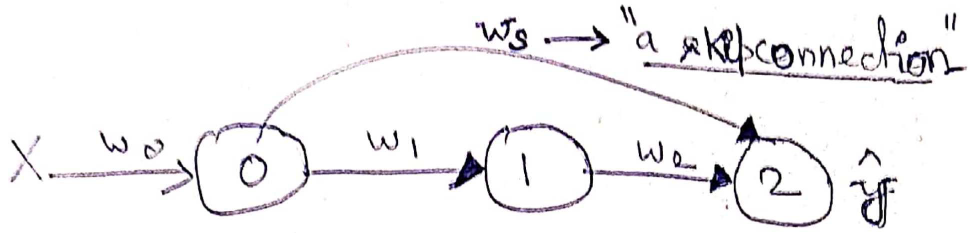
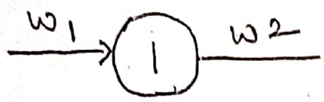
Exercise : Backpropagation.



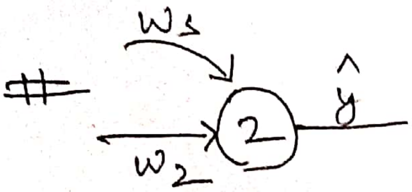
Forward Pass:

$z_0 = w_0 \cdot X$, $h_0 = g_0(z_0) = \begin{cases} 0 , & z_0 < 0 \\ z_0 , & \text{else} \end{cases}$

↑ after activation



$z_1 = w_1 \cdot h_0$, $h_1 = g_1(z_1) = \begin{cases} 0 , & z_1 < 0 \\ z_1 , & \text{else} \end{cases}$



$z_2 = w_2 h_1 + w_s h_0$

In this no activation is applied

So, $h_2 = g_2(z_2) = z_2 = \hat{y}$

Backward Pass:

Since no h_2, we start with z_2;

① $\dfrac{\partial L}{\partial z_2} = \dfrac{\partial L}{\partial \hat{y}} \cdot \dfrac{\partial \hat{y}}{\partial z_2} = \boxed{\dfrac{\partial L}{\partial \hat{y}} \cdot 1} *$

ex: $\dfrac{\partial x}{\partial x}$

$\boxed{\text{Since } \hat{y} = z_2}$

Scanned with CamScanner

$$\frac{\partial L}{\partial h_1} = \frac{\partial L}{\partial z_2} \cdot \underbrace{\frac{\partial z_2}{\partial h_1}}$$

$\downarrow$

$= (\text{known}) \cdot$

$\frac{\partial L}{\partial \hat{y}} \cdot w_2$

change in z_2 is determined by w_2

$$\boxed{\frac{\partial L}{\partial h_1} \Rightarrow \frac{\partial L}{\partial z_2} \cdot w_2} \; *$$

$$\frac{\partial L}{\partial z_1} = \frac{\partial L}{\partial h_1} \cdot \frac{\partial h_1}{\partial z_1} \longrightarrow \underbrace{\begin{cases} z_1, & z_1 \geq 0 \\ 0, & z_1 < 0 \end{cases}}$$

$\hookrightarrow$ depends on given activation function:

$$\boxed{\frac{\partial L}{\partial z_1} = \frac{\partial L}{\partial h_1} \cdot \begin{cases} z_1, & z_1 \geq 0 \\ 0, & z_1 < 0 \end{cases}} \; *$$

$$\frac{\partial L}{\partial h_0} = \frac{\partial L}{\partial z_1} \cdot \frac{\partial z_1}{\partial h_0} + \underbrace{\frac{\partial L}{\partial z_2} \cdot \frac{\partial z_2}{\partial h_0}}$$

$\hookrightarrow$ "because of skip connection"

$$\boxed{\frac{\partial L}{\partial h_0} = \frac{\partial L}{\partial z_1} \cdot w_1 + \frac{\partial L}{\partial z_2} \cdot w_s} \; *$$

$$\frac{\partial L}{\partial z_0} = \frac{\partial L}{\partial h_0} \cdot \underbrace{\frac{\partial h_0}{\partial z_0}} = \frac{\partial L}{\partial h_0} \cdot$$

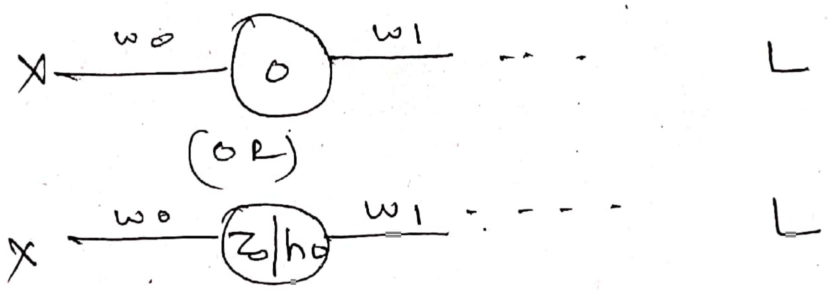
$\hookrightarrow$ activation o/p change based on 'z', can be written as

$$\boxed{\frac{\partial L}{\partial z_0} = \frac{\partial L}{\partial h_0} \cdot \begin{cases} z_0, & z_0 \geq 0 \\ 0, & z_0 < 0 \end{cases}} \; *$$

Scanned with CamScanner

$$\frac{\partial L}{\partial w_0} = \frac{\partial L}{\partial z_0} \cdot \frac{\partial z_0}{\partial w_0} = \frac{\partial L}{\partial z_0} \cdot x \quad \Big\} \; (1)$$

↳ since w_0 directly impacts z_0 in the given Network.



$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_1} = \frac{\partial L}{\partial z_1} \cdot h_0 \quad \Big\} \; (2)$$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial z_2} \cdot \frac{\partial z_2}{\partial w_2} = \frac{\partial L}{\partial z_2} \cdot h_1 \quad \Big\} \; (3)$$

$$\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial z_2} \cdot \frac{\partial z_2}{\partial w_3} = \frac{\partial L}{\partial z_2} \cdot h_0 \quad \Big\} \; 4.$$