Homework 4
Due 2011/4/29

Oded Regev
Dept. of Computer Science
École normale supérieure, Paris

Instructions as before.

1. **Learning juntas with queries:** Show an algorithm for learning k-juntas in time $\text{poly}(n, 2^k)$ using membership queries, without using the Fourier transform.

2. **Weakly learning DNFs [1]:** Show that if f is computable by a DNF of size s then $|\hat{f}(S)| \geq \Omega(1/s)$ for some S with $|S| \leq \log_2(s) + O(1)$. This is the first step in Jackson's algorithm [2]. Hint: deal separately with the case where there are no small terms. If there is a small term, consider a restriction.

3. **Learning noise insensitive functions:** For $f : \{0,1\}^n \to \mathbb{R}$ define the *δ-noise sensitivity* of f as $\mathbb{NS}_\delta(f) = \frac{1}{2}(1 - \langle f, T_{1-2\delta} f \rangle)$.

   (a) Show that for $f : \{0,1\}^n \to \{-1,1\}$, $\mathbb{NS}_\delta(f) = \Pr_{x,w}[f(x) \neq f(x+w)]$ where x is chosen uniformly from $\{0,1\}^n$ and w is chosen according to μ_δ.

   (b) Let $\mathcal{C}_{\delta,\varepsilon}$ be the class of all $f : \{0,1\}^n \to \{-1,1\}$ with $\mathbb{NS}_\delta(f) \leq \varepsilon$. Show that for any $0 < \delta < \frac{1}{2}$ and $\varepsilon > 0$, $\mathcal{C}_{\delta,\varepsilon}$ can be PAC learned under the uniform distribution from random examples to within accuracy $O(\varepsilon)$ in time $\text{poly}(n^{1/\delta}, 1/\varepsilon)$.

   (c) Optional: show that for the majority function, $\mathbb{NS}_\delta(\text{MAJ}_n) = \Theta(\sqrt{\delta})$ assuming n is large enough.

4. **Orthogonal decomposition:** For some $0 < p < 1$, consider the space of functions $f : \{0,1\}^n \to \mathbb{R}$ taken with respect to the measure μ_p, i.e., we define the inner product in this space as $\langle f, g \rangle = \text{Exp}_{x \sim \mu_p}[f(x)g(x)]$.

   (a) Suggest a reasonable choice of an orthonormal basis $\{\chi_S : S \subseteq [n]\}$. Hint: Start with $n = 1$.

   (b) For $f : \{0,1\}^n \to \{-1,1\}$ define the influence of the ith coordinate as
   $$\text{Inf}_i(f) = \Pr_{x \sim \mu_p}[f(x) \neq f(x \oplus e_i)].$$
   Show how to express it using a decomposition of f in your basis (there is more than one possible solution; try to get a simple expression).

5. **A variable that is often much smaller than its expectation has high variance:** Show that if X is a nonnegative random variable with $\Pr[X > K] = \delta$ and $\text{Exp}[X] \geq L > K$ then $\text{Exp}[X^2] \geq (L-K)^2/\delta$.

References

[1] A. Blum, M. L. Furst, J. C. Jackson, M. J. Kearns, Y. Mansour, and S. Rudich. Weakly learning DNF and characterizing statistical query learning using Fourier analysis. In *STOC*, pages 253–262, 1994.

[2] J. C. Jackson. An efficient membership-query algorithm for learning DNF with respect to the uniform distribution. *J. Comput. Syst. Sci.*, 55(3):414–440, 1997. Preliminary version in FOCS'94.