COMP 531: Advanced Theory of Computation (Winter 2014)

Assignment 2

Due February 19th

Instructions

Follow these instructions closely.

You will benefit most if you seriously try solving each problem yourself. You may work with each other but you must write up your own solutions. For each question, you should clearly acknowledge the people you have worked with. You are not allowed to use any resources that contain the solution to an assignment question. However, we value honesty above all. You will get full marks if you happen to find the solution to a question and you write your own solution, **as long as you properly acknowledge your source**. Failure to acknowledge your source can result in 0 points.

Clarity and conciseness of your solutions are as important as correctness. It is important to learn how to write your ideas and solutions clearly and rigorously. You will lose marks for correct solutions that are poorly explained/presented. When writing your solutions, assume that your audience is your class mates rather than the instructor of the course. The high level ideas and an overview of your argument should be presented before any technical details, and all non-trivial claims have to be proven.

If you do not know how to solve a problem, do not answer it. This will earn you 20% of the points. Do not make yourself believe in a wrong proof, this is bad for you. **And definitely do not try to sell it!** If you don't know how to solve a problem but you have some non-trivial ideas, write them down. If you have a solution with gaps, write your argument and clearly indicate the gaps.

Submit your assignments in class or send a copy to `aada@cs.mcgill.ca` before midnight of the due date.

Notation

A Boolean function $f : \{0,1\}^n \to \{0,1\}$ has the range $\{0,1\}$. Often, it is convenient to map 0 to 1 and map 1 to -1, and view the range as $\{1,-1\}$. That is, Boolean functions have the form $f : \{0,1\}^n \to \{1,-1\}$ when convenient. The particular choice of the range does not matter. We could choose *true* and *false* or *apples* and *oranges*. But a clever choice of the range can make the mathematical manipulations easier.

Recall that a threshold function $f : \{0,1\}^n \to \{0,1\}$ is specified by integer weights $w_0, w_1, \ldots, w_n$ and $f(x) = 1$ iff $w_1x_1 + w_2x_2 + \cdots + w_nx_n > w_0$. So

$$\mathrm{AND}(x) = 1 \text{ iff } x_1 + x_2 + \cdots + x_n > n-1,$$

$$\mathrm{OR}(x) = 1 \text{ iff } x_1 + x_2 + \cdots + x_n > 0,$$

$$\mathrm{MAJ}(x) = 1 \text{ iff } x_1 + x_2 + \cdots + x_n > n/2.$$

If $f : \{0,1\}^n \to \{1,-1\}$ is a threshold function, it can be represented as $\mathrm{sign}(w_0 + w_1x_1 + \cdots + w_nx_n)$ for some integer weights. Here, for $t \in \mathbb{R}$, $\mathrm{sign}(t) = 1$ if $t > 0$, $\mathrm{sign}(t) = -1$ if $t < 0$ and $\mathrm{sign}(t) = 0$ if $t = 0$.

Define the $\textsc{mod}_6 : \{0,1\}^n \to \{0,1\}$ function/gate as follows: $\textsc{mod}_6(x) = 0$ if and only if $\sum x_i$ is divisible by 6. A generalized mod 6 function is such that for $A \subseteq \{0,1,2,3,4,5\}$, $\textsc{mod}_6^A(x) = 1$ if and only if $\sum x_i$ modulo 6 is in A. A generalized mod 6 gate is denoted by GMOD_6.

A function $f : \{0,1\}^n \to \{0,1\}$ is called *symmetric* if the output depends only on the number of input bits set to 1, i.e. $f(x) = f(x')$ whenever $\sum_i x_i = \sum_i x'_i$. A symmetric function/gate will be denoted by SYM.

We denote by $f \circ g$ the class of depth 2 circuits whose output gate is f and the gates at the first level are g. For example, $\mathrm{AND} \circ \mathrm{MAJ}$ denotes the class of depth 2 circuits where the output gate is AND and the first level has majority gates.

Every circuit is allowed access to the constants 0 and 1 (on top of the input variables and their negations).

Questions

1. Show that any symmetric function can be computed by linear size $\mathrm{MAJ} \circ \mathrm{MAJ}$ circuits.

2. Let $C = (C_n)$ be a family of circuits constructed with binary AND and OR gates. Assume that C has polynomial size and the graph of each C_n is a tree. Show that the induced boolean function is actually in NC^1.

3. Show that the multiplication of two n-bit integers is not in AC^0.

4. (a) Show that any function $f : \{0,1\}^n \to \{0,1\}$ can be computed by a $\mathrm{MOD}_6 \circ \mathrm{SYM}$ circuit. Note that there is no restriction on the size of the circuit.

   (b) We say that two functions $f, g : \{0,1\}^n \to \mathbb{Z}$ are $\textsc{mod}_6$ equivalent if $f(x)$ is divisible by 6 if and only if $g(x)$ is divisible by 6. Show that if f is computed by "small" size $\mathrm{MOD}_6 \circ \mathrm{SYM}$ circuit, then f is $\textsc{mod}_6$ equivalent to a "low" rank function g over any field (poly(n) would be considered "small" or "low"). The meaning of rank of a function is as follows. Partition the input variables any way you like $x = (x^1, x^2)$ and consider the matrix whose rows are labeled with all the possible values for x^1 and columns are labeled with all the possible values for x^2. The (x^1, x^2) entry contains $g(x^1, x^2)$. The rank of g is defined to be the maximum rank of this matrix, where the maximum is over all possible partitions $x = (x^1, x^2)$. Hint: Use the fact that rank is subadditive.

   (c) Using part (b), find an explicit function that requires exponential size $\mathrm{MOD}_6 \circ \mathrm{SYM}$ circuits.

   (d) (**Open Problem - not for credit**) Find an explicit function that requires superpolynomial size $\mathrm{GMOD}_6 \circ \mathrm{GMOD}_6$ circuits. (Note that a generalized mod 6 function is a symmetric function so part (c) comes close to proving this but it does not.)

5. For this question, all Boolean functions will be ±1 valued. Consider the 2^n dimensional vector space of functions over the reals $V = \{p : \{0,1\}^n \to \mathbb{R}\}$. Obviously this vector space includes all the Boolean functions $f : \{0,1\}^n \to \{1, -1\}$ that we are interested in. Equip this vector space with the following inner product:

$$\langle p, q \rangle = \mathbf{E}_x[p(x)q(x)] = \frac{1}{2^n} \sum_x p(x)q(x),$$

   where the expectation is over a uniformly random $x \in \{0,1\}^n$. For a subset $S \subseteq [n] = \{1, 2, \ldots, n\}$, we define $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$. These are called the *characters* and correspond to the parity function over the input variables in S.

   Question: Show that the set $\{\chi_S : S \subseteq [n]\}$ is an orthonormal basis.

Hence we can write every function $p \in V$ as a linear combination

$$p = \sum_{S \subseteq [n]} \widehat{p}(S) \chi_S,$$

where $\widehat{p}(S)$ are the appropriate real coefficients.

Question: Show that for all $S \subseteq [n]$, $\widehat{p}(S) = \langle p, \chi_S \rangle$.

This way of expanding the function as a sum of the parity functions (characters) is referred to as the Fourier expansion of p and $\widehat{p}(S)$ are called the Fourier coefficients. We will refer to the elements of V as polynomials. To justify this terminology, observe that if we consider the domain to be $\{1, -1\}^n$ rather than $\{0, 1\}^n$, we see that

$$p = \sum_{S \subseteq [n]} \widehat{p}(S) \chi_S = \sum_{S \subseteq [n]} \widehat{p}(S) \prod_{i \in S} x_i,$$

which is a multilinear polynomial (every variable has exponent 0 or 1).

For $p, q \in V$, and a distribution μ over $\{0, 1\}^n$, define the correlation of p and q under μ as follows:

$$\mathrm{Cor}_\mu(p, q) \stackrel{\text{def}}{=} \left| \mathbf{Pr}_{x \sim \mu} \left[p(x) = q(x) \right] - \mathbf{Pr}_{x \sim \mu} \left[p(x) \neq q(x) \right] \right|.$$

Question: Show that if $f, g \in V$ are Boolean functions, then

$$\mathrm{Cor}_{\mathcal{U}}(f, g) = |\langle f, g \rangle|,$$

where $\mathcal{U}$ denotes the uniform distribution.

Question: Show that for $p, q \in \mathcal{V}$,

$$\langle p, q \rangle = \sum_{S \subseteq [n]} \widehat{p}(S) \widehat{q}(S).$$

This forms the bridge between the usual representation of a function in terms of the values $\{p(x) \mid x \in \{0, 1\}^n\}$ and the Fourier representation in terms of the Fourier coefficients $\{\widehat{p}(S) \mid S \subseteq [n]\}$.

Question: Show that for boolean functions, $\sum_S \widehat{f}(S)^2 = 1$.

In other words, the squares of the Fourier coefficients of a boolean function can be thought as a probability distribution over the subsets S.

In many different settings, the *hardness* of a function exposes itself in the function's Fourier expansion. In other words, different analytic measures associated with the Fourier coefficients of f can be good approximations to how *complex* the function is e.g., in circuit complexity, learning theory, communication complexity, etc.