Lecture 11: 查询优化

Database

Author: Forliage

Email: masterforliage@gmail.com

Date: June 6, 2025

College: 计算机科学与技术学院

Abstract

本讲笔记系统地介绍了关系型数据库查询优化的核心原理与实现方法，主要内容包括三个方面：首先，在“引言”部分阐述了查询评估计划的概念及其重要性，强调了等价表达式生成与多种操作算法选择对查询性能的影响，并简要说明了基于成本的优化流程包括关系表达式转换、统计信息标注与成本估算、以及评估计划的选择三个阶段。接着，“关系表达式的转换”章节详细列举了常用的等价规则（如选择谓词拆分与下推、投影与连接的分配律、连接操作的交换与结合性等），并说明了查询优化器如何通过递归应用这些规则枚举所有可能的等价表达式，以便后续的成本评估。然后，“用于成本估算的统计信息”部分重点介绍了查询优化器所需的元数据信息，包括关系的元组数、块数、属性不同值数等基本统计量，以及利用这些统计信息对选择操作、连接操作、投影与分组聚合操作的基数与不同值数进行估算的方法，分别推导了等值选择、范围选择、合取/析取谓词、自然连接与等值连接的大小估计公式，并讨论了投影与分组聚合后不同值数的近似估计策略。最后，“评估计划的选择”部分分析了独立为各个操作选择最优算法可能无法产生全局最优方案，介绍了启发式与基于代价的动态规划方法，说明如何在指数级的连接顺序空间中通过子问题最优子结构原理高效地寻找最低成本连接顺序。通过本讲笔记的学习，读者能够掌握查询优化器在逻辑重写、统计估算与成本驱动决策方面的关键技术，为后续深入研究查询执行引擎与性能调优打下坚实基础。

（该Abstract由ChatGPT-o4-mini-high生成）

Contents

1 引言

评估给定查询的替代方法：

- 等效表达式
- 每种操作的不同算法

评估计划精确地定义了每个操作所使用的算法，以及操作执行的协调方式。

了解如何在你喜欢的数据库上查看查询执行计划。

查询评估计划之间的成本差异可能非常大：例如，在某些情况下，可能是几秒和数天的差别。

基于成本的查询优化步骤：

1. 使用等价规则生成逻辑等价的表达式规则
2. 对结果表达式进行注释以获取替代查询计划
3. 根据估计成本选择最便宜的计划

基于以下因素估计计划成本：

1. 关于关系的统计信息，例如，元组数量，属性的不同值数量
2. 中间结果的统计估计，计算复杂表达式的成本
3. 使用统计方法计算的算法成本公式

2 关系表达式的转换

2.1 关系表达式的转换

如果两个关系代数表达式在每个合法的数据库实例上生成相同的元组集，则称这个两个表达式是等价的。

- 注意：元组的顺序无关紧要
- 我们不关心它们在违反完整性约束的数据库上是否产生不同的结果

在SQL中，输入和输出是元组的多重集：如果关系代数多重集版本中的两个表达式在每个合法的数据库实例上生成相同的元组多重集，则称这两个表达式是等价的。

等价规则表明，两种形式的等价的等价的。可以用第二种形式的表达式替换第一种形式的表达式，反之亦然。

2.2 等价规则

1.合取（并且）谓词的拆分

$$\sigma_{\theta_1 \wedge \theta_2}(E) = \sigma_{\theta_1}(\sigma_{\theta_2}(E))$$

2.选择运算的交换性（Commutativity）

$$\sigma_{\theta_1}(\sigma_{\theta_2}(E)) = \sigma_{\theta_2}(\sigma_{\theta_1}(E))$$

3.多重投影只保留最外层

$$\pi_{L_1}(\pi_{L_2}(...(\pi_{L_n}(E))...)) = \pi_{L_1}(E)$$

4.选择与笛卡尔积/Theta-Join的结合

a

$$\sigma_{\theta}(E_1 \times E_2) = E_1 \bowtie_{\theta} E_2$$

b

$$\sigma_{\theta_1}(E_1 \bowtie_{\theta_2} E_2) = E_1 \bowtie_{\theta_1 \wedge \theta_2} E_2$$

5.Theta-Join和自然连接的交换性

$$E_1 \bowtie_{\theta} E_2 = E_2 \bowtie_{\theta} E_1$$

6.连接的结合性

a 自然连接的结合性

$$(E_1 \bowtie E_2) \bowtie E_3 = E_1 \bowtie (E_2 \bowtie E_3)$$

b Theta-Join的结合性

$$(E_1 \bowtie_{\theta_1} E_2) \bowtie_{\theta_2} E_3 = E_1 \bowtie_{\theta_1 \wedge \theta_3} (E_2 \bowtie_{\theta_2} E_3)$$

其中θ_2只涉及E_2和E_3的属性，θ_1只涉及E_1与E_2，θ_3只涉及E_1与E_3。

7.选择对Theta-Join的分配律

a 当谓词θ_0只涉及E_1中的属性时：

$$\sigma_{\theta_0}(E_1 \bowtie_{\theta} E_2) = (\sigma_{\theta_0}(E_1)) \bowtie_{\theta} E_2$$

b 当谓词θ_1只涉及E_1的属性，θ_2只涉及E_2的属性时：

$$\sigma_{\theta_1 \wedge \theta_2}(E_1 \bowtie_{\theta} E_2) = (\sigma_{\theta_1}(E_1)) \bowtie_{\theta} (\sigma_{\theta_2}(E_2))$$

通常也记为“先对各自表Select，再做Join”。

8.投影对Theta-Join的分配律

a 若连接条件θ只涉及属性集合$L_1 \cup L_2$，

$$\pi_{L_1 \cup L_2}(E_1 \bowtie_\theta E_2) = (\pi_{L_1}(E_1)) \bowtie_\theta (\pi_{L_2}(E_2))$$

b 更一般地，若要在两个关系$E_1 \bowtie_\theta E_2$之后做投影到属性L，令

- $L_1 \subset E_1$中最终要投影的属性
- $L_2 \subset E_2$中最终要投影的属性
- $L_3 \subset E_1$中仅在连接条件θ中出现但不在$L_1 \cup L_2$的属性集
- $L_4 \subset E_2$中仅在连接条件θ中出现但不在$L_1 \cup L_2$的属性集

则：

$$\pi_{L_1 \cup L_2}(E_1 \bowtie_\theta E_2) = \pi_{L_1 \cup L_2}((\pi_{L_1 \cup L_3}(E_1)) \bowtie_\theta (\pi_{L_2 \cup L_4}(E_2)))$$

这意味着：先对每一侧保留在投影或连接条件中所需的属性，再做Join，最后再投影。

9.集合运算的交换性

$$E_1 \cup E_2 = E_2 \cup E_1, E_1 \cap E_2 = E_2 \cap E_1$$

（注意：集合差$E_1 - E_2 \neq E_2 - E_1$，无交换性）

10.并与交的结合性

$$(E_1 \cup E_2) \cup E_3 = E_1 \cup (E_2 \cup E_3), (E_1 \cap E_2) \cap E_3 = E_1 \cap (E_1 \cap E_3)$$

11.选择对集合运算的分配律

- 对差运算：

$$\sigma_\theta(E_1 - E_2) = \sigma_\theta(E_1) - \sigma_\theta(E_2)$$

  同样对并、交也成立。
- 注意特殊式：

$$\sigma_\theta(E_1 - E_2) = \sigma_\theta(E_1) - E_2$$

  当θ只与E_1相关时可写成这样，但不能把选择条件推进到右侧的E_2。

12.投影对并运算的分配律

$$\pi_L(E_1 \cup E_2) = \pi_L(E_1) \cup \pi_L(E_2)$$

2.3 等效表达式的枚举

查询优化器使用等价规则系统地生成与给定表达式等效的表达式。

可以按如下方式生成所有等效表达式：

- 重复
  - 对到目前为止找到的每个等效表达式的每个等效表达式的每个子表达式应用所有适用的等价规则
  - 将新生成的表达式添加到等效表达式集合中
- 直到上述过程不再生成新的等效表达式

上述方法在空间和时间上的开销都非常大

- 基于转换规则的优化计划生成
- 仅含选择、投影和连接操作的查询的特殊处理方法

3 用于成本估算的统计信息

3.1 成本估算的统计信息

- n_r:关系r中的元组数量
- b_r:包含r元组的块数量
- l_r:r的元组大小
- f_r:r的块因子，即一个块中能容纳的r元组数量
- $V(A,r)$:属性A在r中出现的不同值的数量；与$\prod_A(r)$的大小相同
- 如果r的元组在文件中物理上存储在一起，那么：$b_r = \lceil \frac{n_r}{f_r} \rceil$

3.2 选择大小估计

$\sigma_{A=v}(r)$:

- $n_r/V(A,r)$:将满足选择条件的记录数量
- 键属性上的相等条件：大小估计=1

$\sigma_{A\leq V}(r)$（$\sigma_{A\geq V}(r)$的情况是对称的）

- 设c表示满足该条件的元组的估计数量
- 如果min(A,r)和max(A,r)在目录中可用：若$v < \min(A,r)$，则$c = 0$；否则$c = n_r \times \frac{v-\min(A,r)}{\max(A,r)-\min(A,r)}$
- 如果有直方图可用，则可以完善上述估计
- 在缺乏统计信息的情况下，假定$c = \frac{n_r}{2}$

3.3 复杂选择的规模估计

- 条件θ_i的选择性是关系r中的一个元组满足θ_i的概率：如果s_i是r中满足条件的元组数量，那么θ_i的选择性由s_i/n_r给出。
- 合取$\sigma_{\theta_1\wedge\theta_2\wedge\ldots\wedge\theta_n}(r)$。假设相互独立，对结果中的元组估计为$n_r\times\frac{s_1*s_2*\ldots*s_n}{n_r^n}$
- 析取$\sigma_{\theta_1\vee\theta_2\vee\ldots\vee\theta_n}(r)$。估计的元组数量：$n_r*\left(1-(1-\frac{s_1}{n_r})*(1-\frac{s_2}{n_r})*\ldots(1-\frac{s_n}{n_r})\right)$
- 否定$\sigma_{\neg\theta}(r)$。估计的元组数量：$n_r-size(\sigma_\theta(r))$

3.4 连接大小的估计

- 笛卡尔积$r\times s$包含n_r*n_s个元组；每个元组占用s_r+s_s字节
- 如果$R\cap S=\phi$，那么$r\times s$和$s\times r$相同
- 如果$R\cap S$是R的键，那么S的一个元组最多与r中的一个元组进行连接。因此，$r\times s$中的元组数量不超过s中的元组数量
- 如果S中的$R\cap S$是S中引用R的外键，那么$r\bowtie s$中的元组数量与S中的元组数量完全相同：$R\cap S$作为S的外键是情况是对称的

如果$R\cap S=\{A\}$不是R或S的键。如果我们假设R中的灭个元组t都会在$R\bowtie S$中生成元组，那么$R\times S$中的元组数量估计为：$\frac{n_r*n_s}{V(A,s)}$ 如果反之成立，则得到的估计值将为：$\frac{n_r*n_s}{V(A,r)}$这两个估计值中较小的那个可能更准确。

3.5 不同值估计

3.5.1 对选择后不同值数的估计

设有关系 r，其属性 A 在 r 中的不同值数量为 $V(A,r)$，且关系 r 的元组总数为 $|r|$。对选择操作 $\sigma_\theta(r)$，若谓词 θ 只涉及属性 A，则常用以下几种估计方法：

1. 如果 θ 将 A 强制为某个常数值，例如 $A=7$，则

$$V\Big(A,\ \sigma_{A=7}(r)\Big)\ =\ 1.$$

2. 如果 θ 强制 A 在一组指定常数值中取值，例如

$$\theta:\ A\in\{1,3,4\},$$

则有

$$V\Big(A,\ \sigma_{A\in\{1,3,4\}}(r)\Big)\ =\ 3.$$

3. 如果 θ 是一个范围类谓词（如 $A<100$、$A\leq 500$ 或 $10\leq A\leq 20$），假定属性 A 在 r

中近似均匀分布，选择率为 s（满足 θ 的元组数占 $|r|$ 的比例），则

$$V\big(A,\ \sigma_\theta(r)\big)\ \approx\ V(A,\, r)\ \times\ s.$$

例如，当 r 中 A 的值域是 $[1,\, 1000]$，θ 为 $A < 100$，可近似认为 $s = 0.1$，于是

$$V\big(A,\ \sigma_{A<100}(r)\big) \approx V(A,\, r) \times 0.1.$$

4. 在其他更复杂的情况下（如 θ 同时涉及多个属性或函数表达式），难以精确估计时可采用保守估计：

$$V\big(A,\ \sigma_\theta(r)\big)\ \approx\ \min\big(V(A,\, r),\ n_\sigma(r)\big),$$

其中 $n_\sigma(r)$ 表示 $\sigma_\theta(r)$ 的基数估计（筛选后元组总数）。

3.5.2 对连接后不同值数的估计

设要对关系 r 与 s 做自然连接（或等值连接）$r \bowtie s$，连接属性集为 $\{A_1, \ldots, A_k\}$，连接结果的基数估计为 $n_{r\bowtie s}$。如果要估计连接结果中某属性 A 的不同值数 $V(A,\, r \bowtie s)$，常见估计方法如下：

1. 若 A 来自 r（即 A 在 s 中不存在），则

$$V\big(A,\ r \bowtie s\big)\ \approx\ \min\big(V(A,\, r),\ n_{r\bowtie s}\big).$$

因为连接后不可能出现比原来 r 中更过的不同值，也无法超过连接结果的行数。

2. 若属性 A 是由 r 的子属性集合 A_1 与 s 的子属性集合 A_2 组合而成，则可用经验公式

$$V\big(A,\ r \bowtie s\big)\ \approx\ \min\Big(V(A_1,\, r) \times V(A_2 - A_1,\, s),\ V(A_1 - A_2,\, r) \times V(A_2,\, s),\ n_{r\bowtie s}\Big),$$

其中

- $V(A_1,\, r)$：r 中属性集合 A_1 的不同值数；
- $V(A_2,\, s)$：s 中属性集合 A_2 的不同值数；
- $V(A_2 - A_1,\, s)$：s 中仅在 A_2 中且不属于 A_1 的属性组合的不同值数；
- $V(A_1 - A_2,\, r)$：r 中仅在 A_1 中且不属于 A_2 的属性组合的不同值数；
- $n_{r\bowtie s}$：连接结果的基数估计。

该公式表示连接后组合属性的不同值数不会超过“各自参与属性数量的乘积”，也不会超过连接后的行数，取三者最小。

3.5.3 对投影与分组/聚合后不同值数的估计

1. **投影操作** 对关系 r 做投影 $\pi_L(r)$，只保留属性集 L 中的列。若 $A \in L$，则通常认为：

$$V\big(A,\ \pi_L(r)\big)\ =\ V\big(A,\, r\big),$$

因为投影不会合并本身已不同的值。若仅保守估计，也可直接沿用 $V(A,r)$。

2. **分组 + 聚合操作** 对关系 r 按属性集 G 分组，再对某列 A 进行聚合，例如计算 $\min(A)$ 或 $\max(A)$。
   - 当聚合函数为 $\min(A)$ 或 $\max(A)$ 时，分组后每组会产生一个最小值或最大值。估计不同值数时，最常用经验公式是：

     $$V\big(\min(A),\,r\big) \;=\; \min\Big(V(A,\,r),\;V(G,\,r)\Big), \qquad V\big(\max(A),\,r\big) \;=\; \min\Big(V(A,\,r),\;V(G,\,r)\Big).$$

     其中 $V(G,\,r)$ 是分组属性集 G 在 r 中的不同值数。直觉上结果数不会超过分组数，也不会超过原来 A 的不同值数。
   - 对于其他聚合函数（如 $\mathrm{SUM}(A)$、$\mathrm{AVG}(A)$、$\mathrm{COUNT}(*)$ 等），一般保守假设“每个分组都产生不同的聚合值”，于是估计：

     $$V\big(\text{其他聚合结果},\,r\big) \;\approx\; V\big(G,\,r\big).$$

     即聚合结果的不同值数近似等于分组属性集 G 的不同值数。

3.6 评估计划的选择

选择评估计划时必须考虑评估技术的相互作用：为每个操作独立选择最便宜的算法可能无法产生最佳的整体算法，例如：

- 合并连接的成本可能高于哈希连接，但可能提供排序后的输出，从而降低外层聚合的成本
- 嵌套循环连接可能为流水线操作提供机会

实际的查询优化器结合了以下两种广泛方法的要素：

- 搜索所有方案，并以基于成本的方式选择最佳方案
- 使用启发式方法选择方案

3.7 基于成本的优化

考虑为$r_1 \times r_2 \times ... \times r_n$找到最佳连接顺序。

上述表达式有$\frac{(2(n-1))!}{(n-1)!}$种不同的连接顺序。当$n = 7$时，数量为665280；当$n = 10$时，数量超过1760亿。

无需生成所有的连接顺序。使用动态规划，$\{r_1, r_2, ..., r_n\}$的任何子集的最低成本连接顺序只需计算一次，并存储以供将来使用。