CSC 244 Programming Assignment-2
NATURAL JOIN
RAMYA VELLORE RAMESH
SAC ID:219123243

1.The description of your algorithm.

We have Relation R(A,B,Y) & S(Y,Z,X) where secondary index is assigned based on the Y value.Relation S is sorted based on Y value.
Load the blocks of R and S into the Memory.
Load the index of R as key in the hash map i.e the Memory.
For each tuples in the block of Relation S
   If secondary index of R matches with Y value of the tuples of Relation S
   Then we join the tuples of R and S
   We output the joined tuples into the output buffer
WriteToOutputBuffer

 If size(outputbuffer)< 3

   put the result to the outputbuffer

 Else

   Move the current buffer value to disk

   put the result to outputbuffer

2.The description of the simulation of the memory and disk storage

Assuming Relation R and S are in disk,S has 6 tuples which is more than the memory(size=5),R has 8 tuples which is more than the size of the memory(size=5) .Relation S and R are assumed to be in two-dimensional array.Relation R and S contains 3 Attributes for each tuple.(B(R) > M and B(S)>M)

- 2-Dimensional arrays are used for both the relations R and S for disk storage.
- ArrayList is used as memory to store the block values of Relation R and S from the disk.
- HashMap is used to store the tuples of Relation R as a key-value pair.The secondary index is stored as key and the tuple of the block is stored as the value.The secondary index size of R which is stored in hash map as key is less than M.
- If the primary key of S i.e the Y value of each tuple in Relation S is equal to secondary index of R which is the key in hashmap ,then we join the tuples of the two relations R and S
- An ArrayList is used to store the Natural Join of tuples of Relation R and S to the output buffer.
- When the output buffer is full, the buffer values are copied to the diskoutput .ArrayList is used as disk output.

3.Example used to test the program
R={{2,4,1},{3,9,5},{6,2,5},{2,5,9},{9,1,2},{13,2,4},{7,5,8},{1,5,6}};
S= {{1,3,2},{2,4,6},{2,6,0},{3,2,4},{4,8,6},{5,2,3}};

4. The input and output of your program using this example. Input can be either hardcoded or from the command line.

Input is hardcoded.
Output:

```
Buffer contents after writing to Outputbuffer:
[2, 4, 1, 3, 2]
Buffer contents after writing to Outputbuffer:
[2, 4, 1, 3, 2]
[9, 1, 2, 4, 6]
Buffer contents after writing to Outputbuffer:
[2, 4, 1, 3, 2]
[9, 1, 2, 4, 6]
[9, 1, 2, 6, 0]
Buffer is full writing the buffer contents to disk:
[2, 4, 1, 3, 2]
[9, 1, 2, 4, 6]
[9, 1, 2, 6, 0]
Buffer contents after writing to Outputbuffer:
[13, 2, 4, 8, 6]
[6, 2, 5, 2, 3]
The Natural Join of R and S
[[2, 4, 1, 3, 2], [9, 1, 2, 4, 6], [9, 1, 2, 6, 0], [13, 2, 4, 8, 6], [6, 2, 5, 2,3]]
```