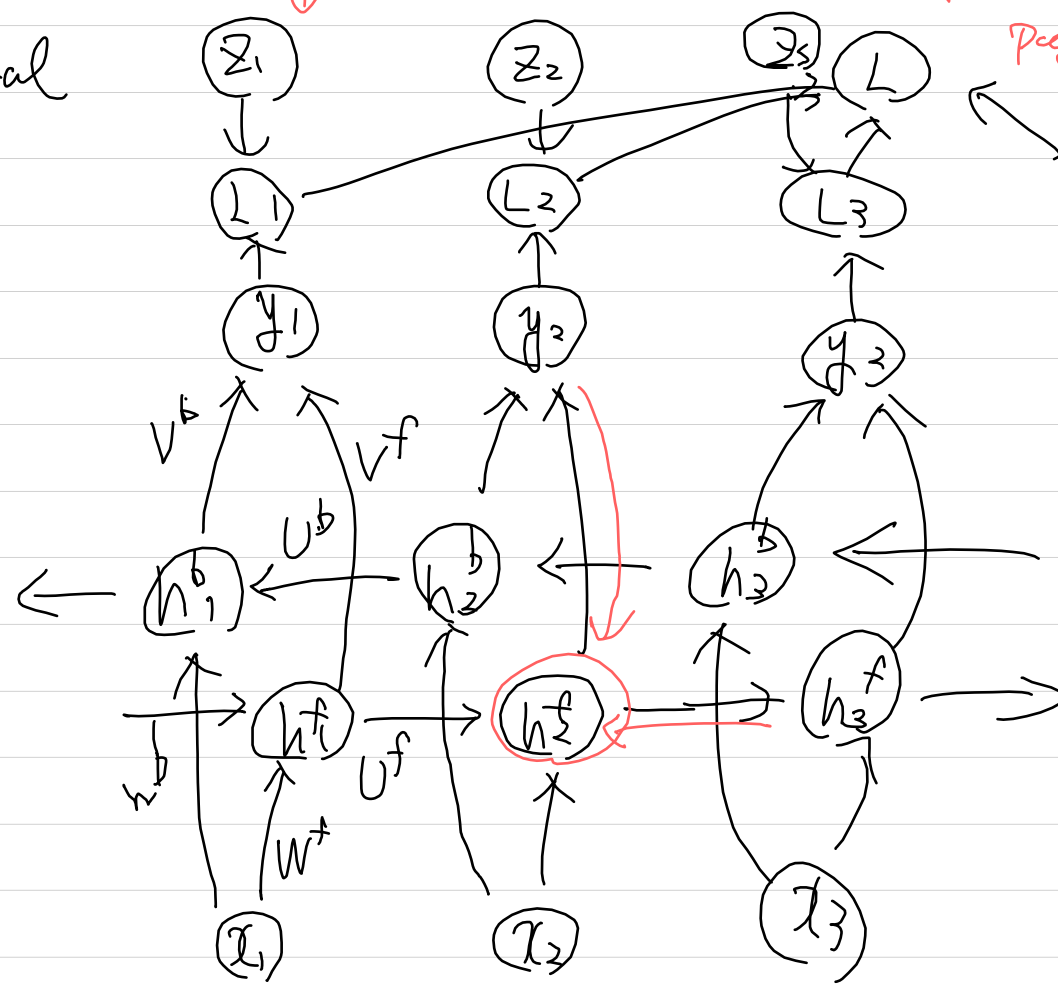
2014 Q4 (Bidirectional RNN) April 20th

[1] Draw computational graph

This includes mistake go to page 6



[2]
$$\nabla_{h_t^f} L = \frac{\partial L}{\partial h_t^f} = \frac{\partial L}{\partial h_{t+1}^f}\frac{\partial h_{t+1}^f}{\partial h_t^f} + \frac{\partial L}{\partial L_t}\frac{\partial L_t}{\partial h_t^f}$$

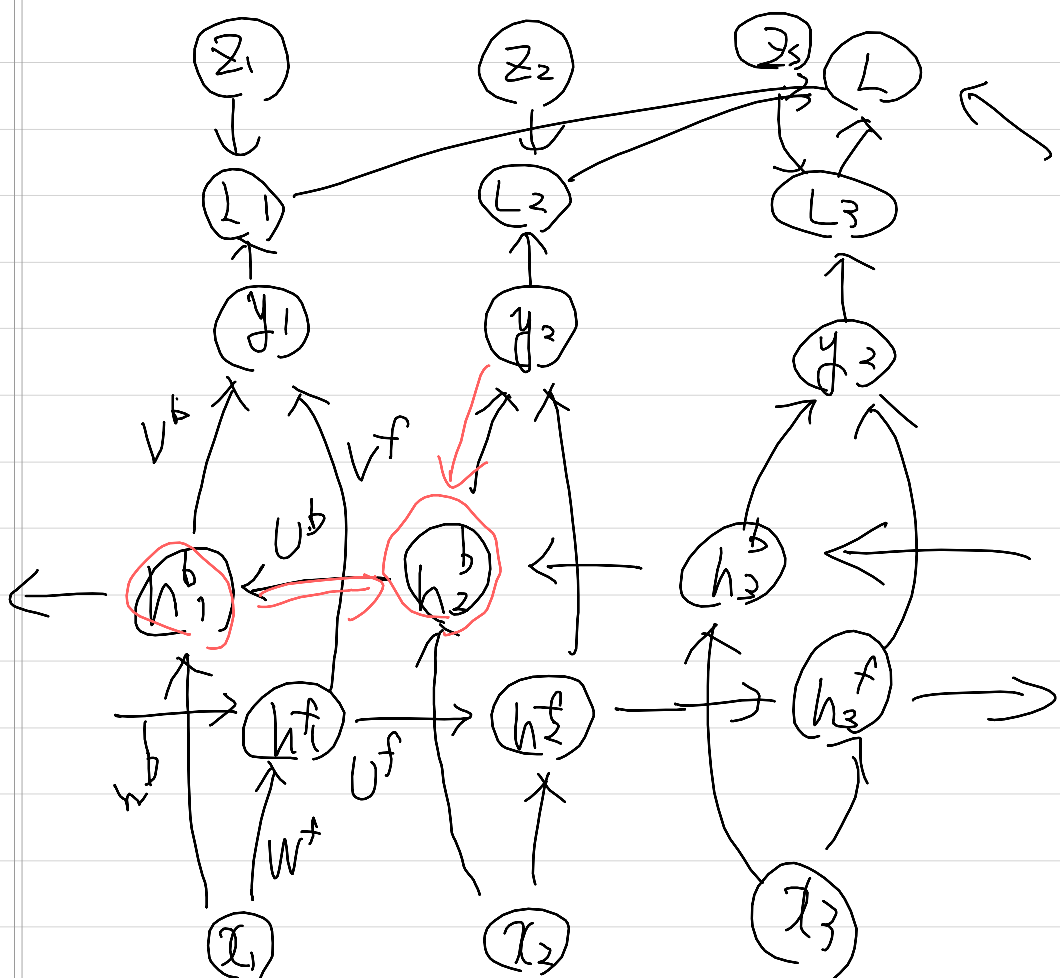
$$\frac{\partial L}{\partial L_t} = 1$$

$$= \nabla_{h_{t+1}^f} L \left(\frac{\partial h_{t+1}^f}{\partial h_t^f}\right) + \nabla_{h_t^f} L_t$$

20H_Q4

$$\nabla_{h_t^f} L_{t+1} = \frac{\partial L_{t+1}}{\partial h_t^f} = \frac{\partial h_{t+1}^f}{\partial h_t^f} \frac{\partial L_{t+1}}{\partial h_{t+1}^f}$$

$$= \left(\frac{\partial h_{t+1}^f}{\partial h_t^f} \right) \nabla_{h_{t+1}^f} L_{t+1}$$



$$\nabla_{h_t^b} L = \frac{\partial L}{\partial h_t^b} = \frac{\partial L}{\partial h_{t-1}^b} \frac{\partial h_{t-1}^b}{\partial h_t^b} + \frac{\partial L}{\partial L_t} \frac{\partial L_t}{\partial h_t^b}$$

$$= \nabla_{h_{t-1}^b} L \left(\frac{\partial h_{t-1}^b}{\partial h_t^b} \right) + \nabla_{h_t^b} L_t$$

$$\nabla_{h_t^b} L_t = \frac{\partial L_t}{\partial h_t^b} = \frac{\partial L_t}{\partial y_t} \frac{\partial y_t}{\partial h_t^b} = -2y_t \frac{\partial y_t}{\partial h_t^b}$$

2UH_Q4

$$\nabla_{h_t^b} L_t = 2y \cdot \frac{\partial y_t}{\partial h_t^b} = -2y \cdot V^b$$

$$\left(y_t = V^f h^f_{t-1} + V^b h_t^b \right)$$

$$h_t = \sigma(g_t)$$

$$g_t^f = W^f x_t + U^f h^f_{t-1}$$

$$h_{t+1} \leftarrow g_t \leftarrow h_t$$

$$\frac{\partial h_{t+1}^f}{\partial h_t^f} = \frac{\partial h_{t+1}^f}{\partial g_{t+1}^f} \frac{\partial g_{t+1}^f}{\partial h_t^f}$$

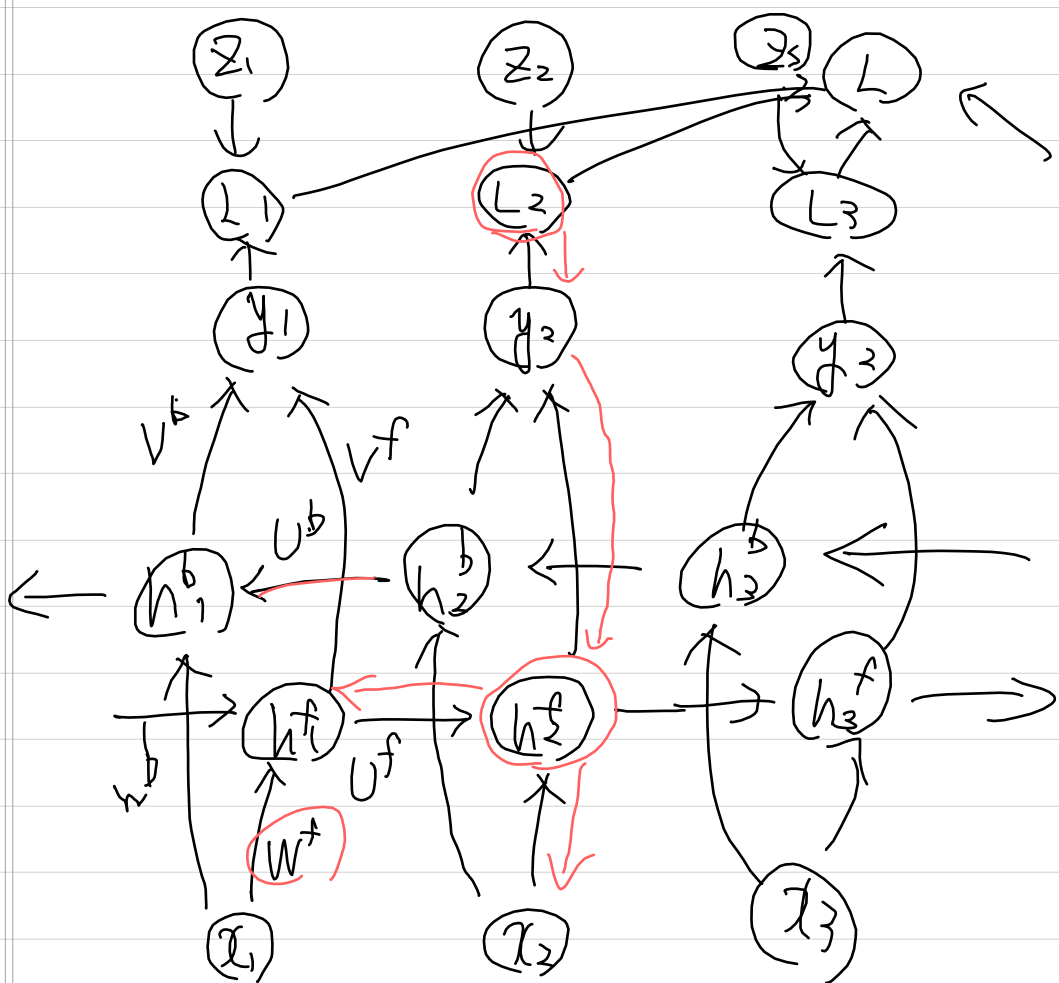
$$= (1 - h_{t+1}^f) h_{t+1}^f \cdot U^f$$

$$g_t^b = W^b x_t + U^b h_{t+1}^b$$

$$\frac{\partial h_{t-1}^b}{\partial h_t^b} = \frac{\partial h_{t-1}^b}{\partial g_{t-1}^b} \frac{\partial g_{t-1}^b}{\partial h_t^b}$$

$$= (1 - h_{t-1}^b) h_{t-1}^b \cdot U^b$$

$$\nabla_{w^f} L = \frac{\partial L}{\partial w^f} = \sum_{t=1}^{\tau} \frac{\partial L_t}{\partial w^f}$$



then focus $\frac{\partial L_t}{\partial w^f} = \frac{\partial L_t}{\partial y_t} \cdot \frac{\partial y_t}{\partial h^f_t} \cdot \frac{\partial h^f_t}{\partial w^f}$

$$\frac{\partial L_t}{\partial w^f} = \frac{\partial L_t}{\partial y_t} \frac{\partial y_t}{\partial h^f_t} \cdot \left(\frac{\partial h_t^f}{\partial w^f} + \frac{\partial h_t^f}{\partial h_{t-1}^f} \frac{\partial h_{t-1}^f}{\partial w^f} \right)$$

2014-Q4

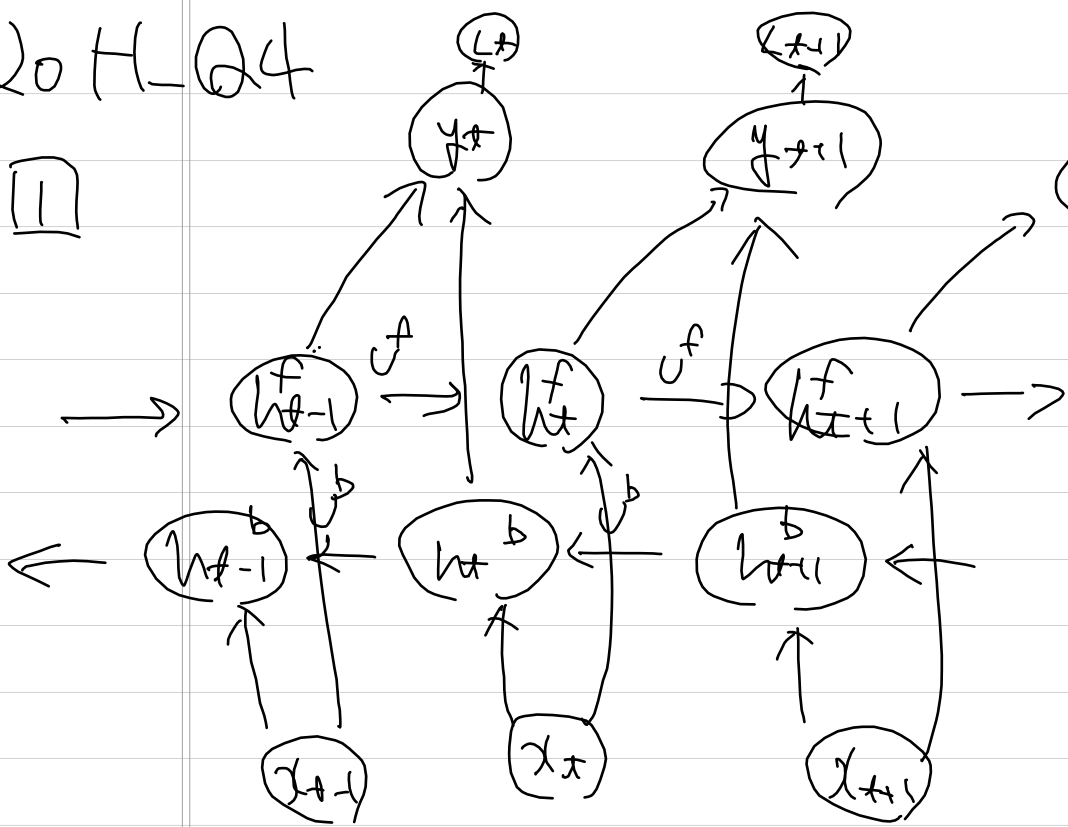
$$\frac{\partial L_t}{\partial w^f} = \frac{\partial L_t}{\partial y_t} \frac{\partial y_t}{\partial h_{t-1}} \left(\frac{\partial h_{t-1}^f}{\partial w^f} + \frac{\partial h_{t-1}^f}{\partial h_{t-2}^f} \frac{\partial h_{t-2}^f}{\partial w^f} \right)$$

$$-2y_t \cdot T^f \cdot \left(\quad \cdots \quad \right)$$

to be continued...

20H_Q4

$\boxed{1}$



$\boxed{2}$

$$\nabla_{h_t^f} L = \frac{\partial L}{\partial h_t^f} = \frac{\partial L}{\partial h_{t+1}^f}\frac{\partial h_{t+1}^f}{\partial h_t^f}$$

$$+ \frac{\partial L}{\partial L_{t+1}}\frac{\partial L_{t+1}}{\partial h_t}$$

$$= \underline{\nabla_{h_t} L_{t+1}} + \left(\frac{\partial h_{t+1}^f}{\partial h_t^f}\right)^T \nabla_{h_{t+1}^f} L$$