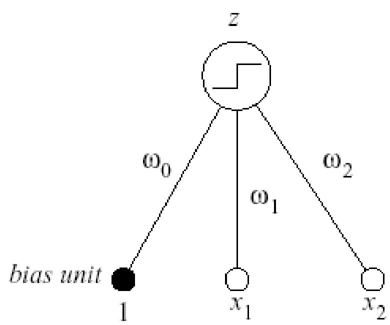
Partial Solution to Homework Set #4

Problem 1:

Consider the simple two-input perceptron below, where the activation function γ for z is the following thresho function:



$$\gamma(u) = \begin{cases} 0 & u \le 0 \\ 1 & u > 0 \end{cases}$$

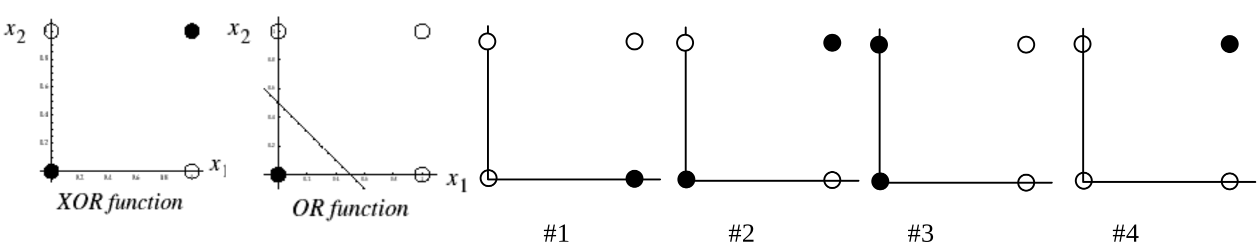
(a) For each data set below, indicate whether or not a set of weights $\{\omega_0, \omega_1, \omega_2\}$ exists which would give the desired mapping.

#1			*#2*			*#3*			*#4*		
x_1	x_2	z	x_1	x_2	z	x_1	x_2	z	x_1	x_2	z
0	0	1	0	0	0	0	0	1	0	0	0
0	1	1	0	1	1	0	1	1	0	1	0
1	0	0	1	0	1	1	0	0	1	0	0
1	1	1	1	1	0	1	1	0	1	1	1

(b) Which (if any) of the data set(s) that is learnable in (a) cannot be learned if $\omega_0 = 0$?

(c) Which (if any) of the data set(s) that is learnable in (a) cannot be learned if $\omega_0 = -20$?

(d) Which (if any) of the data set(s) that is learnable in (a) cannot be learned if $\omega_0 = 20$?



a. Yes, No, Yes, Yes
b. #4 because the line separating the two classes cannot be constrained to go through the origin
c., d. none

Problem 2

Consider a one-hidden layer feedforward neural network, fully connected between layers, with 2 inputs, 2 hidden units in the first layer, and 1 output. Assume sigmoidal activation functions .

a. How many total independent weights are contained in this neural network? Show Work! No Work, No Credit.

Answer: __9__ independent weights. Justification: output given by $z=\gamma\{w_7+w_8\gamma(w_1+w_2x_1+w_3x_2)+w_9\gamma(w_4+w_5x_1+w_6x_2)\}$

b. *Without any local variable*s, how many additions, multiplications, and function evaluations (γ,γ') are required to compute the output of the neural network in terms of the inputs and the weights of the neural network? Show Work! No Work, No Credit.

Justification: $z=\gamma\{w_7+w_8\gamma(w_1+w_2x_1+w_3x_2)+w_9\gamma(w_4+w_5x_1+w_6x_2)\}$ for a total of 6 "+"; 6 "*"; 3 $\gamma(.)$ evaluations
Each γ (.) requires 1 addition, 1 division, and 1 exp evaluation.

Additions:_____6______ Multiplications:_____6______ Function Evaluations:_____3_____

c. Using the backpropagation algorithm, how much total computation is required [additions, multiplications and function evaluations (γ,γ')] to compute the error derivatives with respect to all the weights in the neural network, given a single training pattern $<x_1,x_2,y>$and the error measure $E=½(y-z)^2$? Show Work! No Work, No Credit.

Forward Pass:

Fwd Pass:
$z=\gamma\{Net_3\}$; $Net_3=w_7+w_8h_1+w_9h2$; $h_1=\gamma\{Net_1\}$; $Net_1=(w_1+w_2x_1+w_3x_2)$; $h_2=\gamma\{Net_2\}$; $Net2=(w_4+w_5x_1+w_6x_2)$
Additions:_____6______ Multiplications:_____6______ Function Evaluations:_____3_____

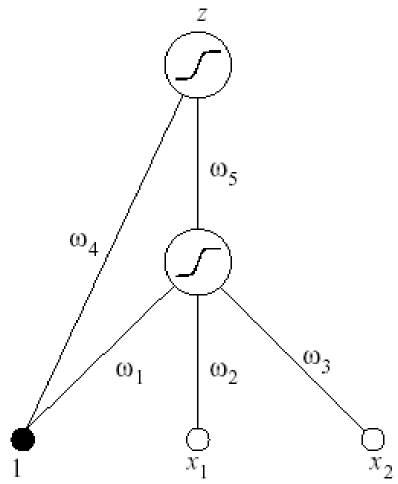
Backward Pass:

Backward Pass:

$\delta_3=(z-y)\gamma'(Net_3)$	A: 1	M: 1	*f*: 1
$\delta_2=\delta_3w_9\gamma'(Net_2)$	A:	M: 2	*f*: 1
$\delta_1=\delta_3w_8\gamma'(Net_1)$	A:	M: 2	*f*: 1
$\delta E/\delta w_9=\delta_3h_2$	A:	M: 1	*f*:
$\delta E/\delta w_8=\delta_3h_1$	A:	M: 1	*f*:
$\delta E/\delta w_7=\delta_3$	A:	M:	*f*:
$\delta E/\delta w_6=\delta_2x_2$	A:	M: 1	*f*:
$\delta E/\delta w_5=\delta_2x_1$	A:	M: 1	*f*:
$\delta E/\delta w_4=\delta_2$	A:	M:	*f*:
$\delta E/\delta w_3=\delta_1x_1$	A:	M: 1	*f*:
$\delta E/\delta w_2=\delta_1x_2$	A:	M: 1	*f*:
$\delta E/\delta w_1=\delta_1$	A:	M:	*f*:

Problem 3:

Consider the neural network with sigmoidal activation functions below:



Assuming initial values for the weights,

$$\omega_i = 0.1 ,\ i \in \{1, 2, 3, 4, 5\} ,$$

a learning rate $\eta = 0.3$, pattern training, and the following training data set,

x_1	x_2	y
1	0	1
0	1	0

compute the value of the weights after two iterations of the backpropagation algorithm (one iteration per training pattern).

$\gamma\{u)\}= 1/(1+\exp(-u)); \gamma'\{u)\}=\exp(-u)/(1+\exp(-u))^2$

$z=\gamma\{w_4+w_5\gamma(w_1+w_2x_1+w_3x_2)\}= \gamma\{Net_2\}; Net_2= w_4+w_5\gamma(Net_1); Net_1= w_1+w_2x_1+w_3x_2$

Forward Pass 1:

Pattern 1: $y=1; x_1=1; x_2=0; w_i=0.1$

$Net_1= w_1+w_2x_1+w_3x_2=0.1+(0.1)\times1+(0.1)\times0=0.2$

$h_1=\gamma(Net_1)= \gamma(0.2)= 0.549833997312$

$Net_2= w_4+w_5\gamma(Net_1)= w_4+w_5h_1=0.1+0.1\times0.549833997312=0.1549833997312$

$z= \gamma\{Net_2\}=\gamma\{0.1549833997312\}=0.538668479964$

Backward Pass 1:

$\delta_2=(z-y)\gamma'(Net_2) = (0.538668479964-1)\times(0.248504748657) = -0.114643073434$

$\delta_1=\delta_2w_5\gamma'(Net_1) = (-0.114643073434) \times (0.1)$ x $(0.247516572712) = -2.837606062154e-03$

$\delta E/\delta w_5=\delta_2h_1=-(-0.114643073434) \times (0.549833997312)= -6.303465933035e-02$

$w_5(t+1) = w_5(t) - \eta\times\delta E/\delta w_5=(0.1)-(0.3) \times (-6.303465933035e-02)= 0.118910397799$

$\delta E/\delta w_4=\delta_2 = -0.114643073434$; $w_4(t+1)=w_4(t) - \eta\times\delta E/\delta w_4 = (0.1)-(0.3)\times(-0.114643073434)=0.13439292203$

$\delta E/\delta w_3=\delta_1x_2=(-2.837606062154e-03) \times 0 = 0$; $w_3(t+1) = w_3(t) - \eta\times\delta E/\delta w_3=(0.1)-(0.3)\times0=0.1$

$\delta E/\delta w_2=\delta_1x_1=(-2.837606062154e-03) \times 1= -2.837606062154e-03$

$w_2(t+1) = w_2(t) - \eta\times\delta E/\delta w_3 = (0.1) - (0.3)\times(-2.837606062154e-03)= 0.100852181819$

$\delta E/\delta w_1=\delta_1=-2.837606062154e-03$; $w_1(t+1)=w_1(t)-\eta\times\delta E/\delta w_1=(0.1)-(0.3)\times(-2.837606062154e-03)=0.100852181819$

New Weights: $\{w_i\}=[0.118910397799, 0.13439292203, 0.1, 0.100852181819, 0.100852181819]$ for $i=\{1,2,3,4,5\}$

Repeat for Pattern 2