

数据库系统 作业

3200105872 庄毅非

16.5 Consider the relations $r_1(A, B, C)$, $r_2(C, D, E)$, and $r_3(E, F)$, with primary keys A , C , and E , respectively. Assume that r_1 has 1000 tuples, r_2 has 1500 tuples, and r_3 has 750 tuples. Estimate the size of $r_1 \bowtie r_2 \bowtie r_3$, and give an efficient strategy for computing the join.

答案：首先，考虑 $r_1 \bowtie r_2$ ，由于 $r_1 \cap r_2 = C$ ，而 C 是 r_2 的主键，所以最终的元组个数不会超过 r_1 的元组的个数，也就是1000。接下来考虑 $(r_1 \bowtie r_2) \bowtie r_3$ ，由于 $(r_1 \bowtie r_2) \cap r_3 = E$ ，而 E 是 r_3 的主键，所以最终的元组的个数不会超过 $(r_1 \bowtie r_2)$ 的元组个数，而 $(r_1 \bowtie r_2)$ 的元组个数不会超过1000，所以最终的结果的元组的个数不会超过1000。

16.16 Suppose that a B+-tree index on (dept name, building) is available on relation department. What would be the best way to handle the following selection?

$\sigma_{(building < \text{"Watson"} \wedge (budget < 55000) \wedge (dept_name = \text{"Music"}))}(Department)$

答案：由于关系在 (dept name, building) 上具有索引，所以可以使用索引执行

$\sigma_{(building < \text{"Watson"} \wedge (dept_name = \text{"Music"}))}(Department)$ 这样一个范围查询，这样可以高效的利用索引进行查询，然后对于上述查询结果（记为 B ）做更进一步查询，也就是 $\sigma_{(budget < 55000)}(B)$ ，这样做可以高效的进行查询。