CS 441/541
Artificial Intelligence
Fall, 2006

Homework 4: Reinforcement Learning and Biologically Inspired AI

Due Tuesday, October 24.

1. Design (i.e., give the weights) for a two-input perceptron that implements the Boolean function $x_1 \wedge \neg x_2$. The activation function is the **sgn** function, and the possible outputs are 1 and -1. Give the equation of the separating line corresponding to this perceptron, and show that this separating line correctly classifies all pairs (x_1, x_2), where $x_1, x_2 \in 0, 1$.

2. Given a two input perceptron with initial weights $w_0 = 0.1$ (bias weight), $w_1 = 0.1$, $w_2 = -0.1$, and the set of training examples given below, run (by hand) the perceptron learning algorithm for two epochs (twice over the entire training set), showing what the new weights would be after each training example is processed. At the end of your two epochs, give the sum-squared error over the training set S, that is,

$$E(\vec{w}) = \frac{1}{2}\sum_{i \in S}(y_i - g(\vec{x}_i))^2.$$

Training set S:

x_1	x_2	y
0	0	1
0	1	-1
1	0	1
1	1	1

3. (OPTIONAL/EXTRA CREDIT!!) Textbook, problem 20.14.

4. Consider the grid world shown in Figure 21.1 in the textbook. Let $\alpha = \gamma = 1$, and let the reward function be as described in the book. The states marked +1 and -1 are terminal states.

For this problem, assume the state transistions are all deterministic—e.g., if you are in state (1,1) and perform action "up", you will go to state (2,1) with probability 1.

Perform Q-learning on this grid world by simulating (by hand) the algorithm of Figure 21.8

for two epochs. That is, start with a $Q(s, a)$ matrix initialized to all zeros, except for terminal states s_{term}, which get $Q(s_{term}, a) = R(s_{term})$ for all actions a.

Each epoch consists of starting at state (1,1) (bottom left corner) and, at each state you visit, take the action that will move you to the neighboring state that is clockwise around the edge of the grid. At each state you visit, apply Q-learning. When a terminal state is reached, the epoch is finished.

(Important: Note that the algorithm in Figure 2.8 says, at each state, to choose action a by calculating a function $f(Q, N_{sa})$. Ignore this—instead, at each state you will be choosing the action that moves clockwise around the edge of the grid.)

After the two epochs have been completed, give the value of $Q(a, s)$ for each state/ action pair.

5. Draw a genetic programming tree that represents the following function:

$$f(x, y) = \frac{1}{1 - x^2} + \frac{1}{1 - y^2}.$$

6. In a few sentences, suggest a problem for which a genetic programming approach might be suitable. Describe a function set and terminal set you might use for this problem.