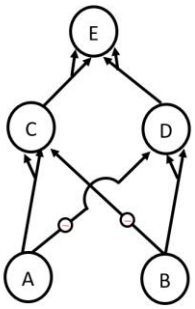
2. Draw an ANN from Figure 10-3 that computes A XOR B



3. Why is it generally preferrable to use Logistic Regression rather than classical Perceptron?

Perceptrons do not output class probability

How to tweak to make equivalent?

Change activation function from step to ReLU

6) MLP: 10,50,3 (All ReLU)

What is shape of input Matrix:

(10x1)

What are shapes of hidden layer's W_h and b_h?:

W_h = (10x50)

b_h = (50x1)

What are the shapes of the output layer's W_h and b_h?:

W_o = (50x3)

b_o = (3x1)

Shape of output matrix

Y = (3x1)

Eq to compute output matrix

$Y = f(W_o*(f(W_h*X + b_h))+b_o)$

7) How many neurons needed in output layer to classify spam or ham?

1

What activation

Step

For MNIST how many output

10

Activation

Softmax

For housing prices

1

Activation

ReLU

9. Hyperparameter for MLP
    a. Layers, neurons, learning rate, activation function, # epochs, optimizer, batch size, dropout

Overfitting?

Cross-validation