00750

Sri Lanka Institute of Information Technology

B. Sc. Honours Degree
in
Information Technology

Final Examination

Year 3, Semester I

IT3020 – Database Systems

Duration: 2 Hours

June 2022

Instructions to Candidates:

- This paper is preceded by a 10-minute reading period. The supervisor will indicate when answering may commence.
- This paper contains 4 questions. Answer All Questions.
- Use the booklets given to provide answers.
- Total marks for the paper will be 100.
- A mark for each question is mentioned in the paper.
- This paper contains 6 pages with the Cover Pager.
- Electronic devices capable of storing and retrieving text, including calculators and mobile phones are not allowed.

Question 1 (25 marks)

Consider the following object relational schema for a database of ABC Airline.

Object types:
Customer_t (*cid*: char(7), *name*: varchar(15), *phone*: varchar(10))
Schedule_t (*flightno*:char(6), *deptime*:char(6), *source*: varchar(12), *destination*: varchar(12), *capacity*: integer)
Booking_t (*passenger*: ref customer_t, ticketno: *char*(10), *fare*: float)
Passenger_list table of booking_t
Flight_t (*flight*: ref schedule_t, *depdate*: date, *passengers*: passenger_list)

Table:
Customers of customer_t (cid primary key)
Schedules of schedule_t (flightno primary key)
Flights of flight_t (flight not null references schedule, depdate not null)
Nested table passengers store as passengerlist_ntab

The **customers table** of customer_t has attributes of customer id (*cid*), *name*, and *phone*. It contains tuples for all customers. **Schedules table** of schedule_t contains tuples for all scheduled flights, and has the attributes flight number (*flightno*), departure time at source (*deptime*), *source* and *destination* of the flight, and seating capacity. The **flights table** of flight_t records the information of an actual flight on a particular date and consists of attributes for flight reference, date of departure (*depdate*), and a nested table of passenger details. The nested table contains passenger reference, ticket number (*ticketno*), and the fare charged for the given flight. The attribute types are specified in the type descriptions above. The primary keys and referential constraints are shown in the table schema.
Note that date literals can be specified in SQL statements as 'DD-MM-YY'

(a) Write Oracle OR-SQL statements for the following queries (use columns of REF type instead of joins to link tables).

(i) Find the cheapest flight from Colombo to London in 15th of October 2021.

(4 marks)

(ii) Find the total fare collected from passengers booked on each flight departing London for Colombo on 25 December 2021. Display the flight number, and the total fare.

(5 marks)

(b) Add a new passenger to flight number 'QR09' departing on 25th of May, 2022. This passenger exists in the customers table with *cid* of 'AW35462'. The ticket number for this passenger is 'LAST10210' and the fare LKR85000.

(5 marks)

(c) It is required to add a member method called *passcount* to calculate the number of passengers on a flight. Write Oracle SQL statements to modify the object type flight_t by adding this method specification.

(7 marks)

(d) Using the method defined above, write an Oracle SQL statement to display the number of available seats on flights from Colombo to Singapore on 1st of June 2022. The number of

available seats is the difference between the capacity of a flight and the current number of passengers booked to travel. Display the flight number, departure time, and the number of available seats.

(4 marks)

Question 2 (25 marks)

(a) Compare and contrast the heap and sequential file organization approaches.

(3 marks)

(b) Consider a employees relation containing records with employee id, name, age, salary and department number. Assume that the records are stored in a sequential file where records are ordered based on the employee's id.
"The aforementioned file makes it inefficient to get records with salaries ranging from Rs. 20,000 to Rs. 30,000."
State whether the statement above is true or false with reasons.

(5 marks)

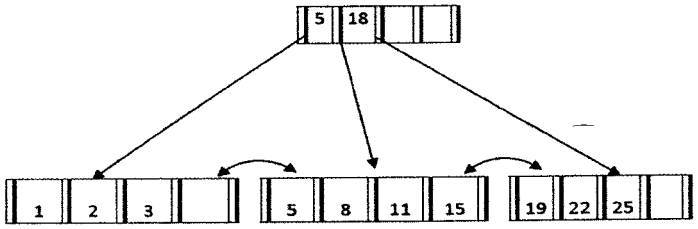
(c) Briefly explain two disadvantages that exist in static hashing.

(3 marks)

(d) Explain the terms dense and sparse index.

(3 marks)

(e) Consider the B+ tree index below.



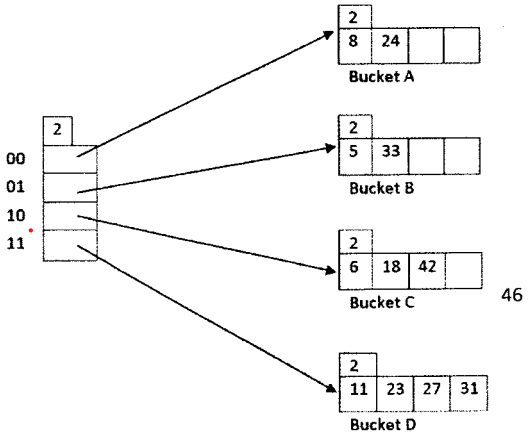
i. Illustrate the tree after inserting 31 and 33.

(3 marks)

ii. illustrate the tree after deleting 2 and 3 from the original tree.

(3 marks)

(f) Consider the Hash index below. Illustrate the hash index after inserting 50.



46

(5 marks)

Question 3 (25 marks)

(a) What is the goal of query optimization? Why is it important?

(3 marks)

(b) Consider the following relational schema and SQL query.

Suppliers (*sid*: integer, *sname*: char(20), *city*: char(20))
Supply (*sid*: integer, *pid*: integer)
Part (*pid*: integer, *pname*: char(20), price: real)

```
SELECT s.sname, p.pname
FROM Suppliers s, Parts p, Supply y
WHERE s.sid = y.sid AND y.pid = p.pid AND p.price ≤ 1000
```

i. How many different join orders, assuming that cross products are disallowed, will a System R style query optimizer consider when deciding how to process the given query? List each of these join orders.

(3 marks)

ii. What indexes might be of help in processing this query? Explain briefly.

(4 marks)

iii. How does adding DISTINCT to SELECT clause affect the plans produced?

(3 marks)

iv. Estimate the I/O cost of retrieving records from Parts table that contains price less than or equal 1000 rupees. Assume that 15% of tuples satisfy the selection criteria. There are 3300 pages in the Parts table, with 50 records each page. The clustered B+ tree index on price attribute is the only index available in the Parts table. This index takes up 1/3 of the table's space.

(8 marks)

v. Estimate the I/O cost of sorting 200 pages of Suppliers table using 10 buffer frames.

(4 marks)

Question 4 (25 marks)

(a) Why do DBMSs interleave actions of multiple transactions? Briefly explain your answer.

(1 marks)

(b) Show a schedule that is unrecoverable? Briefly explain the schedule.

(3 marks)

(c) Briefly explain the terms *deadlock prevention* and *deadlock detection*. Explain approaches for each methodology.

(3 marks)

(d) Consider the following part of the schedule.

T_1	T_2	T_3	T_4
S(A)			
R(A)			
	X(B)		
	W(B)		
S(B)			
			S(D)
			R(D)
	X(D)		
		S(C)	
		R(C)	
			X(A)

Assume that Transaction T_i is higher priority than transaction T_{i+1}(i.e. transaction T_1 has higher priority than T_2; T_2 has higher priority than T_3; and T_3 has higher priority than T_4).

i. Draw a **Wait-For-Graph** for the given schedule above.

(2 marks)

ii. What should do next for deadlocks detection approach to break the deadlock after identify it?

(1 mark)

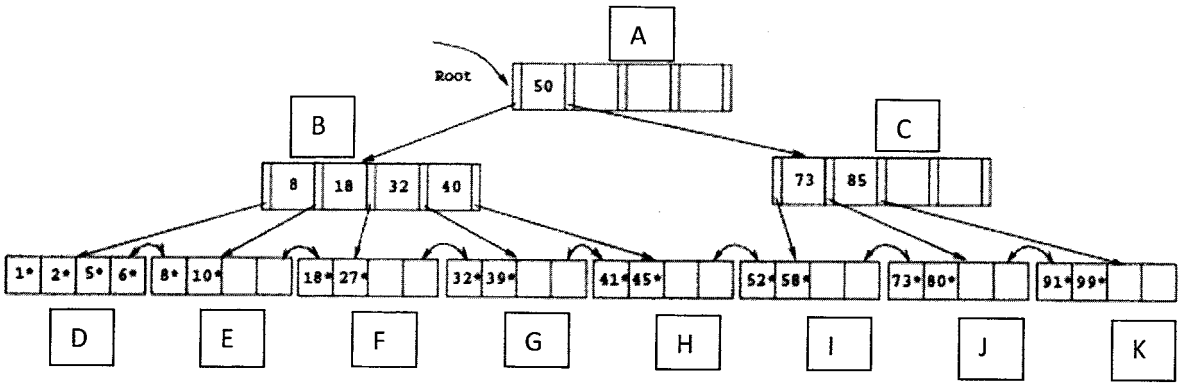
iii. Draw the schedule again considering deadlock prevention algorithm:
Wait-Die approach

(4 marks)

(e) Briefly explain Simple Tree Locking algorithm.

(4 marks)

(f) Consider the following B+ tree. Follow *Simple Tree Locking Algorithm* and specify when and what lock get and release to do the followings.

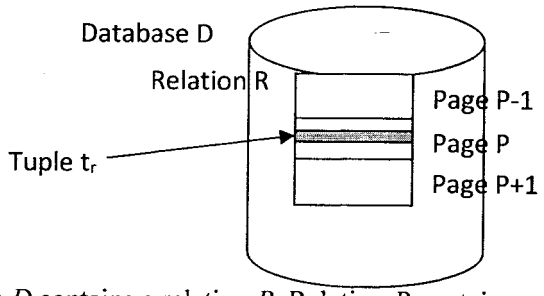


i. Search 11. (1 marks)

ii. Delete 99. (1.5 marks)

iii. Insert 30. (1.5 marks)

(g) Consider the following scenario:



Database *D* contains a relation *R*. Relation *R* contains a page *P*. Page *P* contains a tuple t_r.

Assume that multiple granularity locking scheme is used. Describe the locks acquired when reading all tuples of page P.

(3 marks)

- End of the Question Paper -