19【单选题】如下列表所示，两个关系R1和R2，它们进行_运算后可以得到R3。

关系R1

A	B	C
1	1	x
C	2	y
D	3	y

关系R2

B	E	M
1	m	i
2	n	j
1	m	k

关系R3

A	B	C	E	M
1	1	x	m	i
c	2	y	n	J
1	1	x	m	K

- A、 $R1 \bowtie R2$
- B、 $R1 \times R2$
- C、 $R1 \cup R2$
- D、 $R1 \cap R2$

我的答案：A

20【单选题】自然连接是构成新关系的有效方法。一般情况下,当对关系R和S使用自然连接时,要求R和S含有一个或多个共有的()

A、行 B、元组 C、属性 D、记录

我的答案：C 上一题就是

21【单选题】SQL语言通常称为()。

A、结构化控制语言 B、结构化操纵语言 C、结构化定义语言 D、结构化查询语言

我的答案：D

22【单选题】在通常情况下,下面的关系中不可以作为关系数据库的关系是()。

A、R3(学生号,学生名,宿舍号) B、R1(学生号,学生名,性别) C、R4(学生号,学生名,简历) D、R2(学生号,学生名,班级号)

我的答案：C 简历???

23【单选题】关系代数的5种基本基本运算是()

A、并、差、交、选择、笛卡尔积 B、并、差、选择、投影、笛卡尔积 C、并、差、交、选择、投影 D、并、差、选择、投影、自然连接

我的答案：B

24【单选题】下列传统的集合运算中是复合运算的是() (

A、交B、并C、差D、笛卡尔积

我的答案：A

💡 Tip

复合运算指的是由多个基本运算组合而成的运算。

传统的中只有交运算等价于：

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

25【单选题】关系代数运算是以()为基础的运算。

A、代数运算B、集合运算C、关系运算D、谓词演算

我的答案：B

26【单选题】关系数据库的型也称为关系数据库模式,是对()的描述

A、关系数据库B、数据C、数据库D、数据库管理系统

我的答案：A

27【单选题】对关系R进行投影运算后,得到关系S,则()

A、关系R的元组数小于或等于关系S的元组数

B、[关系R的元组数大于关系S的元组数

C、关系R的元组数等于关系S的元组数

D、*关系R的元组数大于或等于关系S的元组数

我的答案：D

28【单选题】参加差运算的两个关系()

A、属性个数可以不相同B、属性个数必须相同C、一个关系包含另一个关系的属性D、属性名必须相同

我的答案：B

29【填空题】学生表如图所示，用关系代数查询所有在3系就读的且年龄小于21的学生的学号和姓名__ $\Pi_{S\#,Sname}(\sigma_{D\#="03" \wedge Sage < 21}(R))$ 。

S#	Sname	Ssex	Sage	D#	Sclass
98030101	张三	男	20	03	980301
98030102	张四	女	21	03	980301
98030103	张五	男	19	03	980301
98040201	王三	男	18	04	980402
98040202	王四	男	21	04	980402
98050104	孙六	女	19	05	980501

30.图片中的运算结果是（）

teacher					Teach				Course	
T#	Tname	Salary			T#	C#			C#	Cname
001	赵三	1200.00			001	001			001	数学
002	赵四	1400.00			002	002			002	物理
003	赵五	1000.00			004	002			003	化学
004	赵六	1100.00								

T#	TNAME	SALARY	C#	CNAME
001	赵三	1200.00	001	数学
002	赵四	1400.00	002	物理
004	赵六	1100.00	002	物理
null	null	null	003	化学

31.【填空题】已知关系：学生（学号，姓名，性别，年龄）， 课程（课程号，课程名，学分），选课（课程号，学号，成绩）。表达“查询年龄未满20岁的学生姓名及年龄”的关系代数表达式是__

$$\Pi_{姓名,年龄}(\sigma_{年龄 < 20}(学生))$$

“学友”工作室出品

更多学科复习资料、竞赛学习等内容

欢迎关注华理“以学会友ing”公众号