Part A: Common part

Note:
The questions in this part are to be answered by **all candidates**, i.e., **both Computer Science and Systems Science streams.**

1. A box contains 5 red marbles, 8 green marbles, 11 blue marbles, and 15 yellow marbles. We draw marbles uniformly at random without replacement from the box. What is the minimum number of marbles to be drawn to ensure that out of the marbles drawn, at least 7 are of the same colour?

(a) 7

(b) 8

(c) 23

(d) 24 ✓

(e) 39

2. The maximum area of a rectangle inscribed in the unit circle (i.e., all the vertices of the rectangle are on the circle) in a plane is:

(a) 1

(b) 2 ✓

(c) 3

(d) 4

(e) 5

3. Let M be an $n \times m$ real matrix. Consider the following:

- Let k_1 be the smallest number such that M can be factorized as $A \cdot B$, where A is an $n \times k_1$ and B is a $k_1 \times m$ matrix.
- Let k_2 be the smallest number such that $M = \sum_{i=1}^{k_2} u_i v_i$, where each u_i is an $n \times 1$ matrix and each v_i is an $1 \times m$ matrix.
- Let k_3 be the column-rank of M.

Which of the following statements is true?

(a) $k_1 < k_2 < k_3$

(b) $k_1 < k_3 < k_2$

(c) $k_2 = k_3 < k_1$

(d) $k_1 = k_2 = k_3$ ✓

(e) No general relationship exists among k_1, k_2 and k_3.

4. What is the probability that at least two out of four people have their birthdays in the same month, assuming their birthdays are uniformly distributed over the twelve months?

(a) $\frac{25}{48}$
(b) $\frac{5}{8}$
(c) $\frac{5}{12}$
(d) $\frac{41}{96}$ ✓
(e) $\frac{55}{96}$

5. Let n, m and k be three positive integers such that $n \geq m \geq k$. Let S be a subset of $\{1, 2, \ldots, n\}$ of size k. Consider sampling a function f uniformly at random from the set of all functions mapping $\{1, \ldots, n\}$ to $\{1, \ldots, m\}$. What is the probability that f is not injective on the set S, i.e., there exist $i, j \in S$ such that $f(i) = f(j)$?

In the following, the binomial coefficient $\binom{n}{k}$ counts the number of k-element subsets of an n-element set.

(a) $1 - \frac{k!}{k^k}$
(b) $1 - \frac{m!}{m^k}$
(c) $1 - \frac{k!\binom{m}{k}}{m^k}$ ✓
(d) $1 - \frac{k!\binom{n}{k}}{n^k}$
(e) $1 - \frac{k!\binom{n}{k}}{m^k}$

6. A matching in a graph is a set of edges such that no two edges in the set share a common vertex. Let G be a graph on n *vertices* in which there is a subset M of m *edges* which is a matching. Consider a random process where each vertex in the graph is independently selected with probability $0 < p < 1$ and let B be the set of vertices so obtained. What is the probability that there exists at least one edge from the matching M with both end points in the set B?

(a) p^2
(b) $1 - (1 - p^2)^m$ ✓
(c) p^{2m}
(d) $(1 - p^2)^m$
(e) $1 - (1 - p(1 - p))^m$

7. Let d be the number of positive square integers (that is, it is a square of some integer) that are factors of $20^5 \times 21^5$. Which of the following is true about d?

(a) $50 \leq d < 100$
(b) $100 \leq d < 150$
(c) $150 \leq d < 200$ ✓
(d) $200 \leq d < 300$
(e) $300 \leq d$

8. Consider the sequence

$$y_n = \frac{1}{\int_1^n \frac{1}{(1+x/n)^3} dx}$$

for $n = 2, 3, 4, \ldots$ Which of the following is true?

(a) The sequence $\{y_n\}$ does not have a limit as $n \to \infty$.

(b) $y_n \leq 1$ for all $n = 2, 3, 4, \ldots$

(c) $\lim_{n\to\infty} y_n$ exists and is equal to $6/\pi^2$.

(d) $\lim_{n\to\infty} y_n$ exists and is equal to 0. ✓

(e) The sequence $\{y_n\}$ first increases and then decreases as n takes values $2, 3, 4, \ldots$

9. Fix $n \geq 6$. Consider the set $\mathcal{C}$ of binary strings $x_1 x_2 \ldots x_n$ of length n such that the bits satisfy the following set of equalities, all modulo 2: $x_i + x_{i+1} + x_{i+2} = 0$ for all $1 \leq i \leq n-2$, $x_{n-1} + x_n + x_1 = 0$, and $x_n + x_1 + x_2 = 0$. What is the size of the set $\mathcal{C}$?

(a) 1 for all $n \geq 6$

(b) 4 for all $n \geq 6$

(c) 0 for all $n \geq 6$

(d) If $n \geq 6$ is divisible by 3 then $|\mathcal{C}| = 1$. If $n \geq 6$ is *not* divisible by 3 then $|\mathcal{C}| = 4$.

(e) If $n \geq 6$ is divisible by 3 then $|\mathcal{C}| = 4$. If $n \geq 6$ is *not* divisible by 3 then $|\mathcal{C}| = 1$. ✓

10. Lavanya and Ketak each flip a fair coin n times. What is the probability that Lavanya sees more heads than Ketak?

In the following, the binomial coefficient $\binom{n}{k}$ counts the number of k-element subsets of an n-element set.

(a) $\frac{1}{2}$

(b) $\frac{1}{2}\left(1 - \sum_{i=0}^{n} \frac{\binom{n}{i}^2}{2^{2n}}\right)$ ✓

(c) $\frac{1}{2}\left(1 - \sum_{i=0}^{n} \frac{\binom{n}{i}}{2^{2n}}\right)$

(d) $\frac{1}{2}\left(1 - \frac{1}{2^{2n}}\right)$

(e) $\sum_{i=0}^{n} \frac{\binom{n}{i}}{2^n}$

11. Find the following sum.

$$\frac{1}{2^2-1} + \frac{1}{4^2-1} + \frac{1}{6^2-1} + \cdots + \frac{1}{40^2-1}$$

(a) $\frac{20}{41}$ ✓

(b) $\frac{10}{41}$

(c) $\frac{10}{21}$

(d) $\frac{20}{21}$

(e) 1

12. How many numbers in the range $\{0, 1, \ldots, 1365\}$ have exactly four 1's in their binary representation? (Hint: 1365_{10} is 10101010101_2, that is,

$$1365 = 2^{10} + 2^8 + 2^6 + 2^4 + 2^2 + 2^0.)$$

In the following, the binomial coefficient $\binom{n}{k}$ counts the number of k-element subsets of an n-element set.

(a) $\binom{6}{4}$

(b) $\binom{10}{4}$

(c) $\binom{10}{4} + \binom{8}{3} + \binom{6}{2} + \binom{5}{1}$ ✓

(d) $\binom{11}{4} + \binom{9}{3} + \binom{7}{2} + \binom{5}{1}$

(e) 1024

13. What are the last two digits of 7^{2021}?

(a) 67

(b) 07 ✓

(c) 27

(d) 01

(e) 77

14. Five married couples attended a party. In the party, each person shook hands with those they did not know. Everyone knows his or her spouse. At the end of the party, Shyamal, one of the attendees, listed the number of hands that other attendees including his spouse shook. He got every number from 0 to 8 once in the list. How many persons shook hands with Shyamal at the party?

(a) 2

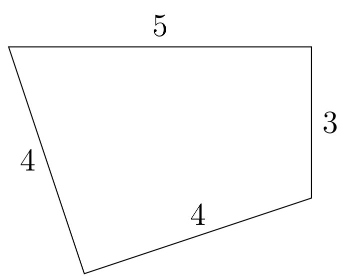
(b) 4 ✓

(c) 6

(d) 8

(e) Insufficient information

15. Let P be a convex polygon with sides 5, 4, 4, 3. For example, the following:



Consider the shape in the plane that consists of all points within distance 1 from some point in P. If ℓ is the perimeter of the shape, which of the following is always correct?

(a) ℓ cannot be determined from the given information.

(b) $20 \leq \ell < 21$

(c) $21 \leq \ell < 22$

(d) $22 \leq \ell < 23$ ✓

(e) $23 \leq \ell < 24$

—End of Common Section—

(CS section begins in the following page)

Part B: Computer Science

Note:
Only for Computer Science stream candidates.

If you are applying for the **Systems Science** stream, please skip this section and attempt questions from **Part C** on page 11.

1. Consider the following statements about propositional formulas.

(i) $(p \wedge q) \rightarrow r$ and $(p \rightarrow r) \wedge (q \rightarrow r)$ are *not* logically equivalent.

(ii) $(\neg a \rightarrow b) \wedge (\neg b \vee (\neg a \vee \neg b))$ and $\neg(a \leftrightarrow b)$ are *not* logically equivalent.

(a) Both (i) and (ii) are true.

(b) (i) is true and (ii) is false. ✓

(c) (i) is false and (ii) is true.

(d) Both (i) and (ii) are false.

(e) Depending on the values of p and q, (i) can be either true or false, while (ii) is always false.

2. Let L be a singly-linked list and X and Y be additional pointer variables such that X points to the first element of L and Y points to the last element of L. Which of the following operations cannot be done in time that is bounded above by a constant?

(a) Delete the first element of L.

(b) Delete the last element of L. ✓

(c) Add an element after the last element of L.

(d) Add an element before the first element of L.

(e) Interchange the first two elements of L.

3. What is the prefix expression corresponding to the expression:
$((9+8)*7+(6*(5+4))*3)+2$?

You may assume that $*$ has precedence over $+$.

(a) $* + + 9\,8\,7 * * 6 + + 5\,4\,3\,2$

(b) $* + + + 9\,8\,7 * * 6 + 5\,4\,3\,2$

(c) $+ * + + 9\,8\,7 * * 6 + 5\,4\,3\,2$

(d) $+ + * + 9\,8\,7 * * 6 + 5\,4\,3\,2$ ✓

(e) $+ * + * 9\,8\,7 + + 6 * 5\,4\,3\,2$

4. Consider the following two languages.

$$\text{PRIME} = \{1^n \mid n \text{ is a prime number } \},$$
$$\text{FACTOR} = \{1^n 0 1^a 0 1^b \mid n \text{ has a factor in the range } [a, b]\}.$$

What can you say about the languages PRIME and FACTOR?

(a) PRIME is in P, but FACTOR is not in P.

(b) Neither PRIME nor FACTOR are in P.

(c) Both PRIME and FACTOR are in P. ✓

(d) PRIME is not in P, but FACTOR is in P.

(e) None of the above since we can answer this question only if we resolve the status of the NP vs. P question.

5. For a language L over the alphabet $\{a, b\}$, let $\overline{L}$ denote the complement of L and let L^* denote the Kleene-closure of L. Consider the following sentences.

(i) $\overline{L}$ and L^* are both context-free.

(ii) $\overline{L}$ is not context-free but L^* is context-free.

(iii) $\overline{L}$ is context-free but L^* is regular.

Which of the above sentence(s) is/are true if $L = \{a^n b^n \mid n \geq 0\}$?

(a) Both (i) and (iii)

(b) Only (i) ✓

(c) Only (iii)

(d) Only (ii)

(e) None of the above

6. Consider the following pseudocode:

```
procedure HowManyDash(n)
    if n = 0 then
        print '-'
    else if n = 1 then
        print '-'
    else
        HowManyDash(n − 1)
        HowManyDash(n − 2)
    end if
end procedure
```

How many '-' does HowManyDash(10) print?

(a) 9

(b) 10

(c) 55

(d) 89 ✓

(e) 1024

7. Which of the following regular expressions defines a language that is different from the other choices?

(a) $b^*(a + b)^*a(a + b)^*ab^*(a + b)^*$

(b) $a^*(a+b)^*ab^*(a+b)^*a(a+b)^*$

(c) $(a+b)^*ab^*(a+b)^*a(a+b)^*b^*$

(d) $(a+b)^*a(a+b)^*b^*a(a+b)^*a^*$

(e) $(a+b)^*b^*a(a+b)^*b^*(a+b)^*$ ✓

8. Let A and B be two matrices of size $n \times n$ and with real-valued entries. Consider the following statements.

1. If $AB = B$, then A must be the identity matrix.
2. If A is an idempotent (i.e. $A^2 = A$) nonsingular matrix, then A must be the identity matrix.
3. If $A^{-1} = A$, then A must be the identity matrix.

Which of the above statements MUST be true of A?

(a) 1, 2, and 3

(b) Only 2 and 3

(c) Only 1 and 2

(d) Only 1 and 3

(e) Only 2 ✓

9. Let L be a context-free language generated by the context-free grammar $G = (V, \Sigma, R, S)$ where V is the finite set of variables, Σ the finite set of terminals (disjoint from V), R the finite set of rules and $S \in V$ the start variable. Consider the context-free grammar G' obtained by adding $S \to SS$ to the set of rules in G. What must be true for the language L' generated by G'?

(a) $L' = LL$

(b) $L' = L$

(c) $L' = L^*$

(d) $L' = \{xx | x \in L\}$

(e) None of the above ✓

10. Let G be a connected bipartite simple graph (i.e., no parallel edges) with distinct edge weights. Which of the following statements on MST (minimum spanning tree) need not be TRUE?

(a) G has a unique MST.

(b) Every MST in G contains the lightest edge.

(c) Every MST in G contains the second lightest edge.

(d) Every MST in G contains the third lightest edge.

(e) No MST in G contains the heaviest edge. ✓

11. Suppose we toss a fair coin repeatedly until the first time by which at least *two* heads *and* at least *two* tails have appeared in the sequence of tosses made. What is the expected number of coin tosses that we would have to make?

(a) 8

(b) 4

(c) 5.5 ✓

(d) 7.5

(e) 4.5

12. Let G be an undirected graph. For any two vertices u, v in G, let $\mathsf{cut}(u, v)$ be the minimum number of edges that should be deleted from G so that there is no path between u and v in the resulting graph. Let a, b, c, d be 4 vertices in G. Which of the following statements is impossible?

(a) $\mathsf{cut}(a, b) = 3$, $\mathsf{cut}(a, c) = 2$, and $\mathsf{cut}(a, d) = 1$

(b) $\mathsf{cut}(a, b) = 3$, $\mathsf{cut}(b, c) = 1$, and $\mathsf{cut}(b, d) = 1$

(c) $\mathsf{cut}(a, b) = 3$, $\mathsf{cut}(a, c) = 2$, and $\mathsf{cut}(b, c) = 2$

(d) $\mathsf{cut}(a, c) = 2$, $\mathsf{cut}(b, c) = 2$, and $\mathsf{cut}(c, d) = 2$

(e) $\mathsf{cut}(b, d) = 2$, $\mathsf{cut}(b, c) = 2$, and $\mathsf{cut}(c, d) = 1$. ✓

13. Let A be a 3×6 matrix with real-valued entries. Matrix A has rank 3. We construct a graph with 6 vertices where each vertex represents a distinct column in A, and there is an edge between two vertices if the two columns represented by the vertices are linearly independent. Which of the following statements MUST be true of the graph constructed?

(a) Each vertex has degree at most 2.

(b) The graph is connected.

(c) There is a clique of size 3. ✓

(d) The graph has a cycle of length 4.

(e) The graph is 3-colourable.

14. Consider the following greedy algorithm for colouring an n-vertex undirected graph G with colours $c_1, c_2, \ldots$: consider the vertices of G in any sequence and assign the chosen vertex the first colour that has not already been assigned to any of its neighbours. Let $m(n, r)$ be the minimum number of edges in a graph that causes this greedy algorithm to use r colours. Which of the following is correct?

(a) $m(n, r) = \Theta(r)$

(b) $m(n, r) = \Theta(r\lceil \log_2 r \rceil)$

(c) $m(n, r) = \binom{r}{2}$ ✓

(d) $m(n, r) = nr$

(e) $m(n, r) = n\binom{r}{2}$

15. Let $A[i] : i = 0, 1, 2, \ldots, n-1$ be an array of n distinct integers. We wish to sort A in ascending order. We are given that each element in the array is at a position that is at most k away from its position in the sorted array, that is, we are given that $A[i]$ will move to a position in $\{i-k, i-k+1, \ldots, i, \ldots, i+k-1, i+k\}$ after the array is sorted in ascending order. Suppose insertion sort is used to sort this array: that is, in the i-th iteration, $A[i]$ is compared with the elements in positions $A[i-1], A[i-2], \ldots$ until one that is smaller is found and $A[i]$ is inserted after that element. Note that elements can be moved back when later insertions are made before them. Let $t(n)$ be the worst-case number of comparisons made by insertion sort for such inputs. Then,

(a) $t(n) = \Theta(n^2)$

(b) $t(n) = \Theta(n \log_2 n)$

(c) $t(n) = \Theta(nk \log k)$

(d) $t(n) = \Theta(n \log_2 k)$

(e) $t(n) = \Theta(nk)$ ✓

—End of CS Section—

(SS section begins in the following page)

Part C: Systems Science

Note:
Only for Systems Science stream candidates.

1. Consider a system with input $x(t)$ and output $y(t)$ such that

$$y(t) = t\ x(t).$$

Consider the following statements:

1. The system is linear.
2. The system is time-invariant.
3. The system is causal.

Then which of the following is TRUE?

(a) Only statement 1 is correct.
(b) Only statement 2 is correct.
(c) Only statement 3 is correct.
(d) Only statements 1 and 3 are correct. ✓
(e) All three statements 1, 2, and 3 are correct.

2. Given a fixed perimeter of 1, among the following shapes, which one has the largest area?

(a) Square
(b) A regular pentagon
(c) A regular hexagon
(d) A regular septagon
(e) A regular octagon ✓

3. Consider the following statements:

1. $\lim_{x\to 0} \frac{\sin x}{x} = 1.$
2. $\lim_{x\to 0} \frac{1-\cos x}{x^2} = 1.$
3. $\lim_{x\to 0} \frac{1-\cos x}{x} = 1.$

Which of the following is TRUE?

(a) Only Statement 1 is correct. ✓
(b) Only Statements 1 and 2 are correct.
(c) Only Statements 1 and 3 are correct.
(d) All of Statements 1, 2, and 3 are correct.

(e) None of the three Statements 1,2, and 3 are correct.

4. The first-order differential equation $\frac{dy(t)}{dt}+2y(t) = x(t)$ describes a particular continuous-time system initially at rest at origin i.e., $x(0) = 0$. Consider the following statements?

(1) System is memoryless.

(2) System is causal.

(3) System is stable.

Which of the following fact about the statements above is true? (*Hint: You may try answering by finding the impulse response*)

(a) Only (1) is correct.

(b) Only (1) and (2) are correct.

(c) All (1), (2) and (3) are correct.

(d) Only (2) and (3) are correct. ✓

(e) None of the above

5. Recall that

$$\text{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$$

and convolution of functions $x(t)$ and $y(t)$ is defined as

$$x(t) \star y(t) = \int_{-\infty}^{\infty} x(t-\tau)y(\tau)d\tau.$$

What is the necessary and sufficient condition on positive real numbers f and a such that the following is true for some *non-zero* real number K (which may depend on f and a)?

$$\text{sinc}^2(at) \star \cos(2\pi ft) = K\cos(2\pi ft), \qquad \text{for all real } t.$$

(a) $f < a$ ✓

(b) $f > a$

(c) $f < a^{-1}$

(d) $f > a^{-1}$

(e) None of the above

6. Consider a fair coin (i.e., both heads and tails have equal probability of appearing). Suppose we toss the coin repeatedly until both sides have been seen. What is the expected number of times we would have seen heads?

(a) 1

(b) 5/4

(c) 3/2 ✓

(d) 2

(e) None of the above

7. Consider the function

$$f(y) = \int_1^y \frac{1}{1+x^2}dx - \log_e(1+y)$$

where $\log_e(x)$ denotes the natural logarithm of x.

Which of the following is true:

(a) The function $f(y)$ is non-positive for all $y \geq 1$. ✓

(b) The function $f(y)$ first increases and then decreases with y for $y \geq 1$.

(c) The function $f(y)$ first decreases and then increases with y for $y \geq 1$.

(d) The function $f(y)$ oscillates infinitely often between negative and positive value for $y \geq 1$.

(e) The derivative of function $f(y)$ does not exist at $y = 1$.

8. The maximum area of a parallelogram inscribed in the ellipse (i.e. all the vertices of the parallelogram are on the ellipse) $x^2 + 4y^2 = 1$ is:

(a) 2

(b) 4

(c) 1 ✓

(d) 5

(e) 3

9. A stick of length 1 is broken at a point chosen uniformly at random. Which of the following is false?

(a) Twice the length of the smaller piece is greater than the length of the larger piece with positive probability.

(b) One half of the length of the larger piece is greater than the length of smaller piece with positive probability.

(c) The product of the length of the smaller piece and the larger piece is less than $1/4$ in expectation.

(d) The ratio of the length of larger piece to the smaller piece is greater than 100 with positive probability.

(e) The product of the length of the smaller piece and the larger piece is greater than $1/4$ with positive probability. ✓

10. Suppose $\vec{u}, \vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$. Let the real number a_1^* be such that it solves the following optimization problem

$$d_1 = \min_{a_1 \in \mathbb{R}} \|\vec{u} - a_1\vec{v}_1\|,$$

where we denote the length $\sqrt{\vec{w}^T\vec{w}}$ of a vector $\vec{w} \in \mathbb{R}^n$ by $\|\vec{w}\|$. Let the pair of real numbers (b_1^*, b_2^*) be such that they solve the following optimization problem

$$d_2 = \min_{b_1, b_2 \in \mathbb{R}} \|\vec{u} - (b_1\vec{v}_1 + b_2\vec{v}_2)\|.$$

Let

$$\vec{p}_1 = a_1^* \vec{v}_1$$
$$\vec{p}_2 = b_1^* \vec{v}_1 + b_2^* \vec{v}_2$$

Compute $(\vec{p}_2 - \vec{p}_1)^T \vec{p}_1$.

(a) $\|\vec{u} - \vec{p}_1\|$

(b) $\|\vec{u} - \vec{p}_2\|$

(c) $\|\vec{u} - \vec{p}_2 - \vec{p}_1\|$

(d) $\|\vec{u} - (\vec{p}_2 - \vec{p}_1)\|$

(e) 0 ✓

11. Suppose that X_1 and X_2 denote the output of rolls of two independent dices that can each take integer values $\{1, 2, 3, 4, 5, 6\}$ with probability 1/6 for each outcome. Further, U denotes a continuous random variable that is independent of X_1 and X_2 and is uniformly distributed in the interval $[0, 1]$.

Suppose that the sum of the three random variables, that is, $X_1 + X_2 + U$, equals 6.63. Conditioned on this sum what is the probability that X_1 equals 2?

(a) 2.21

(b) 3

(c) 1/6

(d) 1/5 ✓

(e) 1/3

12. An ant does a random walk in a two dimensional plane starting at the origin at time 0. At every integer time greater than 0, it moves one centimeter away from its earlier position in a random direction independent of its past. After 4 steps, what is the expected square of the distance (measured in centimeters) from its starting point?

(a) 4 ✓

(b) 1

(c) 2

(d) π

(e) 0

13. Consider a unit Euclidean ball in 4 dimensions, and let V_n be its volume and S_n its surface area. Then S_n/V_n is equal to:

(a) 1

(b) 4 ✓

(c) 5

(d) 2

(e) 3

14. A tourist starts by taking one of the n available paths, denoted by $1, 2, \cdots, n$. An hour into the journey, the path i subdivides into further $1 + i$ subpaths, only one of which leads to the destination. The tourist has no map and makes random choices of the path and the subpaths. What is the probability of reaching the destination if $n = 3$?

(a) $\frac{10}{36}$

(b) $\frac{11}{36}$

(c) $\frac{12}{36}$

(d) $\frac{13}{36}$ ✓

(e) $\frac{14}{36}$

15. We have the sequence, $a_n = \frac{1}{n \log^2 n}$, $n \geq 2$, where log is the logarithm to the base 2 and let $A = \sum_{n=2}^{\infty} a_n$ be the sum of the sequence. We define a random variable and the corresponding distribution, $P(X = n) = p_n = \frac{a_n}{A}, n \geq 2$. Entropy or information of the random variable X is defined as $H(X) = \sum_{n=2}^{\infty} -p_n \log p_n$, where log is the logarithm to the base 2. Which of the following is true about the entropy $H(X)$?

(a) $H(X) \leq 3$

(b) $H(X) \in (3, 5]$

(c) $H(X) \in (5, 10]$

(d) $H(X) > 10$ but finite

(e) $H(X)$ is unbounded ✓

— End of paper —