2024

COMPUTER SCIENCE

Paper : CSMC-201

(Advanced Database Management System)

Full Marks : 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

1. Answer *any five* of the following : 2×5

(a) Distinguish between immediate update and deferred update.

(b) Explain BCNF with the help of a suitable example.

(c) Why cascading schedule is not desirable?

(d) Define view serializability.

(e) A relation R(A, B, C, D, E) and FDs are A → BC, CD → E, B → D, E → A. Find all the candidate keys.

(f) State two disadvantages of the two-phase locking protocol.

(g) What is fragmentation transparency in a distributed database management system? Write one importance of it.

2. Answer *any five* of the following : 4×5

(a) What is a deadlock in a transaction? How do we prevent such deadlock?

(b) Differentiate between heap and sorted file organization. Also, give an example of each of them.

(c) Suppose a book file contains 20000 records stored in 4000 blocks. For nonlinear search, assume the level is 4. Find out cost of following select operations for any two searching techniques :

(i) $\sigma_{ID=003}$ (CATALOG)

(ii) $\sigma_{year>1995}$ (CATALOG)

where *year* and *ID* are attributes, and *CATALOG* is the relation.

(d) What do you understand by unstructured data handling? Explain one of the techniques for it.

(e) Write the following query for local transparency and location transparency level :

Select NAME from EMP where EMPNUM = 'E003'.

Where relation is :

EMP(<u>EMPNUM</u>, NAME, SAL, TAX, MGRNUM, DEPTNUM).

(f) Explain the Key-value database with the help of a suitable example.

(g) Insert the following data using an extendible hashing technique :

16, 4, 6, 22, 24, 10, 31, 7, 9, 20, 26.

Answer *any four* questions.

3. (a) Explain Armstrong's Axioms with suitable examples.

(b) Write the rule of 3NF. Consider the relation R = (A, B, C, D, E, F, G, H, I, J) and the Functional dependencies are following : FD = {AB → C, A → DE, B → F, F → GH, D → IJ}. Decompose R into 3NF. 5+5

4. (a) Discuss the deferred update techniques of recovery.

(b) What are the advantages and disadvantages of this technique?

(c) Why is it called the NO-UNDO/REDO method? 5+3+2

5. (a) What is conflict serializability? Explain with an example.

(b) Consider the three transactions, $T1$, $T2$, and $T3$, and the schedules S below. Draw the serializability (precedence) graph for S and state whether each schedule is serializable or not. Write the equivalent serial schedule(s) if a schedule is serializable.

$T1$: $r1(X)$; $r1(Z)$; $w1(X)$;

$T2$: $r2\ (Z)$; $r2\ (Y)$; $w2\ (Z)$; $w2\ (Y)$;

$T3$: $r3\ (X)$; $r3\ (Y)$; $w3\ (Y)$;

S : $rl\ (X)$; $r2\ (Z)$; $r1\ (Z)$; $r3\ (X)$; $r3\ (Y)$; $wl\ (X)$; $w3\ (Y)$; $r2\ (Y)$; $w2\ (Z)$; $w2\ (Y)$. 5+5

6. (a) What is heuristic query optimization?

(b) Optimize the following query using heuristic query optimization technique :

Select *Lname* from EMPLOYEE, PROJECT, WORKS_ON where *Pname* = *'Aquaries'* and *PROJECT. PNo.* = *WORKS_ON. PNo.* and *EMPLOYEE. SSN* = *WORKS_*ON. *SSN* and *Bdate>* *'1957-12-31'*.

Where relations are as follows :

EMPLOYEE (Fname, Lname, SSN, Bdate, Add, Gender, Salary)

PROJECT (Pname, PNo., Plocation, Dnum)

WORKS_ON (SSN, Pno., Hours). 2+8

7. (a) Explain the reference architecture of the distributed database with the help of a suitable diagram.

(b) Explain the advantages of a distributed database over a centralized database. 5+5

8. Consider the global relations :

PATIENT (NUMBER, NAME, SSN, AMOUNT-DUE, DEPT, DOCTOR, MED-TREATMENT)

DEPARTMENT (DEPT, LOCATION, DIRECTOR)

STAFF (STAFFNUM, DIRECTOR, TASK)

Define their fragmentation as follows :

(a) DEPARTMENT has a horizontal fragmentation by LOCATION, with two locations; each department is conducted by one DIRECTOR.

(b) There are several staff members for each department, led by the department's director. STAFF has a horizontal fragmentation derived from that of DEPARTMENT and a semi-join on the DIRECTOR attribute. Which assumption is required in order to assure completeness and disjointness?

(c) PATIENT has a mixed fragmentation : attributes NUMBER, NAME, SSN, and AMOUNT-DUE constitute a vertical fragment used for accounting purposes; attributes NUMBER, NAME, DEPT, DOCTOR, and MED-TREATMENT constitute a vertical fragment used for describing cares. This last fragment has a horizontal fragmentation derived from that of DEPARTMENT and a semi-join on the DEPT attribute. Which assumption is required in order to assure completeness and disjointness?

(d) Give also the reconstruction of global relations from fragments. 10

S(2nd Sm.)-Computer Science/CSMC-202

2024

COMPUTER SCIENCE

Paper : CSMC-202

(Advanced Operating Systems)

Full Marks : 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Answer ***question nos.*** 1 and 2 and ***any four*** questions from the rest.

1. Answer ***any five*** questions : 2×5

(a) Define access transparency and name transparency for a distributed system.

(b) State the condition for termination detection in a distributed system.

3 2 1 1

(c) Arrange the time stamps taken using the Lamport's logical clock model [⟨3,8⟩, ⟨3,5⟩, ⟨2,7⟩, ⟨5,1⟩].

(d) Which of the following is the smallest time stamp according to Vector clock model?

(i) [2,5,1,3] (ii) [1,3,0,2] (iii) [1,5,2,3] (iv) [1,3,1,2].

(e) Which of the following defines Lamport's second implementation rule for clock, where C(X) is clock for node X and ∂ is arbitrary delay for transmission? If A is the event of sending message *m* from node S and B is the event of receiving the same message *m* at D, then

(i) C(D) := max(C(D), C(S) + ∂)

(ii) C(D) := min(C(D), C(S) + ∂)

(iii) C(D) := max(C(D) + ∂, C(S))

(iv) C(D) := min(C(D) + ∂, C(S) + ∂)

(f) Which of the following is true for the Ricart-Agrawala algorithm for mutual exclusion?

(i) This is not a symmetric algorithm.

(ii) A process currently in CS that receives requests, store them in a local queue

(iii) A process currently in CS that receives requests, store them in a local stack

(iv) A process currently in CS that receives requests, store them in a local list.

(g) Why are multiple racks suggested for replication of blocks in case of Hadoop?

2. Comment on the correctness of the following statements and justify your opinion (*any five*) : 4×5

(a) "One of the major deficiencies for the Hadoop File System is the slow process of creating multiple replicas of blocks in different data nodes."

(b) "It's easier to achieve semantic transparency than syntactic transparency for a distributed system."

(c) "In case of SUN RPC, there is no system-wide binding mechanism maintaining distribution transparency."

(d) "A clock cannot be taken back for the sake of synchronization in a distributed system."

(e) "Name nodes are not expected to create performance bottleneck for Hadoop File System."

(f) "In consistent state recording (CSR), for every message that is recorded as sent, the corresponding state recording in the receiver node must reflect that the message has been received."

(g) "Migration of resources is a greater concern than address-space migration."

3. (a) In Chandy-Lamport's state recording algorithm, on receipt of the first marker from A, the recipient process B records state of the channel CH_{AB} through which it received the first marker as empty. Subsequently, a new message is placed on the channel CH_{AB} by A. Do you consider the algorithm to be correct in this context? Justify your opinion.

(b) "Two unrelated events X and Y occurs in two different nodes. The logical clock time-stamp values for the two events are TS(X) and TS(Y), respectively, such that TS(X) > TS(Y). It cannot be inferred from these statements that physically Y has occurred before X." — Give your opinion on the validity of the statements above and justify the same.

(c) Compare the performance of Vector clock with Lamport's logical clock model. 4+4+2

4. (a) Describe a token-based algorithm to ensure mutual exclusion of processes run from multiple nodes in a distributed system connected using a hierarchical topology.

(b) What would be the worst-case complexity for the above algorithm for a system with N processes running in that many nodes in the system?

(c) Compare performances of symmetric algorithms with token-based algorithms for mutual exclusion. 6+1+3

5. (a) Describe the Ho-Ramamurthy's deadlock detection algorithm for distributed environment. Illustrate the same with an example.

(b) Comment on the safety and liveness properties of Ho-Ramamurthy's deadlock detection algorithm. 6+4

6. (a) State at least two different motivations behind process migration.

(b) Describe the sender-initiated process migration approach.

(c) What is stability? What is done to improve the stability of the system for sender-initiated process migration?

(d) Define preemptive and non-preemptive process migrations. 2+4+2+2

7. (a) How the RPCs with large data arguments may be handled?

(b) Name the IDL for SUN RPC environment. What compiler is used for IDL in SUN RPC. Which call semantic is followed for SUN RPC?

(c) Suggest the appropriate call semantics to be used among may-be, last-of-many, at-least-once or exactly-once for the following application :

(i) To update an employee phone number in the employee database.

(ii) To increase the salary of an employee by 10% of the basic salary. 2+4+4

8. (a) Describe the Mitchell-Merritt's deadlock detection algorithm for a distributed environment.

(b) Explain the correctness aspect of Mitchell-Merritt's deadlock detection algorithm. Illustrate the same with an example. 6+4

———————

2024

COMPUTER SCIENCE

Paper : CSMC-203

(Automata and Compiler Design)

Full Marks : 70

The figures in the margin indicate full marks.
Candidates are required to give their answers in their own words as far as practicable.

Answer ***question nos.*** **1, 2** and ***any four*** questions from the rest.

1. Answer ***any five*** questions : 2×5

(a) Verify the following identity :

$$(0 * 01 + 10) * 0 * = (0 + 01 + 10)*$$

(b) "Every unambiguous grammar is LL(1)."— Comment on the truth/falsehood of the statement.

(c) Draw the transition diagram to recognize a signed exponential number.

(d) What do you mean by pass of a compiler? How can you reduce the number of passes?

(e) Comment on : "Equivalence of PDA and CFL".

(f) Write down the conditions to be satisfied for a CFG to be in CNF.

(g) Define the dominators of a node. When is a flow graph said to be reducible?

2. Answer ***any five*** questions : 4×5

(a) Explain the meaning of handle and viable prefixes with suitable examples.

(b) Eliminate the left recursion from the following grammar :

$$S \rightarrow ABC ;\ A \rightarrow Aa \mid d ;\ B \rightarrow Bb \mid e ;\ C \rightarrow Cc \mid f$$

(c) Generate the triple and indirect triples for the following statement :

IF A > B then C = B + D*3
else C = A + D*4

(d) For what condition a grammar can be LR(K)? How does it differ from a LL(K) grammar?

(e) What is a reserved word strategy? How is it handled in lexical analysis?

(f) What is a Right Recursive grammar? Will it create any problems in Recursive Descent Parsing? Give reasons for your answer.

(g) What input buffering concept is used in the lexical analyzer? How do the sentinels help the input buffering problem?

3. (a) Write a context-free grammar that generates all the strings of balanced parentheses.

(b) Consider the following statements :

G := C*(A + B) + (A + B)

C := A + B

A := (C * D) + (E – F)

(i) Draw the DAG for the above statements.

(ii) What is DAG's optimal ordering for optimizing the code?

(c) Write down some of the errors that a compiler should detect.

2+6+2

4. (a) Define Useful and Useless symbols in CFG. Write an algorithm to eliminate all productions containing useless symbols from the grammar.

(b) Identify Useful and Useless Symbols for the following grammar :

$S \rightarrow aA \mid bB$

$A \rightarrow aA \mid a$

$B \rightarrow bB$

$D \rightarrow ab \mid Ea$

$E \rightarrow aC \mid d$

(2+3)+5

5. (a) Write down the regular expression for the following :

'Set of strings consisting of an even number of a's followed by an odd number of b's.'

(b) Draw the NDFA of the above expression.

(c) Convert the above NDFA to its corresponding minimal DFA.

3+2+5

6. (a) Give the formal definition of TYPE-II grammar. Write down the TYPE-II grammar for deriving the language $\{WW^R \mid W \, \varepsilon \, (a, b)^*\}$.

(b) Simplify the following CFG and convert it to CNF :

$S \rightarrow AaB \mid aaB$

$A \rightarrow \varepsilon$

$B \rightarrow bbA \mid \varepsilon$

(c) What is a unit production? Why do we need to eliminate unit production from grammar?

(1+3)+3+3

7. (a) Answer the following with respect to Mealy machine :

(i) "For the input string of length n, the output sequence consists of n symbols, not $n + 1$." Why is it true?

(ii) "There are no accepted states in the Mealy machine." Why is that so?

(b) Construct a Mealy machine to print out 1's complement of an input bit-stream.

(c) Consider the grammar :

$$S \rightarrow AB \mid Abad$$
$$A \rightarrow d$$
$$E \rightarrow d$$
$$E \rightarrow b$$
$$D \rightarrow b \mid \varepsilon$$
$$B \rightarrow c$$

Construct the predictive parsing table using FIRST and FOLLOW sets. Show whether the given grammar is LL(1) or not. 2+4+4

8. (a) What are the themes behind optimization techniques? Explain the flow of control optimization with an example.

(b) Construct a Syntax Directed Translation Scheme (SDTS) that translates arithmetic expressions in infix notation into arithmetic expressions in infix notation having no redundant parenthesis. Show the annotated parse tree for the input (((1 + 2) * (3 + 4)) + 5).

(c) Write down the observations of the Turing Machine in case of the following :

(i) ε — transition

(ii) Halting Problem

(iii) Infinite Loop

Give examples wherever you need them. 2+5+3

Type-II :- $\alpha \rightarrow \beta$ $\left\{ \begin{array}{l} \alpha = V_n \\ \beta = (V_n \cup \Sigma)^* \end{array} \right\}$ [V_n - Set of veriable] [Σ - Set of Terminal]

2024

COMPUTER SCIENCE

Paper : CSMC-204

(Cryptography and Network Security)

Full Marks : 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Answer ***question nos.*** **1** and **2**, and ***any four*** questions from the rest.

1. Answer ***any five*** questions : 2×5

(a) Test the primality of integer 19 using the square root test.

(b) State the difference between the Chosen PT attack and the Known PT-CT attack.

(c) Find the values of $\Phi(77)$ and $\Phi(49)$.

(d) Alice has a long message to send. She is using the monoalphabetic substitution cipher. She thinks that if she compresses the message, it may protect the text from a single-letter frequency attack by Eve. Should she compress the message before or after the encryption? Defend your answer.

(e) Why is the deterministic algorithm of the primality test not feasible?

(f) State the unique contribution of challenge-response authentication scheme over password-based authentication.

(g) Define trapdoor one-way function.

2. Answer ***any five*** questions : 4×5

(a) State the Birthday paradox problem and in this context, comment on the feasibility of the collision resistance attack.

(b) In $GF(2^8)$, find the inverse of (x^5) modulo $(x^8 + x^4 + x^3 + x + 1)$.

(c) Find the orders of all elements in the group $G = \langle Z_9^*, x \rangle$.

(d) Show an LFSR with the characteristic polynomial $x^5 + x^2 + 1$. What is the period?

(e) The Miller-Rabin test can determine if a number is not prime but cannot determine if a number is prime. How can such an algorithm be used to test for primality?

(f) In the elliptic curve E(l, 2) over the GF(11) field : (i) find the equation of the curve (ii) find all points on the curve and draw the figure for the curve.

(g) Describe how authentication is achieved in RSA.

3. (a) Define the terms "Confusion" and "Diffusion".

(b) How are DES algorithm addressing the "confusion" and "diffusion" issues? Discuss with the necessary explanation.

(c) Is DES a Feistel Cipher? Why or why not?

(d) State the issues (with a brief justification) on which the performance of a Feistel Cipher depends. 2+5+1+2

4. (a) State the motivation behind proposing Elliptic Curve Cryptography even if the performance of RSA is satisfactory.

(b) What is the one-way function used in ECC?

(c) Discuss the key generation and encryption process.

(d) What trapdoor is used, and how is the trapdoor utilized for decryption? 2+2+4+2

5. (a) Alice and Bob want to establish a secret key using the Diffie-Hellman Key exchange protocol. The values of n, g, x and y are 11, 5, 2 and 3, respectively. Find out the values of the secret key.

(b) Discuss the attack(s) that may happen against the above-said protocol.

(c) State the principle difference between HMAC and MD5. 4+4+2

6. (a) Discuss the role of KDC for key distribution.

(b) Show that in RSA, the encryption and decryption processes are computationally easy for authenticated users, whereas decrypting the PT from CT becomes computationally infeasible for an intruder. 3+7

7. (a) Is there any specific advantage(s) for executing the IPSec protocol in tunnel mode compared to transport mode? Explain your answer.

(b) Consider the plaintext as "paymoremoney" and find out the ciphertext through Hill Cipher using the given key as follows. Also, show the decryption process (the computation for finding the PT from CT for the first three letters is sufficient).

K = 17 17 05
21 18 21
02 02 19

3+7

8. (a) Is it possible to use the algorithmic mode (without changing the encryption algorithm) to reduce the predictability within cipher text? Explain your answer.

(b) Discuss the operation mode that can speed up by parallel processing.

(c) Write a note on possible "factorization attack" in RSA. 4+3+3

———————