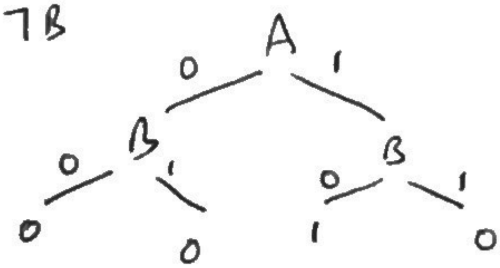
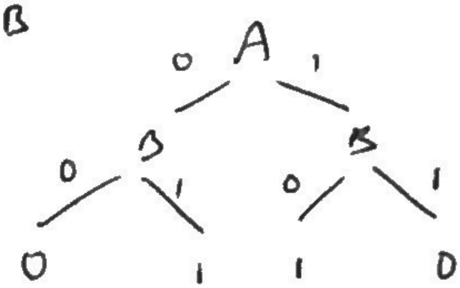
HW-1

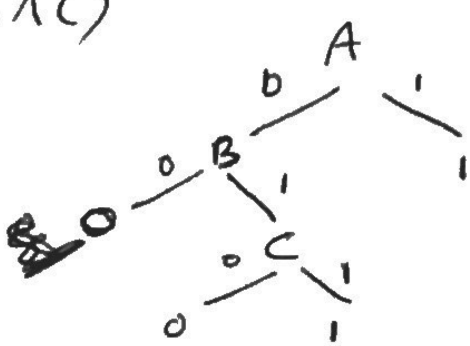
1. $A \wedge \neg B$



$A \oplus B$



$A \vee (B \wedge C)$



2. $w_0 + w_1 \cdot A + w_2 \cdot B$ is the equation of the perceptron

for $A \vee B$, $w_0 = -1.0$, $w_1 = 1.5$, $w_2 = 1.5$

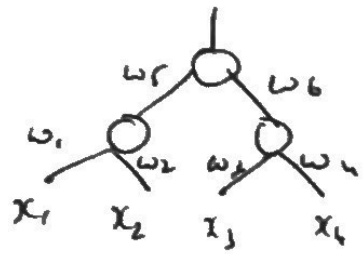
for $\neg A \vee B$, ~~we cannot find a seperating line~~

~~$w_0 = -1.0$, $w_1 = -1.5$, $w_2 = 2.5$~~

$w_0 = 1.0$, $w_1 = -1.5$, $w_2 = 2.0$

for $A \oplus B$, we cannot find a seperating line to distinguish between the 0's & 1's

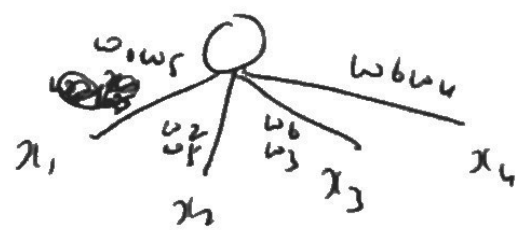
3) a) False



If all nodes are linear units, we can write the output as.

$w_5(w_1x_1 + w_2x_2) + w_6(w_3x_3 + w_4x_4)$

we can represent this as a single linear unit



∴ expresiveness is same as a single perceptron.

In general, linear combinations of linear functions only give linear functions.

b) False, the number of nodes can be exponentially larger than n. Eg. represent Xor function as a decision tree.

c) True, we can represent all boolean functions.

d) True, convergence depends on error function & gradient descent (given optimal learning rate) will always converge

4). I will show the first step, & the others should be the same.

Condition on Price

Conditional entropy $= \left(\frac{2}{4}\right) \cdot E(1,0) + \frac{2}{4} E(0.5, 0.5)$

$= 1/2$

For fast

$\frac{2}{4} E(1,0) + \frac{2}{4} E(0.5, 0.5)$

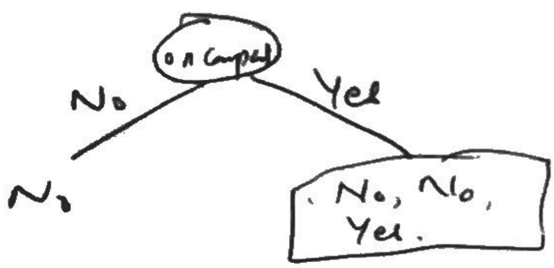
$= 1/2$

On Campus

$\frac{1}{4} E(1,0) + \frac{3}{4} E(2/3, 1/3)$

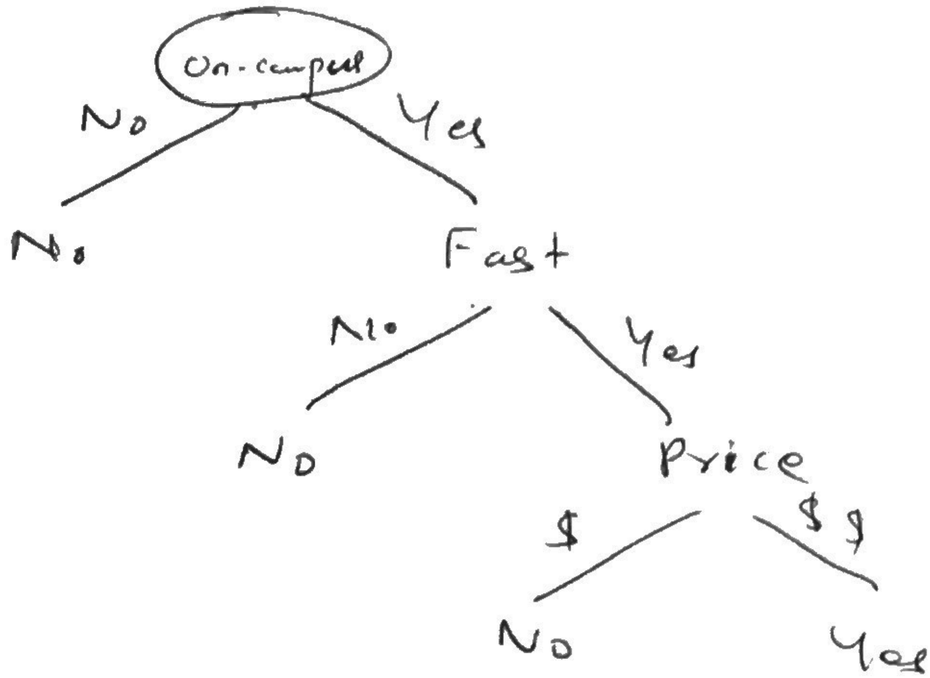
$= 0.47.$

So choose on-campus.

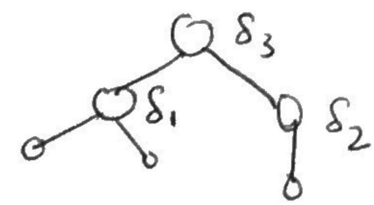


Continue the above process for these examples

Final tree would look like.



5).



<u>Forward prop.</u>

$O_1 = S(-1*w_1 + 1*w_2) = S(3)$

$O_2 = S(-1*w_3) = S(1)$

$O_3 = S(w_4 O_1 + w_5 O_2) = S(O_1 + 2O_2)$

<u>Back-propagation</u>

$\delta_3 = O_3(1-O_3)(0 - O_3)$ [since expected output is 0]

$= -(O_3)(1-O_3)(O_3)$

$\delta_1 = \delta_3 \cdot w4 \cdot O_2(1-O_2)$

$\delta_2 = \delta_3 \cdot w_5 \cdot O_1(1-O_1)$

Updates

$$w_4 = w_4 + \eta \cdot \delta_3 \cdot O_1$$

$$w_5 = w_5 + \eta \cdot \delta_3 \cdot O_2$$

$$w_1 = w_1 + \eta \cdot \delta_1 \cdot x_1$$

$$w_2 = w_2 + \eta \cdot \delta_1 \cdot x_2$$

$$w_3 = w_3 + \eta \cdot \delta_2 \cdot x_3 .$$