# Back Propagation:- @t = 1

$\Delta_1 = \partial_x E = 0.771 - 1.25 = -0.470$

$\Delta out_1 = 0$ because there are no future time steps.

$$\delta out_1 = \Delta_1 + \Delta out_1 = -0.47 + 0 = -0.47$$

$$\delta state_1 = \delta out_1 \odot o_1 \odot (1 - \tanh^2(state_1)) + \delta state_2 \odot$$

$$f_2 = -0.478 \times 0.84 \times (1 - \tanh^2(1.517)) + 0 \times 0 = -0.071$$

$$\delta a_1 = \delta state_1 \odot i_1 \odot (1 - a_1^2) = -0.071 \times 0.98 \times (1 - 0.84)^2 = -0.1019$$

$$\delta i_1 = \delta state_1 \odot a_1 \odot i_1 \odot (1 - i_2) = -0.071 \times 0.849 \times 0.981 \times (1 - 0.981) = -0.001$$

$$\delta f_1 = \delta state_1 \odot state_0 \odot f_1 \odot (1 - f_1) = 0.071 \times 0.785 \times 0.870 \times (1 - 0.8703) = -0.00631$$

$$\delta o_1 = \delta out_1 \odot \tanh(state_1) \odot o_1 \odot (1 - o_1) = -0.47 \times \tanh(1.5776) \times 0.849 \times (1 - 0.849) = -0.055$$

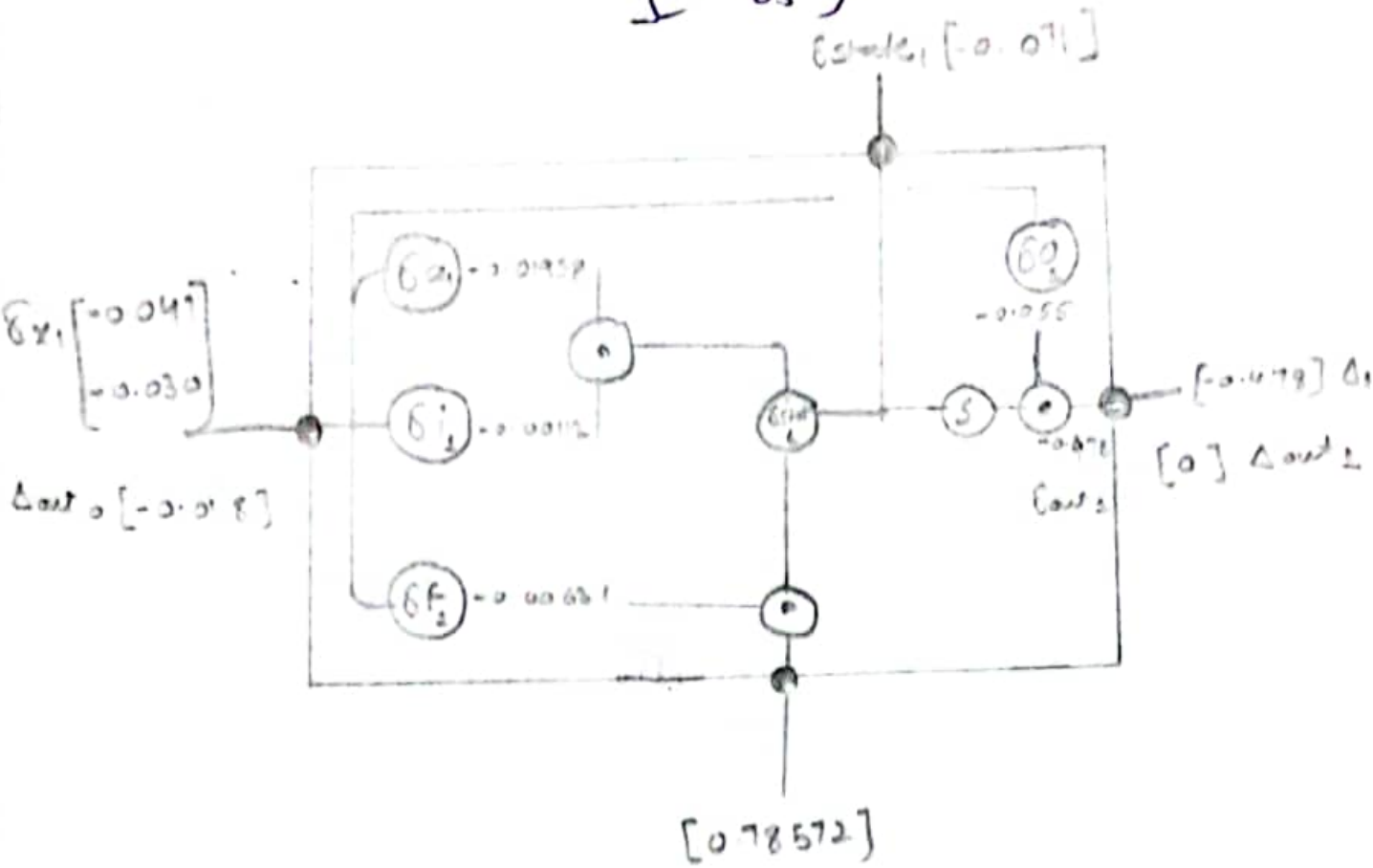
$$\delta x_1 = W^T \cdot \delta gates_1$$

$$= \begin{bmatrix} 0.45 & 0.95 & 0.70 & 0.60 \\ 0.25 & 0.80 & 0.45 & 0.40 \end{bmatrix} \begin{bmatrix} -0.0193 \\ -0.001 \\ -0.00631 \\ -0.055 \end{bmatrix} = \begin{bmatrix} -0.0474 \\ -0.0307 \end{bmatrix}$$

Scanned with CamScanner

$\Delta out_0 = U^T \cdot \delta gates_1$

$$= \begin{bmatrix} 0.15 & 0.80 & 0.10 & 0.25 \end{bmatrix} \begin{bmatrix} -0.019 \\ -0.0011 \\ -0.0063 \\ -0.05 \end{bmatrix} = -0.018228$$




Scanned with CamScanner

## Backward @t = 0

$\Delta_0 = \partial_x E = 0.536 - 0.5 = 0.0363$

$\Delta out_0 = -0.018$, passed back from T = 1

$\Delta out_0 = \Delta_0 + \Delta out_0 = 0.03631 + -0.018 = 0.01803$

$\delta state_0 = \delta out_0 \odot o_0 \odot (1 - \tanh^2(state_0)) + \delta state_1 \odot f_1 =$
$0.01803 \times 0.81757 \times (1 - \tanh^2(0.785))$
$+ -0.071 \times 0.870 = -0.053$

$\delta a_0 = \delta state_0 \odot i_0 \odot (1 - a_0^2) = -0.053 \times 0.960 \times (1 - 0.817^2) = -0.017$

$\delta i_0 = \delta state_0 \odot a_0 \odot i_0 \odot (1 - i_0) = -0.053 \times 0.817 \times 0960 \times (1 - 0.96)$
$= -0.00165$

$\delta f_0 = \delta state_0 \odot state_{-1} \odot f_0 \odot (1 - f_0) = -0.053 \times 0 \times 0.851 \times$
$(1 - 0.851) = 0$

$\delta o_0 = \delta out_0 \odot \tanh(state_0) \odot o_0 \odot (1 - o_0) = 0.018 \times \tanh($
$0.785) \times 0.817 \times (1 - 0.817)$
$= 0.0017$

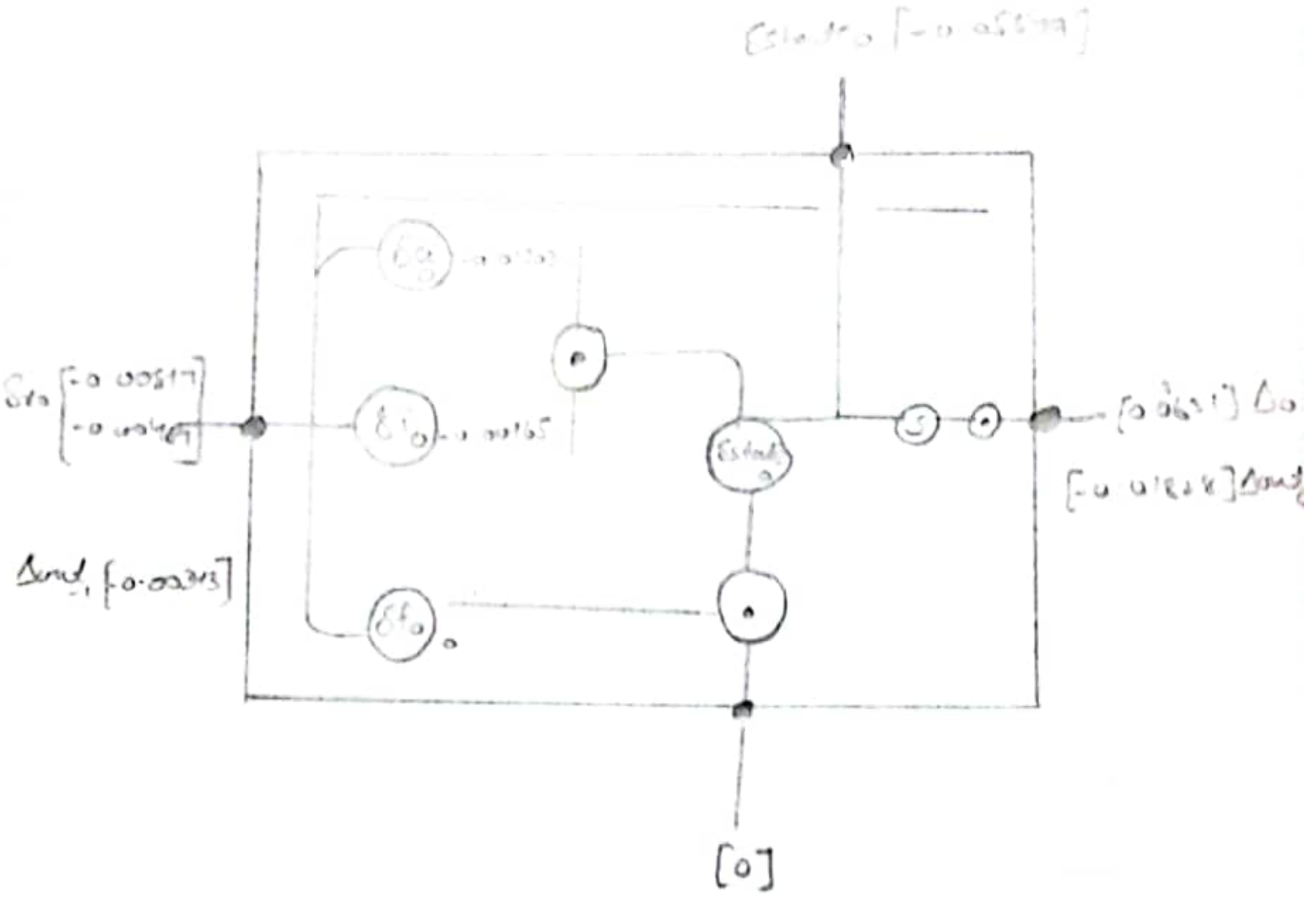
$$\delta x_0 = W^T \cdot \delta gates_0$$

$$= \begin{bmatrix} 0.45 & 0.95 & 0.70 & 0.60 \\ 0.25 & 0.80 & 0.45 & 0.40 \end{bmatrix} \begin{bmatrix} -0.0170 \\ -0.00165 \\ 0 \\ 0.0017 \end{bmatrix} = \begin{bmatrix} -0.00817 \\ -0.00487 \end{bmatrix}$$

$$\Delta out_{-1} = U^T \cdot \delta gates_1$$

$$= \begin{bmatrix} 0.15 & 0.80 & 0.10 & 0.25 \end{bmatrix} \begin{bmatrix} -0.0170 \\ -0.00165 \\ 0 \\ 0.0017 \end{bmatrix} = -0.00343$$

Scanned with CamScanner




Scanned with CamScanner

$$\delta W = \sum_{t=0}^{T-1} \delta gates_t \otimes x_t$$

$$= \begin{bmatrix} -0.01703 \\ -0.00165 \\ 0 \\ 0.00176 \end{bmatrix} [1.0 \quad 2.0] + \begin{bmatrix} -0.01963 \\ -0.00112 \\ -0.00631 \\ -0.05538 \end{bmatrix} [0.5 \quad 3.0] =$$

$$\begin{bmatrix} -0.02672 & - & 0.0922 \\ -0.00221 & - & 0.00666 \\ -0.00316 & - & 0.01893 \\ -0.02593 & - & 0.16262 \end{bmatrix}$$

$$\delta U = \sum_{t=0}^{T-1} \delta gates_{t+1} \otimes out_t$$

$$= \begin{bmatrix} -0.01938 \\ -0.00112 \\ -0.00631 \\ -0.05538 \end{bmatrix} [0.53631] = \begin{bmatrix} -0.01039 \\ -0.00060 \\ -0.00338 \\ -0.02970 \end{bmatrix}$$

$$\delta b = \sum_{b=0}^{T} \delta gates_{t+1}$$

$$= \begin{bmatrix} -0.01703 \\ -0.00165 \\ 0 \\ 0.00176 \end{bmatrix} + \begin{bmatrix} -0.01938 \\ -0.00112 \\ -0.00631 \\ -0.05538 \end{bmatrix} = \begin{bmatrix} -0.03641 \\ -0.00277 \\ -0.00631 \\ -0.05362 \end{bmatrix}$$

And updating our Parameters based on the SGD update function.

$$W^{new} = W^{old} - \lambda * \delta W^{old}$$

Scanned with CamScanner

$$W_a = \begin{bmatrix} 0.45267 \\ 0.25922 \end{bmatrix}, \quad U_a = [0.15104], \quad b_a = [0.20364]$$

$$W_i = \begin{bmatrix} 0.95022 \\ 0.80067 \end{bmatrix}, \quad U_i = [0.8006], \quad b_i = [0.65028]$$

$$W_f = \begin{bmatrix} 0.76031 \\ 0.45189 \end{bmatrix}, \quad U_f = [0.10034], \quad b_f = [0.15063]$$

$$W_o = \begin{bmatrix} 0.60259 \\ 0.41629 \end{bmatrix}, \quad U_o = [0.25297], \quad b_o = [0.10536]$$

Scanned with CamScanner