Themistoklis Gouleakis **Winter 2021**

Property Testing, Exercise Sheet 3

https://www.mpi-inf.mpg.de/departments/algorithms-complexity/teaching/winter20/property-testing

Total Points: 40 + 5   Due: Wednesday, **December 16**, 2020

You are allowed to collaborate on the exercise sheets, but you have to write down a solution on your own, ***using your own words****. Please indicate the names of your collaborators for each exercise you solve. Further, cite all external sources that you use (books, websites, research papers, etc.). You need to collect at least 50% of all points on exercise sheets to be admitted to the exam.*
Please send your solutions directly to Philip (`wellnitz@mpi-inf.mpg.de`*).*

Exercise 1 — 10 points

Recall the three steps of the POT for monotone dictatorships from the lecture: First we test the given function f for linearity, then we check that f is closed under bitwise conjunction, and finally we test whether f is not the all-zero function.

Consider the following variation of said algorithm: Instead of running all three tests, we first uniformly select an integer $i \in [3]$ and only run the i-th test. Show that this modified algorithm (using only four queries) is a one-sided error POT for monotone dictatorship with rejection probability $\rho(\delta) = \Omega(\delta)$, where δ is the distance of f to being a monotone dictatorship.

Exercise 2 — 10 points

Let $h : \{0,1\}^\ell \to \{0,1\}$ denote a Boolean function. Show that testing whether the set $h^{-1}(1)$ is an $(\ell-k)$-dimensional affine space can be reduced to testing whether for a Boolean function $h' : \{0,1\}^\ell \to \{0,1\}$ the set $\{x : h'(x) = 1\}$ is an $(\ell-k)$-dimensional linear space, where the reduction introduces an additive overhead of $O(2^k)$ queries.

Exercise 3 — 10 points

Let $\Pi = \bigcup_{n\in\mathbb{N}} \Pi_n$ denote a property, where Π_n contains functions from $[n]$ to an ordered set R_n. Further, let $q : \mathbb{N} \times [0,1) \to \mathbb{N}$ denote a function.

Suppose that for some $\varepsilon > 0$ and $n \in \mathbb{N}$, there is a distribution f of functions from $[n]$ to R_n such that for every deterministic oracle machine M that makes at most $q(n,\varepsilon)$ queries, we have

$$\mathsf{Pr}[\, f \in \Pi_n \wedge M(f) \neq 1\,] + \mathsf{Pr}[\, f \in \Gamma_\varepsilon(\Pi_n) \wedge M(f)) \neq 0\,] > \frac{1}{3},$$

where $\Gamma_\varepsilon(\Pi_n)$ is the set of functions (from $[n]$ to R_n) that are ε-far from Π_n.

Show that then, there are a distribution f_1 of functions in Π_n and a distribution f_0 of functions in $\Gamma_\varepsilon(\Pi_n)$, such that for every deterministic oracle machine M that makes at most $q(n,\varepsilon)$ queries we have that

$$|\mathsf{Pr}[\, M(f_1) = 1\,] - \mathsf{Pr}[\, M(f_0) = 1\,]| < \frac{1}{2}.$$

Exercise 4 — 2 + 2 + 3 + 3 points + 5 bonus points

In this exercise, we prove a lower bound for testing monotonicity of functions, thereby complementing the upper bound that was discussed in the lecture. In particular, in this exercise, we prove that any non-adaptive 1-sided error tester for monotonicity requires $\Omega(n^{1/4})$ queries.

To that end, let $f : \{0,1\}^n \to \{0,1\}$ denote a function and let $i \in [n]$ denote an integer. Define the *truncated anti-dictator* $f_i : \{0,1\}^n \to \{0,1\}$ as

$$f(x) := \begin{cases} 1, & \text{if } \sum_{j\in[n]} x_j \geq n/2 + \sqrt{n}; \\ 0, & \text{if } \sum_{j\in[n]} x_j < n/2 - \sqrt{n}; \text{ and} \\ \overline{x_i}, & \text{otherwise.} \end{cases}$$

a) Show that there is a constant $c > 0$, such that at least $c \cdot 2^n$ points $x \in \{0,1\}^n$ satisfy $n/2 - \sqrt{n} \leq \sum_{i\in[n]} x_i < n/2 + \sqrt{n}$.

b) Show that for any $i \in [n]$, the function f_i disagrees with any monotone function in $\Omega(2^n)$ points. (*Hint:* For an $i \in [n]$, consider the pairs (z, z'), where $z = (z_1, \ldots, z_{i-1}, 1, z_{i+1}, \ldots, z_n)$, with $f_i(z) = 0$ and $z' = (z_1, \ldots, z_{i-1}, 0, z_{i+1}, \ldots, z_n)$, with $f_i(z) = 1$.)

c) Let $\mathbb{A}$ denote any non-adaptive q-query algorithm. Show that for some $i \in [n]$, the algorithm $\mathbb{A}$ reveals a violation on f_i with probability at most $2q^2/\sqrt{n}$. (*Hint:* Fix $\mathbb{A}$ to be any q-query non-adaptive algorithm, and show that the queries that it makes can reveal violations of f_i for at most $2q^2 \cdot \sqrt{n}$ many $i \in [n]$.)

d) Deduce that no non-adaptive 1-sided tester that makes $o(n^{1/4})$ queries exists. *Hint:* Note that any 1-sided error tester must output "ACCEPT" if no violation of the monotonicity property is detected, since otherwise it would have a non-zero probability of rejecting a monotone function (and thus have 2-sided error).

e) (Bonus) Can this bound be improved further?