Exercises 7-10

Circuit Complexity, Autumn 2024, University of Chicago
Instructor: William Hoza (williamhoza@uchicago.edu)

Submission. Solutions are due **Wednesday, November 6** at 5pm Central time. Submit your solutions through Gradescope. You are encouraged, but not required, to typeset your solutions using a LaTeX editor such as Overleaf.

The policies below can also be found on the course webpage.

Collaboration. You are encouraged to collaborate with your classmates on homework, but you must adhere to the following rules.

- Work on each exercise on your own for at least fifteen minutes before discussing it with your classmates.
- Feel free to explain your ideas to your classmates in person, and feel free to use whiteboards/chalkboards/etc. However, do not share any written/typeset solutions with your classmates for them to study on their own time. This includes partial solutions.
- Write your solutions on your own. While you are writing your solutions, do not consult any notes that you might have taken during discussions with classmates.
- In your write-up, list any classmates who helped you figure out the solution. The fact that student A contributed to student B's solution does not necessarily mean that student B contributed to student A's solution.

Permitted Resources for Full Credit. In addition to discussions with me and discussions with classmates as discussed above, you may also use any slides or notes posted in the "Course Timeline" section of the course webpage, and you may also use Wikipedia. If you wish to receive full credit on an exercise, you may not use any other resources.

Outside Resources for Partial Credit. If you wish, you may use outside resources (ChatGPT, Stack Exchange, etc.) to solve an exercise for partial credit. If you decide to go this route, you must make a note of which outside resources you used when you were working on each exercise. You must disclose using a resource even if it was ultimately unhelpful for solving the exercise. Furthermore, if you consult an outside resource while working on an exercise, then you must not discuss that exercise with your classmates.

Impagliazzo's Hard-Core Lemma and Yao's XOR Lemma provide sufficient conditions under which a function of interest is strongly average-case hard with respect to a dense distribution and the uniform distribution, respectively. In this exercise, you will study necessary and sufficient conditions under which a function of interest is strongly average-case hard with respect to *some* distribution, which is not necessarily dense or uniform or anything else.

Exercise 7 (10 points)**.** Let $\mathcal{C}$ be a class of functions $C\colon \{0,1\}^n \to \{0,1\}$, let $f\colon \{0,1\}^n \to \{0,1\}$, and let $t \in \mathbb{N}$.

(a) Assume that $f \in \mathsf{MAJ}_t \circ \mathcal{C}$, and assume that t is odd for simplicity's sake. Prove that for every distribution μ over $\{0,1\}^n$, there exists $C \in \mathcal{C}$ such that

$$\Pr_{x\sim\mu}[C(x) = f(x)] \geq \frac{1}{2} + \Omega\left(\frac{1}{t}\right).$$

Hint: What can you say about the function $s(x) := ((-1)^{C_1(x)} + \cdots + (-1)^{C_t(x)}) \cdot (-1)^{f(x)}$?

(b) Assume that for every distribution μ over $\{0,1\}^n$, there exists $C \in \mathcal{C}$ such that

$$\Pr_{x\sim\mu}[C(x) = f(x)] \geq \frac{1}{2} + \frac{1}{t}.$$

Prove that $f \in \mathsf{MAJ}_{O(n\cdot t^2)} \circ \mathcal{C}$.

Hint: Use von Neumann's minimax theorem.

The version of Impagliazzo's Hard-Core Lemma that we prove in class guarantees the existence of a dense hard-core *distribution*. In this exercise, you will show that the existence of such a distribution implies the existence of a large hard-core *set*. For simplicity, we focus on the case of AC^0 circuits.

Exercise 8 (10 points)**.** Let $h\colon \{0,1\}^n \to \{0,1\}$, let $\varepsilon, \delta \in (2^{-0.1n}, 1)$, let H be a δ-dense distribution over $\{0,1\}^n$, let $S, d \in \mathbb{N}$ where $S \leq 2^{0.1n}$, and assume that for every AC^0_d circuit C of size at most S, we have

$$\Pr_{x\sim H}[C(x) = h(x)] \leq 1/2 + \varepsilon.$$

Prove that there exists a set $H' \subseteq \{0,1\}^n$ such that $|H'| \geq \Omega(\delta \cdot 2^n)$ and for every AC^0_d circuit C of size at most S, we have

$$\Pr_{x\in H'}[C(x) = h(x)] \leq 1/2 + O(\varepsilon).$$

Hint: Use the probabilistic method. Independently for each $x \in \{0,1\}^n$, include x in H' with probability proportional to $H(x)$. For the analysis, apply Hoeffding's inequality to a suitable sum of independent random variables, some $\{0,1\}$-valued and others $\{0,-1\}$-valued.

Recall that Yao's XOR Lemma says (roughly speaking) that if h is moderately hard for $\mathsf{MAJ} \circ \mathcal{C}$, then $h^{\oplus k}$ is very hard for $\mathcal{C}$. In this exercise, we will see a special case in which the MAJ gate can be avoided, namely, the case that $\mathcal{C}$ consists of *shallow decision trees*.

A *depth-D decision tree* is a depth-D tree in which each internal node is labeled with a variable from among $\{x_1, \ldots, x_n\}$; each internal node has two outgoing edges labeled 0 and 1; and each leaf is labeled with an output value. Given an input $x \in \{0,1\}^n$, we start at the root. Whenever we reach an internal vertex v, say labeled with the variable x_i, we traverse the outgoing edge of v that is labeled with the value $x_i \in \{0,1\}$. Finally, when we reach a leaf, we output its value. For example, the function $f(x_1, x_2, x_3) = (x_1 \wedge x_2) \vee (\neg x_1 \wedge x_3)$ can be computed by a depth-2 decision tree: first we query x_1 and then, depending on the value of x_1, we query either x_2 or x_3. We can also consider decision trees that output values other than 0 and 1, such as ± 1.

For a function $h\colon \{0,1\}^n \to \{\pm 1\}$ and a number $D \in \mathbb{N}$, we define $\mathsf{Corr}_D(h)$ to be the maximum, over all depth-D decision trees $T\colon \{0,1\}^n \to \{\pm 1\}$, of the quantity

$$\mathop{\mathbb{E}}_{x \in \{0,1\}^n}[h(x) \cdot T(x)].$$

Furthermore, for each $i \in [n]$ and each $b \in \{0,1\}$, we define

$$h^{i \leftarrow b}(x) = h(x_1, x_2, \ldots, x_{i-1}, b, x_{i+1}, \ldots, x_n).$$

Exercise 9 (10 points)**.**

(a) Let $h\colon \{0,1\}^n \to \{\pm 1\}$ and let D be a positive integer. Prove that

$$\mathsf{Corr}_D(h) = \frac{1}{2} \cdot \max_{i \in [n]} \sum_{b \in \{0,1\}} \mathsf{Corr}_{D-1}(h^{i \leftarrow b}).$$

(b) Let $h_1, \ldots, h_k\colon \{0,1\}^n \to \{\pm 1\}$, let $h(x^{(1)}, \ldots, x^{(k)}) = \prod_{i=1}^k h_i(x^{(i)})$, and let $D \in \mathbb{N}$. Prove that

$$\mathsf{Corr}_D(h) \leq \prod_{i=1}^k \mathsf{Corr}_D(h_i).$$

Hint: Prove it by induction on D. For the inductive step, use part (a) twice.

(c) Let $h\colon \{0,1\}^n \to \{0,1\}$. Assume that for every depth-D decision tree $T\colon \{0,1\}^n \to \{0,1\}$, we have

$$\Pr_{x \in \{0,1\}^n}[T(x) = h(x)] \leq \frac{1}{2} + \varepsilon.$$

Prove that for every $k \in \mathbb{N}$ and every depth-D decision tree $T\colon \{0,1\}^{nk} \to \{0,1\}$, we have

$$\Pr_{x \in \{0,1\}^{nk}}[T(x) = h^{\oplus k}(x)] \leq \frac{1}{2} + \frac{1}{2} \cdot (2\varepsilon)^k.$$

Let $S_1, S_2, \ldots, S_n \subseteq [s]$, where $|S_i| = r$ for every i and $|S_i \cap S_j| < \log n$ for every $i \neq j$. Define a pseudorandom generator $G\colon \{0,1\}^s \to \{0,1\}^n$ by the formula

$$G(x) = \left(\bigoplus_{i \in S_1} x_i, \ldots, \bigoplus_{i \in S_n} x_i \right).$$

In class, we show that if we choose a suitable value $r = (\log(S/\varepsilon))^{O(d)}$, then G fools size-S AC^0_d circuits with error ε. In this exercise, you will show that G satisfies a different "pseudorandomness" property, namely, the output bits of G are *k-wise uniform*.

Definition 1 (k-wise uniformity)**.** Let X be a distribution over $\{0,1\}^n$, and let $k \in [n]$. We say that X is *k-wise uniform* if, for every $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, the substring $X_{i_1} X_{i_2} \ldots X_{i_k}$ is distributed uniformly over $\{0,1\}^k$.

Exercise 10 (7 points)**.** Show that $G(U_s)$ is k-wise uniform where $k = r/\log n$.