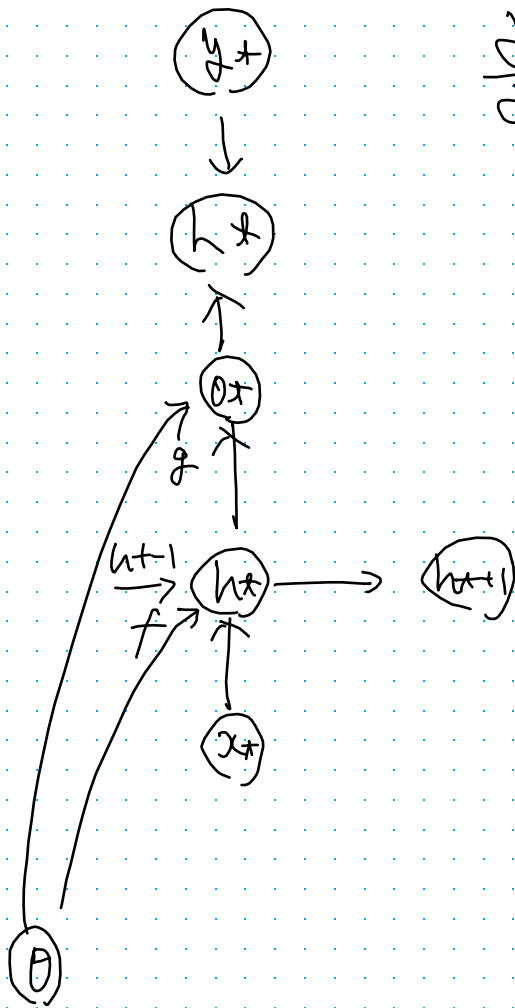




$$\frac{\partial \mathcal{L}}{\partial \theta} = \frac{\partial \mathcal{L}_1}{\partial \theta} + \dots + \underbrace{\frac{\partial \mathcal{L}_t}{\partial \theta}} + \dots + \frac{\partial \mathcal{L}_T}{\partial \theta}$$

Focus

$$\frac{\partial \mathcal{L}_t}{\partial \theta} = \frac{\partial \mathcal{L}_t}{\partial o_t} \frac{\partial o_t}{\partial \theta}$$

$$= \frac{\partial \mathcal{L}_t}{\partial o_t} \left\{ \left(\frac{\partial o_t}{\partial \theta} \right)_{o_t} + \left(\frac{\partial o_t}{\partial \theta} \right)_{h_t} \right\}$$

$$= \frac{\partial \mathcal{L}_t}{\partial o_t} \left\{ \nabla_{\theta} g + \frac{\partial o_t}{\partial h_t} \frac{\partial h_t}{\partial \theta} \right\}$$

$$= \frac{\partial \mathcal{L}_t}{\partial o_t} \left\{ \nabla_{\theta} g + \nabla_{g f} f \left(\left(\frac{\partial h_t}{\partial \theta} \right)_{h_t} + \left(\frac{\partial h_t}{\partial \theta} \right)_{h_{t+1}} \right) \right\}$$

$$= \nabla_{g_t} \mathcal{L} \left\{ \nabla_{\theta} g + \nabla_{g f} f \left(\nabla_{\theta} f + \frac{\partial h_t}{\partial h_{t+1}} \frac{\partial h_{t+1}}{\partial \theta} \right) \right\}$$

$$= \nabla_{g_t} \mathcal{L} \left\{ \nabla_{\theta} g + \nabla_{g f} f \left\{ \nabla_{\theta} f + \frac{\partial h_t}{\partial h_{t+1}} \left(\left(\frac{\partial h_{t+1}}{\partial \theta} \right)_{h_{t+1}} + \left(\frac{\partial h_{t+1}}{\partial \theta} \right)_{h_{t+2}} \right) \right\} \right\}$$

recurrence ...