Tutorial 5
Complexity Theory

Time Complexity

1. For each of the following statements, answer _True_, _False_ or _Open-Question_ according to our current state of knowledge of complexity theory, as described in class. Give brief justifications for your answers.

   (a) $\mathbf{P} \subseteq \mathbf{TIME}\left(n^{2026}\right)$?

   (b) SAT $\leq_P$ $\overline{\text{SAT}}$?

   (c) HAMPATH $\leq_P$ PATH ?

   (d) PATH $\leq_P$ $\overline{\text{PATH}}$?

2. Prove that the following languages (defined over graphs) are in **P**.

   (a) BIPARTITE : the set of all bipartite graphs, i.e., $G = (V, E) \in$ BIPARTITE if V can be partitioned into two sets V_1, V_2 such that every edge in E is adjacent to a vertex in V_1 and a vertex in V_2 (no edge falls inside V_1 or V_2).

   (b) TRIANGLE-FREE : the set of all graphs that do not contain a triangle (where triangle is a set of three distinct vertices that are mutually connected).

3. Normally, we assume that numbers are represented as strings using the binary basis. That is, a number n is represented by the sequence $x_0, x_1, \ldots, x_{\log n}$ such that $n = \sum_{i=0}^{\log n} x^i 2^i$. However, we could have used other encoding schemes. If $n \in \mathbb{N}$ and $b \geq 2$, then the representation of n in base b, denoted by $\llcorner n \lrcorner_b$ is obtained as follows: first represent n as a sequence of digits in $\{0, \ldots, b-1\}$, and then replace each digit by a sequence of zeroes and ones. The unary representation of n, denoted by $\llcorner n \lrcorner_1$ is the string 1^n (that is, a sequence of n ones).

   (a) Show that choosing a different base of representation (other than unary) will make no difference to the class **P**. That is, show that for every subset S of the natural numbers, if we define $L_S^b = \left\{ \llcorner n \lrcorner_b \mid n \in S \right\}$, then for every $b \geq 2$, $L_S^b \in \mathbf{P}$ iff $L_S^2 \in \mathbf{P}$.

   (b) Show that choosing the unary representation makes a difference by showing that the following language is in **P**.

   $$\text{UNARY-FACTORING} = \left\{ \langle \llcorner n \lrcorner_1, \llcorner k \lrcorner_1 \rangle \mid \text{there is a } j \leq k \text{ dividing } n \right\}$$

4. Let DOUBLE-SAT $= \left\{ \langle \varphi \rangle \mid \varphi \text{ is a CNF formula having at least two satisfying assignments} \right\}$. Show that DOUBLE-SAT is **NP-complete**.

5. Consider a variant M-3SAT (minimal 3SAT) of 3SAT defined as follows: a 3-CNF Boolean formula φ is in M-3SAT if there exists atleast one satisfying assignment for φ which makes exactly one literal in every clause true. Show that M-3SAT is **NP-complete**.

6. Let φ be a Boolean formula in CNF. An assignment to the variables of φ is called _not-all-equal_ if in each clause, at least one literal is assigned 1 and at least one literal is 0. Show that the following language is **NP-complete**. Assume no constants are allowed in the formulae.

   $$\text{NESAT} = \{\varphi \mid \varphi \text{ is a CNF formula having a satisfying } \textit{not-all-equal} \text{ assignment}\}$$

7. Consider the following *solitaire* game. You are given an $m \times m$ board where each one of the m^2 positions may be empty or occupied by either a red stone or a blue stone. Initially, some configuration of stones is placed on the board. Then, for each column you must remove either all of the red stones in that column or all of the blue stones in that column. (If a column already has only red stones or only blue stones in it then you do not have to remove any further stones from that column.) The objective is to leave at least one stone in each row. Finding a *solution* that achieves this (mentioned) objective may or may not be possible depending upon the initial configuration. Prove that the following language is **NP-complete**.

$$\text{SOLITAIRE} = \{\langle G \rangle \mid G \text{ is a game configuration with a solution}\}$$

8. Let S be a finite set and $C = \{C_1, C_2, \ldots, C_k\}$ be a collection of subsets of S, for some $k > 0$. We say S is 2-colourable with respect to C if we can colour each element of S in red or blue, such that every C_i contains at least one red element and at least one blue element. Show that the language, $\text{2COLOUR} = \{\langle S, C \rangle \mid S \text{ is 2-colourable with respect to } C\}$ is **NP-complete**.

9. Let $n(G)$ denotes the number of vertices in G and $\lfloor\ \rfloor$ is the floor function.

   Prove / Disprove: The following languages are **NP-complete**.

   (a) $\text{SMALLCYCLE} = \left\{\langle G \rangle \mid \text{The longest cycle in the directed graph } G \text{ is of length } \leq \left\lfloor \frac{n(G)}{2} \right\rfloor\right\}$

   (b) $\text{BIGCYCLE} = \left\{\langle G \rangle \mid G \text{ is a directed graph having a cycle of length } \geq \left\lfloor \frac{n(G)}{2} \right\rfloor\right\}$

   You may make use of the assumption that **NP** ≠ **co-NP**, if necessary. Clearly indicate where you require this assumption.

10. A *vertex cover* in a graph $G = (V, E)$ is a set of vertices $S \subseteq V$ such that every edge of G is incident on at least one vertex in S. Show that the following language is **NP-complete**.

$$\text{VERTEX-COVER} = \{(G, k) \mid \text{graph } G \text{ has a vertex cover of size} \leq k\}$$

11. Let S be a set and let $C = \{X_1, X_2, \ldots, X_n\}$ be a collection of n subsets of S (for each $i \in [1, n], X_i \subseteq S$). A set S', with $S' \subseteq S$, is called a *hitting set* for C if every subset in C contains at least one element in S', i.e., $|X_i \cap S'| \geq 1$ for each $i \in [1, n]$. Let $\text{HITSET} = \{\langle C, k \rangle \mid C \text{ has a hitting set of size } k\}$. Prove that HITSET is **NP-complete**.

    Example: $S = \{a, b, c, d, e, f, g\}$, $C = \{\{a, b, c\}, \{d, a\}, \{d, e, f\}, \{g\}\}$

    - $k = 2$, no hitting sets exist.
    - $k = 3$, $S' = \{a, d, g\}$ (other choices exist).

    Hint: Try reducing from VERTEX-COVER.

12. Prove that **P** = **coP** and **P** ⊆ **NP** ∩ **coNP**.

13. Assuming **NP** ≠ **coNP**, show that no **NP-complete** problem can be in **coNP**.

14. Show that the halting problem is **NP-hard**.

15. **[Scaling Resource Bounds]** Let $\mathbf{CL}_1, \mathbf{CL}_2$ denote some time/space complexity classes. Show that, if $\mathbf{CL}_1(f(n)) \subseteq \mathbf{CL}_2(g(n))$, then $\mathbf{CL}_1(f(n^c)) \subseteq \mathbf{CL}_2(g(n^c))$.

16. The following two classes are exponential time analogues of **P** and **NP**.

$$\mathbf{EXP} = \bigcup_{c \geq 1} \mathbf{DTIME}(2^{n^c}) \quad \text{and} \quad \mathbf{NEXP} = \bigcup_{c \geq 1} \mathbf{NTIME}(2^{n^c})$$

    Clearly, **P** ⊆ **NP** ⊆ **EXP** ⊆ **NEXP**. Show that, if **EXP** ≠ **NEXP**, then **P** ≠ **NP**.

    Hint: Consider padding strings in **EXP**/**NEXP** *languages with exponentially sized strings in order to "scale down" to* **P**/**NP**.

Space Complexity

1. Is $\mathbf{NSPACE}(n^{2026}) \subseteq \mathbf{PSPACE}$? Answer stating *True*, *False* or *Open-Question* according to our current state of knowledge of complexity theory, as described in class. Justify.

2. Show that,

   (a) If every **NP-hard** language is **PSPACE-hard**, then **PSPACE** = **NP**.

   (b) **PSPACE** is closed under union, intersection, complement and Kleene star operations.

3. A *ladder* is a sequence of strings $s_1, s_2, \ldots, s_k$, wherein every string differs from the preceding one in exactly one character. For example, the following is a ladder of English words, starting with "$head$" and ending with "$free$":

   $$head \to hear \to near \to fear \to bear \to beer \to deer \to deed \to feed \to feet \to fret \to free$$

   Let, $\mathsf{LADDER_{DFA}} = \left\{ \langle \mathcal{M}, s, t \rangle \mid \mathcal{M} \text{ is a DFA defined over } \Sigma \text{ and } \mathcal{L}(\mathcal{M}) \text{ consists of a ladder of strings starting with } s \text{ and ending with } t, \text{ where } s, t \in \Sigma^* \right\}$,

   Show that $\mathsf{LADDER_{DFA}}$ is in **PSPACE**. [*Hint: Use the fact that* **PSPACE** = **NPSPACE**.]

4. In the generalized version of the game Tic-Tac-Toe, 2 players places marks X (crosses) and O (noughts) on an $m \times n$ grid. A player wins if she is the first to place k marks in a row, column or diagonal. The game ends in a draw if no such sequence is present when all the mn cells of the grid are filled. Assuming that X always starts, show that the following language,

   $\mathsf{GTICTACTOE} = \left\{ \langle m, n, k, c \rangle \mid c \text{ is an intermediate configuration on the } m \times n \text{ board with next move by } X \text{ and } \exists \text{ a winning strategy for } X \right\}$,

   is in **PSPACE**.

5. An undirected graph is k-colorable if the vertices of the graph can be assigned a colour from a given fixed set of k colors such that no two adjacent vertices (sharing a common edge in between) receive the same color. Consider the following language,

   $\mathsf{3COLOR} = \left\{ \langle G \rangle \mid \text{The undirected graph } G \text{ is 3-colorable} \right\}$

   Note that, 3COLOR is known to be **NP-complete**. Prove that 3COLOR is **PSPACE-complete** if and only if **PSPACE** = **NP**.