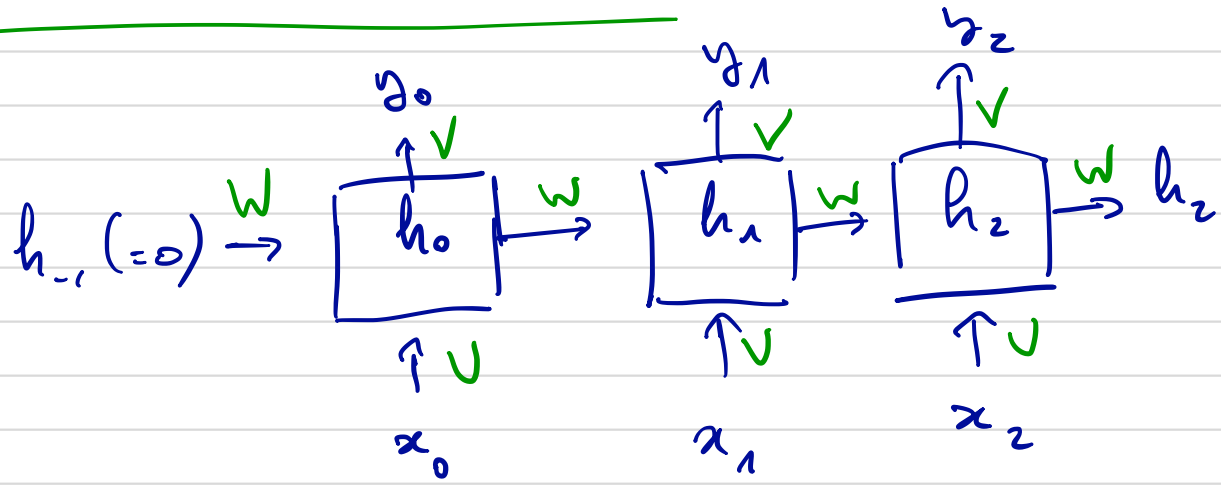
Correction TD3 exercice 1

1)

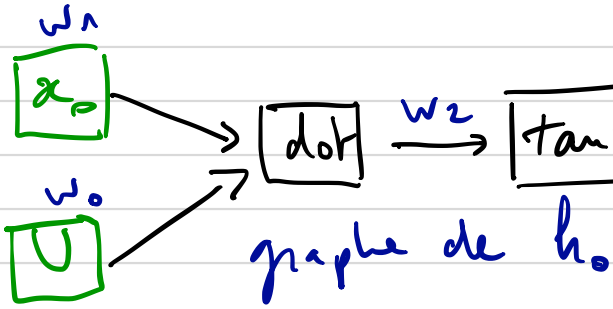


2) $y_2 = \text{softmax}(V \tanh(U x_2 + W h_1))$

$y_2 = \text{softmax}(V \tanh(U x_2 + W(\tanh(U x_1 + W h_0))))$

$= \text{softmax}(V \underbrace{\tanh(U x_2 + W(\tanh(U x_1 + W \tanh(U x_0))))}_{= h_2})$

3)



graphe de $h_0 = \tanh(U x_0)$

$w_0 = U$, $w_1 = x_0$

$w_2 = w_0 w_1$

$w_3 = \tanh(w_2)$

$\bar{w}_0 = \bar{w}_2 \times \frac{\partial w_2}{\partial w_0} = \bar{w}_2 w_1$

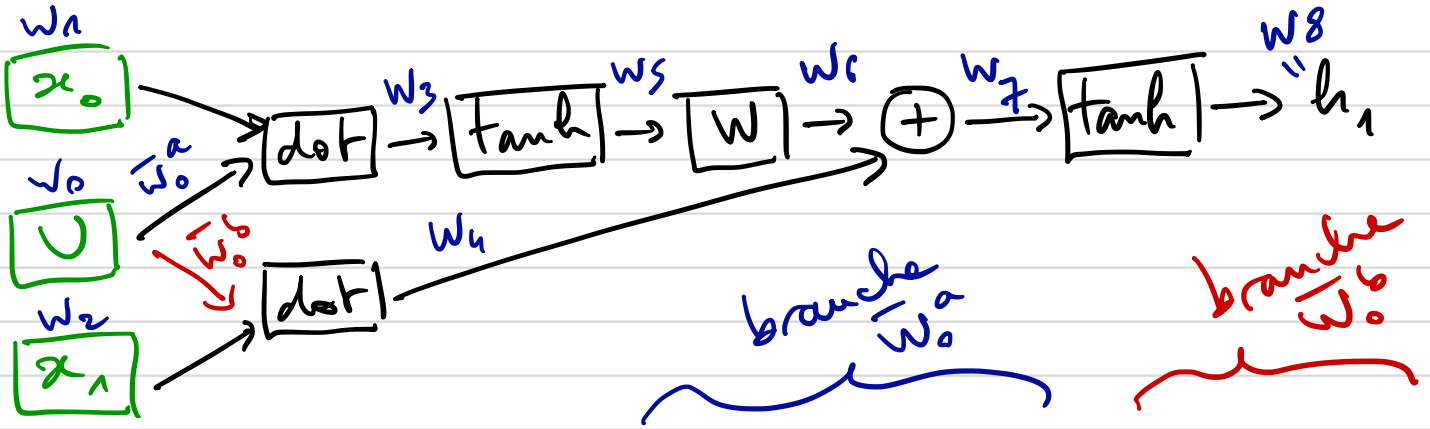
$\bar{w}_2 = \bar{w}_3 \frac{\partial w_3}{\partial w_2} = 1 \times (1 - \tanh^2(w_2))$

$\bar{w}_3 = \frac{dw_3}{dw_3} = 1$ seed

"dU" $= \frac{\partial h_0}{\partial U} = \bar{w}_0 = (1 - \tanh^2(U x_0)) \times x_0$

$\Rightarrow$ "dU" $= x_0 (1 - h_0^2)$

Graphe $h_1 = \tanh(U x_1 + W \tanh(U x_0))$



$w_0 = U,\ w_1 = x_0,\ w_2 = x_1$

$w_3 = w_0 w_1$

$w_4 = w_0 w_2$

$w_5 = \tanh(w_3) = h_0$

$w_6 = W w_5$

$w_7 = w_4 + w_6$

$w_8 = \tanh(w_7) = h_1$

$\bar{w}_0^a = \bar{w}_3 \frac{\partial w_3}{\partial w_0} = \bar{w}_3 w_1$

$\bar{w}_3 = \bar{w}_5 (1 - \tanh^2 w_3)$
$= \bar{w}_5 (1 - h_0^2)$

$\bar{w}_5 = \bar{w}_6 W$

$\bar{w}_6 = \bar{w}_7 \times 1$

$\bar{w}_7 = 1 - w_8^2 = 1 - h_1^2$

$\bar{w}_8 = \frac{dw_8}{dw_8} = 1$

$\bar{w}_0^b = \bar{w}_4 \frac{\partial w_4}{\partial w_0}$
$= \bar{w}_4 w_2$

$\bar{w}_4 = \bar{w}_7 \times 1$

$\bar{w}_7 = 1 - w_8^2$

$\bar{w}_8 = 1$

$\bar{w}_8 = 1$: "seed"

Au final, on a $\bar{w}_0 = \bar{w}_0^a + \bar{w}_0^b$

Soit "dU" $= \frac{\partial h_2}{\partial U} = \bar{w}_0 = (1 - h_1^2) W (1 - h_0^2) x_0$
$+ (1 - h_1^2) x_1$

Soit $dU = (1 - h_1^2)(x_1 + W(1 - h_0^2) x_0)$

En déduire $\frac{\partial h_2}{\partial U} = ?$

On a $\frac{\partial h_0}{\partial U} = (1 - h_0^2)\, x_0$

$$\frac{\partial h_1}{\partial U} = (1 - h_1^2)\left(x_1 + W \underbrace{(1 - h_0^2) x_0}_{\partial h_0 / \partial U}\right)$$

$$\Rightarrow \frac{\partial h_2}{\partial U} = (1 - h_2^2)\left[x_2 + W \underbrace{(1 - h_1^2)\left(x_1 + W(1 - h_0^2) x_0\right)}_{\partial h_1 / \partial U}\right]$$