KØBENHAVNS UNIVERSITET

Computability and Complexity: Problem Set 4

Due: Friday April 12 at 23:59 AoE.
Submission: Please submit your solutions via *Absalon* as a PDF file. State your name and e-mail address at the top of the first page. Solutions should be written in LaTeX or some other math-aware typesetting system with reasonable margins on all sides (at least 2.5 cm). Please try to be precise and to the point in your solutions and refrain from vague statements. Never just state an an answer, but make sure to also explain your reasoning. *Write so that a fellow student of yours can read, understand, and verify your solutions.* In addition to what is said below, the general rules for problem sets stated on the course webpage always apply.
Collaboration: Discussions of ideas in groups of two to three people are allowed—and indeed, encouraged—but you should always write up your solutions completely on your own, from start to finish, and you should understand all aspects of them fully. It is not allowed to compose draft solutions together and then continue editing individually, or to share any text, formulas, or pseudocode. Also, no such material may be downloaded from or generated via the internet to be used in draft or final solutions. Submitted solutions will be checked for plagiarism. You should also clearly acknowledge any collaboration. State close to the top of the first page of your problem set solutions if you have been collaborating with someone and if so with whom. *Note that collaboration is on a per problem set basis, so you should not discuss different problems on the same problem set with different people.*
Reference material: Some of the problems are "classic" and hence it might be possible to find solutions on the internet, in textbooks or in research papers. It is not allowed to use such material in any way unless explicitly stated otherwise. Anything said during the lectures or in the lecture notes, or any material found in Arora-Barak, should be fair game, though, unless you are specifically asked to show something that we claimed without proof in class. All definitions should be as given in class or in Arora-Barak and cannot be substituted by versions from other sources. It is hard to pin down 100% watertight, formal rules on what all of this means—when in doubt, ask the main instructor.
Grading: A total score of 70 points will be enough for grade 02, 110 points for grade 4, 150 points for grade 7, 190 points for grade 10, and 230 points for grade 12 on this problem set. Any revised versions of the problem set with clarifications and/or corrections will be posted on the course webpage `jakobnordstrom.se/teaching/CoCo24/`.
Questions: Please do not hesitate to ask the instructors or TA if any problem statement is unclear, but please make sure to send private messages when using Absalon—sometimes specific enough questions could give away the solution to your fellow students, and we want all of you to benefit from working on, and learning from, the problems. Good luck!

1 (10 p) Prove that any (non-constant) monotone Boolean function can be computed by a monotone Boolean circuit.

2 (30 p) For a language $L \subseteq \{0,1\}^*$, let $L_k = \{x \in L; |x| \leq k\}$ denote all strings in L of length at most k. We say that L is *downward self-reducible* if there is a polynomial-time algorithm A that given x and oracle access to $L_{|x|-1}$ decides correctly whether $x \in L$ or not.

Prove that if a language L is downward self-reducible, then it must hold that $L \in \mathsf{PSPACE}$.

3 (40 p) Let R_t denote the set of all restrictions of subsets of exactly t out of n variables, where n is supposed to be large and $t \geq n/2$. When proving Håstad's switching lemma, we argued that the set $B \subseteq R_t$ of *bad* restrictions for which the conclusion of the lemma does not hold is very small compared to R_t, and hence it is very unlikely that a random restriction will be bad.

More formally, we constructed a one-to-one mapping from B to $R_{t+s} \times \{0,1\}^\ell$ for some $\ell = \mathrm{O}(s \log k)$, and claimed this showed that the probability to get a bad restriction is

$$\frac{|B|}{|R_t|} \leq \frac{\left|R_{t+s} \times \{0,1\}^\ell\right|}{|R_t|} = n^{-\Omega(s)} \,.$$

The purpose of this problem is to fill in the details in these calculations and show that one gets a failure probability for the restriction as small as the one claimed in the Arora-Barak textbook.

That is, just trusting that the one-to-one map $m : B \to R_{t+s} \times \{0,1\}^\ell$ constructed in class was correct, show that the qoutient $\left|R_{t+s} \times \{0,1\}^\ell\right|/|R_t|$ is small enough to give the probability bound in the switching lemma as stated in the textbook.

Hint: Show that for $t > n/2$ it holds that

$$\binom{n}{t+s} \leq \binom{n}{t} \left(\frac{e(n-t)}{n}\right)^s$$

by first proving

$$\binom{n}{t+s} = \binom{n}{t}\binom{n-t}{s} \Big/ \binom{t+s}{t}$$

(and try to find a nice combinatorial proof for this latter equality). You can use the well-known inequalities

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k$$

without proof.

4 (30 p) When proving $\textsc{Parity} \notin \mathsf{AC}^0$, our starting point was a bounded-depth polynomial-size circuit C that supposedly computed the parity of its input bits. Before proving the actual lower bound, we claimed that such a circuit C can be preprocessed to get an equivalent circuit C' such that:

1. All gates in C' have fan-out 1 (i.e., it is what is known as a *formula*, with a DAG structure that is a tree).
2. All NOT ($\neg$) gates are at the input level of C' (i.e., they only apply to variables).
3. The AND ($\wedge$) and OR ($\vee$) gates alternate, so that at each level of C' all gates are either AND or OR.
4. The bottom level has AND gates of some small, bounded fan-in (for the purposes of this problem, let us say some global constant K).

Show how these modifications can be done without increasing the circuit depth by more than a constant and the size more than polynomially (so that C' is also a bounded-depth polynomial-size circuit computing the same function as C). If C is a circuit of size S and depth d, what size and depth do you get for C'?

5 (30 p) In this problem, you will prove a special case of Håstad's switching lemma, which is already quite interesting. Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function that can be represented as a k-DNF formula F. Let R_t be the set of all possible restrictions on the n Boolean variables of f that set t variables.

Given a restriction $\rho \in R_t$, we will say that the restricted function $f\restriction_\rho$ is a constant function if it takes the same value (either 0 or 1) on all possible settings of its remaining $n - t$ input variables. (For example, if f is the OR function on n variables, then $f\restriction_\rho$ is a constant function if the restriction ρ sets any variable to 1.) Prove that for a random restriction ρ sampled uniformly at random from R_t it holds that

$$\Pr_\rho [f\restriction_\rho \text{ is not a constant function}] = O\left(\frac{(n-t)k}{n}\right) .$$

The proof should ideally be much simpler than the full-blown switching lemma you saw in class!

Hint: Follow the encoding proof of the switching lemma that you have already seen. Show how to map the set of bad restrictions B injectively into $R_{t+1} \times A$, where A is a set of size $O(k)$.

6 (50 p) A half-space $h : \mathbb{R}^d \to \{0,1\}$ is a function of the form

$$f(x) = f(x_1, \ldots, x_n) = \begin{cases} 1 & \text{if } \sum_{i=1}^n a_i x_i \geq b \\ 0 & \text{otherwise} \end{cases}$$

for some $a_1, \ldots, a_n, b \in \mathbb{R}$. Let $X \subset \mathbb{R}^d$ be a finite set of points. Let $f_1, \ldots, f_m : \mathbb{R}^d \to \{0,1\}$ be a collection of half-spaces. Think of each f_i as function from X to $\{0,1\}$. Prove that the VC dimension of the concept class $\{f_i : i \in [m]\}$ is at most $d + 1$.

7 (50 p) For every integer n and each $i \in \{0, 1, \ldots, n+1\}$, let $f_{n,i} : [n] \to \{0,1\}$ be defined by

$$f_{n,i}(x) = \begin{cases} 0 & x < i\,, \\ 1 & x \geq i\,. \end{cases}$$

Prove from first principles (i.e., directly from the definition) that there is a constant $C > 0$ such that for all integers n the sample complexity of PAC learning the class $\{f_{n,i} : i \in \{0, 1, \ldots, n+1\}\}$ is at most C.

8 (40 p) Prove from first principles (i.e., directly from the definition) that there is a constant $C > 0$ such that the sample complexity of PAC learning a class $H \subseteq \{0,1\}^n$ of size $|H| = m$ is at most $C \log_2 m$.

9 (50 p) Prove from first principles (i.e., directly from the definition) that the sample complexity of PAC learning $\{0,1\}^d$ is at least $\frac{d}{4}$.