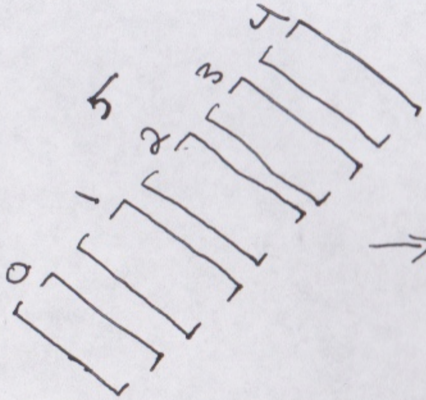
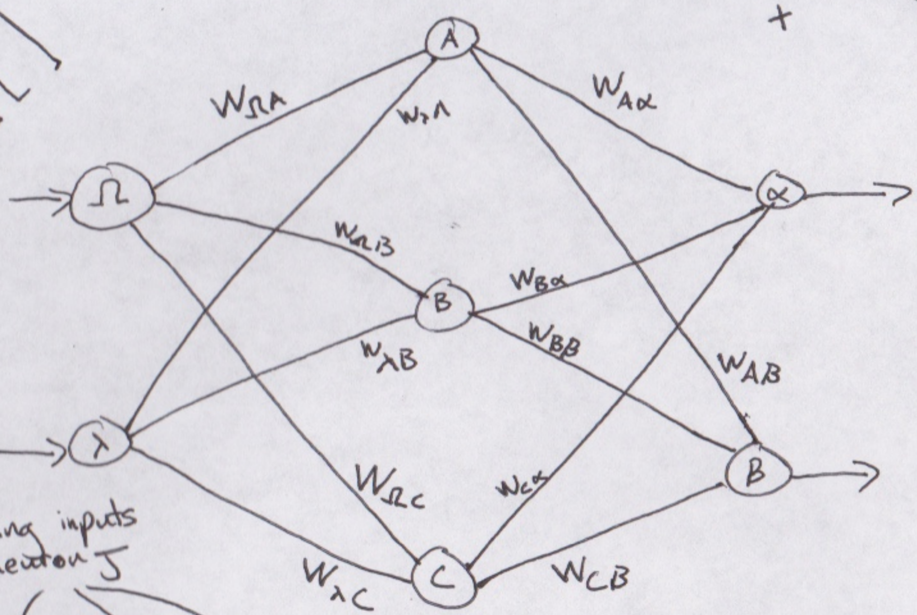




$$E = \frac{1}{2}(\underset{t}{\text{target}} - \underset{y}{\text{output}})^2$$

J = neuron #

$$\delta_\alpha = \cancel{\frac{dE}{dout_\alpha}} \cdot \cancel{\frac{dout_\alpha}{d\alpha}}$$

$$= \frac{\partial E}{\partial o_J} \cdot \frac{\partial o_J}{\partial net_J} \cdot \frac{\partial net_J}{\partial w_{iJ}}$$

$$\frac{\partial net_J}{\partial w_{iJ}} = \frac{\partial}{\partial w_{iJ}}\left(\sum_{k=1}^{n} w_{kJ} o_k\right)$$

$$= \frac{\partial}{\partial w_{iJ}} w_{iJ} o_i$$

$$= o_i$$

$L = u, v, \ldots, w$ receiving inputs from neuron J

<u>Inner node</u>

$$\frac{\partial E(o_J)}{\partial o_J} = \frac{\partial E(net_u, net_v, \ldots, net_w)}{\partial o_J}$$

$$\frac{\partial E}{\partial o_J} = \sum_{\ell \in L}\left(\frac{\partial E}{\partial net_\ell} \cdot \frac{\partial net_\ell}{\partial o_J}\right)$$

$$= \sum_{\ell \in L}\left(\frac{\partial E}{\partial o_\ell} \frac{\partial o_\ell}{\partial net_\ell} w_{J\ell}\right)$$

$$\boxed{\frac{\partial o_J}{\partial net_J}} = \frac{\partial}{\partial net_J}\left(\sigma(net_J)\right)$$

<u>for sigmoid</u> $\rightarrow$ $= \sigma(net_J)(1 - \sigma(net_J))$

<u>output node</u> $\rightarrow$ $\boxed{\frac{\partial E}{\partial o_J}} = \frac{\partial}{\partial y} \frac{1}{2}(t - y)^2 = -(t - y)$

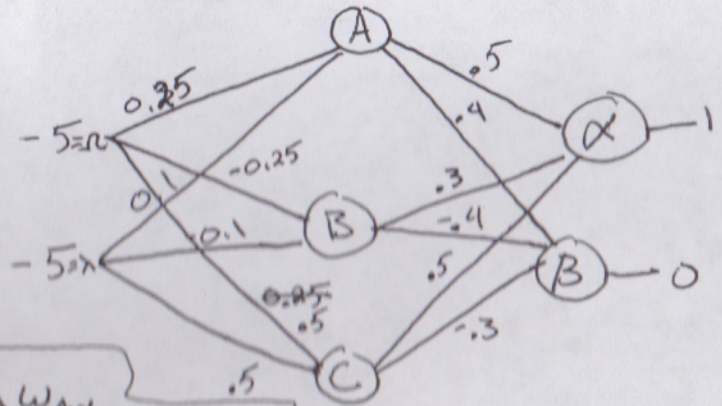
$$= (y - t)$$

$$\frac{\partial E}{\partial w_{iJ}} = o_i \delta_J$$

with $\delta_J = \frac{\partial E}{\partial o_J} \cdot \frac{\partial o_J}{\partial net_J} = \begin{cases} (o_J - t_J)\, o_J (1 - o_J) \rightarrow \text{if } J \text{ is output} \\ \left(\sum_{\ell \in L} \delta_\ell w_{J\ell}\right) o_J (1 - o_J) \rightarrow \text{if } J \text{ is inner} \end{cases}$

$$\Delta w_{iJ} = -\eta \frac{\partial E}{\partial w_{iJ}}$$

$-5 \; \frac{1}{0}$



Relu

$f(x) = max(x, 0)$

$\frac{\partial f(x)}{\partial x} = \begin{cases} 1 & if\ x > 0 \\ 0 & o/w \end{cases}$

$\delta_\alpha = (out_\alpha - target_\alpha) \cdot \frac{\partial f(x)}{\partial x}$

$= (0.525 - 1) \cdot 1$

$= -0.475$

$\delta_\beta = (0 - 0) \cdot 0$

$= 0$

Forward

$W_{\Omega A} = 0.35$

$W_{\lambda A} = 0.1$

$net_A = -5 \cdot 0.35 + -5 \cdot .1$

$= -1.75$

$out_A = max(0, -1.75) = 0$

$W_{\Omega B} = -0.25$

$W_{\lambda B} = -0.1$

$net_B = -5 \cdot -0.25 + -5 \cdot -0.1$

$= 1.25 + .5 = 1.75$

$out_B = max(0, 1.75) = 1.75$

$W_{\Omega C} = .5$

$W_{\lambda C} = .5$

$net_C = .5 \cdot -5 + .5 \cdot -5$

$= -2.5 + -2.5 = -5$

$out_C = max(0, -5) = 0$

$W_{A\alpha} = .5 \cdot 0 + .3 \cdot 1.75 + .5 \cdot 0$

$net_\alpha = 0.525$

$out_\alpha = max(0, 0.525) = 0.525$

$\cancel{W_{AB} = .4 \cdot 0}$

$net_\beta = .4 \cdot 0 + -.4 \cdot 1.75 + 0$

$= -0.7$

$out_\beta = max(0, -0.7) = 0$

$\Delta w_{A\alpha}$

$\Delta w_{A\alpha} = O_A \cdot \delta_\alpha = 0 \cdot -0.475 = 0$

$\Delta w_{B\alpha} = O_B \cdot \delta_\alpha = 1.75 \cdot -0.475 = -0.83125 \cdot -1 = \boxed{0.83125}$

$\Delta w_{C\alpha} = O_C \cdot \delta_\alpha = 0 \cdot -0.475 = 0$

$\Delta w_{A\beta} = O_A \cdot \delta_\beta = 0 \cdot 0 = 0$

$\Delta w_{B\beta} = O_B \cdot \delta_\beta = 1.75 \cdot 0 = 0$

$\Delta w_{C\beta} = O_C \cdot \delta_\beta = 0 \cdot 0 = 0$

$\delta_A = \left(\sum_{\ell \in L} \delta_\ell w_{A\ell}\right) \cdot \frac{\partial f(x)}{\partial x}$

$= \left(\delta_\alpha \cdot w_{A\alpha} + \delta_\beta \cdot w_{A\beta}\right) \frac{\partial f(x)}{\partial x}$

$= 0$

$\delta_B = \left(\delta_\alpha \cdot w_{B\alpha} + \delta_\beta \cdot w_{B\beta}\right) \frac{\partial f(x)}{\partial x}$

$= (-0.475 \cdot .3) \cdot 1$

$= -0.1425$

$\delta_C = 0$

$\Delta w_{\Omega A} = -out_\Omega \cdot \delta_A$

$= 0$

$\Delta w_{\lambda A} = -out_\lambda \cdot \delta_A$

$= 0$

$\Delta w_{\Omega C} = 0, \; \Delta w_{\lambda C} = 0$

$\Delta w_{\Omega B} = -out_\Omega \cdot \delta_B$

$= +5 \cdot -0.1425 = -0.7125$

$\Delta w_{\lambda B} = -out_\lambda \cdot \delta_B = -0.7125$