

Khushveen Kaur Umra

COMS 363

23rd October 2021

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

- a) (10 points) What is the minimal I/O cost of block nested loop join?

→ If R is the outer, then the total cost will be:

$$\begin{aligned} &= b_S + (b_R * \text{ceil}(b_S/B-2)) \\ &= 120 + (20 * \text{ceil}(20 / (10 -2))) \\ &= 120 + (20 + \text{ceil}(15)) \\ &= 120 + 1800 \\ &= \underline{\underline{1920}} \end{aligned}$$

If S is the outer, then the total cost will be:

$$\begin{aligned} &= 40 + (120 * \text{ceil}(40 / (10 -2))) \\ &= 40 + (120 + \text{ceil}(5)) \\ &= 40 + 600 \\ &= \underline{\underline{640}} \end{aligned}$$

- b) (10 points) What is the minimal I/O cost of simple nested loop join?

→ For a simple nested loop join, the total cost will be:

$$\begin{aligned} &= b_S + b_S * b_R \\ &= 40 + (40 * 120) \end{aligned}$$

$$= 40 + 4800$$

$$= \underline{4840}$$

c) (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)

→ For an indexed nested loop join, the total cost will be:

$$= 120 + (120 * 40)$$

$$= 120 + 4800$$

$$= \underline{4920}$$

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

→ For a grace hash join, the total cost will be:

$$= 3(120 + 40)$$

$$= 3(160)$$

$$= \underline{480}$$

e) (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)

→ For a sort-merge join, the total cost will be:

$$= 2bS (1 + \text{ceil}(\log_{b-1} \text{ceil}(bS / B))) + 2bR (1 + \text{ceil}(\log_b \text{ceil}(bR / B))) + bS + bR$$

$$= 2 * 120 (1 + \text{ceil}(\log_{51} (51) (200 / 52))) + 2 * 40 (1 + \text{ceil}(\log_{51} \text{ceil}(1000 / 52))) + 120 + 40$$

$$= 240 (1 + \text{ceil}(\log_{51} 4)) + 80 (1 + \text{ceil}(\log_{51} 20)) + 160$$

$$= 480 + 160 + 160$$

$$= \underline{800}$$