Neha Maddali

***Homework 3-1***

1. (50 points) Derive the I/O costs of different join algorithms of relations R and S given the following variables, which you may or may not use all of them. Suppose that there is 1 page of results for the join. Ignore the CPU time cost. Please write down steps to explain your answer for full credits.

|R|=20: Number of tuples per page in R

|S|=20: Number of tuples per page in S

M=120: Number of pages in R

N=40: Number of pages in S

B=10: Number of available main memory in pages

* 1. (10 points) What is the minimal I/O cost of block nested loop join?

Use R as the outer relation.

Total cost = M + [M/(B-2)]\*N = 120 + (120 / (10-2)) \* 40 = 720

Use S as the outer relation.

Total cost = N + [N/(B-2)]\*M = 40 + (40 / (10-2)) \* 120 = 640 --> minimal cost

* 1. (10 points) What is the minimal I/O cost of simple nested loop join?

Use R as the outer relation

Total cost = M+M\*PR\*N = 120 + 120 \* 20 \* 40 = 96,120

Use S as the outer relation

Total cost = N+N\*PS\*M = 40 + 40 \* 20 \* 120 = 96,040 --> minimal cost

* 1. (10 points) What is the minimal I/O cost of indexed nested Loops Join​? (Suppose the cost of retrieving a matching tuple is 3, for both R and S)

Use R as the outer relation

Total cost = M+M\*PR\*(cost of retrieve) = 120 + 120 \* 20 \* 3 = 7,320

Use S as the outer relation

Total cost = N+N\*PS\*(cost of retrieve) = 40 + 40 \* 20 \*3 = 2,440 --> minimal cost

* 1. (10 points) What is the minimal I/O cost of grace hash join?

Each partition fits in the B-2 pages

I/O cost for a read and write is the same

Ignore the cost of writing the join results

Total cost = 3\*(120 + 40) = 480

* 1. (10 points) What is the minimal I/O cost of Sort-Merge Join​? (Suppose the join is on their primary keys which are sorted already)

Total cost = M+N = 120 + 40 = 160

|  |
| --- |
|  |

**Submission Instruction**

*Please use Microsoft Words or other tools to type your answer. Don't handwrite. Submit your work in pdf through your Canvas account.*