COMP2411 Database Systems

Assignment 2 (due: 21/Nov/2022 @11:59pm!)

Question A.

1) What are the advantages and disadvantages of hash functions relative to B-tree indices? How might the type of index available influence the choice of a query-processing strategy? **[15 marks]**

2) When is it preferable to use a dense index rather than a sparse index? **[8 marks]**

3) What is the difference between primary index and a secondary index? **[7 marks]**

4) If a hash structure is used on a search key for which range queries are likely, what property should the hash functions have? **[10 marks]**

Question B.

Consider the relations:

STUDENT(SNAME, SID, BDATE, ADDRESS, DNUM)
COURSE(CNAME, CID, LEVEL, LECTURER)
COURSE_TAKING(STUDENTID, COURSEID, GRADE)

as well as the following SQL query:

```
SELECT SNAME, CNAME, GRADE
FROM STUDENT, COURSE_TAKING, COURSE
WHERE LEVEL>2 AND STUDENTID=SID
        AND COURSEID=CID AND DNUM = 24;
```

a) Draw a canonical query tree for the above SQL query. **[9 marks]**
b) Apply the optimization rules to the above query tree and come up with the most optimized query tree using those rules. State the necessary assumptions for your decision. **[21 marks]**

Question C.

[30 marks]

1) Do you agree that a nonrecoverable schedule results in a loss of transaction atomicity? Explain your answer.

2) When the system recovers from a crash, in what order must transactions be undone and redone?

3) Suppose that the system crashes during the time it is recovering from a prior crash. When the system again recovers, what action must be taken?