

Model Question Paper

Junior Research Fellowship (JRF) in Computer Science (CS)

Test Code : CSB (Afternoon Examination)

Maximum marks: 60 Time: TWO hours

8th May, 2022

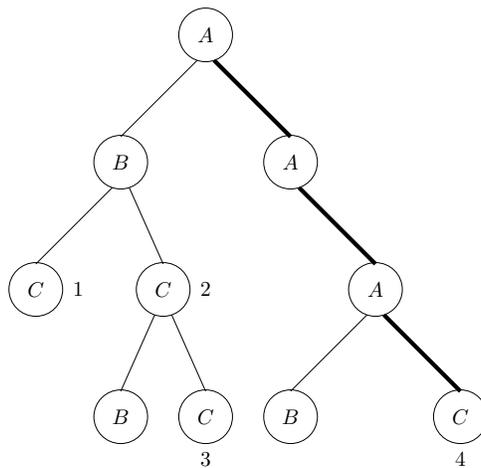
This question paper consists of 16 questions, organized into three groups as indicated below.

Answer any FIVE questions.

I. COMPUTER SCIENCE

- C1. (a) Consider a binary tree with n nodes. Each node is of type A , B , or C . Design an algorithm to find a path from the root of the tree to a C -type node (if such a node exists) that passes through the minimum number of B -type nodes. In case of ties, you may return any one of them. You will get full credit if your algorithm has a time complexity of $O(n)$.

For example, the tree shown below has four C -type nodes, marked 1, 2, 3 and 4. Your algorithm should return the path from the root to the node marked 4, because it does not pass through any B -type node, while the other paths each pass through at least one B -type node. [6]



- (b) Consider a list data structure S that stores n integers. S supports **only** the following two operations:
- $get(S, i)$ returns the i -th element of S (the elements of S are indexed starting from 1);
 - $reverse(S, k)$ modifies S by reversing the order of the first k elements of S .

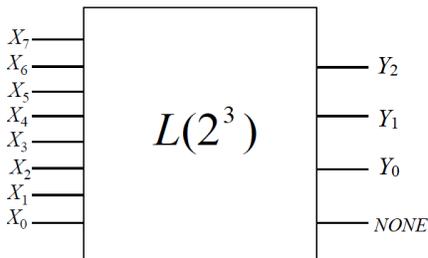
Design an algorithm to sort the elements of S in **increasing** order, using **only** the above operations, comparison operations ($<$, $>$, \leq , \geq , $=$, \neq), and constant additional storage. Your algorithm should use the minimum number of $reverse$ operations in the worst case. [6]

- C2. A Boolean logic circuit $L(2^N)$ takes a 2^N -bit binary number $X = X_{2^N-1}X_{2^N-2}\dots X_0$ as its input, and produces an N -bit binary output $Y = Y_{N-1}Y_{N-2}\dots Y_0$ indicating the most significant bit of X that is TRUE. In case none of the inputs are TRUE, it produces an output $NONE = 1$.

For example, if $N = 2$,

- (a) for $X = 0110$, output $Y = 10$ and $NONE = 0$, and
 (b) for $X = 0000$, output $Y = \text{Don't care}$ and $NONE = 1$.

Consider the block diagram of $L(2^3)$ shown in the figure below. Give a simplified Boolean equation for each output and draw the circuit using the minimum number of gates. Use gates with 4 or less inputs. [12]



- C3. Consider a 6-stage pipelined inorder processor with separate data and instruction memory, and a dual port register file. The pipeline stages are IF (Instruction Fetch), ID (Instruction Decode), RR

(Register Read), EX (Execution / Address Computation), MEM (Memory Access) and RWB (Register Write Back). Every instruction has to go through all stages of the pipe. All stages (except EX) take 1 processor clock cycle. The EX stage takes different clock cycles depending on the instruction as given below.

Operation	Clock Cycles
MUL	2
DIV	3
ADD	1
SUB	1
LOAD	1

MUL	R2,R0,R1	#R2 \leftarrow R0 * R1
DIV	R5,R3,R4	#R5 \leftarrow R3 / R4
LOAD	R1,10[R5]	#R1 \leftarrow MEM[R5 + 10]
ADD	R6,R1,R7	#R6 \leftarrow R1 + R7
SUB	R8,R3,R4	#R8 \leftarrow R3 - R4

- (a) Assume that operand forwarding between stages is available, but no instruction reordering is permitted. Draw the pipeline execution diagram for the above code that *minimises* the number of clock cycles needed. You may insert as many *STALL* cycles as necessary to ensure correct execution. Also, clearly mark the stages between which forwarding needs to be set up, if at all. [8]
- (b) Recompute the *minimum number* of clock cycles, assuming that *instruction reordering* is permitted, provided the final output remains correct. [4]
- C4. (a) Let s and t be two semaphores initialised to 1 and 0, respectively. Two threads T_1 and T_2 each execute the code shown below.

1: P(s)
2: V(s)
3: P(t)
4: V(t)

- (i) Does there exist an interleaving for which at least one of the threads progresses to completion? Clearly justify your answer.
 - (ii) What change(s) in the initial values assigned to the semaphores will guarantee that both threads will always progress to completion? [3 + 3]
- (b) Consider a uni-processor system in which three concurrent processes $P1$, $P2$ and $P3$ are running. The processes share a global variable X stored in shared memory. The initial value of X is 0. The processes update X concurrently as follows (without any synchronisation primitive):

```

Process P1: X = X + 5;
Process P2: X = X - 5;
Process P3: X = X * 5;

```

What will be the maximum and minimum values of X after execution of the three processes? Consider all possible orderings and interleavings of $P1$, $P2$ and $P3$. [6]

C5. Let $R(x, y)$ and $S(y, z)$ be two relations with the following statistics:

- number of records in R and S : $T(R) = 3000$, $T(S) = 200$;
- number of distinct values of x and y in R : $V(R, x) = 500$ and $V(R, y) = 1000$; and
- number of distinct values of y and z in S : $V(S, y) = 50$ and $V(S, z) = 50$.

The records are organized in pages and 20 records fit on each page.

- (a) Consider the following selection query: $\sigma_{x < 500}(R)$, where 100 records match the selection predicate. Calculate how many pages will have to be accessed to answer the query if the relation R is stored in a
 - (i) Heap file
 - (ii) B^+ tree index with search key x . Assume all index pages are available in memory. [3 + 3]
- (b) Consider the following SQL query:

```

SELECT * FROM R, S WHERE R.y = S.y AND S.z = 10

```

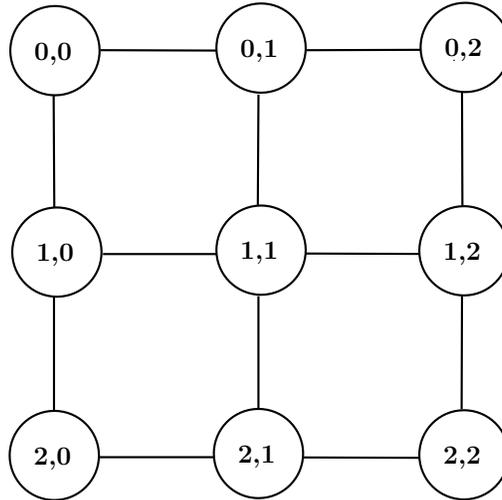
For this query, two equivalent logical plans in relational algebra are given below:

- $\sigma_{z=10}(R \bowtie_{y=y} (S))$
- $R \bowtie_{y=y} \sigma_{z=10}(S)$.

Explain which one is more efficient. [3]

- (c) Assuming that values are uniformly distributed, estimate the size of the relation that is the result of the SQL query in (b). [3]

C6. Consider an $N \times N$ mesh network. All links are full-duplex, i.e., data may flow in both directions simultaneously. A node in row- i and column- j is represented by a tuple (i, j) . For example, a 3×3 mesh network is shown in the figure below.



- (a) Design an algorithm that any node (i, j) may apply uniformly to route a packet to a destination (i', j') via shortest path, for $0 \leq i, j, i', j' < N$. What will be the *maximum packet latency* L ? Assume transmission of one packet through a link takes unit time, and the nodes are synchronised. [6]
- (b) Suppose at some time instant t , each of the N^2 nodes has a single data packet to be delivered to a destination. No two packets have the same destination.
- (i) Modify your routing algorithm proposed in (a) so that all the packets arrive at their respective destinations within time $(t + L)$.

- (ii) How many buffers will each node need to implement your algorithm? [4 + 2]

II. MATHEMATICS FOR COMPUTER SCIENCE

- MC1. (a) A 3×3 magic square is an arrangement of the numbers from a set of odd integers $\{1, 3, 5, \dots, 17\}$ in a 3×3 square grid, where the numbers in each row, in each column, and in the main and secondary diagonals, all add up to 27. Prove that the element at the center of the grid is 9. [7]
- (b) Let $D = \{d_1, d_2, \dots, d_k\}$ be the set of distinct divisors of a positive integer n (D includes 1 and n). Show that

$$\sum_{i=1}^k \sin^{-1} \sqrt{\log_n d_i} = \frac{\pi}{4} \times k.$$

HINT: $\sin^{-1} x + \sin^{-1} \sqrt{1 - x^2} = \frac{\pi}{2}$ [5]

- MC2. Consider the sets X and Y given below:

$$\begin{aligned} X &= \{x_1^{(a)}, x_1^{(b)}, x_2^{(a)}, x_2^{(b)}, \dots, x_k^{(a)}, x_k^{(b)}\} \\ Y &= \{y_1^{(a)}, y_1^{(b)}, y_1^{(c)}, y_2^{(a)}, y_2^{(b)}, y_2^{(c)}, \dots, y_m^{(a)}, y_m^{(b)}, y_m^{(c)}\}. \end{aligned}$$

Construct an undirected graph $G = (X \cup Y, E)$ by including the following edges in E :

- for each $i = 1, 2, \dots, k$, join $x_i^{(a)}$ and $x_i^{(b)}$ by an edge;
- for each $i = 1, 2, \dots, m$, join $y_i^{(a)}$ and $y_i^{(b)}$ by an edge; also join $y_i^{(b)}$ and $y_i^{(c)}$ by an edge, and $y_i^{(c)}$ and $y_i^{(a)}$ by an edge;
- for each $y \in Y$, join y to exactly one element x of X .

- (a) Show that a vertex cover of G cannot have less than $k + 2m$ elements. [6]
- (b) Suppose G' is obtained from G by deleting all edges of the form (x, y) where $x \in X$ and $y \in Y$. Compute the number of distinct vertex covers for G' . [6]

[Recall that, for a graph $G = (V, E)$, a subset V' of V is a *vertex cover*, if for every $e = (v_1, v_2) \in E$, either $v_1 \in V'$ or $v_2 \in V'$.]

MC3. Routes A and B are the only two escape routes from a state prison. Prison records show that 40% of the prisoners who tried to escape used route A. These records also show that 80% of those who tried to escape via A, and 70% of those who tried to escape via B were captured.

- (a) What is the expected number of attempts that a prisoner needs to make to successfully escape from the prison? [4]
- (b) Given that two prisoners have independently and successfully escaped from the prison, what is the probability that they have used the same route to escape? [8]

MC4. Define a language $L \subseteq \{0, 1\}^*$ as follows:

$$L = \{x \mid \text{The number of 1s in } x \text{ that are not immediately preceded by a 0 is even}\}.$$

For example, $011011 \in L$ because it has **two** 1s that are not immediately preceded by a 0, whereas $111 \notin L$ because it has **three** 1s that are not immediately preceded by a 0. Draw a Deterministic Finite Automaton (DFA) to recognize L . You will get full credit if your DFA has no more than 4 states. [12]

III. MATHEMATICS

- M1. (a) Show that each of the equations $\sin(\cos x) = x$ and $\cos(\sin y) = y$ has exactly one root in $[0, \pi/2]$. Moreover, if x_1 and x_2 are the roots of the former and the latter equations in $[0, \pi/2]$, respectively, then show that $x_1 < x_2$. [6]
- (b) Let f be a real-valued continuous function on \mathbb{R} satisfying the inequality

$$f(x) \leq \frac{1}{2h} \int_{x-h}^{x+h} f(y) dy, \quad \forall x \in \mathbb{R}, \quad \forall h > 0.$$

Prove that for any bounded closed interval, the maximum of f on that interval is attained at one of its end points. [6]

- M2. Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ have continuous partial derivatives and satisfy $\left| \frac{\partial f}{\partial x_j}(x) \right| \leq K$ for all $x = (x_1, x_2)$, $j = 1, 2$. Prove that

$$|f(x) - f(y)| \leq \sqrt{2}K\|x - y\|.$$

Here, $\|(x_1, x_2)\| = \sqrt{x_1^2 + x_2^2}$. [12]

- M3. (a) Let G be a group of order n , H a subgroup of G of order m , $k = \frac{n}{m}$ and S_k the symmetric group on k symbols. Show that there is a nontrivial group homomorphism $\phi : G \rightarrow S_k$. [5]
- (b) Let \mathbb{C}^* denote the group of non-zero complex numbers under multiplication. Show that \mathbb{C}^* can be expressed as a direct product of two proper subgroups. [7]

- M4. (a) Let R be a commutative ring with 1 and P be a prime ideal of R . Consider the polynomial ring $R[x]$ and let $P[x]$ be the ideal of $R[x]$ consisting of polynomials whose coefficients all belong to P . Show that the ideal

$$P[x] + \langle x \rangle := \{f(x) + xg(x) : f(x) \in P[x], g(x) \in R[x]\},$$

is a prime ideal of $R[x]$. [5]

- (b) Let R be an integral domain and K be a subring of R . If K is a field and the dimension of R as a vector space over K is finite, then show that R is a field. [7]

- M5. Let A and X be invertible complex matrices such that $XAX^{-1} = A^2$. Prove that there exists a natural number m such that each eigenvalue of A is an m -th root of unity. [12]

- M6. Consider the metric space (\mathbb{R}^2, d) where d is the Euclidean metric $d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$, $(x_i, y_i) \in \mathbb{R}^2$, $i \in \{1, 2\}$.

Let

$$S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

Define a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ by

$$f((x, y)) = \inf \{d((x, y), (a, b)) \mid (a, b) \in S^1\}.$$

Show that f is continuous, and $f((x, y)) = 0$ if and only if $(x, y) \in S^1$. [12]