Problem Set 2

- Due date: **25 Mar, 2025** (released on 7 Mar, 2025).
- The points for each problem is indicated on the side. The total for this set is **70** points
- The problem set has a fair number of questions so please do not wait until close to the deadline to start on them. Try and do one question every couple of days.
- Turn in your problem sets electronically (PDF; either LaTeXed or scanned etc.) via email. If you submit a LaTeX document, you will get an additional **10 points**.
- Collaboration with other students taking this course is encouraged, but collaboration with others is not allowed. Irrespective of this, all writeups must be done individually and must include names of all collaborators (if any).
- Referring to sources other than the text book and class notes is STRONGLY DISCOURAGED. But if you do use an external source (eg.,other text books, lecture notes, or any material available online), ACKNOWLEDGE all your sources (including collaborators) in your writeup. This will not affect your grades. However, not acknowledging will be treated as a serious case of academic dishonesty.

---

1. **[Improving the time-hierarchy theorem]** (15)

   You may assume that any 'reasonable-looking' function is time-constructible. And whenever you see functions like $n^2 \log^{3/4}(n)$, assume that there is an implicit ceiling to make sure this is an integer etc. (Basically don't worry about technicalities!)

   In class we proved that the deterministic time hierarchy theorem that stated the following:

   > Suppose $t_1, t_2 : \mathbb{N} \to \mathbb{N}$ are non-decreasing time-constructible functions with $t_1(n), t_2(n) \geq n$. If we have $t_1(n) \log t_1(n) = o(t_2(n))$, then we have $\mathsf{DTIME}(t_1) \subsetneq \mathsf{DTIME}(t_2)$.

   (a) Let $t_1, t_2, f : \mathbb{N} \to \mathbb{N}$ be time-constructible non-decreasing functions that satisfy $t_1(n), t_2(n), f(n) \geq n$. Show that $\mathsf{DTIME}(t_1(n)) = \mathsf{DTIME}(t_2(n))$ implies

   $$\mathsf{DTIME}(t_1(f(n))) = \mathsf{DTIME}(t_2(f(n))).$$

   [Hint: Padding.]

   (b) Show that $\mathsf{DTIME}(n^2) \subsetneq \mathsf{DTIME}(n^2 \log^{3/4}(n))$.

   [Hint: You may have to use the above part multiple times.]

   (c) Extend this to show that for any rational number a, ε satisfying $a > 1$ and $0 < \varepsilon < 1$, we have

   $$\mathsf{DTIME}(n^a) \subsetneq \mathsf{DTIME}(n^a (\log n)^{\varepsilon}).$$

2. **[Sparse NP-hard languages probably don't exist]** (10)

   Recall that a language L is said to be sparse if there is a constant c such that $\left|L \cap \{0,1\}^n\right| = O(n^c)$.

   Show that if L is a sparse NP-hard language, then $\mathsf{P} = \mathsf{NP}$.

[Hint: Think of a method to amplify the number of satisfiable formulas in a bag of formulas. Specifically, suppose you had a bag B of formulas, can you think of some operation you can do on these formulas and build a new bag B' with the following two properties — If B had no satisfiable formula, then neither does B'; and if B had even one satisfiable formula, then B' has *many* satisfiable formulas.]

3. **[Equivalence of quantified formulas]** (10)

Consider the following task where you are given two quantified expressions, Ψ_1 and Ψ_2:

$$\Psi_1 = \exists x \in \{0,1\}^n \, \forall y \in \{0,1\}^n \ : \ \varphi_1(x,y)$$
$$\Psi_2 = \forall x \in \{0,1\}^n \, \exists y \in \{0,1\}^n \ : \ \varphi_2(x,y)$$

Consider the following language L:

$$L = \left\{ (\Psi_1, \Psi_2) \ : \ \begin{array}{c} \Psi_1 \text{ is a } \Sigma_2\text{-expression, and } \Psi_2 \text{ is a } \Pi_2\text{-expression} \\ \text{with } \Psi_1 \equiv \Psi_2 \end{array} \right\}$$

(recall that $A \equiv B$ is just shorthand for "either both are true, or both are false.")

What level of the polynomial hierarchy is this language in?

[Hint: $x \equiv y$ is just $(x \wedge y) \vee (\neg x \wedge \neg y)$.]

4. **[Addition and multiplication in logspace]** (3 + 7 + 0)

(a) Consider the task where you are given two n-bit numbers and wish to compute the sum of these two numbers. Show that this is in LOGSPACE.

(b) Consider the task where you are given n integers of n-bits each, and wish to compute the sum of these n numbers. Show that this is also in LOGSPACE. As a consequence, show that the task computing the product of two given n-bit numbers is in LOGSPACE.

(c) What is your guess about computing the *product* of n-integers of n-bits each? Do you think this is in LOGSPACE?

5. **[Equivalence of regular expressions]** (10)

Consider the task where you are provided two regular expressions R_1 and R_2 and wish to check if they both give the same language.

Show that this problem is in PSPACE.

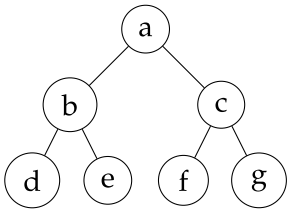
[Hint: It might actually be easier to show that it is in coNPSPACE (i.e., construct a non-deterministic space-bounded TM that checks if two regular expressions are *different*.)

As a first step, suppose the two regular expressions of size at most s each are different, what is the length of the shortest string that is accepted by one but not the other?]

6. **[Boolean Formula Evaluation]** (7 + 8)

(i) Prove that computing the DFS order (the order of vertices visited, including repetitions, in a DFS traversal that starts at the root and ends at the root) of an *undirected binary tree* $T = (V, E)$ can be done in L (logspace).

For example, the DFS order of the tree below is $a, b, d, b, e, b, a, c, f, c, g, c, a$.



[Hint: (a) You may assume that the tree is described as follows: For every vertex $v \in V$, there is a function $next_v : V \to V \cup \{\perp\}$ which gives a clockwise ordering of the edges around the vertex V. I.e., For every vertex v, there is a cyclic ordering among the neighbours of v and $next_v(u)$ is the next neighbour in this cyclic ordering if u is a neighbour of v and $\perp$ otherwise. Finally, check that one can make this assumption without loss of generality.]

(ii) A Boolean formula φ on n inputs is a directed tree with n *sources* (vertices with no incoming edges) and one *sink* (vertex with no outgoing edges). All nonsource vertices are called *gates* and are labeled with one of $\vee, \wedge$ or $\neg$. The vertices labeled with $\vee$ or $\wedge$ have fan-in 2 and the vertices labeled with $\neg$ have fan-in 1. Let $x \in \{0, 1\}^n$ be some input. The output of φ on x, denoted by $\varphi(x)$, is defined in the natural way. The Boolean formula evaluation problem deals with, given a formula φ on n inputs and $x \in \{0, 1\}^n$, computing the value of $\varphi(x)$. Show that formula evaluation can be done in logspace. More precisely, define

$$\text{FormulaEval} = \{\langle \varphi, x \rangle \ : \ \varphi \text{ is a Boolean formula and } \varphi(x) = 1\}$$

Prove that FormulaEval $\in$ L.