Assignment 3

This assignment is due back by Thursday April 8. You can hand it in at the beginning of class.

*Note:* Make sure you write and justify all your statements in a precise, formal way.

**Problem 1**

Solve problem 7.3 on page 153 of the text by Kearns and Vazirani.

**Problem 2**

Solve problem 7.2 on page 153 of the text by Kearns and Vazirani.

*Notes:* (1) The reduction used for this problem does not satisfy the definition of reduction seen in class. The consistent concept transformation may depend now on the input sample $S$. We look now for a concept in the class of *exclusive-or of two halfspaces* that classifies correctly all examples in $S$. This notion of reduction is sufficient, however, for the purpose of showing the implication of their learnability. (2) Think first of the simpler case where boundary points such that $\sum u_i x_i = 0$ or $\sum v_i x_i = 0$ are ignored.

**Problem 3**

The class of $k$-LA$_n$ is a restricted class of DFAs accepting strings in $\{0,1\}^n$. A concept in $k$-LA$_n$ is defined by an array of $k \times (n + 1)$ nodes. Nodes at level (column) $i$ where $i \leq n$ are marked with $x_i$, and each node at level $i$ has two outgoing edges, one labelled 0 and one labelled 1. The edges are directed and each edge from level $i$ must be directed at a node at level $i + 1$. One of the nodes at level 1 is marked as the start node and the nodes at level $n + 1$ are marked as accepting nodes (A) or rejecting nodes (R).

For example, the following diagram describes a 3-LA$_4$. In the diagram we have omitted nodes that are not reachable from the start node, and whenever two edges have the same endpoints we drew a single edge marked with both 0 and 1 (i.e., 0/1).
The computation of $k$-LA$_n$ is performed by tracing a path in the graph. For example the graph above accepts the strings 1000, 1001, 1110, and rejects 1100, 1101.

Show that the class of DNF expressions is PAC reducible to the class of 3-LA.

*Hint:* Try to implement a conjunction with a 2-LA.

**Problem 4**

Let $K(\cdot, \cdot)$ be a kernel defined over $X \times X$, and let $\hat{X}$ be the set of all finite non-empty subsets of $X$. For $A, B \in \hat{X}$ define

$$\hat{K}(A, B) = \sum_{x \in A, y \in B} K(x, y).$$

Show that $\hat{K}(\cdot, \cdot)$ is a kernel over $\hat{X} \times \hat{X}$.

*Hint:* Let $\{\phi_a\}_{a \in F}$ be the feature set for which $K(\cdot, \cdot)$ is an inner product. Consider a feature set indexed in the same way but where the feature values are suitably modified.

**Problem 5**

Let $K(\cdot, \cdot)$ be a kernel defined over $X \times X$, and let $\hat{X}$ be the set of all finite non-empty subsets of $X$. For $A, B \in \hat{X}$ define

$$\hat{K}(A, B) = \sum_{x \in A \cap B} K(x, x).$$

Show that $\hat{K}(\cdot, \cdot)$ is a kernel over $\hat{X} \times \hat{X}$.

*Hint:* Let $\{\phi_a\}_{a \in F}$ be the feature set for which $K(\cdot, \cdot)$ is an inner product. Expand this feature space so that the new inner product will match examples in the sets $A, B$. 

2