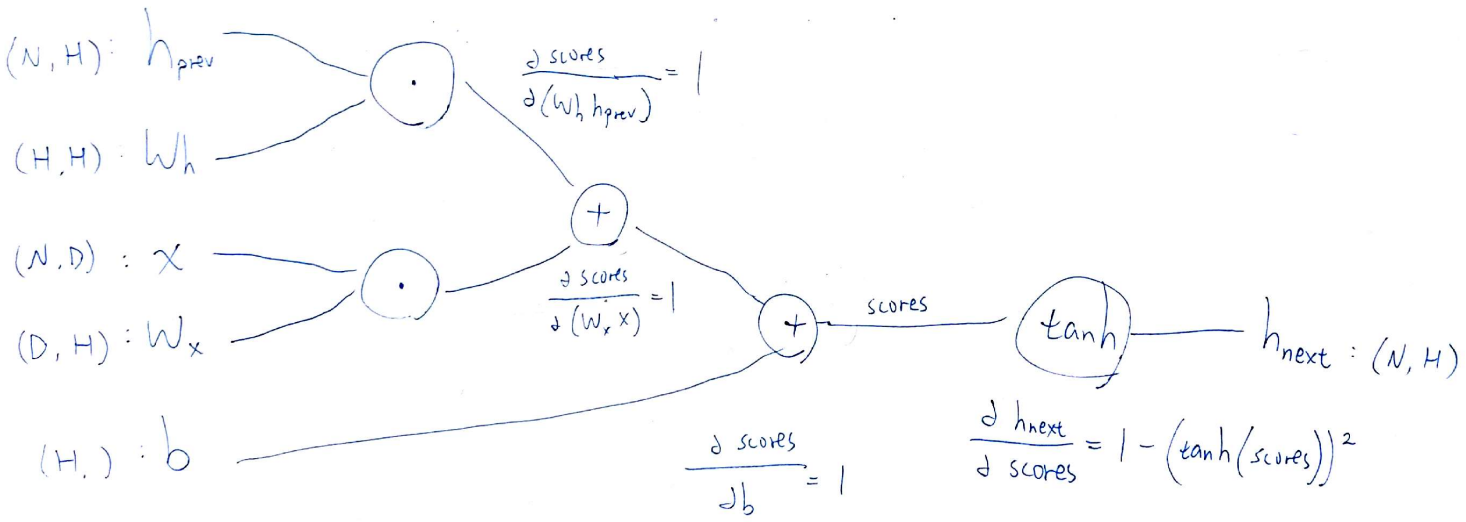
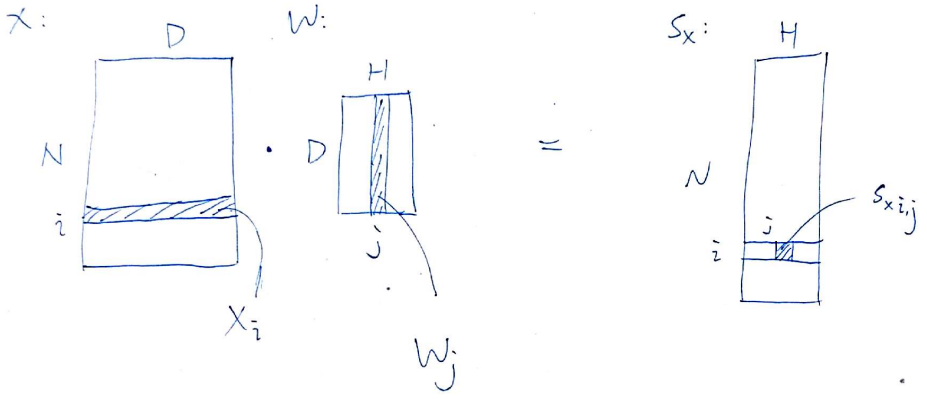
$$h_{next} = \tanh\left(W_h \, h_{prev} + W_x \, x + b \right)$$



$W x = S_x$



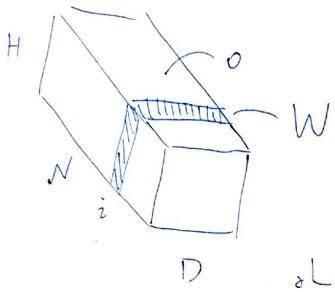
$$X_i \cdot W_j = S_{x\,i,j}$$

$$\frac{\partial S_{x\,ij}}{\partial X_i} = W_j$$

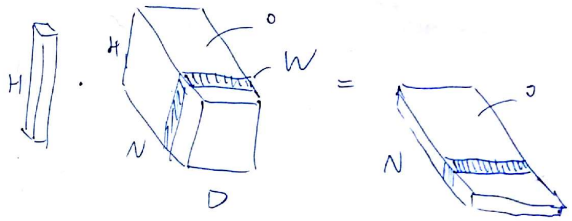
$$\frac{\partial S_{x\,ij}}{\partial W_j} = X_i$$

$$\frac{\partial S_{xi}}{\partial X_i} = W$$

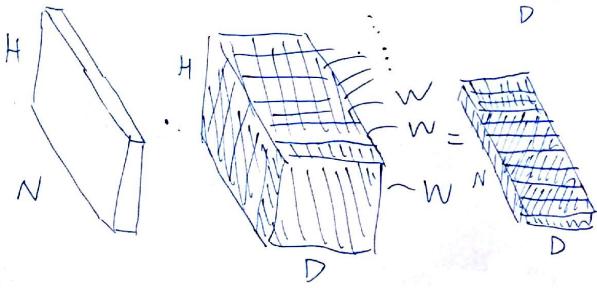
$$\frac{\partial S_{xi}}{\partial X} =$$



$$\frac{\partial L}{\partial S_{xi}} \cdot \frac{\partial S_{xi}}{\partial X} =$$

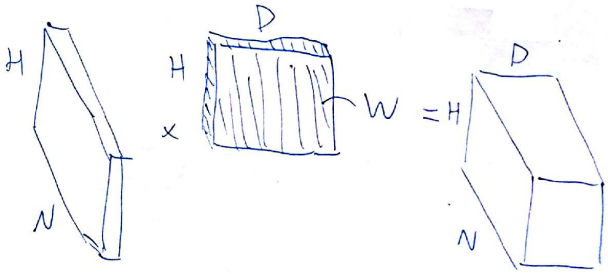


$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial S_x} \cdot \frac{\partial S_x}{\partial X} =$$

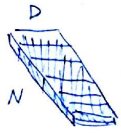


Alternatively,

$$\frac{\partial L}{\partial S_x} \times W =$$



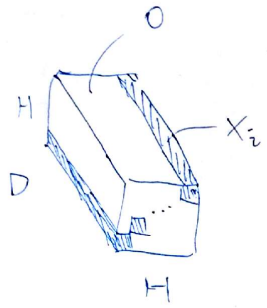
$$\sum_{\text{along axis H}} \left(\frac{\partial L}{\partial S_x} \times W \right) = \sum_{\text{along axis H}} \left(\right) =$$



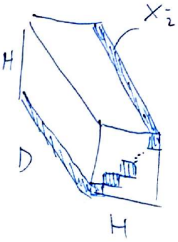
$$= \frac{\partial L}{\partial X}$$

$$\frac{\partial S_{x_{ij}}}{\partial W_j} = X_i$$

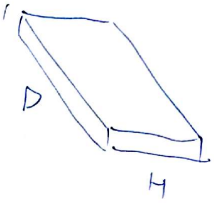
$$\frac{\partial S_{x_i}}{\partial W} =$$



$$\frac{\partial L}{\partial S_{x_i}} \cdot \frac{\partial S_{x_i}}{\partial W} = H \cdot$$

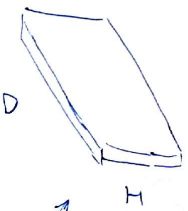


$=$

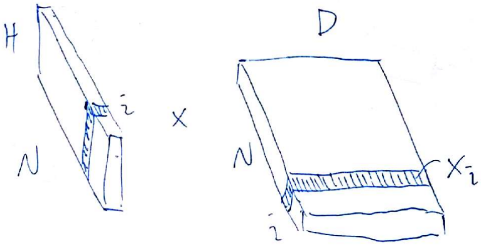


$= H \times$

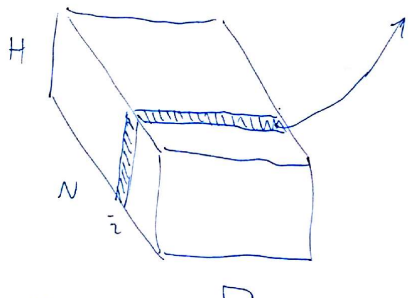
$=$



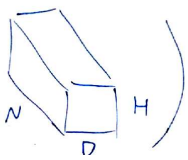
$$\frac{\partial L}{\partial S_x} \times X =$$



$=$



$$\frac{\partial L}{\partial W} = \text{Sum over axis } N \left(\frac{\partial L}{\partial S_x} \times X \right) = \text{Sum over axis } N \left(\right.$$



$\left. \right) =$