

CHENNAI MATHEMATICAL INSTITUTE

M.Sc. / Ph.D. Programme in Computer Science

Entrance Examination, May 2, 2026

- This examination has three parts: Part A, Part B and Part C.
 - Time allowed: 3 hours
 - Total marks: 100
- Part A consists of 10 multiple choice questions, each carrying 3 marks. There may be multiple correct choices. You have to select all the correct options and no incorrect option to get full marks. **There is no partial credit, and no negative marking.**
- Part B has 3 long-answer questions, each carrying 10 marks.
- Part C has 8 long-answer questions, each carrying 10 marks. Students can choose **any 4 out of these 8** questions. Students may attempt more than 4 questions; in such cases, the **best 4 answers** will be considered for the final score.
- For questions in Part A, you have to provide the answers on the computer. You only have to choose the appropriate answer(s) from the choices provided. For example, if the correct choices for a question are options (a) and (c), choose only (a) and (c) to score full marks for that question.

In all questions related to graphs, unless otherwise specified, we use the word “graph” to mean a finite undirected graph with no self-loops, and at most one edge between any pair of vertices.

Part A

1. Six people A, B, C, D, E, F sit around a circular table.

- A sits opposite D
- B is not adjacent to A
- C sits immediately to the right of B
- E is adjacent to D

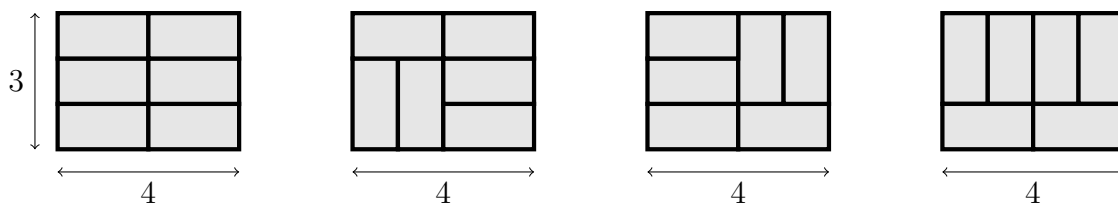
Who sits opposite C ?

- (a) B (b) D (c) E (d) F

2. Dhruva is attending a party. If they play both dumb charades and musical chairs, he will leave the party. If Abhay does not dance or Bharat does not juggle, then also he will leave the party. If Abhay dances and Bharat juggles, then everyone gets free chocolate. Suppose Mahesh, who is also attending the party, did not get free chocolate and musical chairs were played, which of the following can you infer?

- (a) Dumb charades is not played (b) Abhay did not dance
(c) Bharat did not juggle (d) Dhruva left the party

3. Suppose you are writing a 100-question multiple-choice exam with 4 choices per question and exactly one correct answer. You know the answers to 80% of the questions, and you guess the answer to the remaining 20% uniformly at random. What is the expected number of questions you answer correctly?
- (a) 82 (b) 85 (c) 90 (d) 95
4. In a multiple choice test with 100 questions, each question has 5 possible answers with exactly one of them being correct. A student knows the answer to 40 of these questions. For the rest he will guess the answer uniformly at random.
- A question is picked uniformly at random and posed to the student. If he answers it correctly, what is the probability that he actually knew the answer to that question?
- (a) $\frac{2}{5}$ (b) $\frac{5}{13}$ (c) $\frac{13}{25}$ (d) $\frac{10}{13}$
5. 1000 candidates appear in a competitive examination. There are 50 questions, each worth 2 marks. There is no partial credit, and each answer is awarded either 0 marks or 2 marks. Only 10 candidates scored 80 or above. Let n_i ($0 \leq i \leq 100$) denote the number of candidates whose score is i . Which of the following scenarios are possible?
- (a) $n_0 = 100$.
 (b) For all odd numbers $i \in \{0, \dots, 100\}$, $n_i = 0$.
 (c) For all $i \in \{0, \dots, 100\}$, $n_i < 25$.
 (d) There are at most 30 values of $i \in \{0, \dots, 100\}$ such that $n_i = 0$.
6. Let $T(n)$ be the number of ways to tile a $3 \times n$ rectangle with 1×2 dominoes; refer to the figure for some examples for the case $n = 4$.



Four distinct tilings of a 3×4 rectangle with 1×2 dominoes.

Which of the following statements is/are true about the growth of $T(n)$?

- (a) $T(2) = 3$
 (b) $T(n) = 0$ for all odd n .
 (c) $T(n)$ is given by a recurrence of the form $T(n) = T(n-1) + T(n-2)$.
 (d) $T(4) < 9$.
7. A binary string s of length 2026 starts with bit 0 and ends with bit 1. Let N_{00} , N_{01} , N_{10} , and N_{11} denote the number of times that the sequences 00, 01, 10, and 11, respectively, appear as pairs of adjacent bits in s . Which of the following equations are correct?
- (a) $N_{01} = N_{10} + 1$ (b) $N_{10} = N_{01} + 1$
 (c) $N_{01} = 1013 - (N_{00} + N_{11})/2$ (d) $N_{01} = 2026 - N_{00} - N_{11}$

Questions 8 and 9 pertain to the following pseudocode defining the function $f(x)$, where x is assumed to be a nonnegative integer. The code uses the function $\text{rem}(n, d)$, which returns the remainder after dividing n by d .

```
function  $f(x)$ 
|    $a \leftarrow 0$ 
|   while  $x > 0$  do
|       |    $r \leftarrow \text{rem}(x, 2)$ 
|       |    $a \leftarrow (10 \times a) + r$ 
|       |    $x \leftarrow (x - r)/2$ 
|   end while
|   return  $a$ 
end function
```

8. What is the value returned by $f(43)$?
 (a) 101000 (b) 101011 (c) 110101 (d) 110100
9. Which of the following best describes the running time of $f(x)$?
 (a) Linear in x . (b) Quadratic in x .
 (c) Logarithmic in x . (d) Exponential in x .
10. Consider the following two procedures **proc1** and **proc2**, which run in parallel after initializing x to 6 and y to 3. Running in parallel means that between any two lines of code in one process, any number of lines of the other process may run. Both x and y are shared variables which both processes can read/modify.

```
int  $x = 6$ ,  $y = 3$ ;
```

```
proc1 {
     $x = 2*y$ ;
     $x = x - y$ ;
}
```

```
proc2 {
     $y = 2*x$ ;
     $y = y - x$ ;
}
```

Which of the following values are possible for (x, y) after both processes have come to a halt?

- (a) (8,5) (b) (3,3) (c) (0,6) (d) (36,-12)

Part B

1. You are given a long sheet of (ruled) paper with the string MI written on the first line. You are asked to fill in lines 2, 3, 4, ... (in order) with other strings involving the letters M, I and U, subject to the following rules, where x and y stand for any string (possibly empty) formed using the letters M, I and U:

R1. If the previous line has Mx , you can write Mxx .

R2. If the previous line has xI , you can write xIU .

R3. If the previous line has $xUUy$, you can write xy .

R4. If the previous line has $xIIIy$, you can write xUy .

You can stop at any line, at any point of time, and the string on the last line you have filled in is said to be *derived*. Here are two examples (where we write $x_1 \rightarrow x_2 \rightarrow \dots x_{n-1} \rightarrow x_n$ to mean that the lines 1, 2, ..., $n-1, n$ have entries $x_1, x_2, \dots, x_{n-1}, x_n$ (in order)).

E1. MI. (This shows that you can stop even without adding any line.)

E2. $MI \rightarrow MII \rightarrow MIIII \rightarrow MUI \rightarrow MUIUI$.

E3. $MI \rightarrow MII \rightarrow MIIII \rightarrow MIIIIU \rightarrow MUIU$.

Questions:

- (a) For examples E2 and E3 above, write down which rule from R1 to R4 is used in each step.
- (b) Show how to derive the strings MIUIU and MIUIUU. (**Hint:** To derive MIUIUU, first derive MIIIIIII (M followed by 8 occurrences of I) and then proceed.)
- (c) Show that the string MIIMII cannot be derived.
- (d) If x is any string derived using the rules above, and if n is the number of occurrences of I in x , what are the possible values for $n \bmod 3$?
- (e) Can the string MU be derived? If so, provide a derivation. If not, explain why not.

2. For a set X , X^* denotes the set of all strings formed using X , including the empty string. By xy , we mean the *concatenation* of x and y , i.e., the string x followed by the string y . We say that x is a *prefix* of y if there is some $z \in X^*$ such that $xz = y$. Note that every x is a prefix of itself, and ε is a prefix of every x .

A function $f : \Sigma^* \rightarrow \{0, 1\}^*$ is said to be a homomorphism if $f(\varepsilon) = \varepsilon$ and $f(xy) = f(x)f(y)$ for all $x, y \in \Sigma^*$. f is said to be *prefix-free* if $f(a)$ is not a prefix of $f(b)$ for any pair of distinct *letters* $a, b \in \Sigma$. For example, if $f(a) = 01$ and $f(b) = 101$, then $f(aba) = 0110101$. For the following questions, we assume that $|\Sigma| = 8$ and $f : \Sigma^* \rightarrow \{0, 1\}^*$ is a prefix-free homomorphism.

- (a) Is it possible that $f(a)$ is of length at most 3 for all a . If so, define such an f . If not, provide a proof that this is not possible.
 - (b) Is it possible that $f(a)$ is of length 2 for some $a \in \Sigma$ and $f(b)$ is of length 3 for all $b \in \Sigma \setminus \{a\}$. If so, define such an f . If not, provide a proof that this is not possible.
 - (c) Show that there are no distinct $x, y \in \Sigma^*$ such that $f(x) = f(y)$.
3. Suppose $G = (V, E)$ is a graph with $|V| = 2k$ for some $k \geq 1$. Further, suppose there is no cycle of length 3 in G , i.e., G is triangle-free. Prove that $|E| \leq k^2$.

Part C

1. Consider a finite alphabet $\Sigma = \{0, 1\}$. For a string $w \in \Sigma^*$, we write $|w|$ for the length of w . For two strings u, v we write $u \cdot v$ for the string obtained by concatenating u and v .

Let us define an operation $\pi_{\text{even}}(w)$ inductively as follows:

$$\begin{aligned} \pi_{\text{even}}(\varepsilon) &= \varepsilon \\ \forall w \in \Sigma^* \text{ and } \forall a \in \Sigma \quad \pi_{\text{even}}(wa) &= \begin{cases} \pi_{\text{even}}(w) & \text{if } |wa| \text{ is odd} \\ \pi_{\text{even}}(w) \cdot a & \text{if } |wa| \text{ is even} \end{cases} \end{aligned}$$

- (a) What is $\pi_{\text{even}}(001011)$?
 - (b) For a language $L \subseteq \Sigma^*$, define $\pi_{\text{even}}(L) := \{\pi_{\text{even}}(w) \mid w \in L\}$. What is $\pi_{\text{even}}((01)^*)$?
 - (c) Show that if L is regular, $\pi_{\text{even}}(L)$ is also regular.
2. Let Σ be a finite alphabet, and let $L \subseteq \Sigma^*$. Two strings u and v in Σ^* are said to be L -equivalent if for all $z \in \Sigma^*$, $(uz \in L)$ iff $(vz \in L)$.
 - (a) Consider $\Sigma = \{0, 1\}$ and let $L = 0^*1^*0^*$. For the following instances of u and v , explain whether u and v are L -equivalent.
 - (i) $u = 0, v = 1$
 - (ii) $u = 0, v = 00$
 - (b) Let L be a regular language. Show that for any infinite set of strings S , there exist at least two strings in S that are L -equivalent.
 3. You are given a graph on vertex set $V = \{a_1, \dots, a_k\} \cup \{b_1, \dots, b_k\}$. The edge set is $E = \{(a_i, b_j) \mid i \neq j\}$. An edge coloring of a graph is an assignment of colors to edges such that no two edges sharing a common vertex receive the same color.
 - (a) Give an efficient algorithm to color the edges of the graph with minimum number of colors.
 - (b) How many colors does your algorithm use?

4. In the *Pots-of-Gold* game, there is a row of n pots $P_1, \dots, P_n$. Each pot contains some number of gold coins. You remove pots and collect the gold coins from the pots that you remove, subject to the rules below. Your aim is to collect as many gold coins as possible.
- You start with no gold coins and remove pots one at a time. You can remove any pot that is *available* for removal.
 - In the beginning, all pots are available for removal. Removing pot P_i makes the pots P_{i-1} and P_{i+1} (if they exist) permanently unavailable.
- (a) Assuming that we have an array $G[1 \dots n]$ with $G[i]$ storing the number of gold coins in pot P_i , devise an algorithm $\text{POTSOFGOLD}(G, n)$ that computes the maximum number of gold coins that you can collect. Your algorithm should run in $O(n)$ time in the worst case. You can assume that each arithmetic operation involving two integers, and each array access (reading/writing one element) takes constant time. **Hint:** Use dynamic programming.
- (b) Clearly explain why your procedure of part (a) correctly computes the required output.
- (c) Explain why your procedure of part (a) runs in $O(n)$ time in the worst case.
5. In this question, we only consider real-valued functions from $[0, 1] \rightarrow \mathbb{R}$.
- (a) Does there exist a function f that is continuous at exactly 0 and discontinuous at every point in $(0, 1]$?
- (b) Suppose S is a finite subset of $[0, 1]$. Does there exist a function that is continuous at all points in S , and discontinuous everywhere else?

Justify your answers with proofs.

6. Let $x_1, \dots, x_n$ be uniform, independent bits; i.e., each x_i is a random variable that takes value 0 or 1 with equal probability, and is independent of the other random variables. Recall: $(x_1, \dots, x_n)$ are said to be independent if, for all $b_1 \in \{0, 1\}, b_2 \in \{0, 1\}, \dots, b_n \in \{0, 1\}$,

$$\mathbb{P}[x_1 = b_1, x_2 = b_2, \dots, x_n = b_n] = \mathbb{P}[x_1 = b_1] \times \mathbb{P}[x_2 = b_2] \times \dots \times \mathbb{P}[x_n = b_n].$$

A set of bits $y_1, \dots, y_n$ is said to be *pairwise independent* if, for every distinct $i, j \in \{1, 2, \dots, n\}$, the bits y_i and y_j are independent.

- (a) Prove that x_1, x_2 , and $x_1 \oplus x_2$ are pairwise independent when $\oplus$ denotes binary addition (exclusive OR); i.e., $x_1 \oplus x_2 = x_1 + x_2 \pmod{2}$.
- (b) Prove that $x_1 \oplus x_2, x_1 \oplus x_3$, and $x_3 \oplus x_2$ are pairwise independent.
- (c) Can you create 7 pairwise independent bits using x_1, x_2, x_3 ? Justify your answer.
- (d) Generalize this to show that n uniform, independent bits can be used to generate $2^n - 1$ pairwise independent sets. Justify your answer.

7. Let $M_n(\mathbb{R})$ denote the vector space $\mathcal{V}$ of all $n \times n$ matrices over the real numbers $\mathbb{R}$. You may assume that e_{ij} are the basis vectors of this vector space and x_{ij} are the dual basis. Recall that a matrix is said to be singular if its determinant is zero.
- (a) Let E_{ij} be the matrix with 1 in position (i, j) and zeros elsewhere. Describe a linear subspace $\mathcal{L} \subseteq \mathcal{V}$ of dimension $n^2 - n$ containing the elementary matrix E_{ij} such that all matrices in $\mathcal{L}$ are singular. If it helps, you may write down a system of linear equations in the variables x_{ij} whose solution space contains E_{ij} and has the requisite dimension.
 - (b) Let $A \in M_n(\mathbb{R})$ be a matrix of rank less than or equal to $n - 1$. Show that there exists a linear subspace $\mathcal{L} \subseteq \mathcal{V}$ such that $A \in \mathcal{L}$, every matrix in $\mathcal{L}$ is singular and $\dim \mathcal{L} \geq n^2 - n$.
8. Consider a bivariate function $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$.
- (a) What does it mean to say f is continuous at a point $(a, b) \in \mathbb{R} \times \mathbb{R}$? Can you define it formally?
 - (b) Consider any straight line L passing through the point (a, b) . Show that f restricted to the line is a real function $f_L : \mathbb{R} \rightarrow \mathbb{R}$.
 - (c) If f_L is continuous at the point (a, b) for each line L passing through (a, b) then is the bivariate function f continuous at (a, b) ?

Justify answers with proofs.