Attention Mechanism

Decoder State　　　**Context**

$$s_t \longrightarrow c_t$$

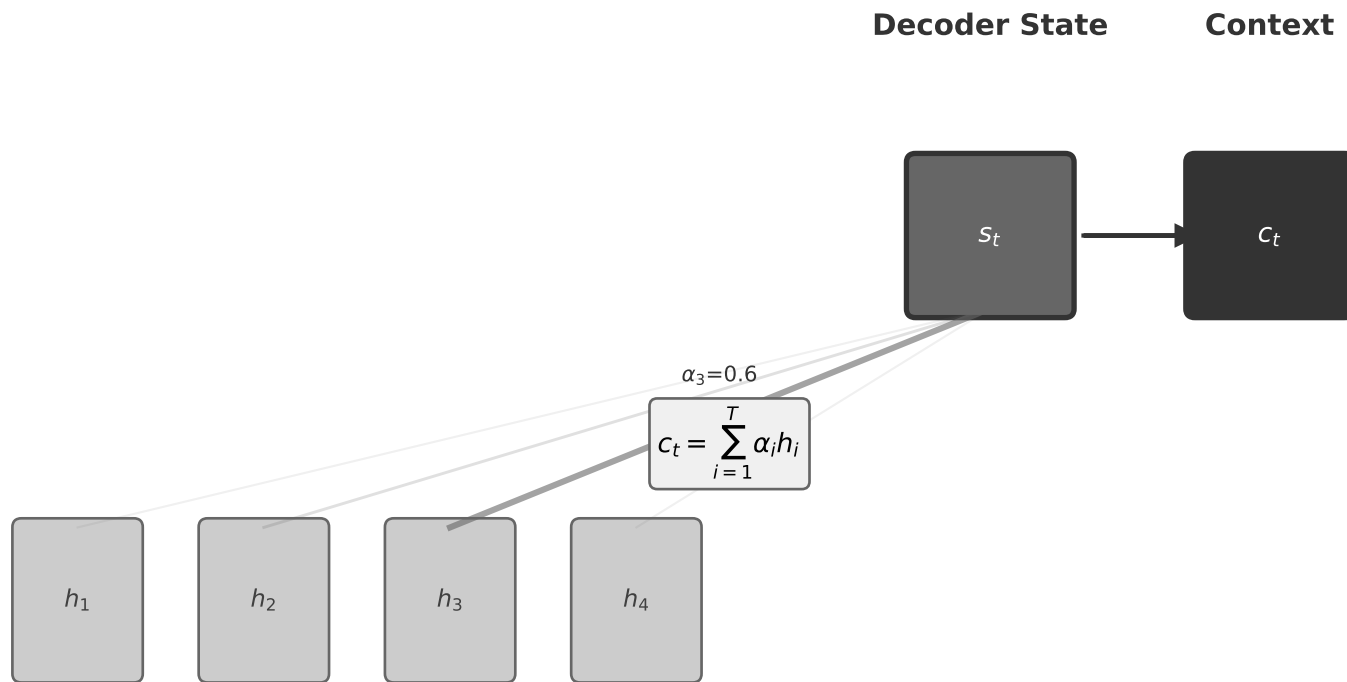$\alpha_3 = 0.6$

$$c_t = \sum_{i=1}^{T} \alpha_i h_i$$

h_1　　h_2　　h_3　　h_4

Encoder Hidden States *Weighted sum of encoder states based on relevance*