Assignment 3: Query Processing and Query Optimization

Question 1: Two relational algebra expressions are said to be equivalent if the two expressions generate the same set of tuples. To obtain a relational algebra expression from another, we use equivalence rules, as discussed in *ch13.pptx*.

Consider the following SQL, which is to show the student name and the course

titles if the student is from the SEEM department and took the courses in 2021.

一張含有 文字 的圖片

自動產生的描述

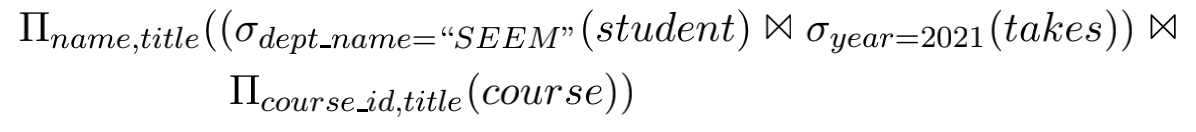
1. Show the initial relational algebra for this SQL. We discussed it in *Chapter 6*.

(2) Let the relational algebra you get in (1) be E. Show how you get the relational algebra E’ below from E using equivalence rules step by step.

一張含有 文字 的圖片

自動產生的描述

(3) Show how you get the relational algebra E’’ below from E’ using equivalence rules step by step.



一張含有 文字 的圖片

自動產生的描述

1. Estimate the number of tuples and the number of blocks for the result of each of the 3 natural joins. (You have already known them for the first join.)
2. Assume that (a) the block nested-loop join algorithm is used, (b) the relation on the right-hand side of  is selected as the outer relation in the algorithm, and (c) 3 out of the 5 blocks in the buffer are used for the outer relation. Suppose you need to read blocks to process a join, and you have to write the result back to the disk. Give the estimated number of block accesses for reading and writing for each of the 3 natural joins and explain how you get it in detail.
3. Consider that you can develop a new block nested-loop join algorithm that can do 3 natural joins for 4 input relations together. Show its pseudo-code.
4. Based on your new algorithm, give the estimated number of block accesses for reading and writing over the given buffer M for doing the new join, and explain how you get it in detail.
5. Now you have two options to process this sequence of joins. One is to do natural joins 3 times, with each being on two input relations. The other is to do 3 natural joins together by one join algorithm. Which is better? Explain your answer.

Hint: You may have some reasonable assumptions for your answer.