V. PRANEETH CHANDR

Assignment-2

180101085

1. Given, $B(R) = B(S) = 10000$, $M = 1000$

Here $B(R) = B(S)$, so we can take any of R and S into main memory and compare it with the tuples of the other relation.

Let us take a relation S into main memory and compare it with R.

So the outer most loop in nested-loop join takes $\frac{B(S)}{M-1}$ iterations $= \frac{10000}{999} = 10.01$, but number of iterations cannot be a fraction. So number of iterations should be taken as 11.

And the inner loop takes $B(R)$ iterations. $= 10000$.

$\Rightarrow$ Total disk I/O cost $= B(S) + \left\lceil \frac{B(S)}{M-1} \right\rceil \times B(R)$

$= 10000 + 11 \times 10000$

$= 120000$.

2. (3,4).

1. Set union

One pass algorithm for set union is not feasible in this case because both $B(R)$ and $B(S) > M$.

Two-pass algorithm for set union is feasible in this case because $\sqrt{B(S)+B(R)} < M$.

In two-pass algorithm for set union, we do three steps:

1. Read data from disk to main memory, which takes $B(R) + B(S)$ disk I/O's.
2. Make sorted sublists of R and S and write them in disk, which takes disk I/O cost of $B(R) + B(S)$.
3. Repeatedly find the first remaining tuple t among all the buffers. Copy t to the output and remove all the copies of t from the buffers. If a buffer becomes empty, reload it with the next block from its sublist.

   This operation takes disk I/O cost of $B(R) + B(S)$.

So total disk I/O cost $= 3(B(R) + B(S))$.

$= 3(10000 + 10000)$

$= 60000$.

2. Simple sort-join

Simple sort-join is feasible because $\max(B(R), B(S)) < M^2$

In this algorithm, we first use two-phase multi way merge sort with Y as the sort key, if we are joining $R(X, Y)$ and $S(Y, Z)$; this operation takes $4(B(R) + B(S))$ disk I/O operations.

Now, from sort list, find the least value of the join attribute Y that is currently at the front of the block for R and S. Output all the tuples that can be formed by joining tuples from R and S, with a common Y-value.

This operation takes $B(R)+B(S)$ disk I/O operations.

$\Rightarrow$ Total disk I/O operations $= 5(B(R)+B(S))$

$= 5 \times (10000 + 10000)$

$= 100000.$

5.

$$W \bowtie X \bowtie Y \bowtie Z = ((W \bowtie X) \bowtie Y) \bowtie Z.$$

$$\Rightarrow T(W \bowtie X \bowtie Y \bowtie Z) = T((W \bowtie X) \bowtie Y) \bowtie Z)$$

$$T(W \bowtie X) = \frac{T(W) \cdot T(X)}{\max(V(W,b), V(X,b))} = \frac{100 \times 200}{60}$$

$$T((W \bowtie X) \bowtie Y) = \frac{T(W \bowtie X) \cdot T(Y)}{\max(V(X,c), V(Y,c))} = \frac{100 \times 200}{60} \times \frac{300}{100}$$

$$\Rightarrow T(W \bowtie X \bowtie Y \bowtie Z) = \frac{T(W \bowtie X \bowtie Y) \cdot T(Z)}{\max(V(Y,d), V(Z,d))} = \frac{\cancel{100} \times 200}{\cancel{60}} \times \frac{\cancel{300}^{5}}{\cancel{100}} \times \frac{400}{\cancel{80}_{10}}$$

$= 8000$ tuples.