

M.Sc. Semester-IV Examination 2017
Computer Science
Course : MCSC-42
(Advanced DBMS)

Time : 3 Hours

Full Marks : 40

Questions are of value as indicated in the margin

Attempt Question No.1 and **any four** from the rest

1. (a) Draw the State diagram of a transaction.
(b) What do you mean by sublinear scaleup in parallel database system?
(c) How natural join operation of two relations R & S differs from semijoin operation?
(d) How join of two relations R and S placed in two different sites may be computed in one of the two sites without transforming entire relation from one site to other but using semijoin operation. 4×2=8
2. (a) State along with justification whether the following schedule of three transactions (T1, T2, and T3) is conflict serializable, view serializable, recoverable and cascadeless. Further answer along with justification whether it may be generated with Timestamp (TS) without the Thomas Write Rule and Timestamp with the Thomas' Write Rule. Assume $TS(T1) < TS(T2) < TS(T3)$. Finally draw the precedence graph for the schedule.
T1 : R(X), T1: R(Y), T1: W(Y), T2: R(X), T2: W(X), T2: R(Y), T3: R(X), T1: R(Z), T2: R(Z), T2: W(Z), T3: R(Y), T1: W(Z), T1 : Commit, T3: R(Z), T3: W(Z), T2 :Commit, T3: Commit. R and W represent Read and Write , respectively. 4+4=8
3. (a) Explain Hash indices.
(b) Explain log based recovery for concurrent execution of transactions for both immediate and deferred update strategies. 3+5=8
4. (a) Suppose you have a file with 5000 pages and you have five buffer pages. Answer the following questions using external merge sort algorithm. How many runs will you produce in the first pass? How many passes will it take to sort the file completely? What is the total I/O cost of sorting the file?
(b) How to evaluate disjunctive selection (complex selection)?
(c) Briefly Explain the hash join strategy and measure the cost of hash join. 3+2+3=8
5. (a) Briefly explain different steps in query processing.
(b) Estimate the cost of disjunctive selection (complex selection)
(c) Consider the relations $r_1(A, B, C)$, $r_2(C, D, E)$, and $r_3(E, F)$, with primary keys A, C, and E, respectively. Assume that r_1 has 2000 tuples, r_2 has 3000 tuples, and r_3 has 1000 tuples. Estimate the size of $r_1 \bowtie r_2 \bowtie r_3$, and give an efficient join strategy for computing the join. 3+2+3=8
6. (a) Briefly explain different parallel database architectures.
(b) How is 3PC protocol advantageous over 2PC?
(c) Briefly explain evaluation of join operation with parallelism. 3+2+3=8
7. Write short notes on (**any two**) : 2×4=8
 - (a) Equivalence rules for relational algebra
 - (c) Cost-based optimization vs. heuristic optimization
 - (d) Data fragmentation
 - (e) Data warehousing