

16.5解:

连接规模: $r_1 \bowtie r_2 \bowtie r_3 = (r_1 \bowtie r_2) \bowtie r_3$

$\because r_1 \cap r_2 = C$ 是 r_2 的主码, $r_1 \bowtie r_2$ 的最大规模不超过 r_1 的大小, 即 1000.

同理. $r_2 \cap r_3 = E$ 是 r_3 的主码, $(r_1 \bowtie r_2) \bowtie r_3$ 最终的规模不会超过 r_1 , 即 1000.

最终结果最多有 1000 个元组.

计算策略: 为 r_2 中的属性 C 与 r_3 中的属性 E 分别创建一个索引.

对 r_1 中每一个元组, 先通过 r_2 的索引找其中与 r_1 的 C 值匹配的
元组, 再通过 r_3 的索引寻找与中间结果 E 匹配的元组.

16.6 解: 连接规模: 不存在主码.

$$r_1 \bowtie r_2 \text{ 规模为 } \frac{n_{r_1}}{1000} \times \frac{n_{r_2}}{\sqrt{(C, r_2)}} = \frac{15000}{11} = 1364$$

$$\text{~~并~~ } (r_1 \bowtie r_2) \bowtie r_3 \text{ 为 } n_{r_1 \bowtie r_2} \times \frac{n_{r_3}}{\sqrt{(E, r_3)}} = \frac{15000}{11} \times \frac{750}{100} = 10227.$$

即最终大约有 10227 个元组.

策略: $r_1 \bowtie r_2$ 的结果 ~~并未~~ 与 r_1, r_2 本身规模相当, 并未增大多少, 但 $r_2 \bowtie r_3$ 的规模远超他们自身, 导致 r_2 的计算成本.

所以应当先连接 r_1, r_2 , 再将结果与 r_3 连接