16337341(朱志儒)数据库系统作业5

12.1 第一趟：{(emu, 1), (kangaroo, 17), (wallaby, 21)}，

{(lion, 8), (platypus, 3), (wombat, 13)}，

{(meerkat, 6), (warthog, 4), (zebra, 11)}，

{(baboon, 12), (hornbill, 2), (hyena, 9)}；

第二趟：{(emu, 1), (kangaroo, 17), (lion, 8), (platypus, 3), (wallaby, 21), (wombat, 13)}；

{(baboon, 12), (hornbill, 2), (hyena, 9), (meerkat, 6), (warthog, 4), (zebra, 11)}；

第三趟：{(baboon, 12), (emu, 1), (hornbill, 2), (hyena, 9), (kangaroo, 17), (lion, 8), (meerkat, 6), (platypus, 3), (wallaby, 21), (warthog, 4), (wombat, 13), (zebra, 11)}。

12.3 由题可知：r1有800块，r2有1500块。

a.嵌套循环连接

|  |  |  |
| --- | --- | --- |
|  | 块传输 | 磁盘搜索 |
| 最坏的情况下 |  | 20000 + 800 = 20800 |
| 最好的情况下 | 1500 + 800 = 2300 | 2 |

b.块嵌套循环连接

|  |  |  |
| --- | --- | --- |
|  | 块传输 | 磁盘搜索 |
| 最坏的情况下 |  |  |
| 最好的情况下 | 1500 + 800 = 2300 | 2 |

c.归并连接

r1和r2已排序：

|  |  |  |
| --- | --- | --- |
|  | 块传输 | 磁盘搜索 |
| 最坏的情况下 | 800 + 1500 = 2300 | 800 + 1500 = 2300 |
| 最好的情况下 | 800 + 1500 = 2300 | 2 |

r1和r2未排序：

最坏的情况下：

对r1排序的块传输次数：；

磁盘搜索次数：；

对r2排序的块传输次数：；

磁盘搜索次数：；

归并两个关系的块传输次数：800 + 1500 = 2300；

磁盘搜索次数：800 + 1500 = 2300；

总的块传输次数：16000 + 30000 + 2300 = 48300；

磁盘搜索次数：14934 + 28000 + 2300 = 45234；

最好的情况下：

对r1排序的块传输次数：；

磁盘搜索次数：2 + 800 + 800 = 1602；

对r2排序的块传输次数：；

磁盘搜索次数：；

归并两个关系的块传输次数：800 + 1500 = 2300；

磁盘搜索次数：800 + 1500 = 2300；

总的块传输次数：3200 + 6000 + 2300 = 11500；

磁盘搜索次数：1602 + 3002 + 2300 = 6904；

d.散列连接

不需要递归划分：

|  |  |  |
| --- | --- | --- |
|  | 块传输 | 磁盘搜索 |
| 最坏的情况下 |  |  |
| 最好的情况下 |  |  |

需要递归划分：

|  |  |  |
| --- | --- | --- |
|  | 块传输 | 磁盘搜索 |
| 最坏的情况下 |  |  |
| 最好的情况下 |  |  |

12.5 将较小的关系存储在内存中，逐块读取较大的关系，将较大的关系作为外部关系执行嵌套循环连接，I/O操作的次数为，内存要求为页。

13.4 由于C是r2的码，r1参照r2的外码，所以的估计值最大为1000，又由于E是r3的码，r2参照r3的外码，所以的估计值最大为1000。

策略：在r2的属性C和r3的属性E上创建索引，对r1中的每个元组，使用r2的索引查找匹配r1中C值最多的元组，使用r3的索引查找至多一个与r2中的E唯一匹配的元组。

13.5 由，得到的两个估计值，取较小者1364为的估计。由，得到的两个估计值，取较小者10227为的估计。最佳的连接顺序为。

13.6

a. 使用索引查找building = “Watson”的第一个元组，从这个元组开始，按照指针链直到结束，检索所有的元组。

b. 按顺序扫描文件并选择的所有元组。

c. 由于

使用索引查找 building = “Watson”的第一个元组，从这个元组开始，按照指针链直到结束，检索所有的元组，选择的元组。

13.15 使用(dept\_name, building)属性上的索引，找到第一个具有(dept\_name = “Music”, building = “Watson”)属性的元组，然后按照指针检索连续的元组，选择满足条件(building < “Watson”, budget < 55000)的元组。

13.16

a. 根据等价规则1，可得：

b. 根据等价规则1和7.a，可得：

13.20 找到均匀划分两个直方图区间的最大单位u，将每个直方图划分为大小为u的区间，并假设区间内的值均匀分布。用值落入该区间的行数来代替nr和ns，用该区间的不同取值个数代替V(A, r)或V(A, s)，然后使用13.3.3中的方法计算估计值，最后将每个区间大小估计值加起来得到总的大小估计值。