Computational Complexity

Take-home exam

Hand in via Canvas before March 29, 2021, at 17:00

Exercise 1 (*2pt*)**.** Consider the following decision problem TOP3. In this problem, the input consists of a tuple (A, T, u), where $A = \{a_1, \ldots, a_m\}$ is a finite set, $T = (t_1, \ldots, t_k)$ is a sequence of subsets $t_i \subseteq A$ of size exactly 3, and $u \in \mathbb{N}$ is a natural number. Intuitively, each such tuple (A, T, u) describes a collective choice problem, where we are to select a subset of u alternatives from the ones in A, and where k individuals gave their preferences in the form of submitting a top three. The problem is to decide whether there exists a subset $C \subseteq A$ such that $|C| = u$ and such that $C \cap t_i \neq \emptyset$ for each $1 \leq i \leq k$. In other words, the answer is yes if and only if we can choose u alternatives from A such that we pick at least one alternative from everyone's top three.

Prove that the problem TOP3 is complete for some complexity class $\mathrm{K} \in \{\mathsf{NP}, \mathsf{coNP}, \Sigma_2^{\mathrm{p}}, \Pi_2^{\mathrm{p}}, \mathsf{PSPACE}\}$.

Definition 1. Consider the following two languages L_1 and L_2 for specifying sets $B \subseteq \{0,1\}^m$ of binary strings of a particular length m.[1]

- L_1: In language L_1, each expression consists of a propositional logic formula φ over the variables $x_1, \ldots, x_m$ for some $m \in \mathbb{N}$. Each such expression φ over the variables $x_1, \ldots, x_m$ is interpreted as the set $\sigma_1(\varphi)$ of binary strings $b = b_1 b_2 \ldots b_m \in \{0,1\}^m$ such that α_b satisfies φ, where $\alpha_b : \{x_1, \ldots, x_m\} \to \{0,1\}$ is the truth assignment such that $\alpha_b(x_i) = b_i$.
- L_2: In language L_2, each expression consists of a sequence $(\varphi_1, \ldots, \varphi_m)$ propositional logic formulas. Each such expression $(\varphi_1, \ldots, \varphi_m)$ is interpreted as the set $\sigma_2(\varphi_1, \ldots, \varphi_m)$ of binary strings $b = b_1 b_2 \ldots b_m \in \{0,1\}^m$ such that $\bigwedge \Phi_b$ is satisfiable, where $\Phi_b = \{\ \varphi_i \mid 1 \leq i \leq m, b_i = 1\ \} \cup \{\ \neg\varphi_i \mid 1 \leq i \leq m, b_i = 0\ \}$.

Example 1. Consider the set $B = \{0,1\}^3 \setminus 111$. The following two expressions in the languages L_1 and L_2, respectively, express this set B.

- The following expression φ in the language L_1 is such that $\sigma_1(\varphi) = B$:

$$\varphi = \neg(x_1 \wedge x_2 \wedge x_3).$$

- The following expression $(\varphi_1, \varphi_2, \varphi_3)$ in the language L_2 is such that $\sigma_2(\varphi_1, \varphi_2, \varphi_3) = B$:

$$\varphi_1 = x_1 \wedge x_3, \quad \varphi_2 = x_2 \wedge x_4, \quad \varphi_3 = \neg(x_1 \wedge x_2).$$

Exercise 2 (*5pt; a: $1\frac{1}{2}$pt, b: $1\frac{1}{2}$pt, c: 2pt, d: $\frac{1}{2}$pt bonus*)**.**

(a) Show that if there exists a polynomial-time algorithm that, given any expression φ of language L_1 with $\sigma_1(\varphi) \neq \emptyset$, produces an equivalent expression $(\varphi_1, \ldots, \varphi_m)$ of language L_2—i.e., such that $\sigma_1(\varphi) = \sigma_2(\varphi_1, \ldots, \varphi_m)$—then $\mathsf{P} = \mathsf{NP}$.

(b) Show that if there exists a polynomial-time algorithm that, given any expression $(\varphi_1, \ldots, \varphi_m)$ of language L_2, produces an equivalent expression φ of language L_1—i.e., such that $\sigma_1(\varphi) = \sigma_2(\varphi_1, \ldots, \varphi_m)$—then $\mathsf{P} = \mathsf{NP}$.

(c) Show that if there exist a polynomial p and an algorithm[2] that, given any expression $(\varphi_1, \ldots, \varphi_m)$ of language L_2, produces an equivalent expression φ of language L_1—i.e., such that $\sigma_1(\varphi) = \sigma_2(\varphi_1, \ldots, \varphi_m)$—with $|\varphi| \leq p(|(\varphi_1, \ldots, \varphi_m)|)$, then the Polynomial Hierarchy collapses.

  – *Hint:* use the fact that $\mathsf{NP} \subseteq \mathsf{P/poly}$ implies that $\mathsf{PH} = \Sigma_2^{\mathrm{p}}$.

[1] The choice of m is not fixed. In other words, the languages L_1 and L_2 contain expressions for sets $B \subseteq \{0,1\}^m$ of differing lengths m.
[2] Note that we do not put any requirements on the running time of this algorithm.

– *Hint:* consider expressions $(\varphi_1, \ldots, \varphi_{(2\ell)^3}, \varphi_{(2\ell)^3+1})$ of language L_2, for different values of $\ell \in \mathbb{N}$, where $\varphi_i = y_i$ for each $1 \leq i \leq (2\ell)^3$, and where:

$$\varphi_{(2\ell)^3+1} = \bigwedge_{1 \leq i \leq (2\ell)^3} (y_i \to c_i),$$

where $c_1, \ldots, c_{(2\ell)^3}$ is an enumeration of all possible clauses of size 3 over the variables $x_1, \ldots, x_\ell$.

(d) Show that there exist a polynomial p and an algorithm[2] that, given any expression φ of language L_1 such that $\sigma_1(\varphi) \neq \emptyset$, produces an equivalent expression $(\varphi_1, \ldots, \varphi_m)$ of language L_2—i.e., such that $\sigma_1(\varphi) = \sigma_2(\varphi_1, \ldots, \varphi_m)$—with $|(\varphi_1, \ldots, \varphi_m)| \leq p(|\varphi|)$.

Definition 2. Consider a sequence $\mathcal{D} = \{\mathcal{D}_n\}_{n \in \mathbb{N}}$ of probability distributions, where each $\mathcal{D}_n$ is a probability distribution over $\{0,1\}^n$—that is, each $\mathcal{D}_n : \{0,1\}^n \to [0,1]$ is such that $\sum_{x \in \{0,1\}^n} \mathcal{D}_n(x) = 1$.

- $\mathcal{D}$ is *nice* if for each $n \in \mathbb{N}$ and each $x \in \{0,1\}^n$ it holds that $\mathcal{D}_n(x) = \ell/2^n$ for some $\ell \in \mathbb{N}$.
- $\mathcal{D}$ is P-*computable* if there exists a polynomial-time TM that, given $n \in \mathbb{N}$ and $x \in \{0,1\}^n$, computes the value $\mu_{\mathcal{D}_n}(x) = \sum_{y \in \{0,1\}^n, y \leq x} \mathcal{D}_n(y)$, where $\leq$ denotes the lexicographical (i.e., alphabetical) order over strings.
- $\mathcal{D}$ is P-*samplable* if there exists a polynomial-time probabilistic TM $\mathbb{M}$ such that for each $n \in \mathbb{N}$, the random variables $\mathbb{M}(1^n)$ is identically distributed to $\mathcal{D}_n$.

Definition 3. The complexity class PP consists of all decision problems $L \subseteq \{0,1\}^*$ for which there exists a polynomial-time probabilistic TM $\mathbb{M}$ such that for each $x \in \{0,1\}^*$ it holds that $\mathbb{P}[\, \mathbb{M}(x) = L(x) \,] > 1/2$, where $L(x) = 1$ if $x \in L$, and $L(x) = 0$ if $x \notin L$.

Exercise 3 (*3pt; a: $1\frac{1}{2}$pt, b: $1\frac{1}{2}$pt*)**.** Consider sequences $\mathcal{D} = \{\mathcal{D}_n\}_{n \in \mathbb{N}}$ of distributions as described in Definition 2.

(a) Prove that each nice and P-computable sequence $\mathcal{D} = \{\mathcal{D}_n\}_{n \in \mathbb{N}}$ is P-samplable.

(b) Prove that if each P-samplable sequence $\mathcal{D} = \{\mathcal{D}_n\}_{n \in \mathbb{N}}$ is P-computable, then $\mathsf{PP} = \mathsf{P}$.

(Note: the first statement is also true for arbitrary $\mathcal{D}$—i.e., sequences $\mathcal{D}$ that are not necessarily nice. For this assignment, you only need to prove it for nice sequences.)