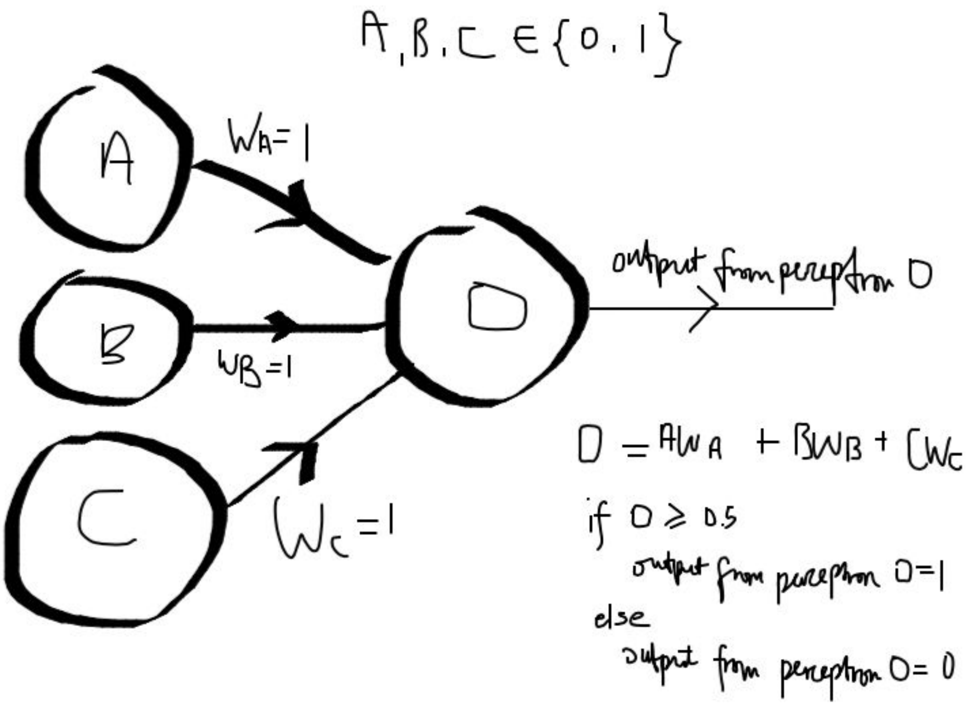
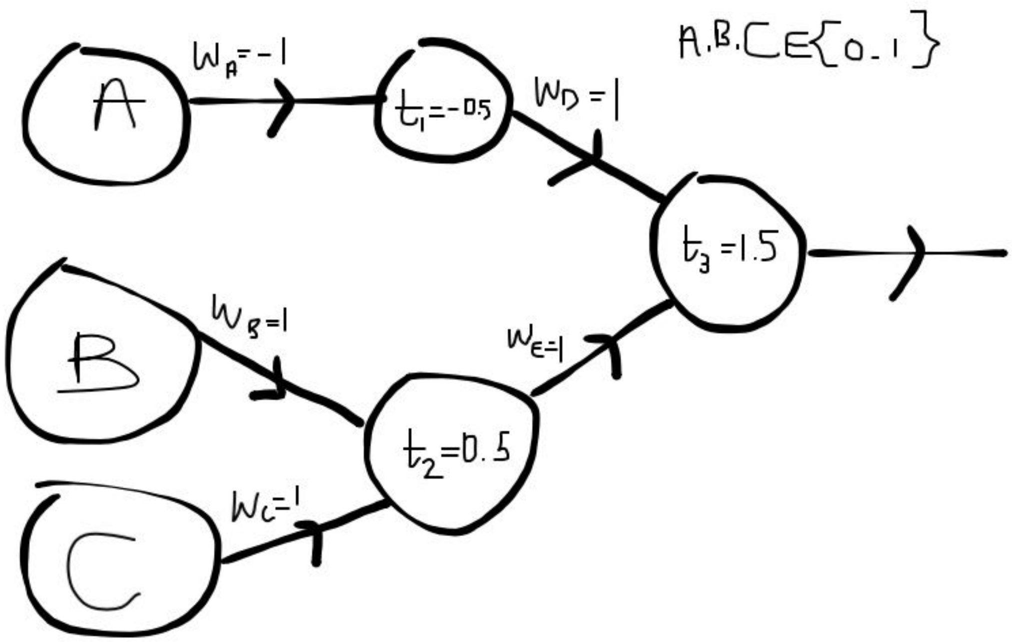
CHOW, Hau Cheung Jasper
Student ID: 20589533
hcjchow@connect.ust.hk
COMP4211 Assignment 1

Q1)



Q2)



Output coming out of perceptron with threshold value t_1 (we'll call it D) is 1 if $A * w_A \geq -0.5$, and 0 otherwise.

CHOW, Hau Cheung Jasper
Student ID: 20589533
hcjchow@connect.ust.hk
COMP4211 Assignment 1

Output coming out of perceptron with threshold value t_2 (we'll call it E) is 1 if $B * w_B + C * w_C \geq 0.5$, and 0 otherwise.

Output coming out of perceptron with threshold value t_3 is 1 if $D * w_D + E * w_E \geq 1.5$, and 0 otherwise.

Q3)

Apply backpropagation by labeling each edge (as it contains a value in green) with an intermediate variable name from A to H.

Now, we calculate the derivative of f(w, x) wrt each of these intermediate variables:

$$\frac{df(w,x)}{dA} = \frac{df(w,x)}{df(w,x)} = 1$$
$$\frac{df(w,x)}{dB} = \frac{df(w,x)}{dA} * \frac{dA}{dB} = 1 * \frac{d(1/B)}{dB} = -1/B^2$$
$$\frac{df(w,x)}{dC} = \frac{df(w,x)}{dB} * \frac{dB}{dC} = -1/B^2 * \frac{d(C+1)}{dC} = -1/B^2$$
$$\frac{df(w,x)}{dD} = \frac{df(w,x)}{dC} * \frac{dC}{dD} = -1/B^2 * \frac{d(e^D)}{dD} = \frac{-e^D}{B^2}$$
$$\frac{df(w,x)}{dE} = \frac{df(w,x)}{dD} * \frac{dD}{dE} = \frac{-e^D}{B^2} * \frac{d(-E)}{dE} = \frac{e^D}{B^2}$$
$$\frac{df(w,x)}{dF} = \frac{df(w,x)}{dE} * \frac{dE}{dF} = \frac{e^D}{B^2} * \frac{d(F+w_2)}{dF} = \frac{e^D}{B^2}$$
$$\frac{df(w,x)}{dG} = \frac{df(w,x)}{dF} * \left(\frac{\delta F}{\delta G}\right) = \frac{e^D}{B^2} * \left(\frac{\delta(G+H)}{\delta G}\right) = \frac{e^D}{B^2} * 1 = \frac{e^D}{B^2}$$
$$\frac{df(w,x)}{dH} = \frac{df(w,x)}{dF} * \left(\frac{\delta F}{\delta H}\right) = \frac{e^D}{B^2} * \left(\frac{\delta(G+H)}{\delta H}\right) = \frac{e^D}{B^2} * 1 = \frac{e^D}{B^2}$$

Then we calculate the gradients of f wrt x_0 and x_1:

$$\frac{df(w,x)}{dx_0} = \frac{df(w,x)}{dG} * \left(\frac{\delta G}{\delta x_0}\right) = \frac{e^D}{B^2} * w_0$$
$$\frac{df(w,x)}{dx_1} = \frac{df(w,x)}{dH} * \left(\frac{\delta H}{\delta x_1}\right) = \frac{e^D}{B^2} * w_1$$

Plug in the values for B, D, w_0 and w_1 and put them into the diagram. Also, the gradient for w_2 is:

$$\frac{df(w,x)}{dw_2} = \frac{df(w,x)}{dE} = \frac{e^D}{B^2} = 0.196$$