Instructor: Rocco Servedio
Class Manager: Andrew Wan

Computer Science 4252: Introduction to Computational Learning Theory
Problem Set #6 Fall 2006

Due 5:00pm Friday, April 28, 2006

Answer **two** of the following four problems.

Problem 1 A *branching program* is a directed acyclic graph which has two *leaf nodes* which are labelled with 0 and 1, and a collection of *internal nodes* each of which is labelled with a variable from $x_1, \ldots, x_n$. Each internal node has two outgoing edges labeled 0 and 1; there is one internal node (with no incoming edges) which is an initial node. A branching program computes a Boolean function in a fashion similar to a decision tree: given an input $x \in \{0,1\}^n$, we start at the initial node and follow the path indicated by the value of the variables at each node till we arrive at a leaf, whose label then determines the value of the function.

Show that there is probably no polynomial-time PAC learning algorithm for the class of branching programs. (You may use any hardness result stated in class or in the Kearns/Vazirani textbook.)

Problem 2 Let H be a finite concept class on domain X. We say that algorithm A *minimizes disagreements* if, given any sample S of labeled examples, it outputs the hypothesis h in H such that the number of examples in S on which h does not agree with the given label is minimum (over all hypotheses h in H).

Show that if A minimizes disagreements, then A is a (not necessarily computationally efficient) PAC learning algorithm for H in the presence of random classification noise at rate $\eta < \frac{1}{2}$.

Problem 3 Recall that in the malicious noise model with noise rate η, the example oracle flips a biased coin and with probability η returns an arbitrary example in the domain labeled arbitrarily. With probability $1 - \eta$ it draws a random example according to D and passes it with the correct label to the learner.

Suppose that concept class C is learnable to accuracy ϵ and confidence δ in the PAC model by a polynomial time algorithm which has sample size $m(\epsilon, \delta)$. Let $s = m(\epsilon/8, 1/2)$. Show that there is a polynomial time algorithm which PAC learns C in the presence of malicious noise at rate η, where $\eta = \min\{\epsilon/8, \frac{\ln s}{s}\}$.

Problem 4 Let C be any concept class. Suppose that algorithm A is an efficient proper PAC learning algorithm for C in the noise-free setting (i.e. given access to an example oracle $EX(c, \mathcal{D})$, algorithm A outputs a hypothesis h which belongs to C and satisfies the PAC criteria). Suppose that moreover there is an efficient PAC learning algorithm B for concept class C in the presence of random classification noise, but B is not a proper PAC learning algorithm (i.e. the hypotheses which B outputs belong not to C, but to some other hypothesis class H). Show that then there must exist an efficient *proper* PAC learning algorithm for C in the presence of random classification noise.