

CS472 Foundations of Artificial Intelligence

Fall 2002

Assignment 5

Due Wedn., Dec. 11, 5pm, Upson 4148.

Solutions must be typed, although equations, graphs, tables, etc., can be drawn in by hand.

1. (10 pts.) Consider a neural net with step threshold functions.
 - Suppose that you multiply all weights and thresholds by a constant. Will the behavior change?
 - Suppose that you add a constant to all weights and thresholds. Will the behavior change?
2. (20 pts. **Perceptron Learning**) Consider the following training set:

input	output
1 0 0	1
0 1 1	0
1 1 0	1
1 1 1	0
0 0 1	0
1 0 1	1

Train (by hand) a linear threshold element on this training set. Your unit will have four inputs counting the one that implements the threshold. Assume that the initial values of all weights are zero. Train your unit with perceptron learning algorithm discussed in class until it converges to a solution. Give the details of your calculation (i.e., weights after each example). Is the function linearly separable? Draw a sketch of a 3-D cube with the preceding inputs as vertices, and sketch in the separating plane corresponding to the final weight set.

3. (10 pts.) Consider two perceptrons defined by the threshold expression $w_0 + w_1x_1 + w_2x_2 > 0$. Perceptron A has weight values: $w_0 = 1$, $w_1 = 2$, $w_2 = -1$; and perceptron B has the weight values: $w_0 = -1$, $w_1 = 2$, $w_2 = 1$.

Question: True or false? — Perceptron A is *strictly more_general_than* perceptron B?

Note: Given two Boolean values functions h_j and h_k over domain X . We say that h_j is *more_general_than* h_k if and only if $(\forall x \in X)[(h_k(x) = 1) \Rightarrow (h_j(x) = 1)]$. We write $h_j \geq h_k$. (So, any instance classified positively by h_k is also classified positively by h_j .) We say h_j is *strictly more_general_than* h_k iff $h_j \geq h_k$ and $h_k \not\geq h_j$.

4. (10 pts) (Decision Tree Learning) Give an example of training set where the decision tree learning algorithm using information gain produces a tree much bigger than the smallest possible tree. Explain why this happens.
5. (10 pts) Give a formal argument showing a perception cannot represent the XOR function. (Hint: Consider the equations governing the perceptron.)