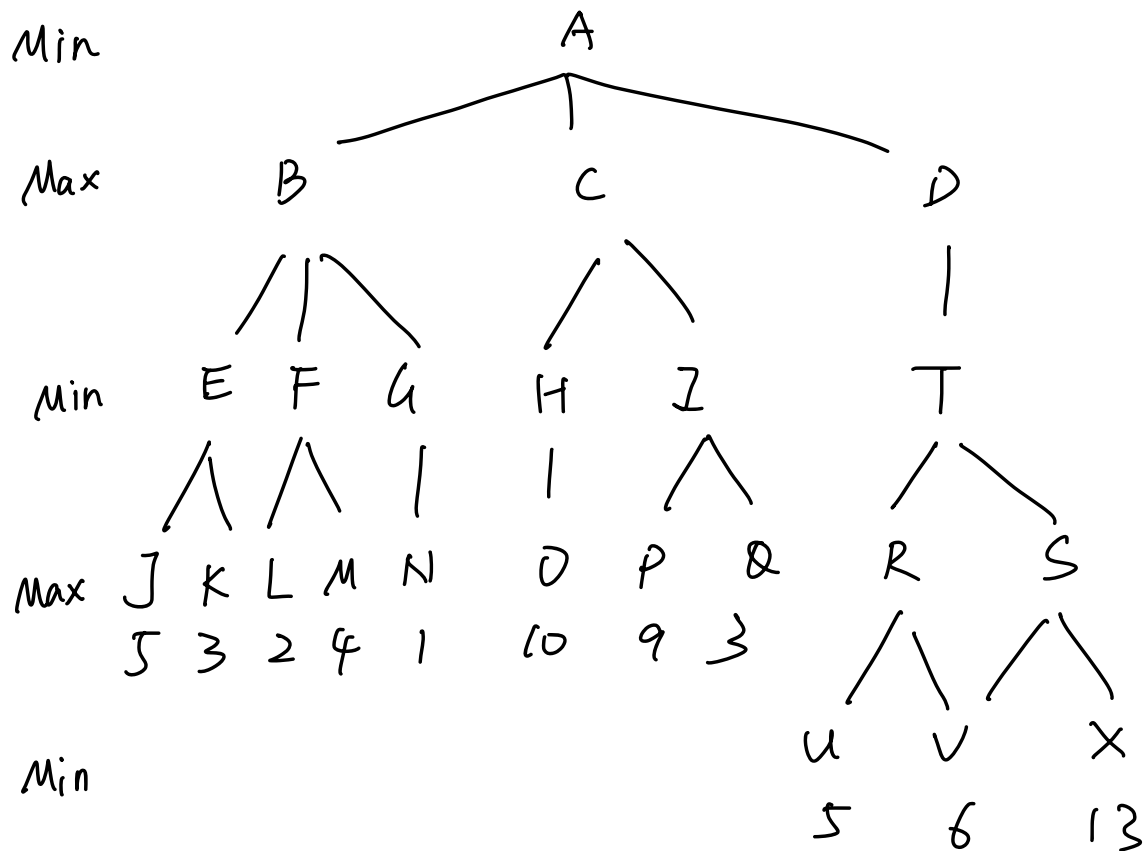


23-S1-Q2

first - minimizer

Qc(i) left \rightarrow right. $\alpha - \beta$.

(ii) not be examined. nodes



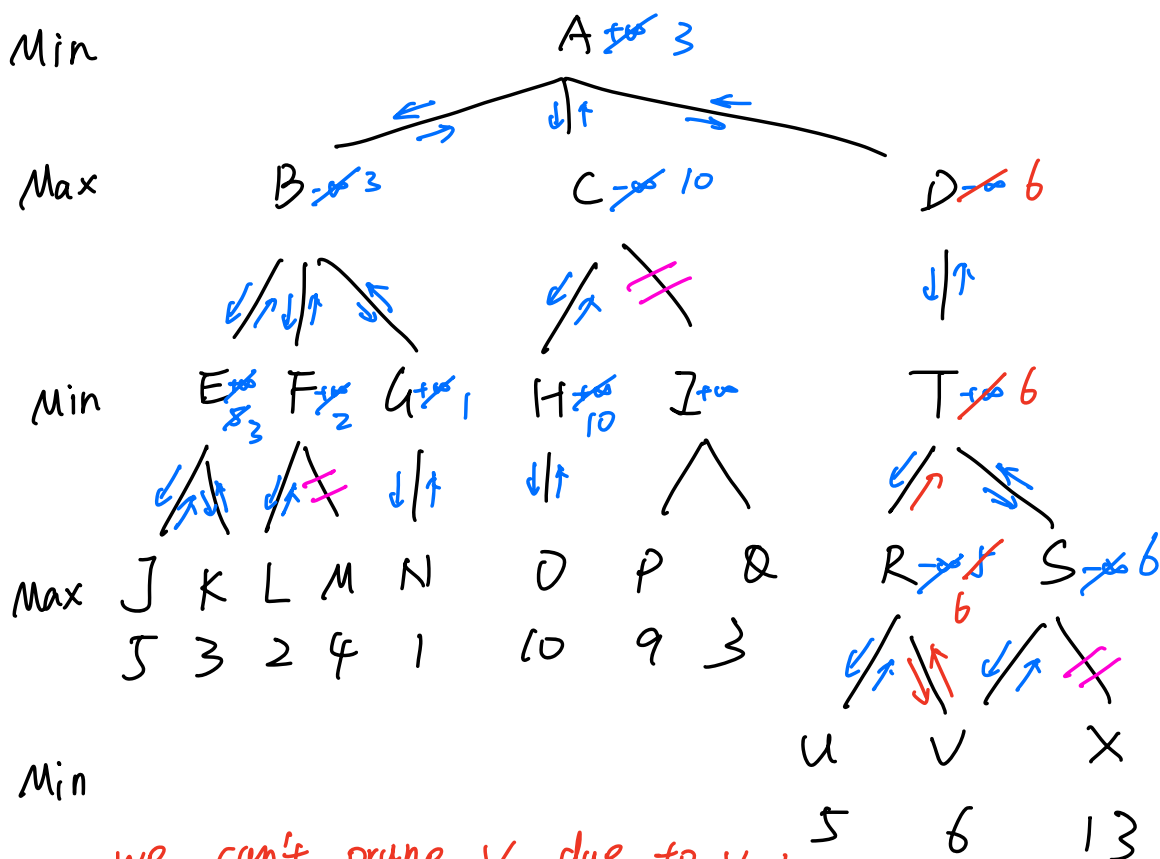
cb) NN

(i) 0j

(ii)

cc) (i) gradient vanishing
(ii) methods

Solution (a)



(ii) not examine nodes

M I P Q X

(b) (i) O_j

$$\text{bias} = [0.2 \quad 0.6 \quad 0.2 \quad -0.9 \quad -1.0 \quad 0.8]$$

unit j	Net input net_j	output O_j
1		1
2	$0.8 \times 1 + 0.2 \times 1 = 1$	1
3	$0.4 \times 1 + 0.6 \times 1 = 1$	1
4	$1 \times 1 + 0.5 \times 1 + 0.3 \times 1 + 0.2 \times 1 = 2$	2
5	$0.5 \times 1 + 0.2 \times 1 + 0.1 \times 1 - 0.9 \times 1 = -0.1$	0
6	$0.3 \times 2 + 0.2 \times 0 - 1 \times 1 = -0.4$	0
7	$0.2 \times 2 + 0.5 \times 0 + 0.8 \times 1 = 1.2$	1.2

$$\text{cii)} \quad \delta_j = \sigma'(\text{net}_j) \sum_k \delta_k w_{kj}$$

$$\delta_k = \sigma'(\text{net}_k) (t_k - O_k)$$

$$\sigma'(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases}$$

$$\begin{aligned} \text{ciii)} \quad \delta_6 &= \sigma'(\text{net}_6) (t_6 - O_6) \\ &= 0 \times (0.8 - 0) \\ &= 0 \end{aligned}$$

$$\begin{aligned} \delta_7 &= \sigma'(\text{net}_7) (t_7 - O_7) \\ &= 1 \times (0.2 - 1.2) \\ &= -1 \end{aligned}$$

$$\delta_4 = \sigma'(\text{net}_4) (\delta_6 \times W_{64} + \delta_7 \times W_{74})$$

$$= 1 \times (0 \times 0.3 + (-1) \times 0.2)$$

$$= -0.2$$

$$(iv) \Delta w_{ji} = \eta \delta_j o_i$$

$$\Delta w_{41} = \eta \delta_4 o_1$$

$$= 0.1 \times (-0.2) \times 1$$

$$= -0.02$$

Cc) (i) Gradients shrink exponentially while BP through many layer, because $\sigma'(x) \leq 0.25$ for a sigmoid.

② Products of many $\sigma'(x)$ terms drive δ toward zero, so earlier layers learn extremely slowly or stop learning altogether.

(ii) ① use ReLU / Leaky-ReLU to keep $\sigma'(x) \approx 1, x > 0$

② maintain variance of activation and gradients by proper weight initialization

③ Batch / Layer Normalization rescales activations keeping them in regions with healthy derivative

④ Use Residual connection

⑤ Use Gradient-clipping or adaptive optimizers e.g. Adam, RMSProp prevent tiny update after many layers.