Theoretical Computer Science Preliminary Exam (June 17th, 2025)

First name : ____________	Student ID : ____________
Last name : ____________	Additional pages : ____________

Instructions:

1. **All problems are worth 10 points.**
2. **Explain your answers clearly. Unclear answers will not receive credit.**
   State results and theorems you are using.
3. **Use the front and back of each page to write the solution of each problem.**
4. **If you need extra pages, please <u>do not</u> use the same sheet for different problems. .**
5. **Write your name and problem number on each additional page you use.**

1. Please answer the following questions. **No justification needed**.

   (a) (2 points) Let G be an undirected connected graph with n vertices and $n + 100$ edges (no self-loops nor parallel edges). What is the minimum number of edges that need to be removed from G so that G has no cycles?

   $\boxed{}$

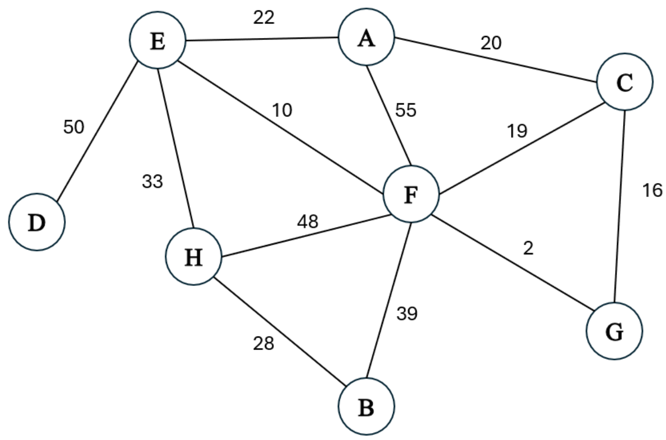
   (b) (2 points) Let $f(n) = a \cdot 2^{n-1} + b$ for $n \geq 1$ and positive *constants* a, b. Define $g(n) = \sum_{i=1}^{n} f(i)$. Which of the following are true? Check **all** that apply.

   ☐ $g(n) = O(f(n))$ ☐ $g(n) = \Omega(f(n))$ ☐ $g(n) = \Theta(f(n))$ ☐ $g(n) = o(f(n))$

   (c) (2 points) State whether the following statement is **True** or **False**: Consider a weighted undirected graph where all edges have **distinct** weights. This graph has only one, i.e., unique, minimum spanning tree.

   ☐ True ☐ False

2. (4 points) Execute **Kruskal's** algorithm for the Minimum Spanning Tree on the following graph with 8 vertices.



Note that the output is a set of $8 - 1 = 7$ edges. Complete Table 1. Write down the edge weight of the i^{th} edge in the i^{th} column, **in the order** they are included in the solution.

1	2	3	4	5	6	7

Table 1: The output edge weights of Kruskal's algorithm on the given graph. The i^{th} column denotes the weight of the i^{th} edge output by the algorithm.

3. (6 points) You are given a *connected* weighted graph $G = (V, E, w)$ on n vertices and m edges. You are also given two positive weights x and y, with $x < y$. The weight of each edge in G is guaranteed to be either x or y. Design an $O(m + n)$ running time algorithm that computes the weight of the minimum spanning tree of G.

   For the full score, analyze the running time of your solution and comment on its correctness. Your argument can be informal and brief but must be **informative**.

4. (10 points) **Automata theory**. Prove that the language $\{1^n | \text{n is prime}\}$ is not regular (the set of unary strings whose lengths are prime). Use either the Myhill-Nerode Theorem or the Pumping Lemma to prove this directly. (In other words, do not assume some other language is not regular and appeal to closure properties of regular languages, nor should you appeal to any known characterization of unary regular languages.) Hint: use the Prime Number Theorem, which states that as n goes to ∞, the number of prime numbers between 1 and n approaches $\frac{n}{\ln n}$.

5. (10 points) **Complexity theory**. A Boolean formula Φ with n variables can be thought to compute a function f_Φ: $\{0,1\}^n \to \{0,1\}$. Let $DIFF = \{(\Phi, \Psi) | f_\Phi \neq f_\Psi\}$ be the set of pairs of Boolean formulas that compute different functions. Prove that DIFF is NP-complete.
Hint 1: use the fact that SAT (the set of satisfiable Boolean formulas) is NP-complete.
Hint 2: What function does an unsatisfiable formula compute?

6. (10 points) **Computability theory.** Prove that the language

$$\mathrm{DIFF} - \mathrm{TM} = \{(M1, M2)|\text{M1 and M2 are Turing machines and } L(M1) \neq L(M2)\}$$

is not decidable (the set of pairs of Turing machines that recognize different languages). Recall that L(M), the language recognized by M, is the set of strings that M accepts. You may assume that the following language is not decidable: $Halts_\epsilon = \{M|\text{M is a Turing machine and } M(\epsilon) \text{ halts}\}$, where ϵ is the empty string.