

学生 (学号, 姓名, 性别, 年龄, 院系)

课程 (课程号, 课程名, 先行课, 学分)

学习 (课程号, 学号, 成绩)

5. 查询选修课程号为1的学生姓名和成绩

$\pi_{\text{姓名, 成绩}} \left(\sigma_{\text{课程号}='1'} \left(\text{学生} \bowtie \text{学习} \right) \right)$

$\pi_{\text{姓名, 成绩}} \left(\text{学生} \bowtie \sigma_{\text{课程号}='1'} \left(\text{学习} \right) \right)$

4 查询同时选修了课程号为1和2的学生的学号学习 (课程号, 学号, 成绩)

方法1: 集合交运算

$$A = \pi_{\text{学号}} (\sigma_{\text{课程号}='1'} (\text{学习}))$$

$$B = \pi_{\text{学号}} (\sigma_{\text{课程号}='2'} (\text{学习}))$$

$$P = A \cap B$$



敲黑板，重点到了

方法2: 采用“自连接”运算

$$\sigma_{i \neq (n+j)} (R \times S)$$

$$\pi_{[2]} (\sigma_{[1]='1' \wedge [4]='2' \wedge [2]=[5]} (\text{学习} \times \text{学习}))$$

查询同时选修了所有课程的学生学号

学习 (课程号, 学号, 成绩)



需要一种能够表示“包含”意义的关系代数

A: 每个学生选课的集合

B: 所有课程的集合

满足条件“A包含B”的学生学号即为所求

5 除： 给定关系 $R(X,Y)$ 和 $S(Y)$ ，其中 X,Y 为属性（组）。 R 除以 S 得到新关系 $P(X)$ ， P 是 R 中满足下列条件的元组在 X 属性列上的投影：

元组在 X 上的分量值 x 的象集 Y_x 包含 S 在 Y 上的投影的集合。记作：

$$R \div S = \{t_r[X] \mid t_r \in R \wedge \pi_y(S) \subseteq Y_x\}$$

其中， Y_x 是分量 x 在 R 中 Y 上的象集。

R

A	B
A1	B1
A1	B2
A2	B1
A3	B2
A1	B3
A3	B1

X

Y

S

B
B1
B2

Y

R ÷ S

A
A1
A3

X

① Y_x X的分量 $x=\{A1, A2, A3\}$ $Y_{A1}=\{B1,B2,B3\}$ $Y_{A2}=\{B1\}$ $Y_{A3}=\{B1,B2\}$ ② $\pi_y(S)$ $=\{B1,B2\}$ ③ $\pi_y(S) \subseteq Y_x$ $Y_{A1} \ Y_{A3}$ $x=\{A1, A3\}$ ④ $R \div S = \{A1, A3\}$

除法的运算
过程

查询哪些**供应商**的零件(同时) 存放在**1,3,5**三个仓库中?

R

仓库号	供应商号
WH1	S1
WH1	S2
WH1	S3
WH2	S3
WH3	S1
WH3	S2
WH5	S1
WH5	S2
WH5	S4
WH6	S1
WH6	S2

S

仓库号
WH1
WH3
WH5

$R \div S$

供应商号
S1
S2

$R(X,Y)$ 和 $S(Y)$ 如何对应?

除法解决带有“包含”关系的问题



敲黑板，重点到了

查询选修了所有课程的学生学号

需要一种能够表示“包含”意义的关系代数

A: 每个学生选课的集合

$$A = \pi_{\text{学号}, \text{课程号}} (\text{学习})$$

B: 所有课程的集合

$$B = \pi_{\text{课程号}} (\text{课程})$$

满足条件“A包含B”的学生的学号即为所求

$$A \div B = \pi_{\text{学号}, \text{课程号}} (\text{学习}) \div \pi_{\text{课程号}} (\text{课程})$$

查询至少选修1号同学选修的所有课程的学生姓名

1号同学选修的所有课程的编号

$$S = \pi_{\text{课程号}} (\sigma_{\text{学号}=1} (\text{学习}))$$

选修课程的学生姓名

$$R = \pi_{\text{课程号}, \text{姓名}} (\text{学生} \bowtie \text{学习})$$

$$\text{结果} = R \div S$$

除法运算的基本代数表示

$$T = \pi_x(R)$$

$$R \div S = T - \pi_x(T \times S - R)$$

R

仓库号	供应商号
WH1	S1
WH1	S2
WH1	S3
WH2	S3
WH3	S1
WH3	S2
WH5	S1
WH5	S2
WH5	S4
WH6	S2

S

仓库号
WH1
WH3
WH5

$$T = \pi_x(R)$$
$$R \div S = T - \pi_x(T \times S - R)$$

仓库号	供应商号
WH1	S1
WH3	S1
WH5	S1
WH1	S2
WH3	S2
WH5	S2
WH1	S3
WH3	S3
WH5	S3
WH1	S4
WH3	S4
WH5	S4

$T = \Pi_{\text{供应商号}}(R) = \{S1, S2, S3, S4\}$

$T \times S = \Pi_{\text{供应商号}}(R) \times S$

R

仓库号	供应商号
WH1	S1
WH1	S2
WH1	S3
WH2	S3
WH3	S1
WH3	S2
WH5	S1
WH5	S2
WH5	S4
WH6	S2

$$T = \pi_x(R)$$

$$R \div S = T - \pi_x(T \times S - R)$$

仓库号	供应商号
WH1	S1
WH3	S1
WH5	S1
WH1	S2
WH3	S2
WH5	S2
WH1	S3
WH3	S3
WH5	S3
WH1	S4
WH3	S4
WH5	S4

$$\pi_{\text{供应商号}}(\pi_{\text{供应商号}}(R) \times S - R)$$

仓库号	供应商号
WH3	S3
WH5	S3
WH1	S4
WH3	S4

供应商号
S3
S4

S

仓库号
WH1
WH3
WH5

R

仓库号	供应商号
WH1	S1
WH1	S2
WH1	S3
WH2	S3
WH3	S1
WH3	S2
WH5	S1
WH5	S2
WH5	S4
WH6	S2

$$T = \pi_x(R)$$

$$R \div S = T - \pi_x(T \times S - R)$$

供应商号
S3
S4

$$T - \Pi_{\text{供应商号}}(\Pi_{\text{供应商号}}(R) \times S - R)$$

$$T = \Pi_{\text{供应商号}}(R) = \{S1, S2, S3, S4\}$$

S

仓库号
WH1
WH3
WH5

供应商号
S1
S2

对九种关系代数进行总结

运算	单/双目	基本运算	复合运算	表示方法
并	双目	\vee		$\mathbf{R \cup S}$
差	双目	$\sqrt{}$		$\mathbf{R-S}$
交	双目		差	$\mathbf{R \cap S}$
笛卡儿积	双目	\vee		$\mathbf{R \times S}$
选择	单目	\vee		$\sigma_{\mathbf{F}}(\mathbf{R})$
投影	单目	\vee		$\pi_{\mathbf{A}}(\mathbf{R})$
连接	双目		笛卡儿积、选择	$\mathbf{R \bowtie_{A \theta B} S}$
自然连接	双目		笛卡儿积、选择、投影	$\mathbf{R \Join S}$
除	双目		笛卡儿积、投影、差	$\mathbf{R \div S}$

关系代数解题过程

- 1 确定题目的结果包含哪些属性，一般用投影运算完成。
- 2 找出题目的检索条件，用选择运算来完成。
- 3 确定结果属性和检索条件相关的关系有哪些。如果关系不只一个，用自然连接运算实现。
- 4 出现“全部”“至少”“包含”的词语时，一般用到除法运算。
- 5 出现“没有”“不”等否定词语时，一般用差运算

设数据库中有三个基本表：

S (SNo(学号), SName(姓名), SSex(性别), SPro(专业方向))

SC (SNo(学号), CNo(课程号), Grade(成绩))

C (CNo(课程号), CName(课程名), CPre(先行课), CCredit(学分))

- (1) 找出选修网络方向女同学名单
- (2) 求选修15164课程的学生姓名和专业方向
- (3) 求选修了数据库原理课程的学生姓名
- (4) 同时选修人工智能及编译技术的学生姓名
- (5) 没有被任何人选修的课程名
- (6) 没有选修任何课程的学生性别和姓名
- (7) 至少选修了002号学生选修的全部课程的学生学号
- (8) 查询课程被选修的情况，列出课程号、课程名、先行课、学分、学号和成绩
- (9) 查询学生没有选修的课程，列出学号、课程号

查询选修了1号同学选修的课程的学生的姓名

查询至少选修1号同学选修的所有课程的学生姓名



区别?

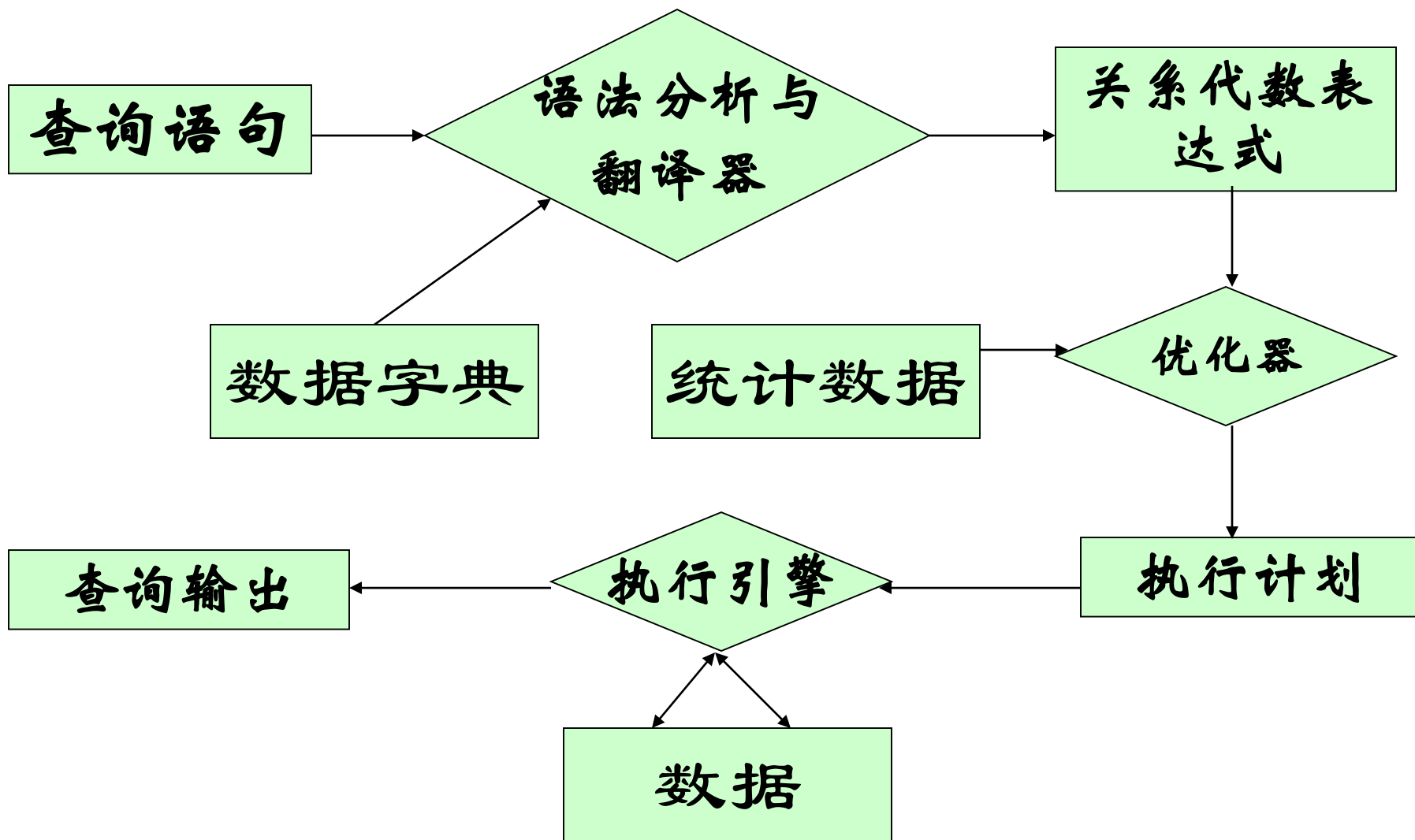
2.4 查询优化

查询优化:

查询优化的组织

查询优化策略、算法

2.4.1 查询优化的必要性



查询开销:

集中式: 总代价=I/O代价+CPU代价

多用户: 总代价=I/O代价+CPU代价+内存代价

关系代数为什么要进行优化?



例: 求选修了2号课程的学生姓名。

$Q1 = \pi_{Sname}(\sigma_{S.Sno=SC.Sno \wedge SC.Cno='2'}(S \times SC))$

$Q2 = \pi_{Sname}(\sigma_{SC.Cno='2'}(S \bowtie SC))$

$Q3 = \pi_{Sname}(S \bowtie \sigma_{SC.Cno='2'}(SC))$

关系代数等价变换规则

1) 连接、笛卡尔积交换律

$$E_1 \times E_2 \equiv E_2 \times E_1$$

$$E_1 \underset{F}{\bowtie} E_2 \equiv E_2 \underset{F}{\bowtie} E_1$$

$$E_1 \underset{F}{\bowtie} E_2 \equiv E_2 \underset{F}{\bowtie} E_1$$

代数优化

2) 连接、笛卡尔积结合律

$$(E_1 \times E_2) \times E_3 \equiv E_1 \times (E_2 \times E_3)$$

$$(E_1 \underset{F}{\bowtie} E_2) \underset{F}{\bowtie} E_3 \equiv E_1 \underset{F}{\bowtie} (E_2 \underset{F}{\bowtie} E_3)$$

$$(E_1 \underset{F1}{\bowtie} E_2) \underset{F2}{\bowtie} E_3 \equiv E_1 \underset{F1}{\bowtie} (E_2 \underset{F2}{\bowtie} E_3)$$

3、投影的串接定律

$$\pi_{A1,A2,\dots,A_n}(\pi_{B1,B2,\dots,B_m}(E)) \equiv \pi_{A1,A2,\dots,A_n}(E)$$

4、选择的串接定律

$$\sigma_{F1 \wedge F2}(E) \equiv \sigma_{F1}(\sigma_{F2}(E))$$

5、选择与投影的交换律

$$\sigma_F(\pi_{A1,A2,\dots,A_n}(E)) \equiv \pi_{A1,A2,\dots,A_n}(\sigma_F(E)) \quad (F \text{ 只涉及 } A1,A2,\dots,A_n)$$

$$\pi_{A1,A2,\dots,A_n}(\sigma_F(E)) \equiv \pi_{A1,A2,\dots,A_n}(\sigma_F(\pi_{A1,\dots,A_n,B1,\dots,B_m}(E)))$$

6、选择与笛卡尔积的交换律

$$\sigma_F(E_1 \times E_2) \equiv \sigma_F(E_1) \times E_2$$

$$\sigma_F(E_1 \times E_2) \equiv \sigma_{F1}(E_1) \times \sigma_{F2}(E_2)$$

$$\sigma_F(E_1 \times E_2) \equiv \sigma_{F2}(\sigma_{F1}(E_1) \times E_2)$$

7、选择与并的分配率

$$\sigma_F(E_1 \cup E_2) \equiv \sigma_F(E_1) \cup \sigma_F(E_2)$$

8、选择与差的分配率

$$\sigma_F(E_1 - E_2) \equiv \sigma_F(E_1) - \sigma_F(E_2)$$

9、投影与笛卡尔积的分配律

$$\pi_{A1,A2,\dots,A_n,B1,B2,\dots,B_m}(E_1 \times E_2) \equiv \pi_{A1,A2,\dots,A_n}(E_1) \\ \times \pi_{B1,B2,\dots,B_m}(E_2)$$

10、投影与并的分配率

$$\pi_{A1,A2,\dots,A_n}(E_1 \cup E_2) \equiv \pi_{A1,A2,\dots,A_n}(E_1) \cup \\ \pi_{A1,A2,\dots,A_n}(E_2)$$

2.4.2 查询优化的策略和算法

- 1、尽可能早地执行选择操作（减少中间运算结果）
- 2、同时进行投影和选择运算
- 3、把投影同其前或后的双目运算结合起来
- 4、合并选择与笛卡尔积组成一个连接运算
- 5、寻找公共子表达式

2、关系代数表达式的优化算法

输入：一个关系表达式的语法树

输出：计算该表达式的程序

方法：

1)把 $\sigma_{F1 \wedge F2 \wedge \dots \wedge F_n}(E)$ 变换为

$$\sigma_{F1}(\sigma_{F2}(\dots(\sigma_{F_n}(E))\dots))$$

[规则4]

2)对每一个选择尽可能把它移到树的叶端。 [规则4-8]

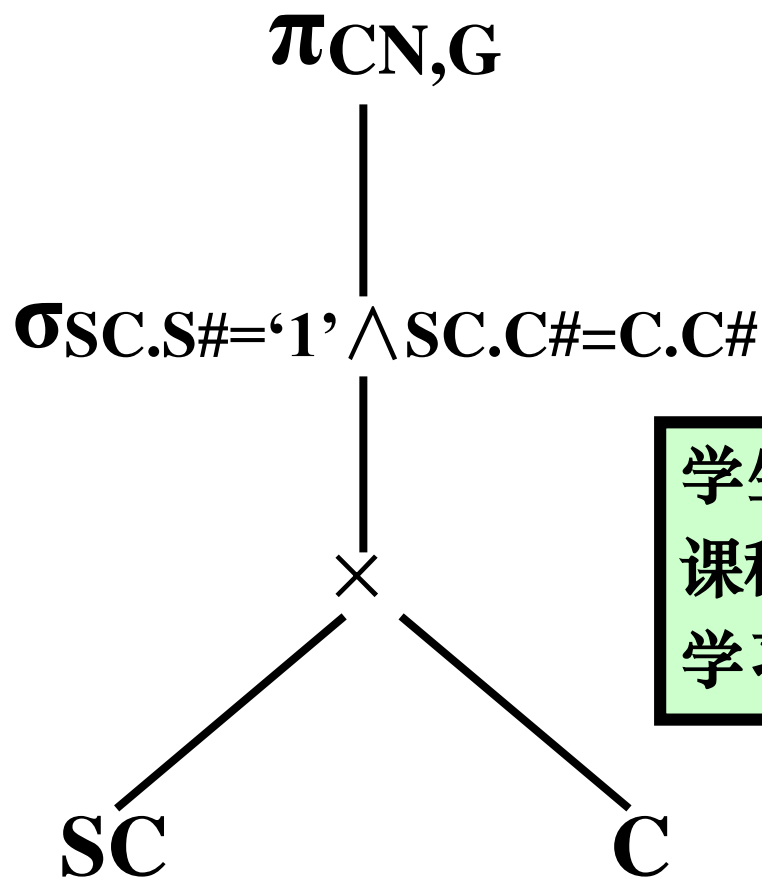
3)对每一个投影尽可能把它移到树的叶端。 [规则3,5,9,10]

4)合并选择和投影或一个选择后跟一个投影。 [规则3-5]

5)将得到的语法树的内节点分组。(每一双目运算和它所有的直接祖先为一组。

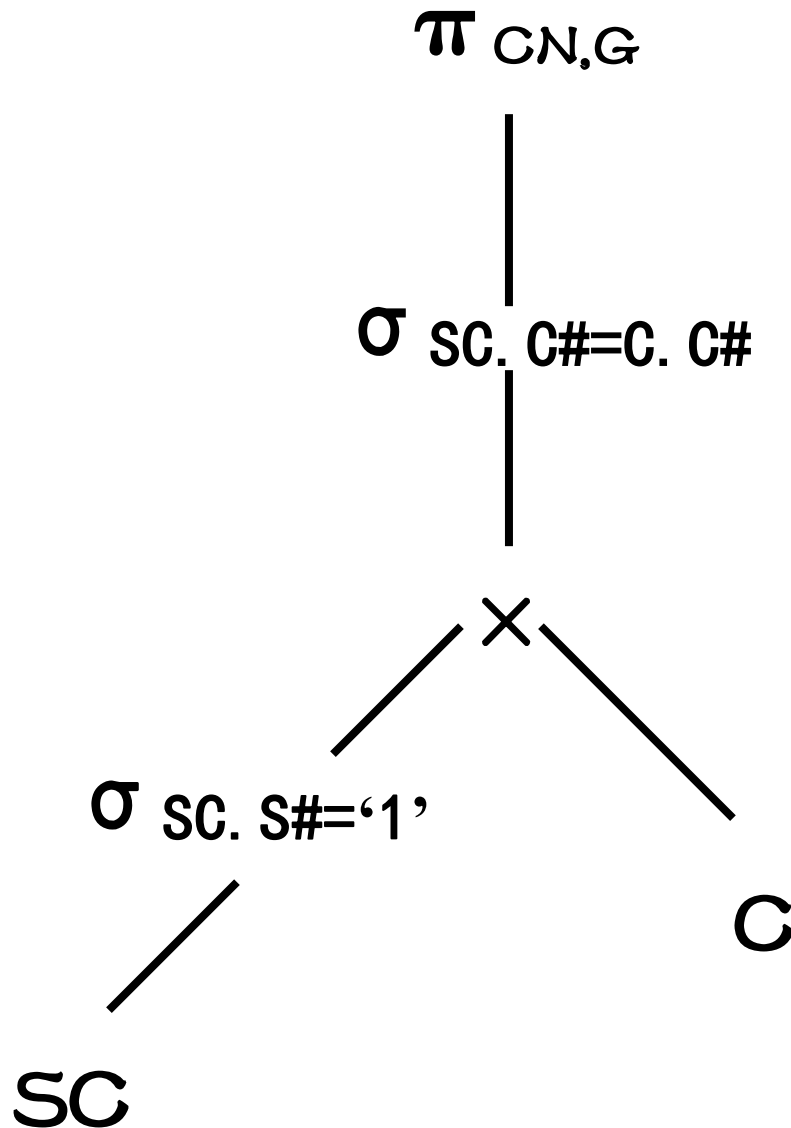
例：求1号学生所选修的课程名及成绩

$\pi_{CN,G} \left(\sigma_{SC.S\#='1' \wedge SC.C\#=C.C\#} (SC \times C) \right)$

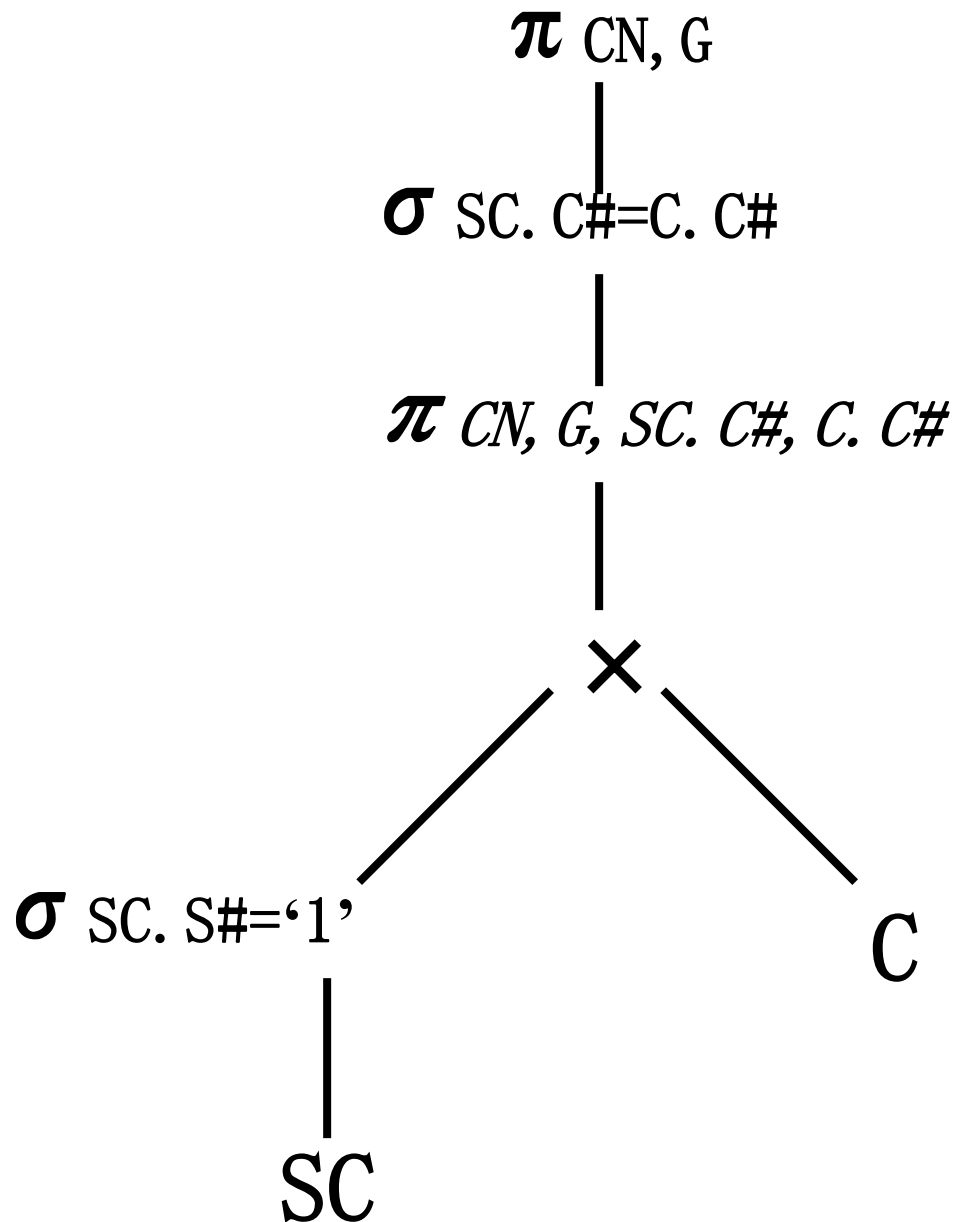


学生（学号，姓名，性别，年龄，院系）
课程（课程号，课程名，先行课，学分）
学习（课程号，学号，成绩）

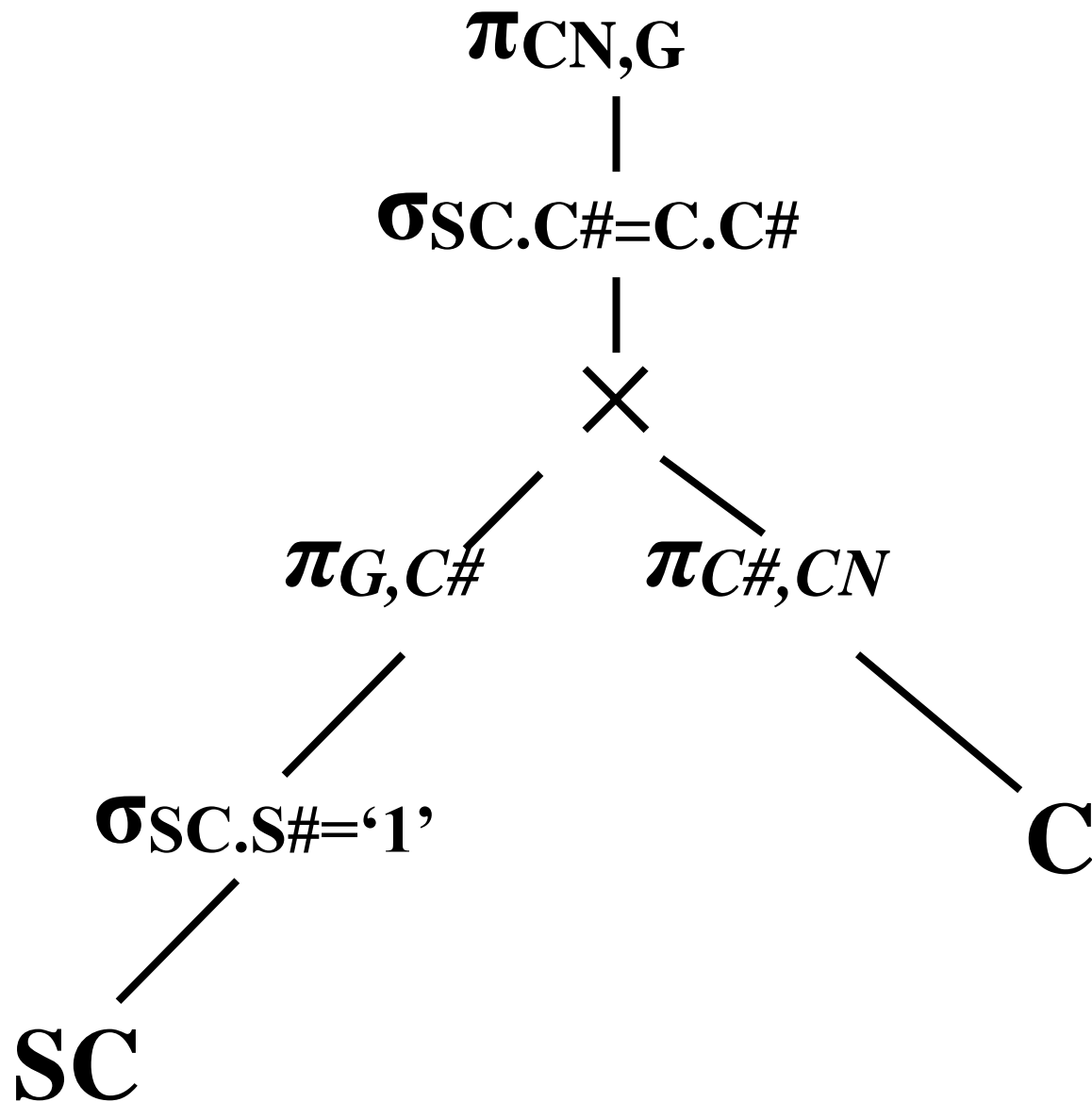
说明： 1.由规则4,6， 可以把选择 $\sigma_{SC.S\#='1'}$ 下移至SC之上



2. 由规则5， 选择与投影交换



3.由规则9， 可以把投影 $\pi_{CN,G,SC.C\#,C.C\#}$ 下移



例：查询选修了数据库原理的学生姓名和成绩

$$\pi_{SN,G} \left(\sigma_{CN='数据库原理'} (S \bowtie SC \bowtie C) \right)$$
$$\pi_{SN,G} \left(\sigma_{CN='数据库原理' \wedge S.S\#=SC.S\# \wedge SC.C\#=C.C\#} (S \times SC \times C) \right)$$

$\Pi_{SN,G}$

$\sigma_{CN='数据库原理' \wedge S.S\#=SC.S\# \wedge SC.C\#=C.C\#}$

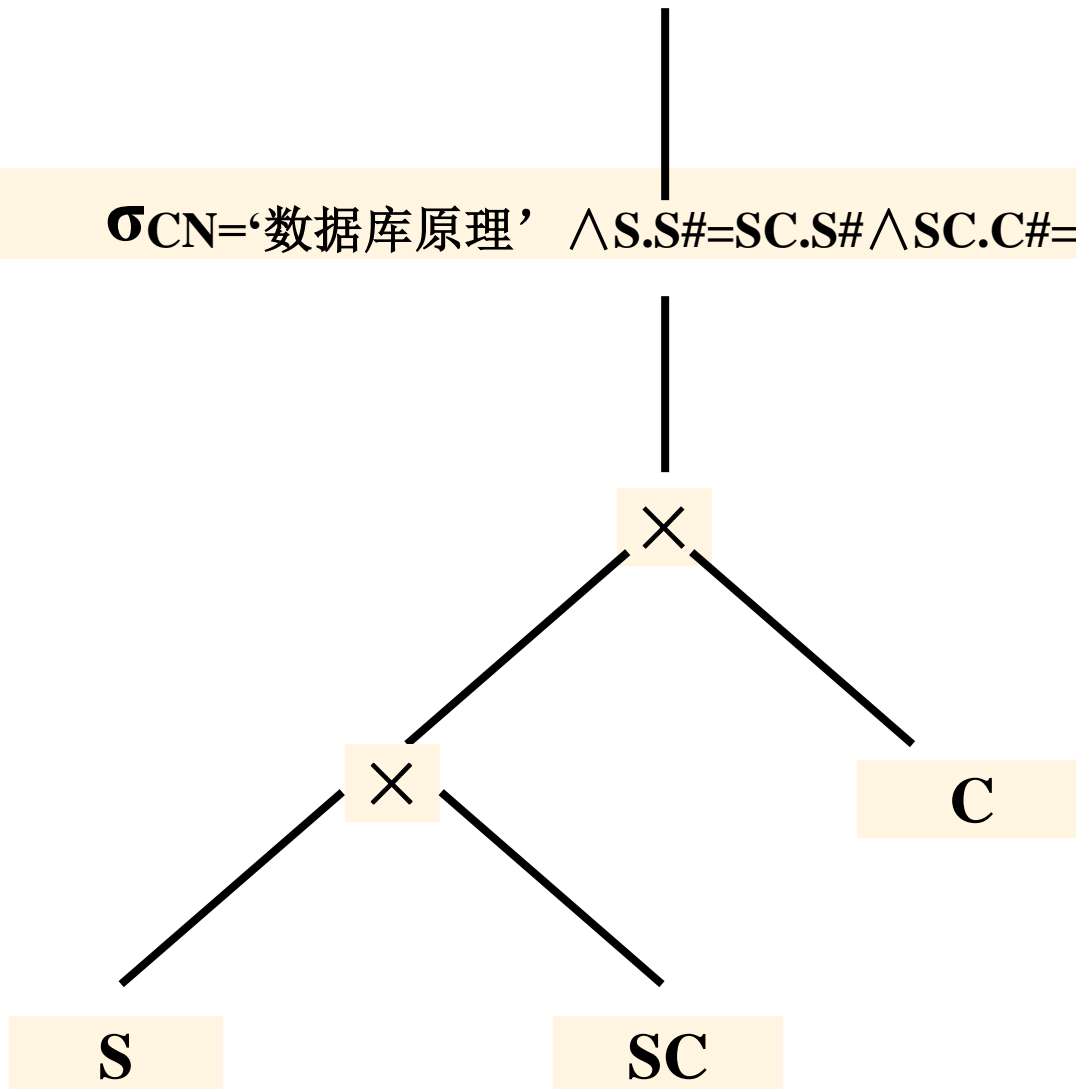
\times

\times

C

S

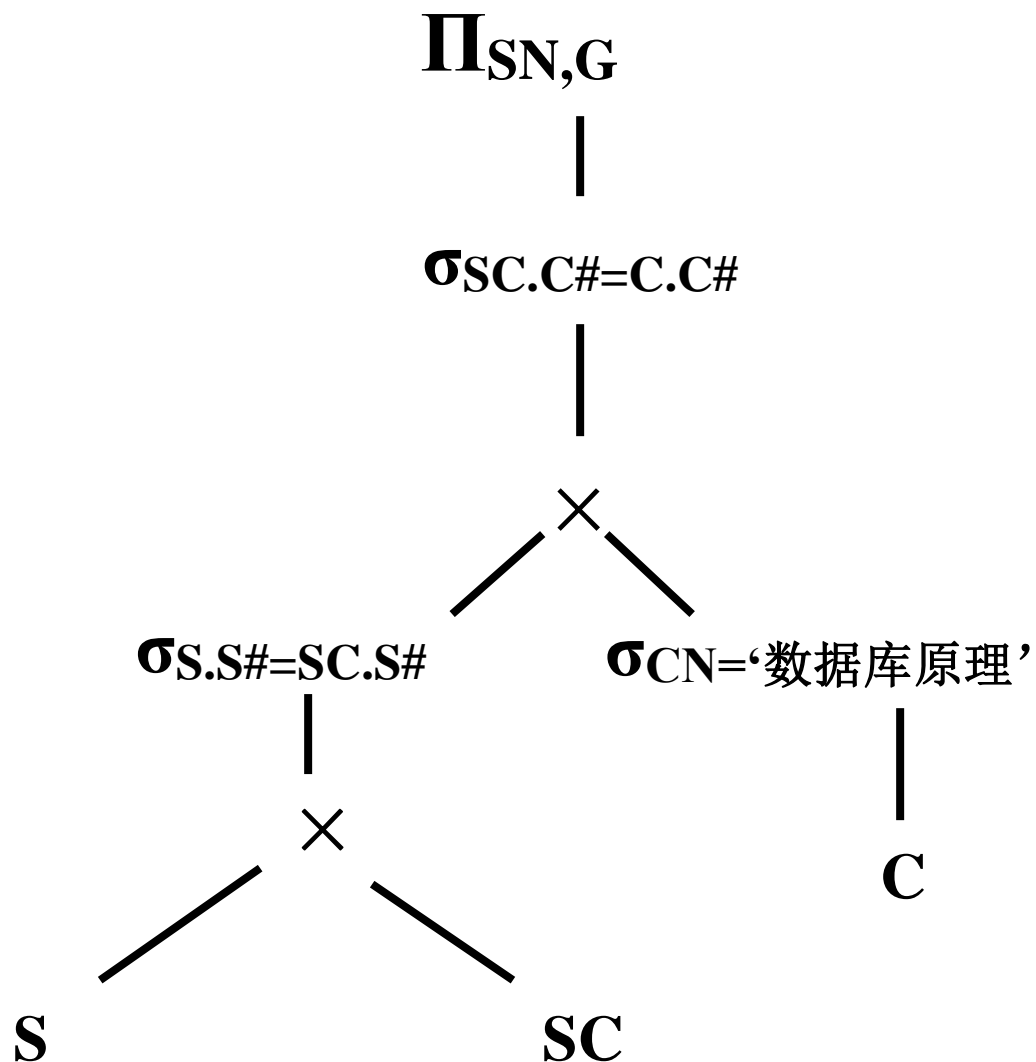
SC

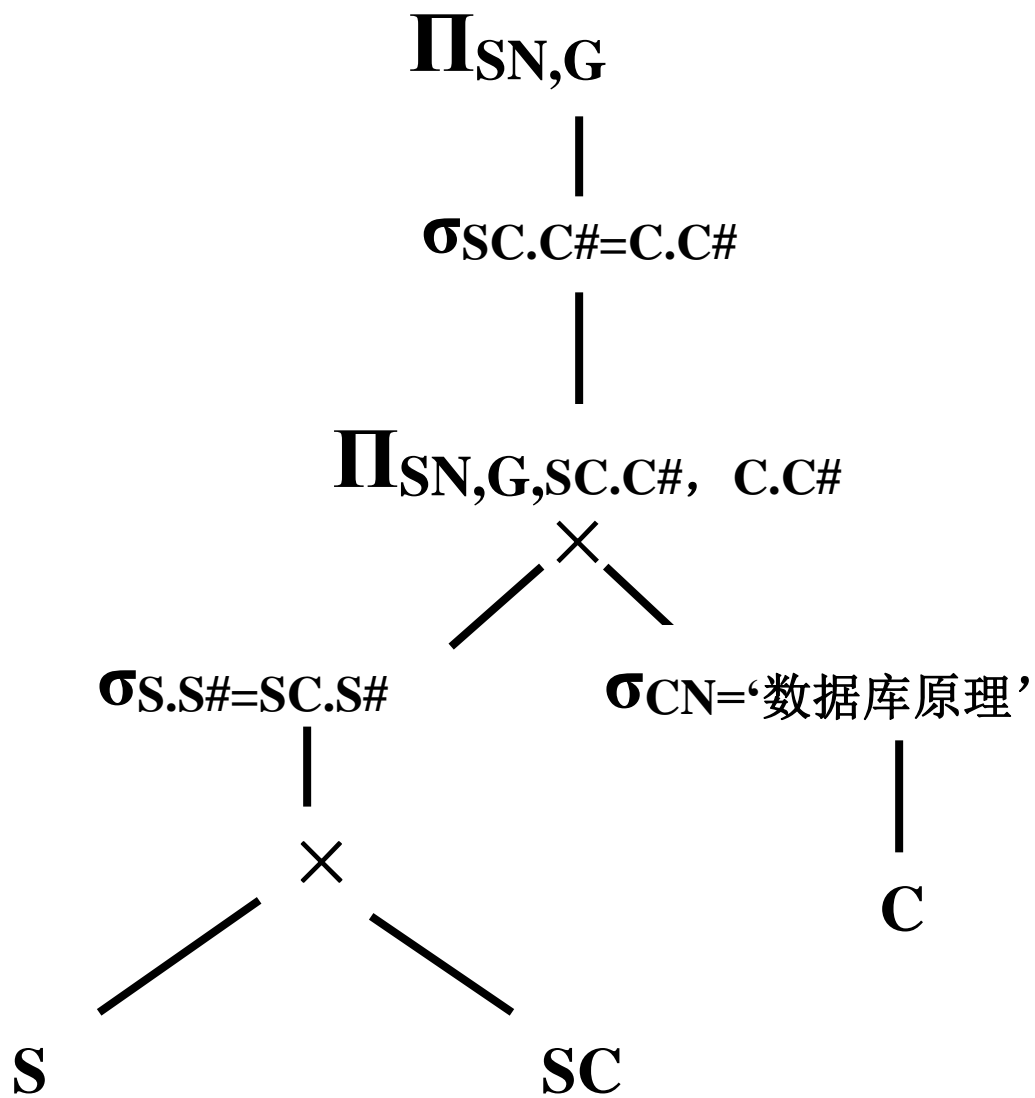


规则 4、6

选择的串接

选择和笛卡
尔积交换



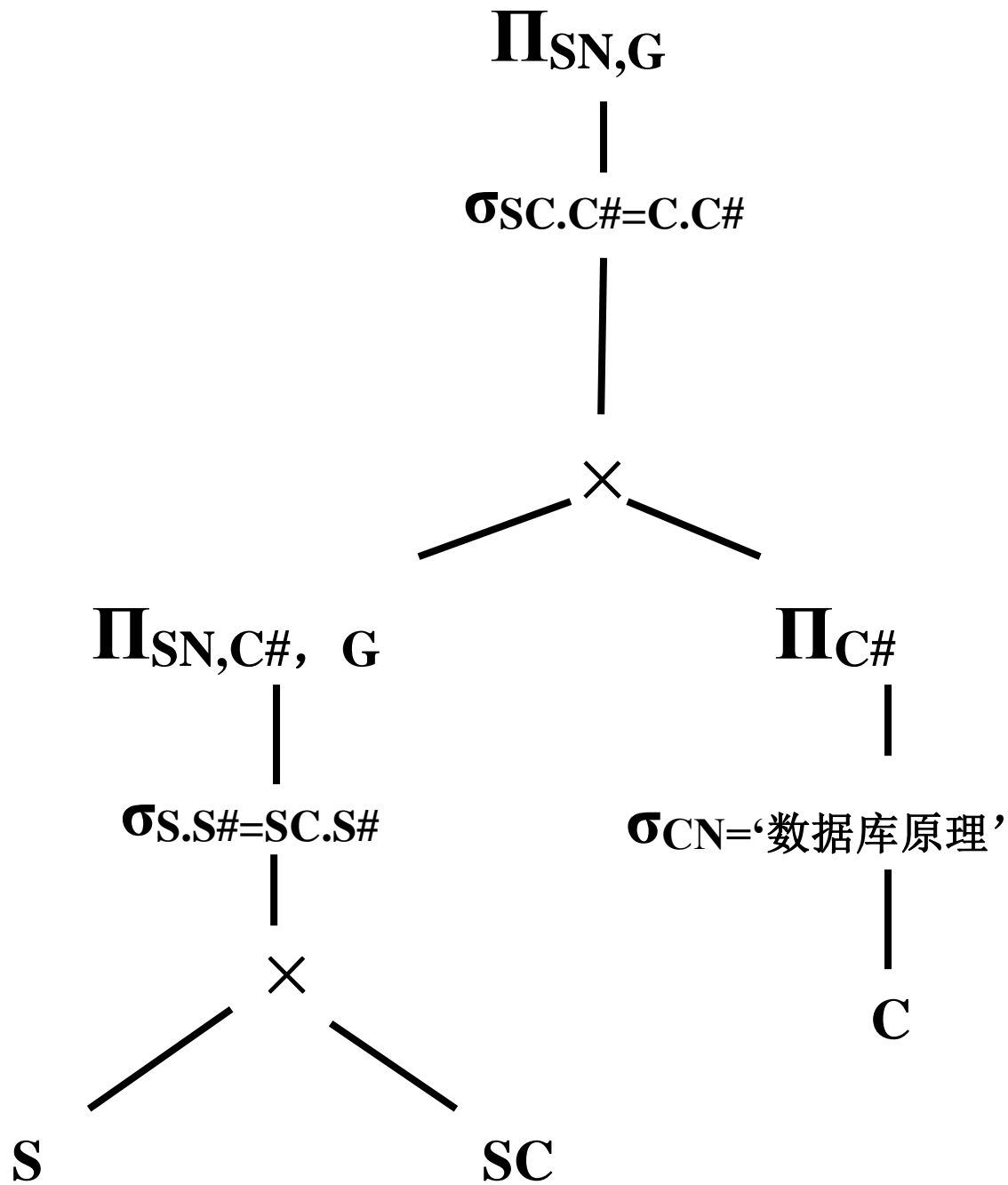


规则 5

选择和投影
交换

规则 9

投影和笛卡
尔积交换



$\Pi_{SN,G}$

|

$\sigma_{SC.C\#=C.C\#}$

|

×

$\Pi_{SN,C\#, G}$

|

$\sigma_{S.S\#=SC.S\#}$

|

$\Pi_{SN,S.S\#, SC.S\#,C\#, G}$

×

↙

S

↘

SC

$\Pi_{C\#}$

|

$\sigma_{CN='数据库原理'}$

|

C

规则 5

选择和投影

交换

规则 9

投影和笛卡
尔积交换

