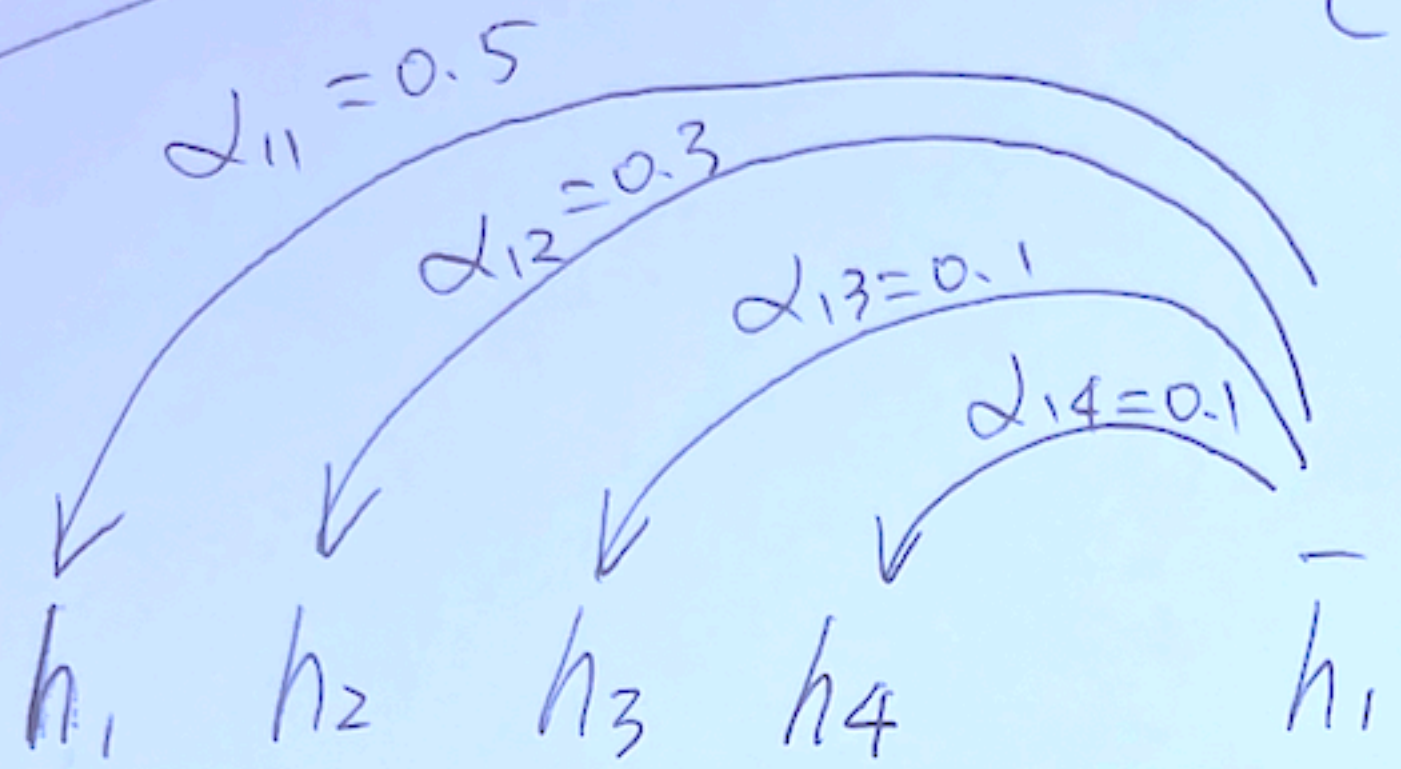


Attention

$$C_1 = 0.5h_1 + 0.3h_2 + 0.1h_3 + 0.1h_4.$$



$$\alpha_{ij} = \text{softmax}(e_{ij})$$

the movie was great
input (encoder)

output (decoder) e.g. dot product: $\bar{h}_i \cdot h_j$

$$\Rightarrow \text{softmax}(W\bar{h}_i)$$
$$\Rightarrow \text{softmax}(W[C_i = \bar{h}_i])$$

↑ ↑
concatenation