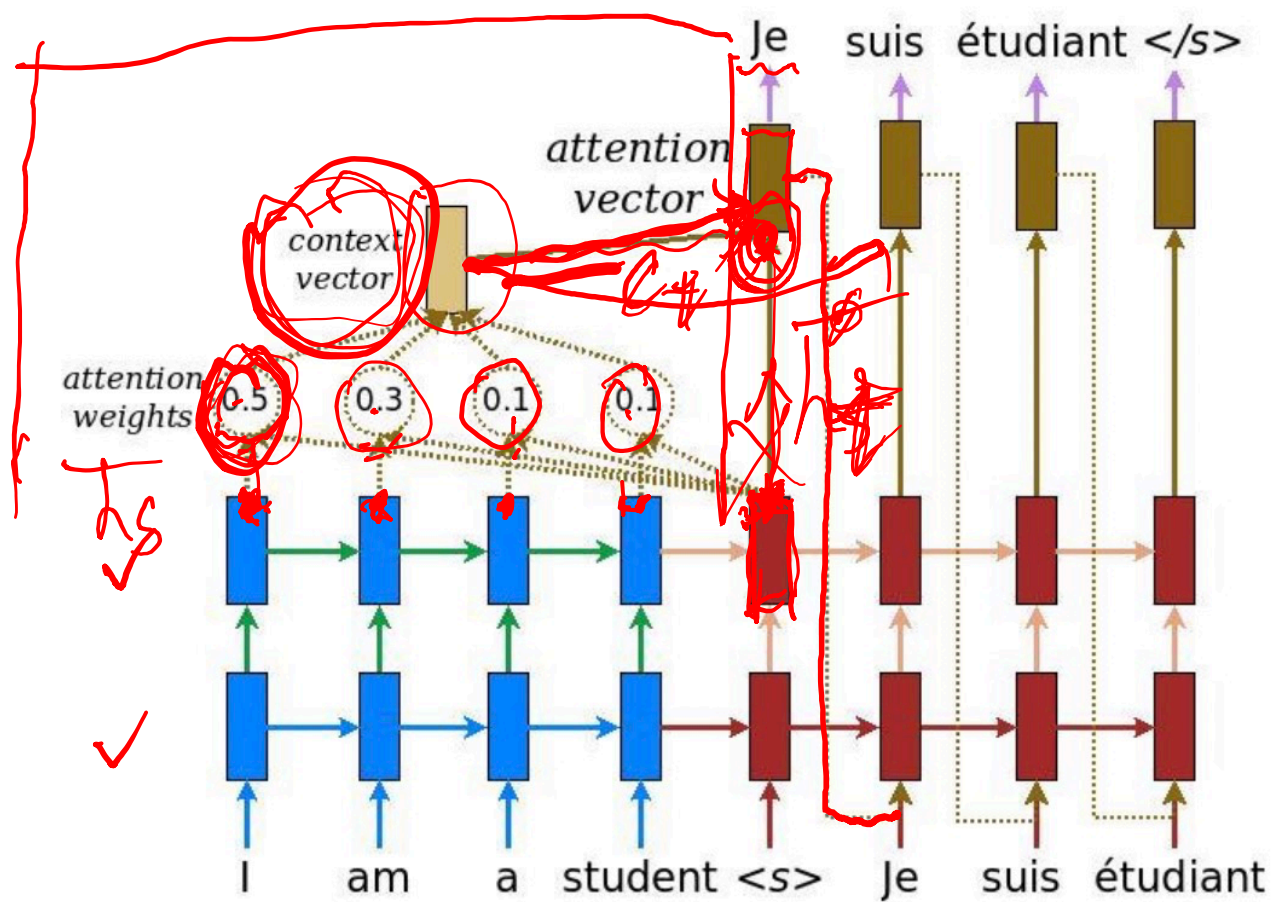


Attention机制



$$C_t = \sum_s \alpha_{ts} \bar{h}_s$$

$$\alpha_t = f(C_t, h_t) = \text{tanh}(W_c [C_t; h_t])$$

$$\text{scores}(h_t, \bar{h}_s)$$

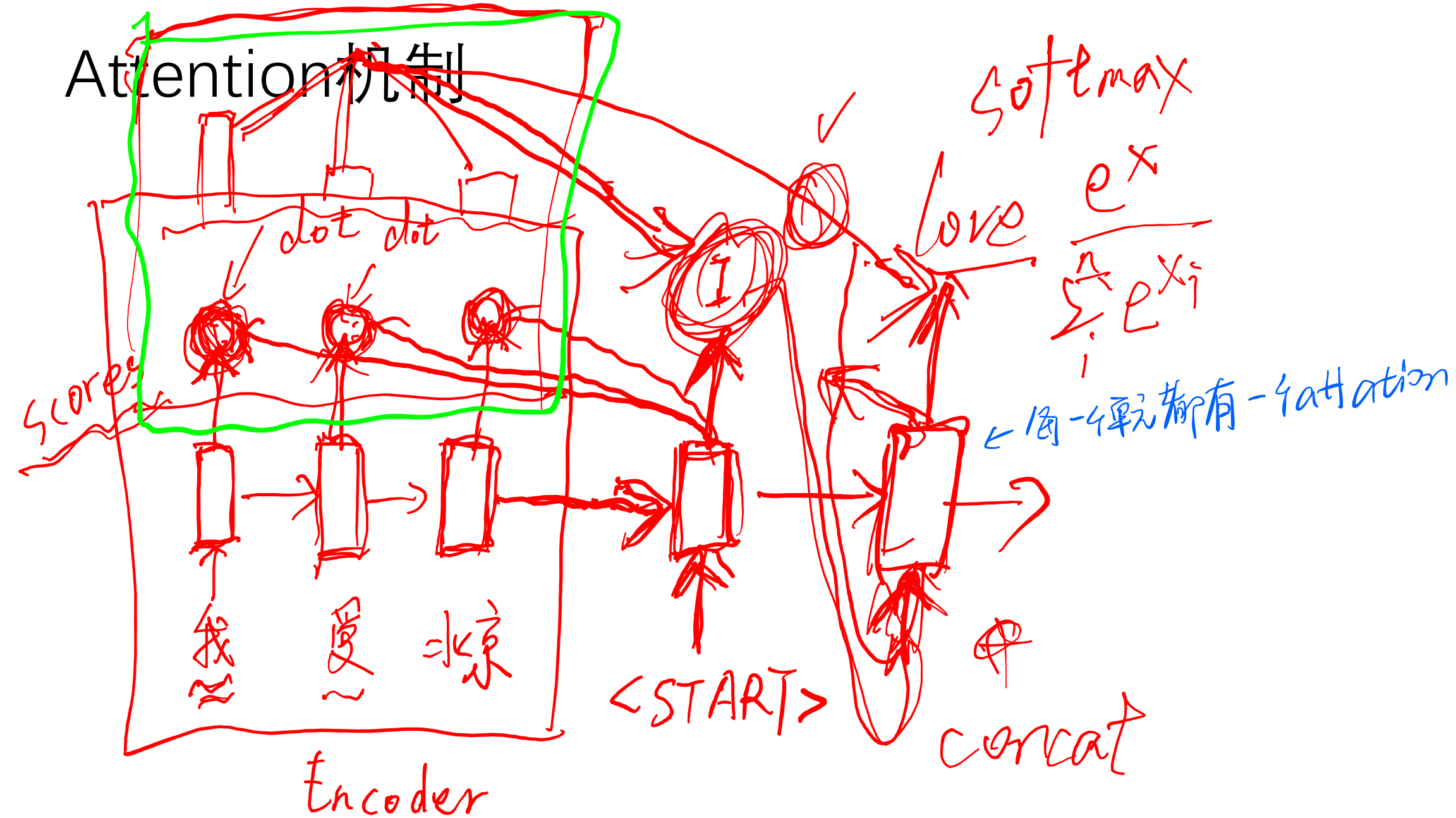
$$= \begin{cases} h_t^T W_h \bar{h}_s & \text{Luong's multiplicative style} \\ \text{tanh}(W_1 h_t + W_2 \bar{h}_s) & \text{Bahdanau's additive style} \end{cases}$$

Bahdanau's additive style

$$\alpha_{ts} = \frac{\exp(\text{scores})}{\sum_{s'=1}^S \exp(\text{scores})}$$

concat

Attention 机制



Attention机制