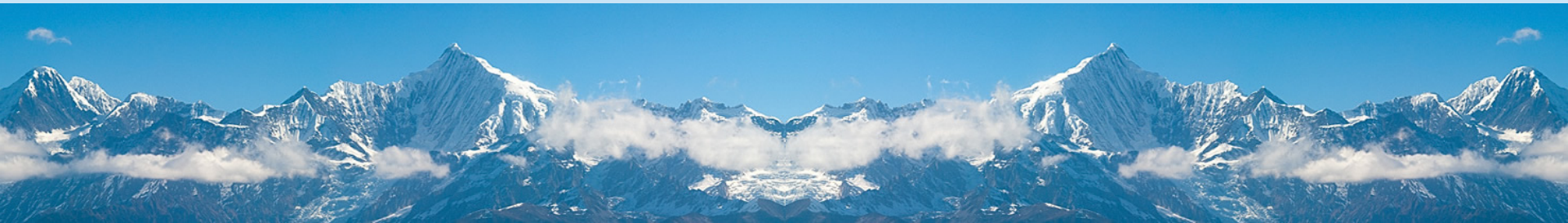
空间数据库2016秋季

查询处理与优化-1215

陈斌

北京大学地球与空间科学学院

gischen@pku.edu.cn



查询处理与优化

- 空间操作计算
- 查询优化
- 关系表达式等价规则
- 选择执行计划

空间操作计算

- 和关系数据运算的区别
  - 空间数据没有公认标准定义的运算，关系数据库中的运算很固定；
  - 空间对象的空间位置和范围在二维平面上定义，不能自然排序成一维数组；
  - 检测空间关系计算代价非常高，不能再假定空间数据库中I/O代价仍然远超过CPU代价

空间操作的基本类型

- 更新操作
  - 空间对象的创建、修改和删除
- 空间选择
  - 点查询：给定点，找出包含它的空间对象
    - PointQuery

$$PQ(p) = \{O \mid p \in O.G \neq \emptyset\}$$

  - 范围查询：给定多边形（矩形时称作“窗口”），找出与之相交的空间对象
    - RangeQuery

$$RQ(P) = \{O \mid O.G \cap P.G \neq \emptyset\}$$

空间操作的基本类型

- 空间连接
  - 空间谓词作为连接条件

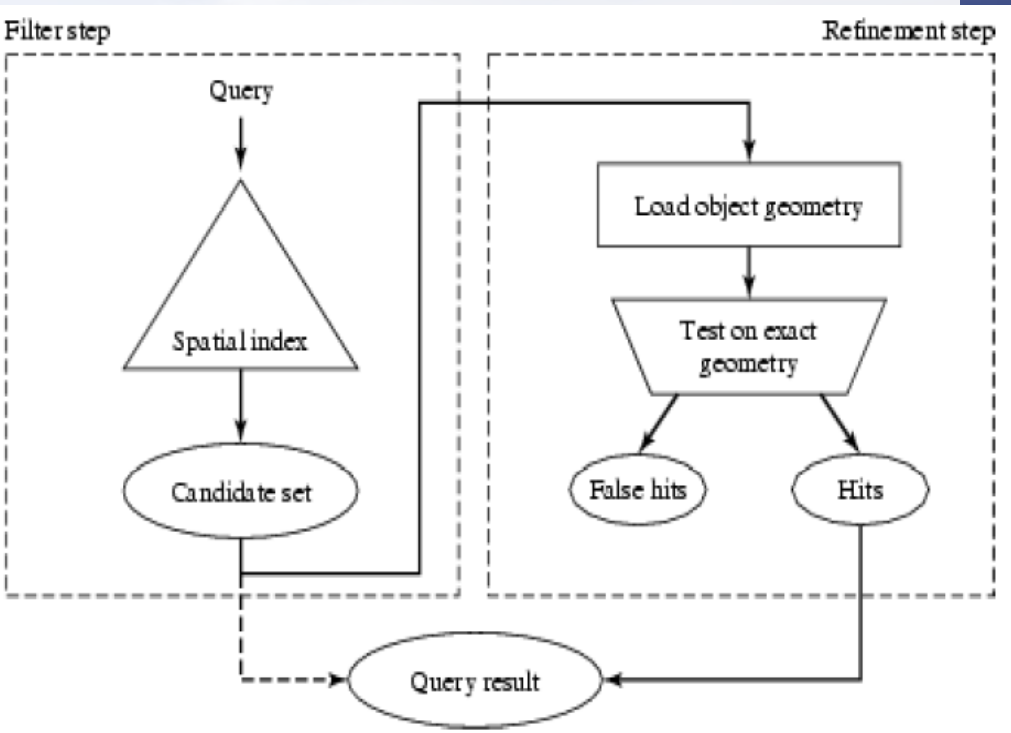
$$R \bowtie_{\theta} S = \{(o, o') \mid o \in R, o' \in S, \theta(o.G, o'.G)\}$$

  - 可以是多种空间关系
    - 相交、包含、距离、西北、邻接、交叠……
- 空间聚集
  - 最近邻居：查找距离给定对象最近的所有对象

$$NNQ(o') = \{o \mid \forall o'' : dist(o'.G, o.G) \leq dist(o'.G, o''.G)\}$$

空间操作的两步处理

- 过滤步骤
  - 空间对象用MBR表示
  - 大大降低计算复杂度
  - 得到近似的结果
- 精炼步骤
  - 采用精确的几何信息进行精确计算
  - 计算代价高
  - 可能在SDB之外实现



空间选择查询

- 对于点查询和范围查询
  - 实现方法依赖于数据文件的存储组织
  - 主要是过滤步骤
- 未排序无索引的数据文件
  - 全表扫描，代价O(n)

空间选择查询

- 采用空间填充曲线散列：
- 将二维空间的点映射为一维
  - Z序和Hilbert曲线
  - 近似实现保持“位置相邻”
- 映射后采用一般的排序树B树或B+树索引
- 点查询代价O(log n)
- 范围查询代价O(log n)+查询结果集大小/记录的聚集程度

空间选择查询

- 具有空间索引的数据文件：
- 普遍采用R树，按照MBR索引，代价O(log n)
- R树的缺点MBR允许交叠，可能导致需要搜索多个子树
- R+树解决了上述问题，但副本的存在可能增加搜索时间和节点溢出

组合空间选择

- 组合条件表示为合取范式
  - 非空间谓词计算代价大致相同，处理顺序不重要
  - 空间谓词计算代价高而且差别大，处理顺序对总代价影响大
- 对每个空间谓词计算代价等级
  - 选择性：输出集合和输入集合数量比，体现I/O代价
  - 差异代价：对单个元组处理的平均代价，体现CPU代价
- 按照代价等级的升序来计算空间谓词

空间连接实现

- 由于空间索引的限制，主要考虑过滤步骤（MBR近似计算）
- 嵌套循环
- 有空间索引的嵌套循环
- 基于分块的空间连接
  - 类似散列连接

空间聚集实现

- 最近邻居算法
- 在空间索引树上进行遍历并修剪子树
- 可扩展为K个最近邻居
  - 保留K个最好候选子树或对象

空间查询优化

- 查询优化的基本步骤与关系数据库相同
  - 代数优化
    - 找出等价的但执行效率更高的一个表达式
  - 执行策略选择（动态规划）
    - 选择执行运算所采用的具体算法
    - 选择将使用的特定索引等等
  - 需要考虑空间谓词特点
    - 计算代价高（可能超过I/O代价）
    - 代价差异大（注意计算顺序）

代数优化：等价变换规则

- 将一个表达式转换为与之等价的另一个表达式的规则
- 规则来源于代数系统的运算性质
  - 交换律
  - 结合律
  - 分配律
- 生成多个等价表达式为可选的执行计划

等价变换规则

- 选择运算的级联：合取选择运算可分解为单个选择运算的序列
  - $\sigma_{\theta_1 \wedge \theta_2}(E)=\sigma_{\theta_1}(\sigma_{\theta_2}(E))$
  - 尽量将所有非空间条件右移
- 选择运算满足交换律
  - $\sigma_{\theta_1}(\sigma_{\theta_2}(E))=\sigma_{\theta_2}(\sigma_{\theta_1}(E))$
  - 先计算非空间条件，然后代价低的空间条件

等价变换规则

- 投影运算的级联：投影运算序列中只有最后一个运算是需要的，其余的可省略
- $\Pi_{L_1}(\Pi_{L_2}(\ldots(\Pi_{L_n}(E))\ldots))=\Pi_{L_1}(E)$

等价变换规则

- 选择与笛卡尔积以及theta连接相结合

  $\sigma_{\theta}(E_1 \times E_2)=E_1 |><|_{\theta} E_2$

  $\sigma_{\theta_1}(E_1 |><|_{\theta_2} E_2)= E_1 |><|_{\theta_1 \wedge \theta_2} E_2$

- theta连接运算满足交换律
  - $E_1 |><|_{\theta} E_2= E_2 |><|_{\theta} E_1$
- 自然连接也满足交换律
  - $E_1 |><| E_2= E_2 |><| E_1$

等价变换规则

- 连接运算的结合律
- 自然连接运算满足结合律
  - $(E_1 |><|_{} E_2) |><|_{} E_3 = E_1 |><|_{} (E_2 |><|_{} E_3)$
- theta连接具有以下方式的结合律

$$(E_1 |><|_{\theta_1} E_2) |><|_{\theta_2 \wedge \theta_3} E_3 = E_1 |><|_{\theta_1 \wedge \theta_3} (E_2 |><|_{\theta_2} E_3)$$

其中θ_2只涉及E_2与E_3的属性

由于任意一个条件都可为空，因此笛卡尔积运算也具有结合律。

等价变换规则

- 选择运算在下面两个条件下对theta连接运算具有分配律

当选择条件θ_0的所有属性只涉及参与连接运算的表达式之一(E_1)时：

$$\sigma_{\theta_0}(E_1 |><|_{\theta} E_2) = (\sigma_{\theta_0}(E_1)) |><|_{\theta} E_2$$

当选择条件θ_1只涉及E_1的属性，选择条件θ_2只涉及E_2的属性时：

$$\sigma_{\theta_1 \wedge \theta_2}(E_1 |><|_{\theta} E_2) = (\sigma_{\theta_1}(E_1)) |><|_{\theta} (\sigma_{\theta_2}(E_2))$$

等价变换规则

- 投影运算对theta连接运算具有分配律

令L_1、L_2分别是E_1、E_2的属性。假设连接条件θ只涉及$L_1 \cup L_2$中的属性，则：

$$\prod_{L_1 \cup L_2}(E_1 |><|_{\theta} E_2)=(\prod_{L_1}(E_1)) |><|_{\theta}(\prod_{L_2}(E_2))$$

考虑连接$E_1 |><|_{\theta} E_2$。令L_1、L_2分别是E_1、E_2的属性；令L_3是E_1中出现在连接条件θ中但不在$L_1 \cup L_2$中的属性；令L_4是E_2中出现在连接条件θ中但不在$L_1 \cup L_2$中的属性。那么：

$$\prod_{L_1 \cup L_2}(E_1 |><|_{\theta} E_2)$$
$$= \prod_{L_1 \cup L_2}((\prod_{L_1 \cup L_3}(E_1)) |><|_{\theta}(\prod_{L_2 \cup L_4}(E_2)))$$

等价变换规则

- 集合运算并与交满足交换律

  $E_1 \cup E_2 = E_2 \cup E_1$

  $E_1 \cap E_2 = E_2 \cap E_1$

  集合差运算不满足交换律

- 集合运算并与交满足结合律

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

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

  集合差运算不满足结合律

等价变换规则

- 选择运算对并、交、差运算具有分配律
  - $\sigma_p(E_1 - E_2) = \sigma_p(E_1) - \sigma_p(E_2)$
  - $\sigma_p(E_1 \cup E_2) = \sigma_p(E_1) \cup \sigma_p(E_2)$
  - $\sigma_p(E_1 \cap E_2) = \sigma_p(E_1) \cap \sigma_p(E_2)$
- 进一步有
  - $\sigma_p(E_1 - E_2) = \sigma_p(E_1) - E_2$
  - $\sigma_p(E_1 \cap E_2) = \sigma_p(E_1) \cap E_2$
  - 但对于∪不成立

等价变换规则

- 投影运算对并运算具有分配律
  - $\prod_L(E_1 \cup E_2) = (\prod_L(E_1)) \cup (\prod_L(E_2))$

等价变换规则

- 规则只说明两个表达式等价，并不说明哪一个更好
- 连接的次序很重要，好的连接次序序列产生小的中间结果
- 对于空间数据库考虑空间谓词计算代价
  - I/O和CPU代价的均衡
  - 好的空间谓词计算次序产生小的计算代价
- 规则的使用会产生大量的等价表达式，优化器要采用适当的技术来减少所产生的表达式的数量

选择执行计划

- 关系代数表达式的基础上，执行计划进一步说明：
  - 每个运算的实现算法
  - 各运算的执行顺序
  - 是否采用流水线技术
- 注意：每个运算的最小代价算法组合起来不一定是整个表达式的最佳算法，必须考虑各个运算之间的相互作用。

选择执行计划的优化算法

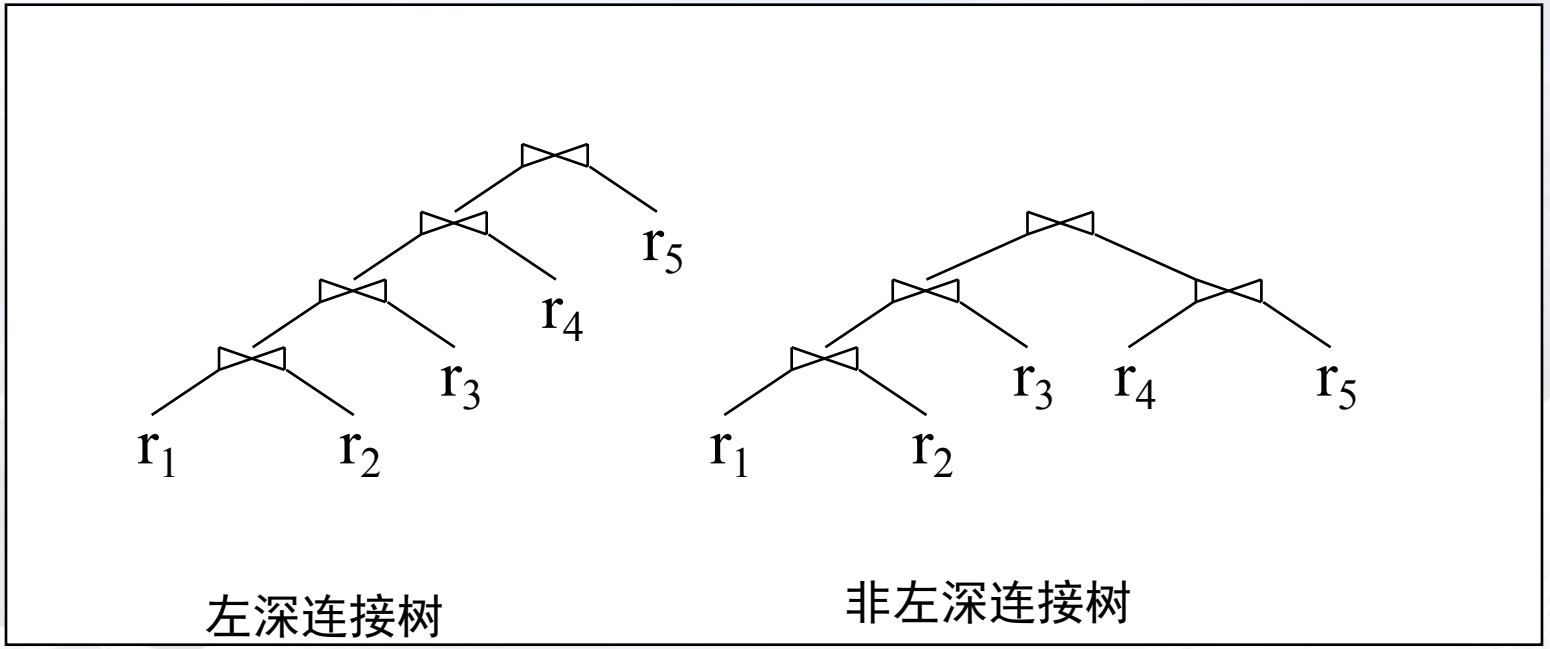
- 基于代价的方法
  - 通过使用等价规则为给定的查询语句产生一系列查询执行计划，并选择其中代价最小或接近最小的
- 启发式方法
  - 运用启发式规则，对关系代数表达式进行等价变换
  - 常用的规则：
    - 尽早进行选择运算
    - 尽早进行投影运算
    - 避免进行笛卡尔积运算……

基于代价的方法

- 等价于给定查询的不同查询计划可能很多，因此优化的代价太大
  - 采用一些预设方法来减少需要考虑的表达式的数目
  - 采用启发式方法来减少需考虑的表达式的数目

基于代价的方法

- 减少需要考虑的表达式
  - 只考虑左深连接次序



基于代价的方法

- 找多个关系的最佳连接顺序时，不是简单地考虑所有的可能顺序，而是为每个子集找出最佳连接顺序，这样能大大减少需要检查的连接顺序的总数
- 如果检查一个表达式的某部分后发现这一部分的最小代价已经比先前已检查过的整个表达式的执行计划的最小代价要大，则可以终止对这个表达式的检查。
  - 没有必要对包含该子表达式的任何完整表达式进行检查

启发式优化

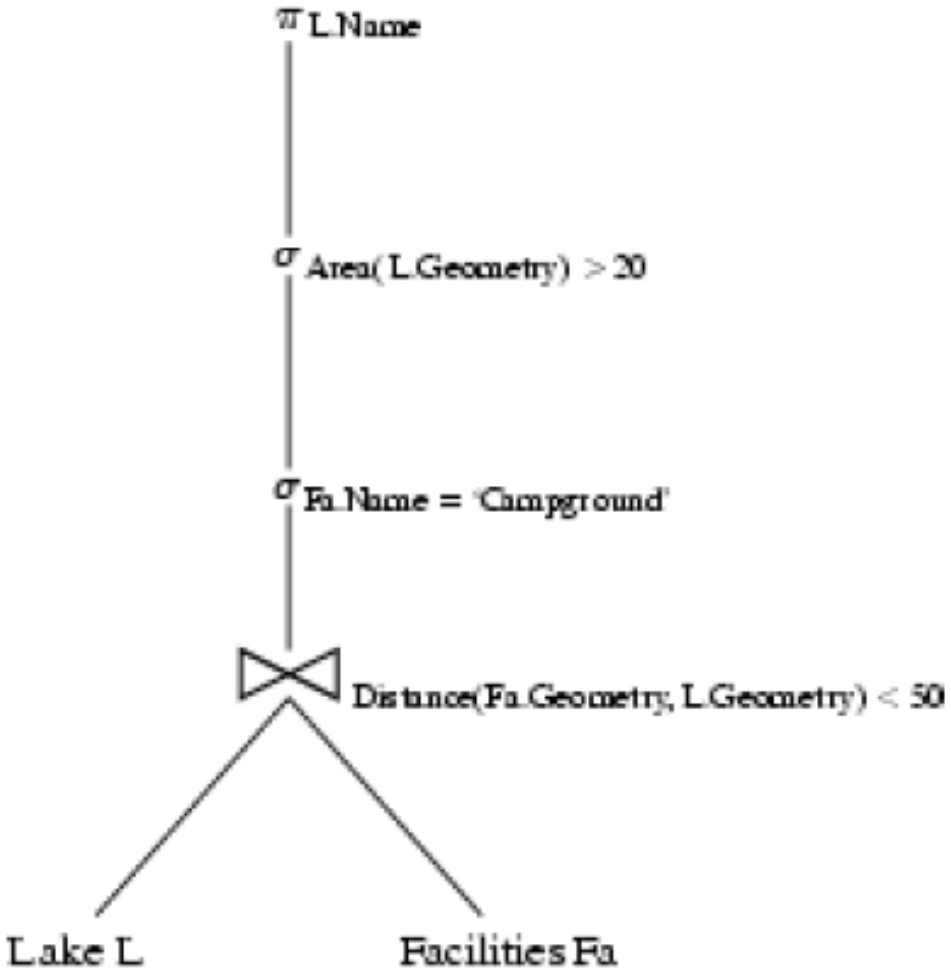
- 将合取选择分解为单个选择运算的序列，这有助于将选择运算往查询树下层移
  - 将非空间选择尽量下移
  - 代价小的空间选择下移
- 把选择运算在查询树上下推到最早可能执行的地方
  - 例如，尽可能将$\sigma_\theta(r|><|s)$转换成$\sigma_\theta(r)|><|s$或$r|><|\sigma_\theta(s)$
  - 考虑CPU代价极高的空间谓词，不一定移到最下

启发式优化（例）

“找出所有面积大于 20 平方公里并且距离露营地小于 50 公里的湖泊。”

```
SELECT  L.Name
FROM    Lake L, Facilities Fa
WHERE   Area(L.Geometry) > 20 AND
        Fa.Name = 'campground' AND
        Distance(Fa.Geometry, L.Geometry) < 50
```

启发式优化（例）


π L.Name
σ Area(L.Geometry) > 20
σ Fa.Name = 'Campground'
Distance(Fa.Geometry, L.Geometry) < 50
Lake L
Facilities Fa

启发式优化（例）

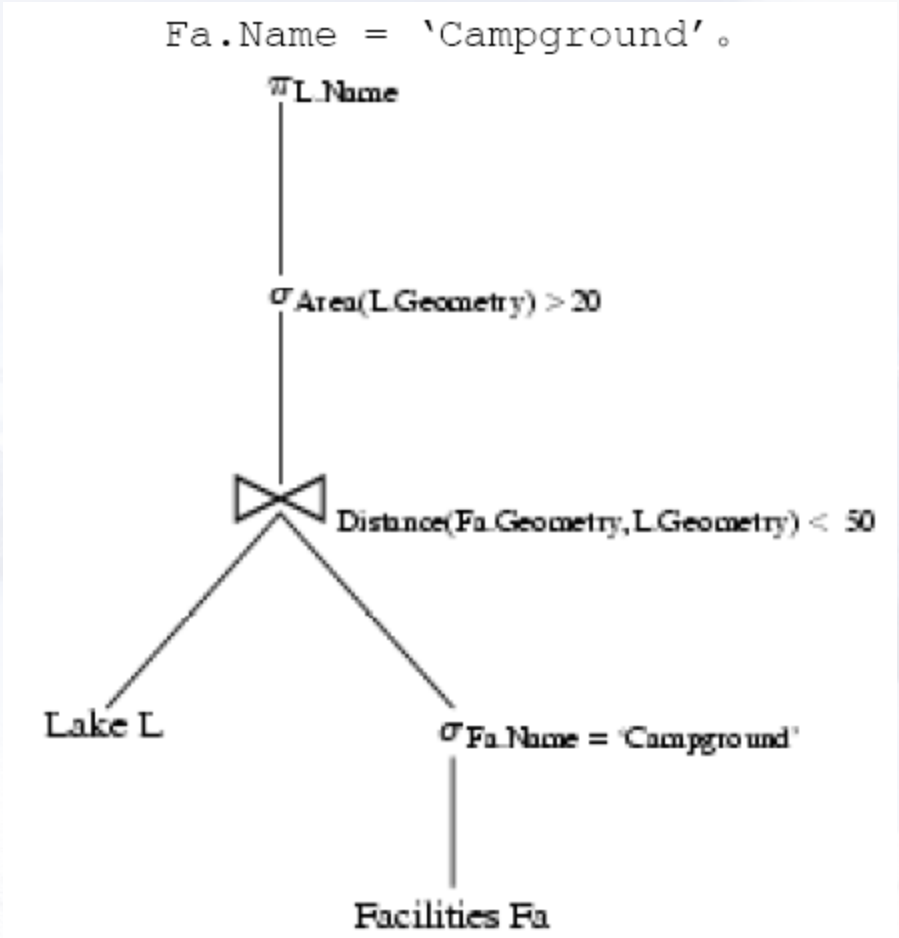


图 5-4 下推：选择操作。

启发式优化（例）

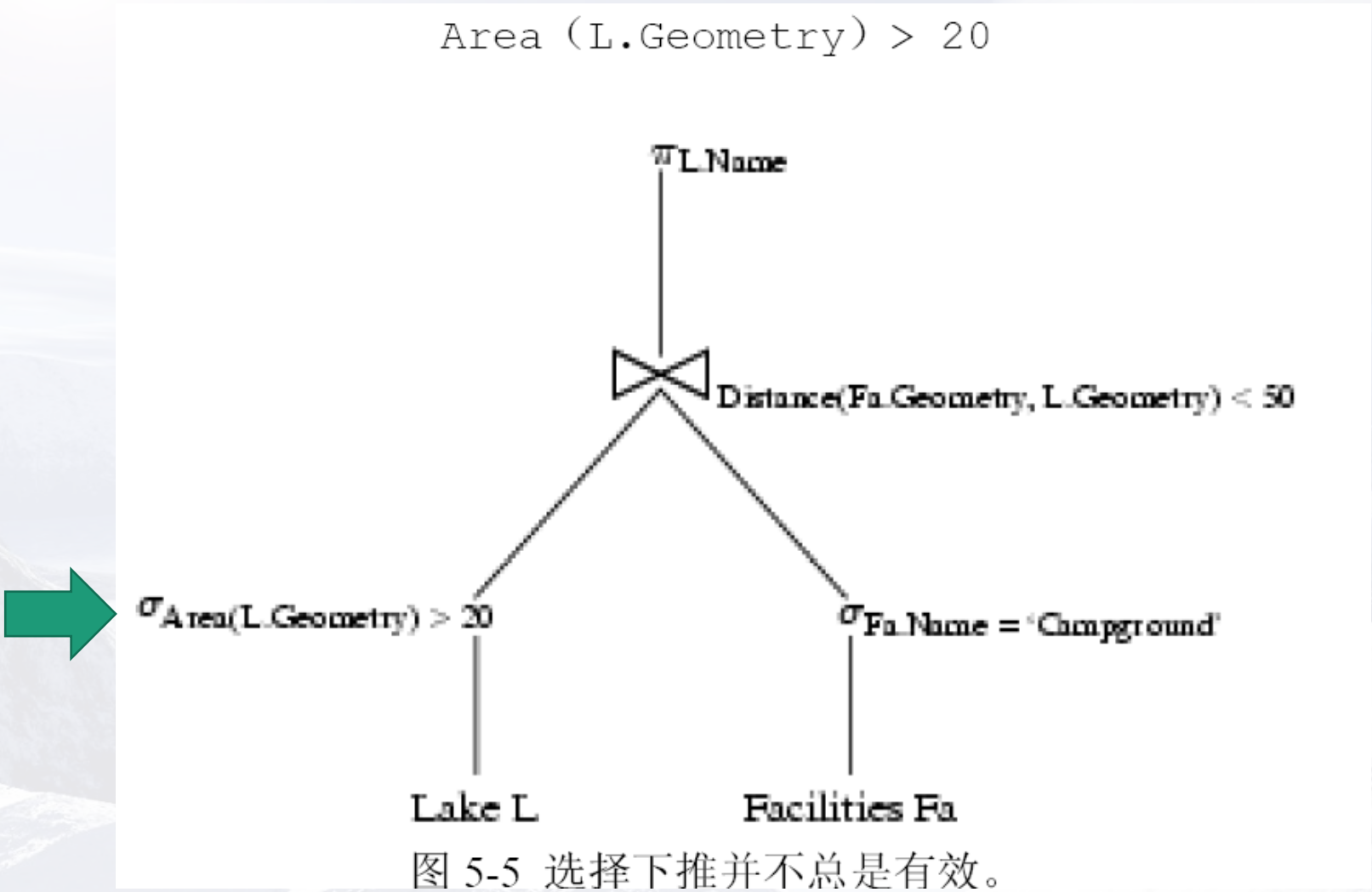


图 5-5 选择下推并不总是有效。

启发式优化

- 通过使用|><|的结合律，重新组织查询树，使得具有限制比较严格的选择运算的叶结点关系首先执行
- 将跟有选择条件的笛卡尔积运算替换成连接运算
- 将投影属性加以分解并在查询树上尽可能往下推，必要时可以引入新的投影运算
- 识别那些可用流水方式执行其运算的子树，并采用流水线方法执行之

分布式数据库的查询优化

- 全局表和本地表
- 分布与冗余
  - 分片策略（垂直分片、水平分片）
  - 空间网格划分与空间对象划分
- 数据的分布性产生了通信代价
  - 尽量减少网络传输
  - 消除查询运算无关的数据传输

分布式数据库的查询优化

- 半连接
  - 选择运算提前，减少需要传输的元组
  - 只传输主码或连接属性的不重复值
  - 连接完成后，只传输输出属性
- 全局空间索引
  - 建立在全局空间对象的索引
  - 采用已有空间索引机制
  - 一致性维护

分布式数据库的查询优化

- 冗余节点选择
  - 全局资源目录
  - 节点状态监控
  - 选择全局表的最小冗余全覆盖
- 动态负载均衡与容错
  - 节点性能评价
  - 全局性能与可用性评价

参考文献

- [TP311.13/261]空间数据库 = Spatial databases a tour (美) Shashi Shekhar, Sanjay Chawla著 谢昆青 ... 等译 北京 机械工业出版社 2004
- 北京大学计算机系数据库教研室
  - 数据库原理与技术讲义（杨冬青）
- 数据库系统概论
  - 岳丽华，丁卫群 编著
  - 科学出版社，000年