

## 第十二章作业答案

XX YYY

Email: xx@scut.edu.cn

华南理工大学软件学院

May 19, 2023



## 12-3

# Current Section

## 1 12-3

3. Let relations  $r1(A, B, C)$  and  $r2(C, D, E)$  have the following properties:  $r1$  has 20,000 tuples,  $r2$  has 45,000 tuples, 25 tuples of  $r1$  fit on one block, and 30 tuples of  $r2$  fit on one block. Estimate the number of block transfers and seeks required, using each of the following join strategies for  $r1 \bowtie r2$ :

- a. Nested-loop join.
- b. Block nested-loop join.
- c. Merge join

r1 needs  $20000 \div 25 = 800$  blocks, and r2 needs  $45000 \div 30 = 1500$  blocks. Let us assume  $M$  pages of memory. If  $M > 800$ , the join can easily be done in  $1500 + 800$  disk accesses, using even plain nested-loop join. So we consider only the case where  $M \leq 800$  pages.

a. Nested-loop join:

Using r1 as the outer relation we need  $20000 * 1500 + 800 = 30000800$  disk accesses, if r2 is the outer relation we need  $45000 * 800 + 1500 = 36001500$  disk accesses.

b. Block nested-loop join:

If  $r_1$  is the outer relation, we need  $\lceil \frac{800}{M-1} \rceil * 1500 + 800$  disk accesses,  
if  $r_2$  is the outer relation we need  $\lceil \frac{1500}{M-1} \rceil * 800 + 1500$  disk accesses.

c. Merge-join:

Assuming that  $r_1$  and  $r_2$  are not initially sorted on the join key, the total sorting cost inclusive of the output is

$B_s = 1500(2\lceil \log_{M-1}(1500/M) \rceil + 2) + 800(2\lceil \log_{M-1}(800/M) \rceil + 2)$  disk accesses. Assuming all tuples with the same value for the join attributes fit in memory, the total cost is  $B_s + 1500 + 800$  disk accesses.