

1. (a) Find the result for the following: X/Y_1 , X/Y_2 and X/Y_3 , where X, Y_1 , Y_2 and Y_3 are relations and '/' indicates division operation performed between two relations. 4

X=	Sno.	Pno.	Y ₁ =	Pno
	N ₁	P ₁		P ₂
	N ₁	P ₂		
	N ₁	P ₃		Pno
	N ₁	P ₄		P ₂
	N ₂	P ₁		P ₄
	N ₂	P ₂		
	N ₃	P ₂		Pno
	N ₄	P ₂		P ₁
	N ₄	P ₄		P ₂
			Y ₂ =	P ₄
			Y ₃ =	P ₄

(b) Consider a Bank database taking care of accounts (saving, current, fixed deposit, flexible etc.) and loans (house, auto, business). Suppose it is necessary to keep track of each account transaction (deposit, withdrawal, checks, ATMs etc.). Draw a graphical object database schema and object database language for the above database. State any assumption you make about additional requirements. 10

(c) Discuss Fragmentation in distributed database with suitable example. 6

2. (a) Consider the following decompositions for the relation schema R.

4+3

$R = \{A, B, C, D, E, F, G, H, I, J\}$ and the set of functional dependencies
 $F = \{\{A, B\} \rightarrow \{C\}, \{A\} \rightarrow \{D, E\}, \{B\} \rightarrow \{F\}, \{F\} \rightarrow \{G, H\}, \{D\} \rightarrow \{I, J\}\}$

- (i) Determine whether each decomposition has the dependency preservation property and the lossless join property with respect to F.
 - (ii) Also determine which normal form each relation in the decomposition is in.
- (b) Illustrate with an simple example how an efficient cost function method can be devolpoed for the JOIN operation.
- (c) Discusss some important advantages and disdvantages of NoSQL database.

7

6

3. (a) A file $r=36,000$ STUDENT records of fixed length. Each record has the following fields: Name(25 bytes), Ssn(8 bytes), Address (40 bytes), PHONE (9 bytes), Birth_date (8 bytes), Sex(1 byte), Major_dept_code(5 bytes), Minor_dept_code(5 bytes), Class_code(3 bytes, integer), and Degree program (3 bytes). An additional byte is used as a deletion marker. The file is stored on disk whose parameters are given in the following: block size $B=1024$ byte, inter block gap size $G=64$ bytes, number of block per track=18, number of tracks per surface=420. A disk pack consists of 20 double-sided disks.

10

- (i) Calculate the record size R in bytes.
- (ii) Calculate the blocking factor bfr and the number of file blocks b , assuming an unspanned organization.
- (iii) Calculate the average time it takes to find a record by doing a linear search on the file if (i) the file blocks are stored contiguously, and double buffering is used. (ii) the file blocks are not stored contiguously.
- (iv) Assume the file is ordered by Ssn, by doing a binary search calculate the time it takes to search for a record given its Ssn values.

(b) How BASE property counters the ACID property.

10

4 (a) Consider the following schema.

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_Ssn	Dno
-------	-------	-------	------------	-------	---------	-----	--------	-----------	-----

DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_Ssn	Mgr_start_date
-------	----------------	---------	----------------

DEPT LOCATIONS

<u>Dnumber</u>	<u>Dlocations</u>
----------------	-------------------

PROJECT

Pname	<u>Pnumber</u>	Plocation	Dnum
-------	----------------	-----------	------

WORKS ON

<u>Essn</u>	<u>Pno</u>	Hours
-------------	------------	-------

DEPENDENT

Essn	<u>Dependent_name</u>	Sex	Bdate	Relationship
------	-----------------------	-----	-------	--------------

Figure 1

Suppose that the *EMPLOYEE* file of an organization has $r_E=10,000$ records stored in $b_E=2000$ disk blocks with blocking factor $bfr_E=5$ records/block and the various access paths (i) A clustering index on salary with levels $x_{salary}=3$ and average selection cardinality $S_{salary}=20$. (ii) A secondary index on the key attribute $x_{Ssn}=4$ ($S_{Ssn}=1$). (iii) A secondary index on Sex with $x_{sex}=1$ and average selection cardinality $S_{sex}=5000$. Illustrate how an efficient cost function method can be developed for the following SQL query $\sigma_{Dno=5 \text{ AND } SALARY>3000 \text{ AND } Sex='F'}(EMPLOYEE)$. 5

(b) Draw the initial query tree for each of these queries and then show how the query tree is optimized by query optimization algorithm.

(i) SELECT Fname, Lname, Address FROM EMPLOYEE, DEPARTMENT WHERE Dname='Research' AND Dnumber=Dno. 5.5

(ii) SELECT E.Fname, E.Lname, S.Fname, E.Lname FROM EMPLOYEE AS E, EMPLOYEE AS S WHERE E.Super_ssn=S.Ssn. 5.5

(c) Discuss Multivalued dependency with an example 4

- 5.(a) What is the difference between persistent and transient objects? 5
How persistence is handled in typical OO database systems?
- (b) What are the trade offs and benefits of using object oriented database management system over relational database management systems. 5
- (c) Discuss in Transaction Processing Monitor architecture in detail. 10