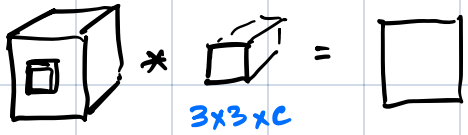
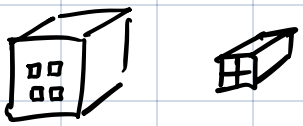
1. Convolution

1. Standard convolution



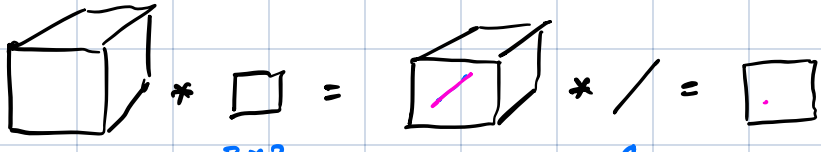
2. Padding, Stride

3. Dilated

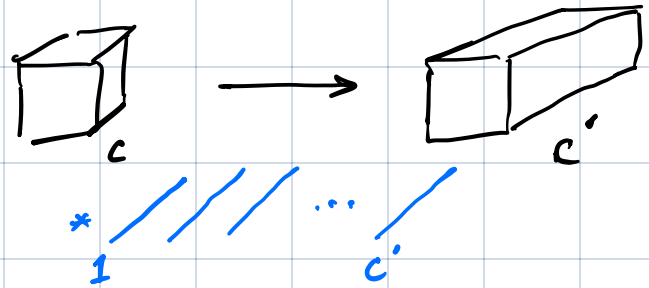


a. Receptive field

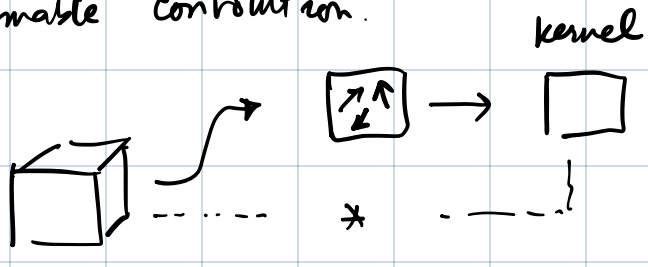
4. Point-wise / Separable convolution



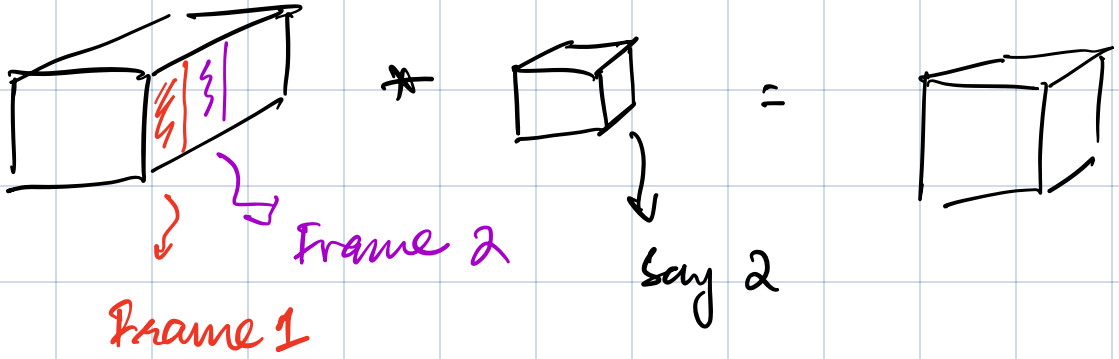
5.



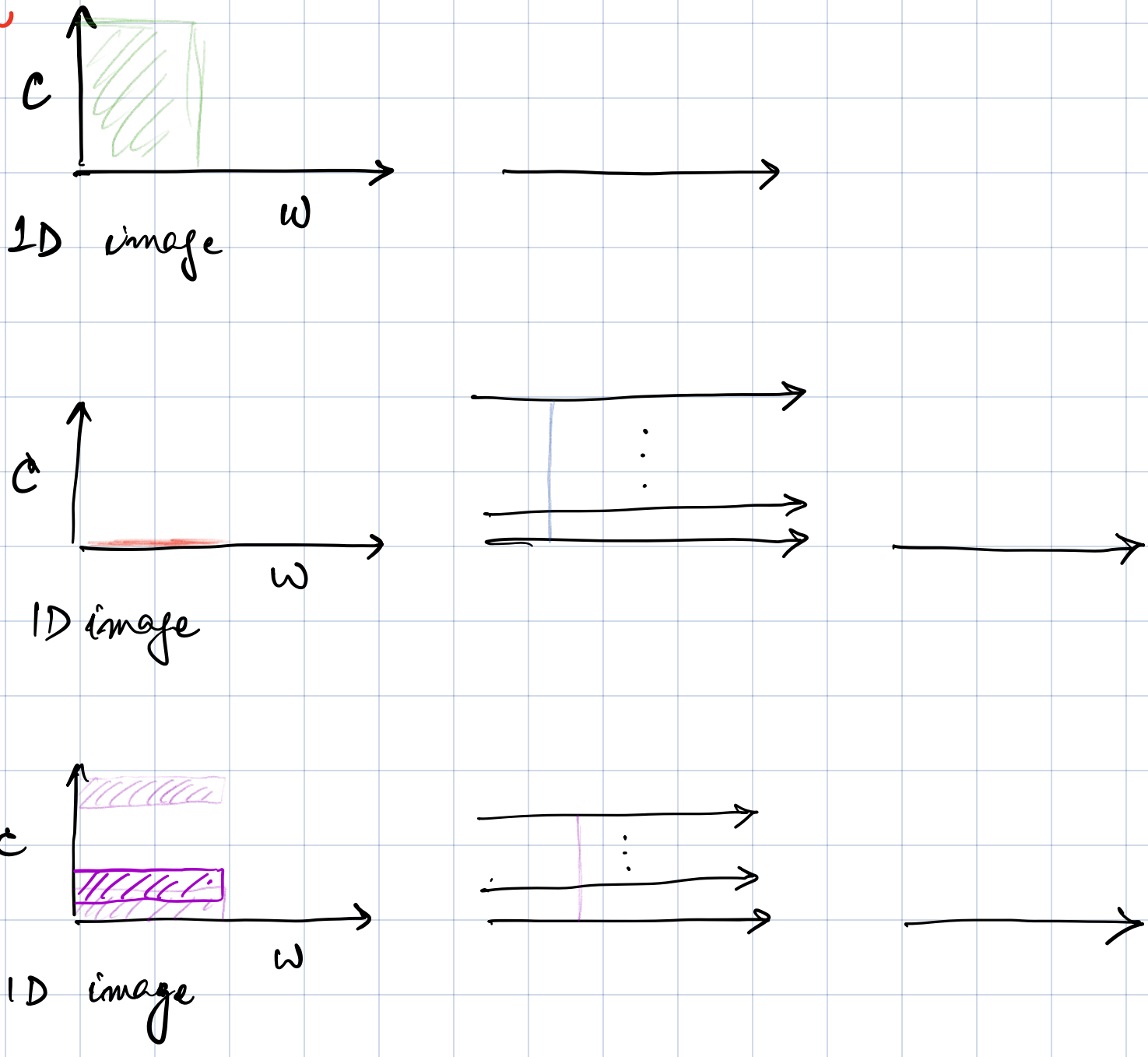
6. Deformable convolution



7. 3D convolution



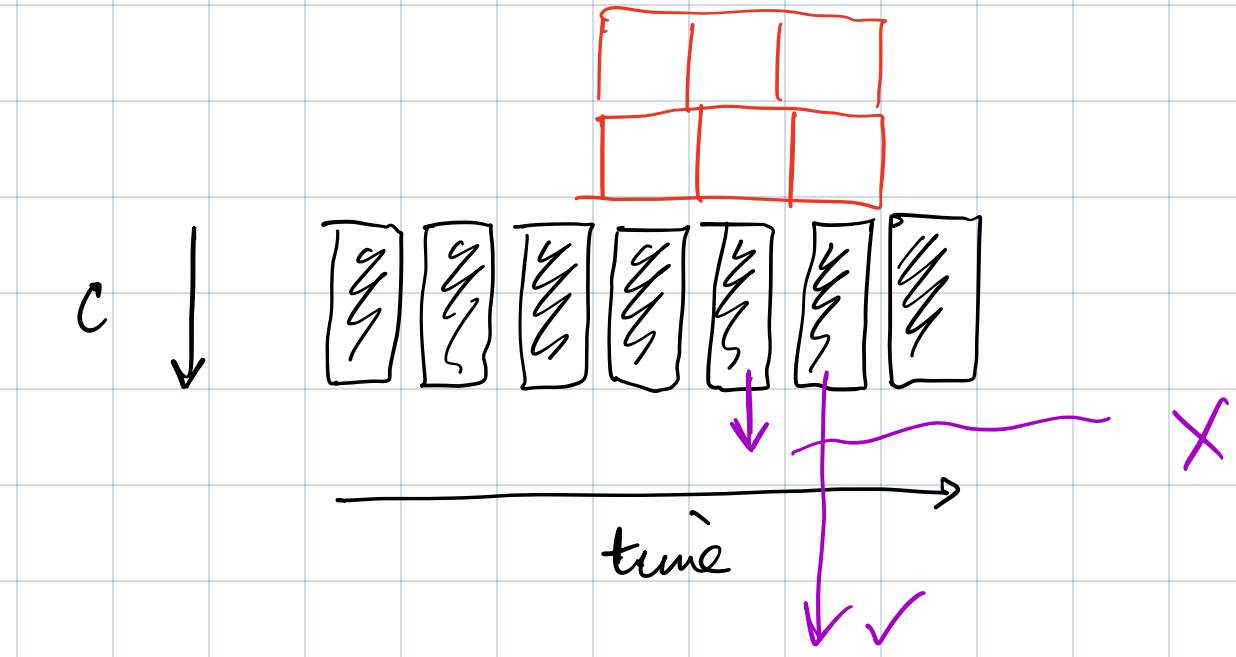
Aside

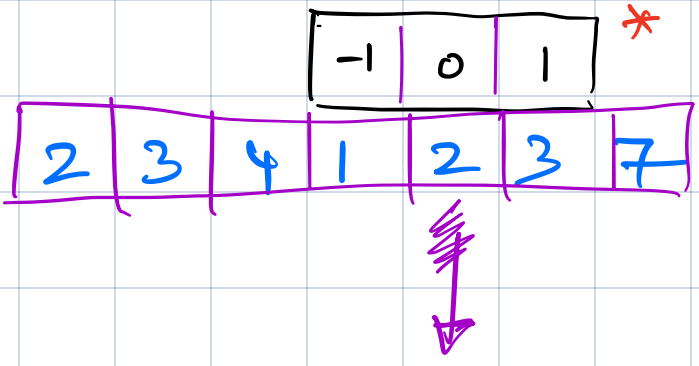


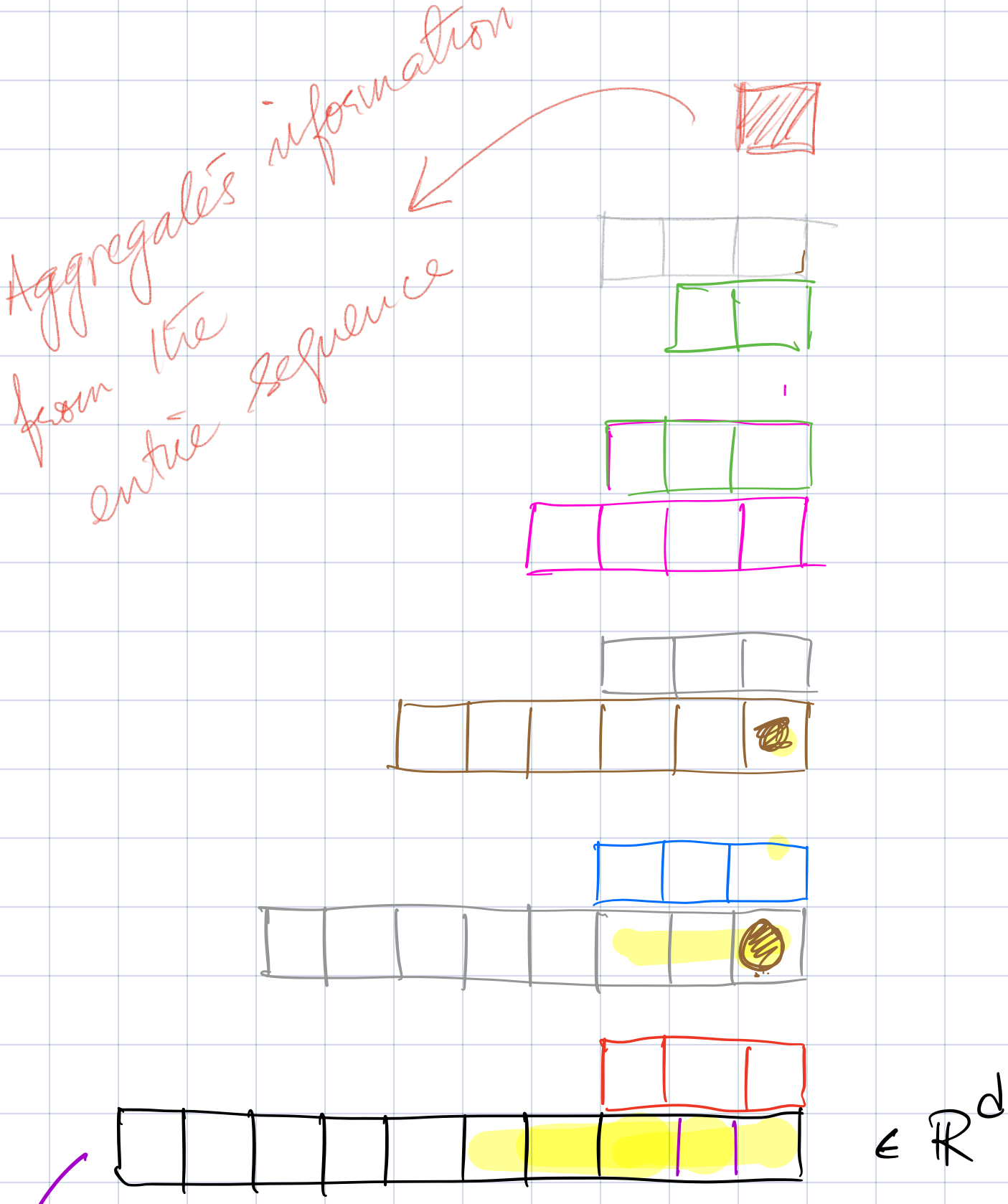
8. Temporal convolutional network

$$\vec{x}_{t-k+1} \quad \cdots \quad \vec{x}_{t-2} \quad \vec{x}_{t-1} \quad \vec{x}_t \in \mathbb{R}^d$$

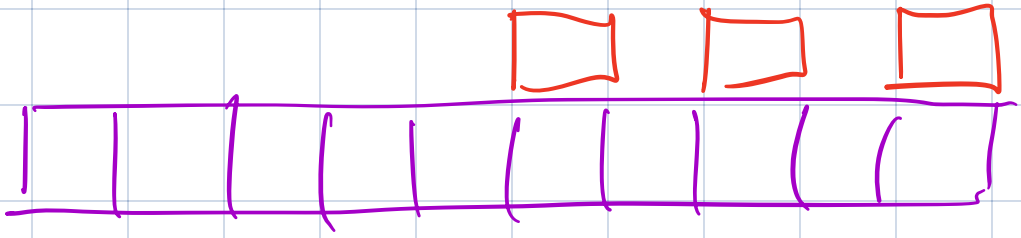
Example: say $d=2$, $k=7$



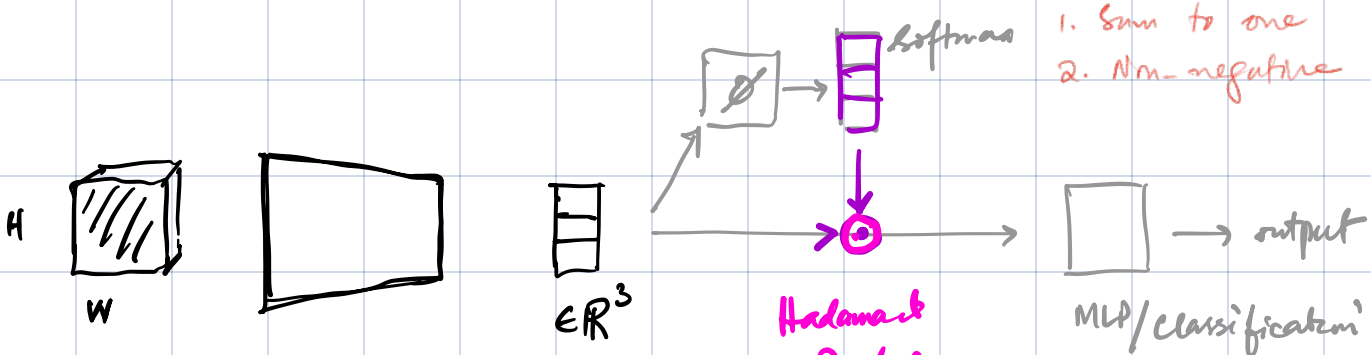


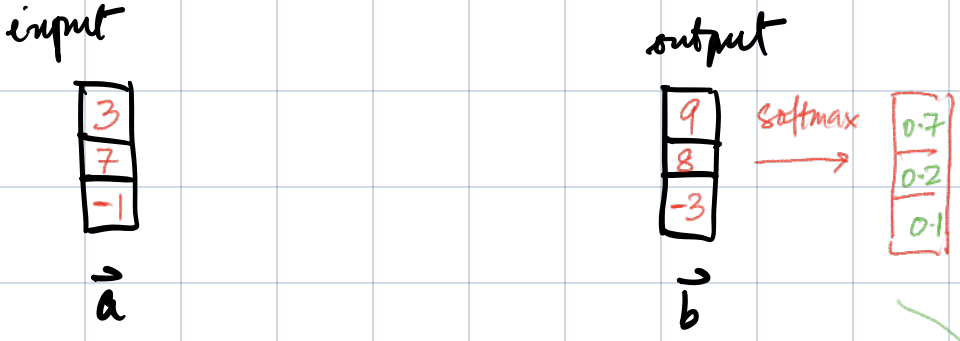

Aggregates information from the entire sequence
$\in \mathbb{R}^d$

increase the receptive field - dilation.



Attention: "All you need is attention" → transformers.

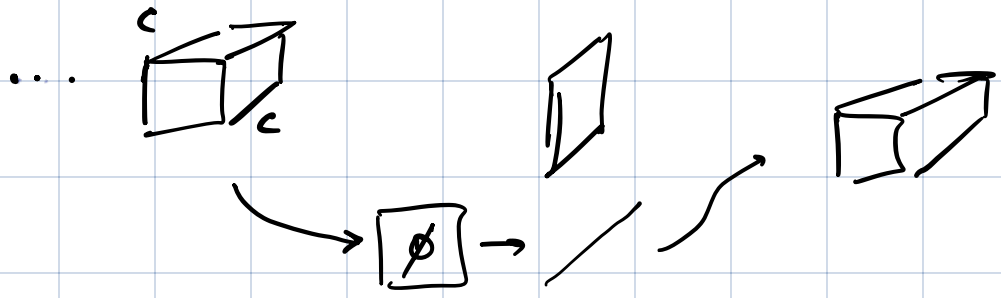


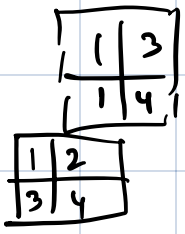


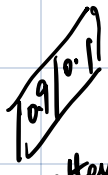
$$\frac{e^{-9}}{e^{-9-8+3}}$$

$$\vec{b} = \text{relu}\left(\left[3\times 3\right]\vec{a}\right) \in \mathbb{R}^3$$

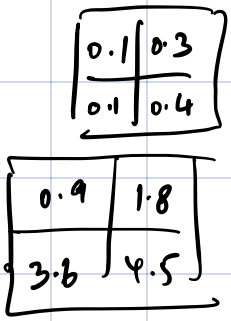
$$\odot \begin{bmatrix} (3)(0.7) \\ (7)(0.2) \\ (-1)(0.1) \end{bmatrix} = \begin{bmatrix} 2.1 \\ 1.4 \\ 0.01 \end{bmatrix}$$

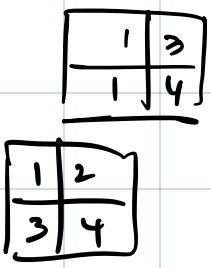




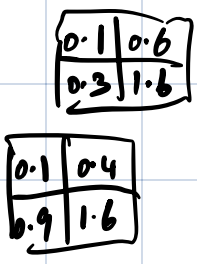


attention channel-wise

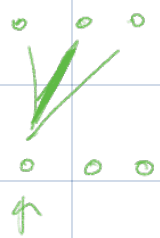


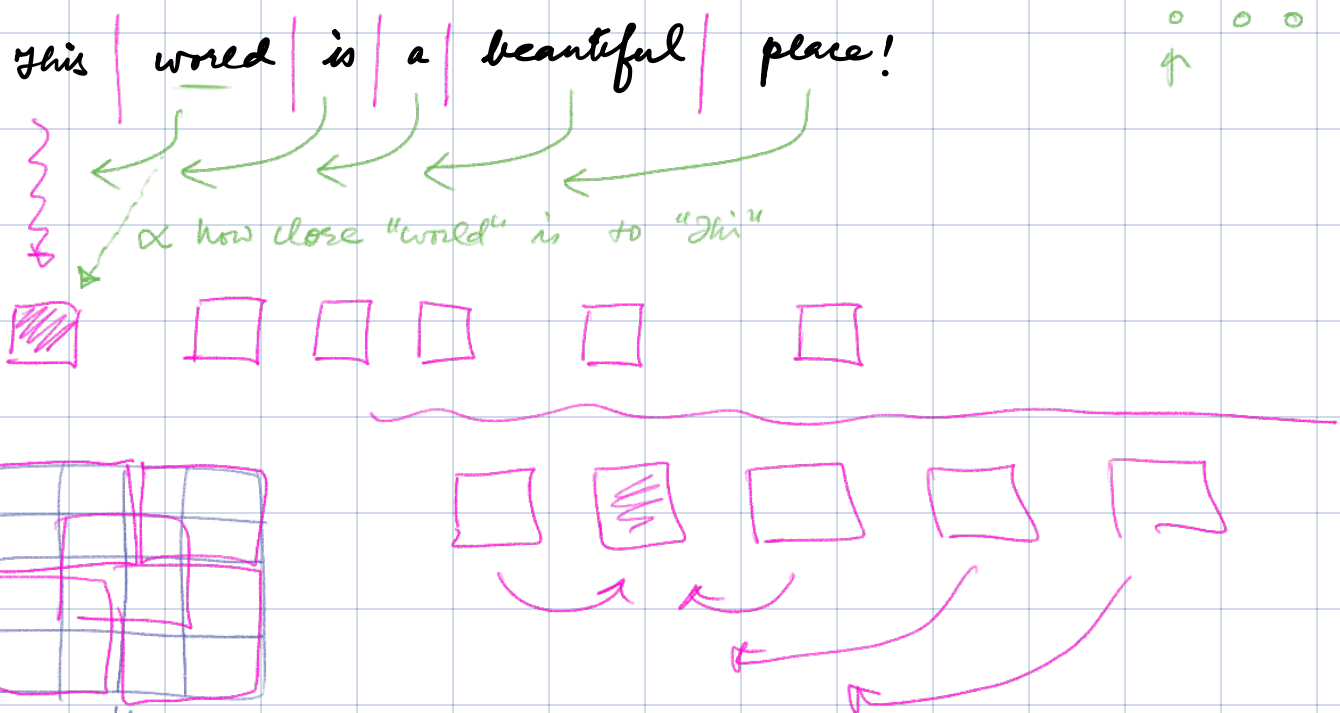


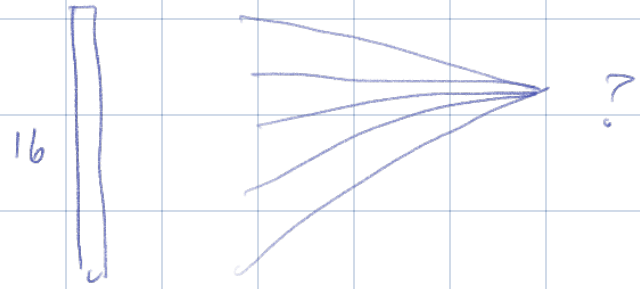
attention spatial



Self-attention:







positional encoding.