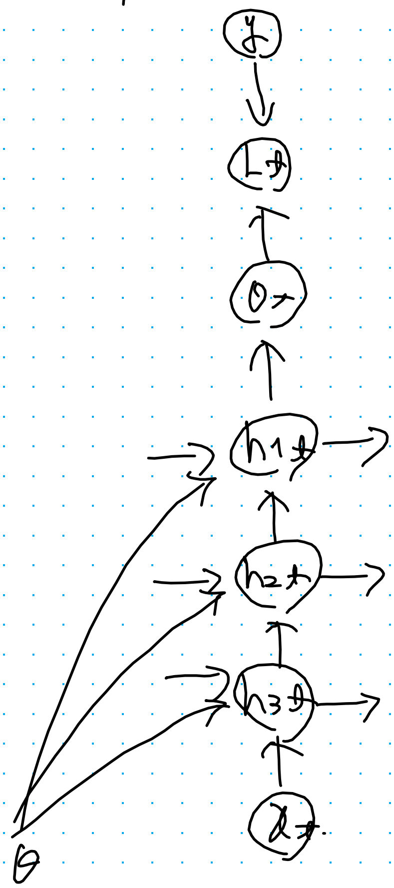
Subject Date 8

Attendants Answer Q-(b)



$$\frac{\partial L_t}{\partial \theta} = \frac{\partial L_t}{\partial O_t}\frac{\partial O_t}{\partial h_{1t}}\frac{\partial h_{1t}}{\partial \theta}$$

$$= \nabla_{g_t} h_t \nabla_{h_{1t}} g_t \left\{ \frac{\partial h_{1t}}{\partial \theta} + \frac{\partial h_{1t}}{\partial h_{1t+1}}\frac{\partial h_{1t+1}}{\partial \theta} + \frac{\partial h_{1t}}{\partial h_{2t}}\frac{\partial h_{2t}}{\partial \theta} \right\}$$

$$= \nabla_{g_t} h_t \nabla_{h_{1t}} g_t \left\{ \nabla_\theta h_{1t} + \underbrace{\left(\frac{\partial h_{1t}}{\partial h_{1t+1}}(\cdots)\right)}_{\text{depends on } t} + \frac{\partial h_{1t}}{\partial h_{2t}}\left\{ \frac{\partial h_{2t}}{\partial \theta} + \frac{\partial h_{2t}}{\partial h_{2t+1}}\frac{\partial h_{2t+1}}{\partial \theta} + \frac{\partial h_{2t}}{\partial h_{3t}}\frac{\partial h_{3t}}{\partial \theta} \right\} \right\}$$

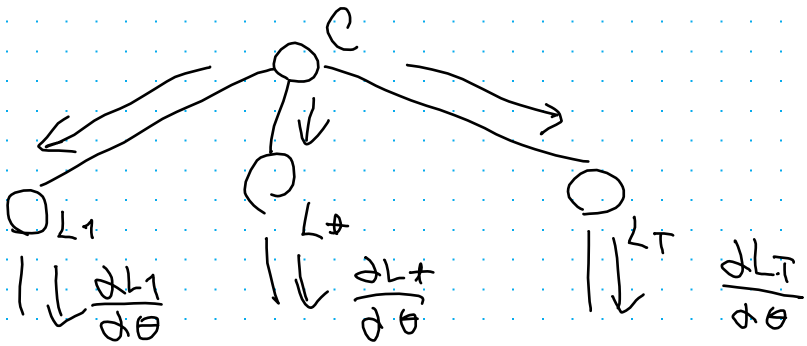
$$= \nabla_{g_t} h_t \nabla_{h_{1t}} g_t \left\{ \nabla_\theta h_{1t} + \left(\frac{\partial h_{1t}}{\partial h_{1t+1}}(\cdots)\right) + \frac{\partial h_{1t}}{\partial h_{2t}}\left\{ \frac{\partial h_{2t}}{\partial \theta} + \left(\frac{\partial h_{2t}}{\partial h_{2t+1}}(\cdots)\right) + \frac{\partial h_{2t}}{\partial h_{3t}}\left\{ \frac{\partial h_{3t}}{\partial \theta} + \underbrace{\frac{\partial h_{3t}}{\partial h_{3t+1}}\frac{\partial h_{3t+1}}{\partial \theta}}_{\text{depends on } t} \right\}_{①} \right\}_{②} \right\}_{③}$$

$$\square \{ \bigcirc + ③^t + \square \{ \bigcirc + ③^t + \square \{ \bigcirc + ③^t \{_{①} \}_{②} \}_{①} + \bigcirc$$

Subject Date 9

Attendants

$$\frac{\partial L_t}{\partial \theta} = \{ h^t + \{ h^t + \{ h^t \}\}\}$$



$$(x \le y \le z)$$

$$(3z)^t \ge x^t + y^t + z^t$$

So, compute $\frac{\partial L_t}{\partial \theta}$ takes $O(3^t)$

as a computational complexity.