



$$\alpha_t(s) = \frac{\exp(\text{score}(h_t, e_s))}{\sum_{s'} \exp(\text{score}(h_t, e_{s'}))}$$

attention weights

$$C_t = \sum_s \alpha_t(s) e_s$$

context vector

