

Машинное обучение

Лекция 10

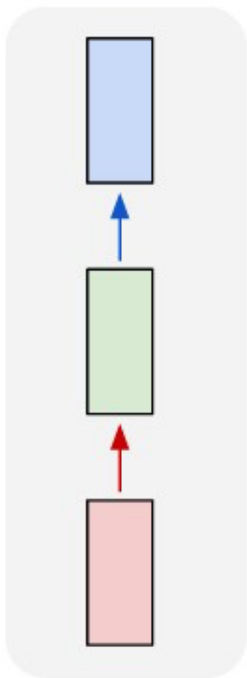
Seq2seq learning, Attention

Власов Кирилл Вячеславович

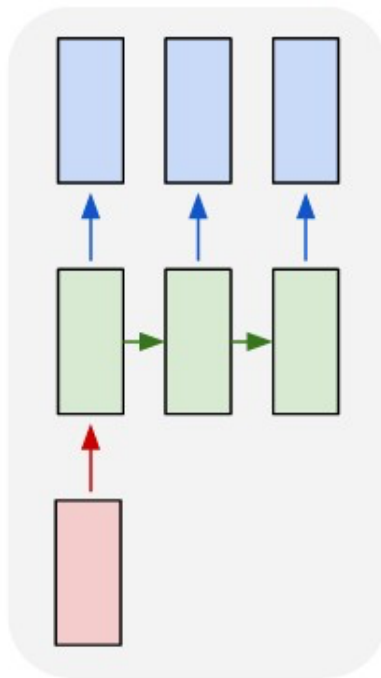


Обработка последовательностей

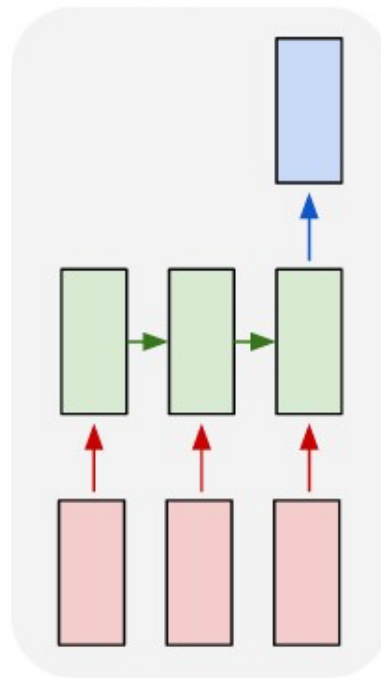
one to one



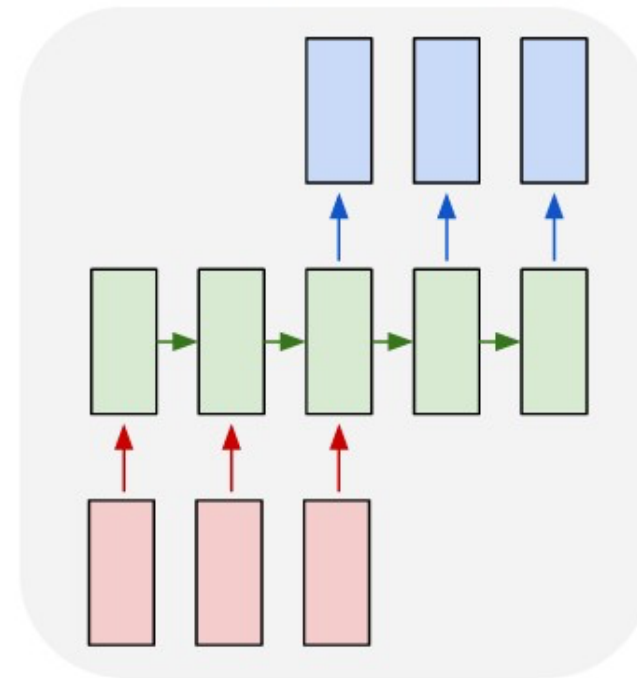
one to many



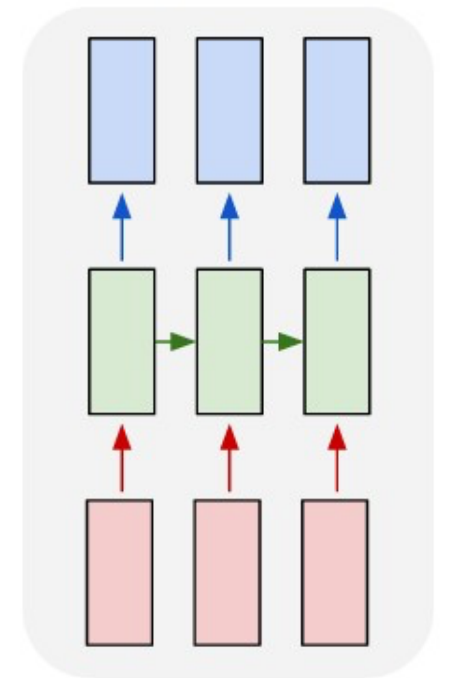
many to one



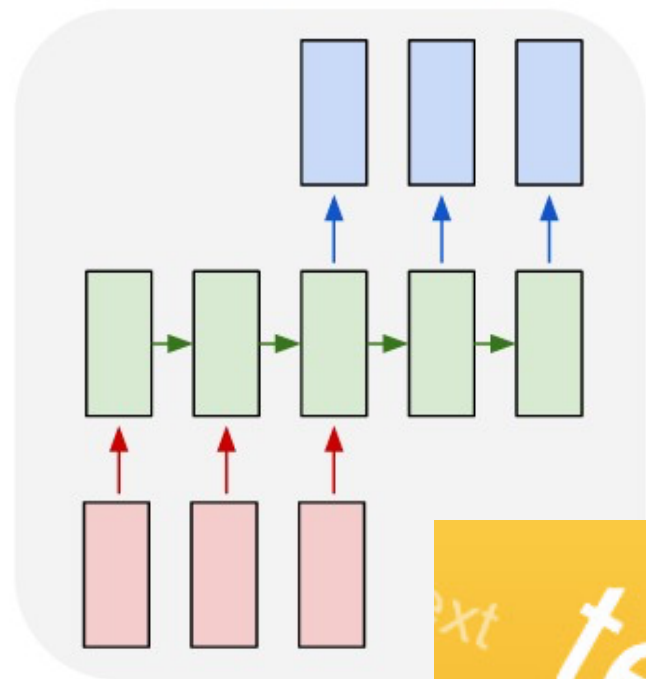
many to many



many to many

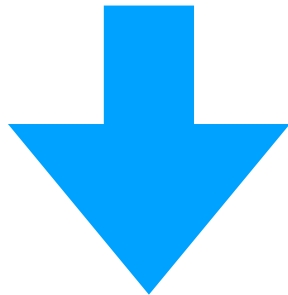


Машинный перевод

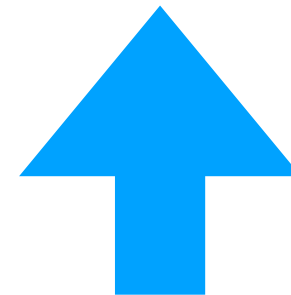


Классический подход к переводу

 *Je ne mangé pas six*
jours.

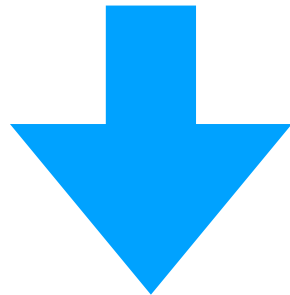


???



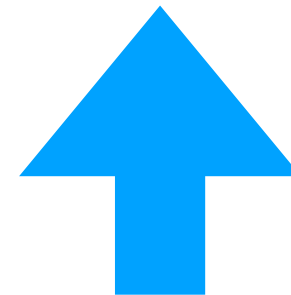
Классический подход к переводу

 *Je ne mangé pas six
jours.*



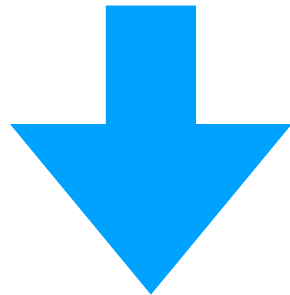
$$e_{best} = \operatorname{argmax}_e P(e | f)$$

*No como seis
días.* 

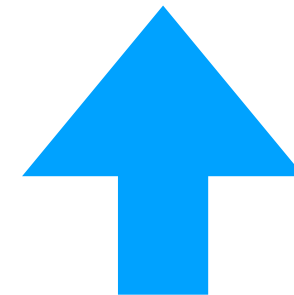


Классический подход к переводу

 *Je ne mangé pas six
jours.*



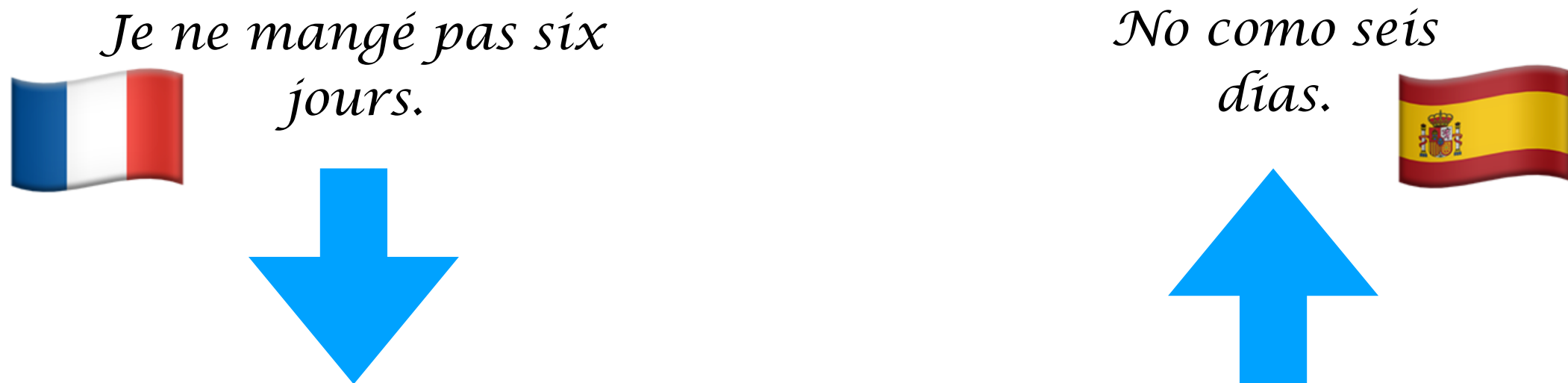
*No como seis
días.*



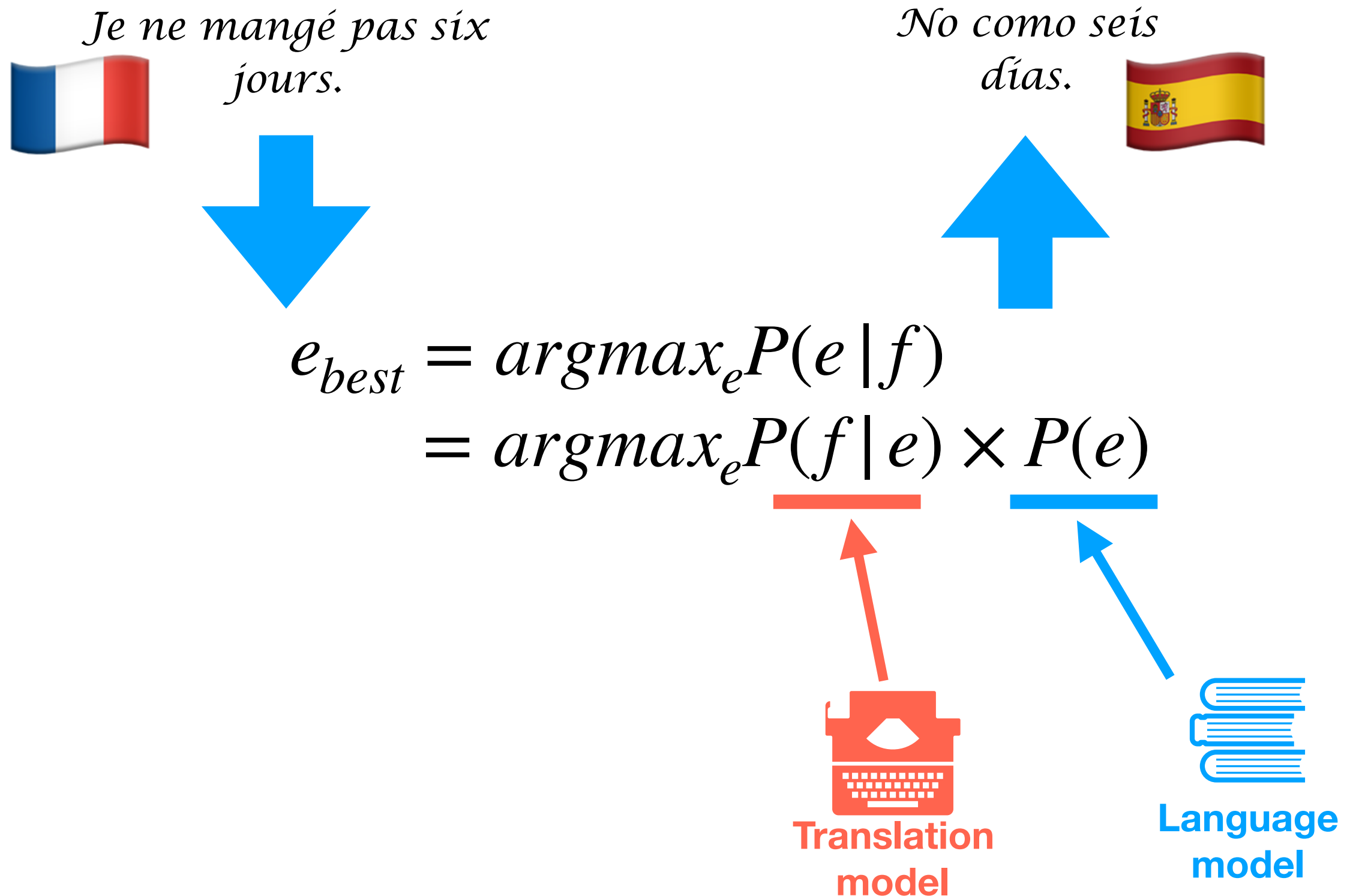
$$\begin{aligned} e_{best} &= \operatorname{argmax}_e P(e | f) \\ &= \operatorname{argmax}_e P(f | e) \times P(e) \end{aligned}$$



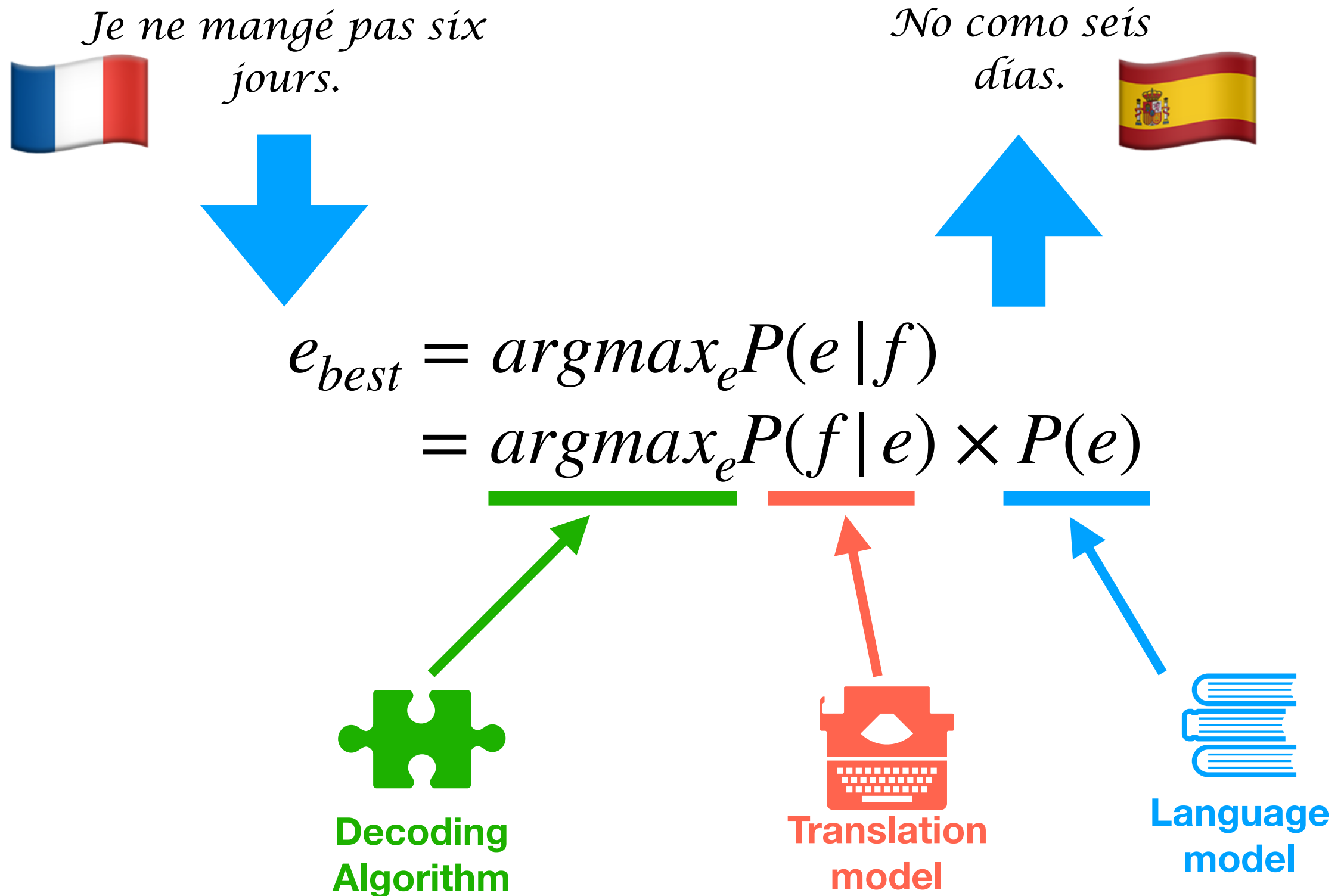
Классический подход к переводу



Классический подход к переводу



Классический подход к переводу



Проблемы

Je mangé



Como

Проблемы

Je mangé



Como

Je ne mangé pas



No como

Проблемы

Je mangé



Como

Je ne mangé pas



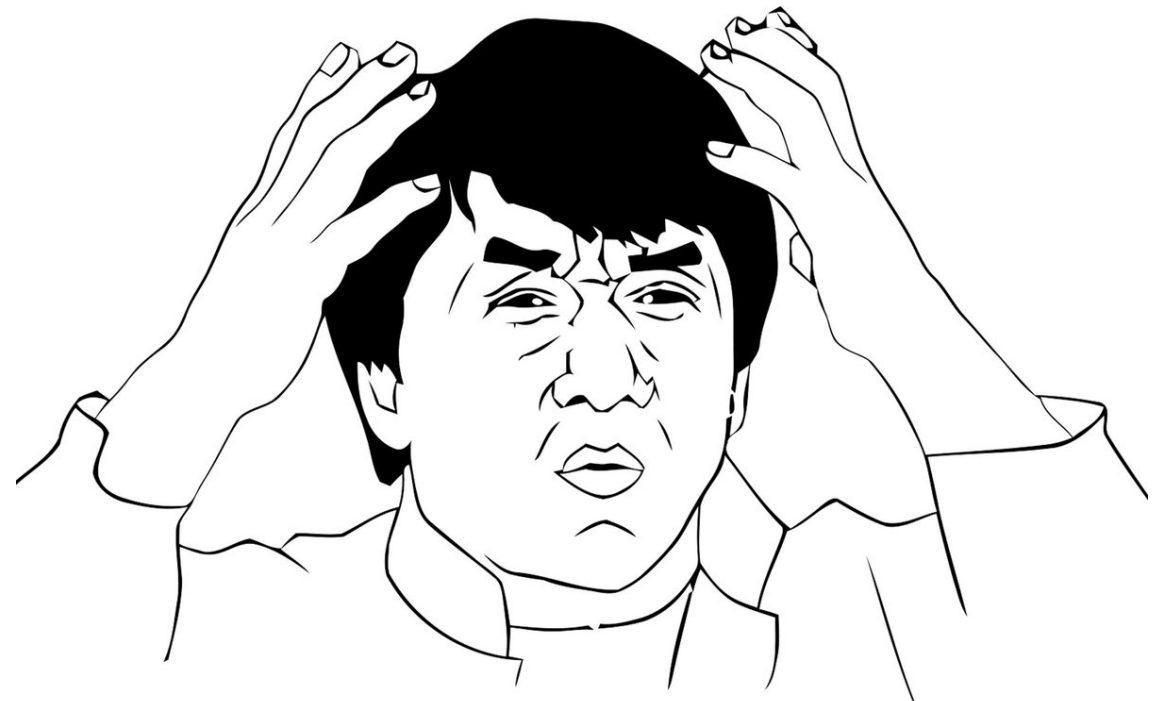
No como

СИНОНИМЫ

ИДИОМЫ

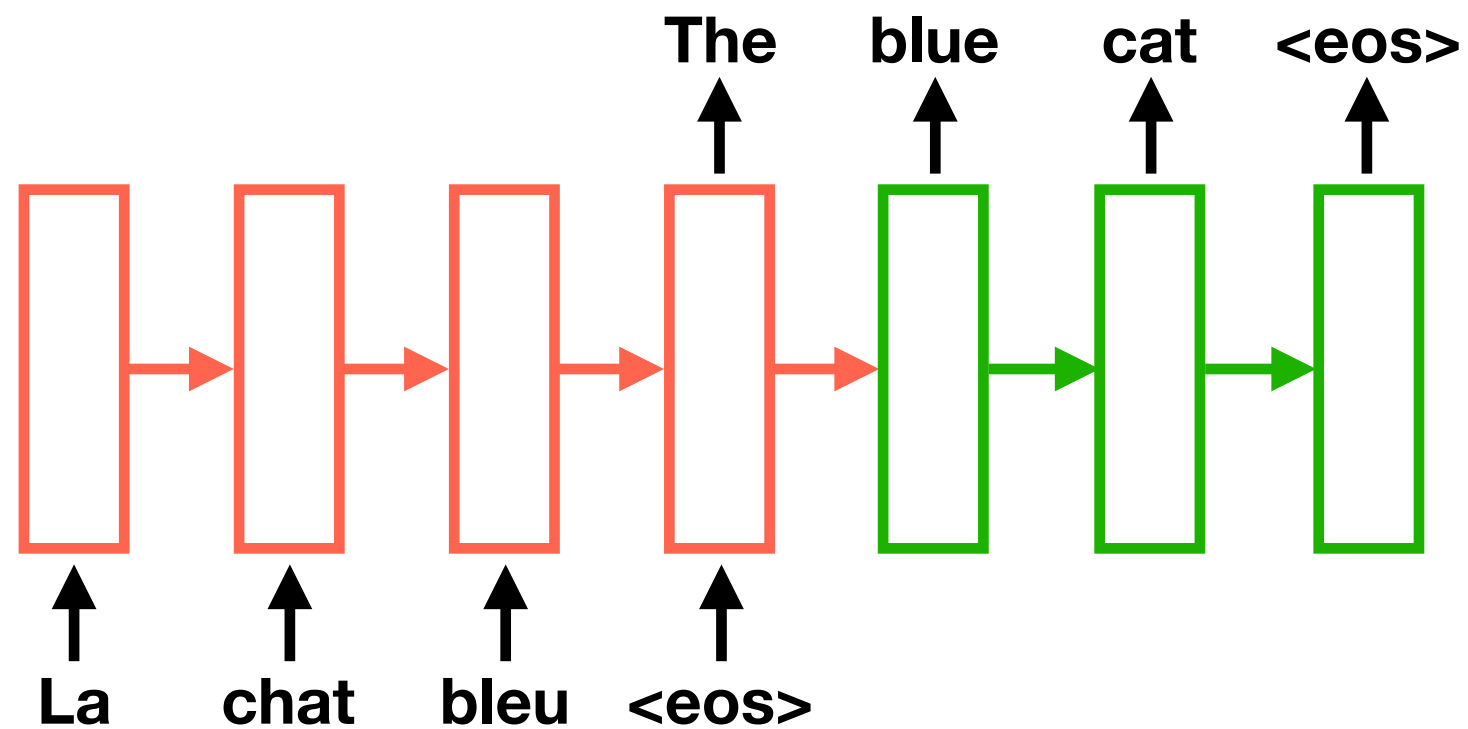
фразовые глаголы

не формальная речь

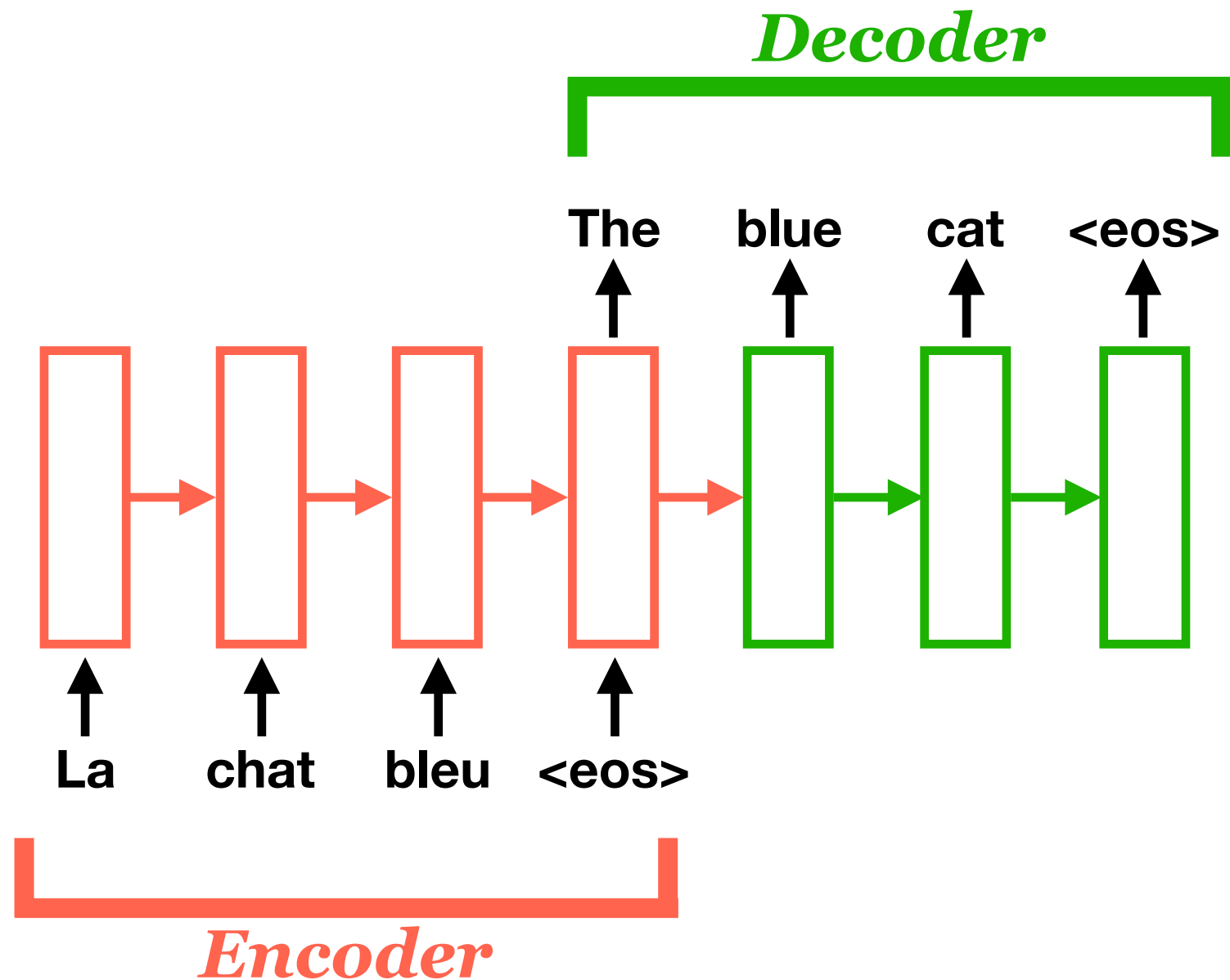


Решение: sequence to sequence

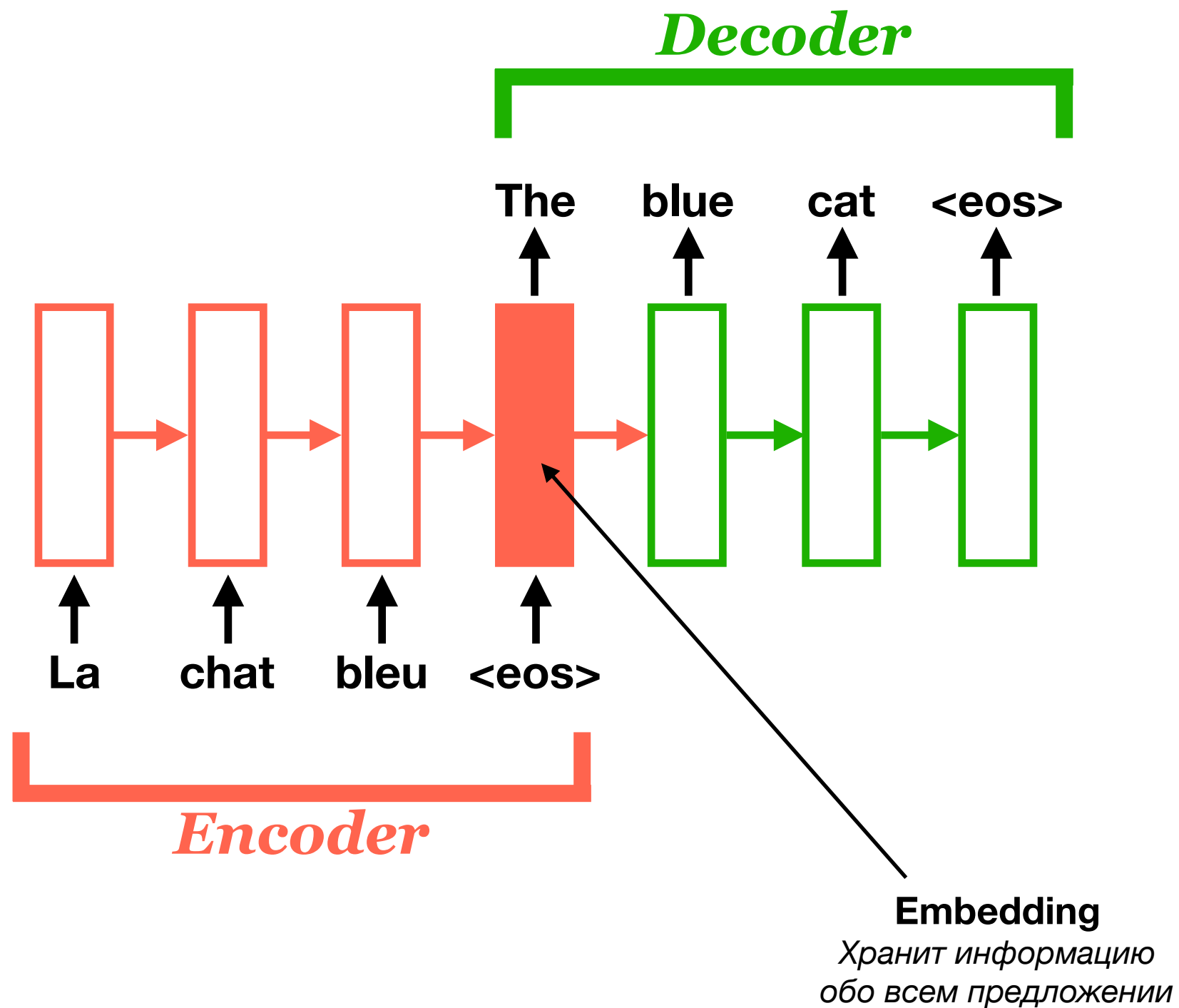
Sequence to sequence



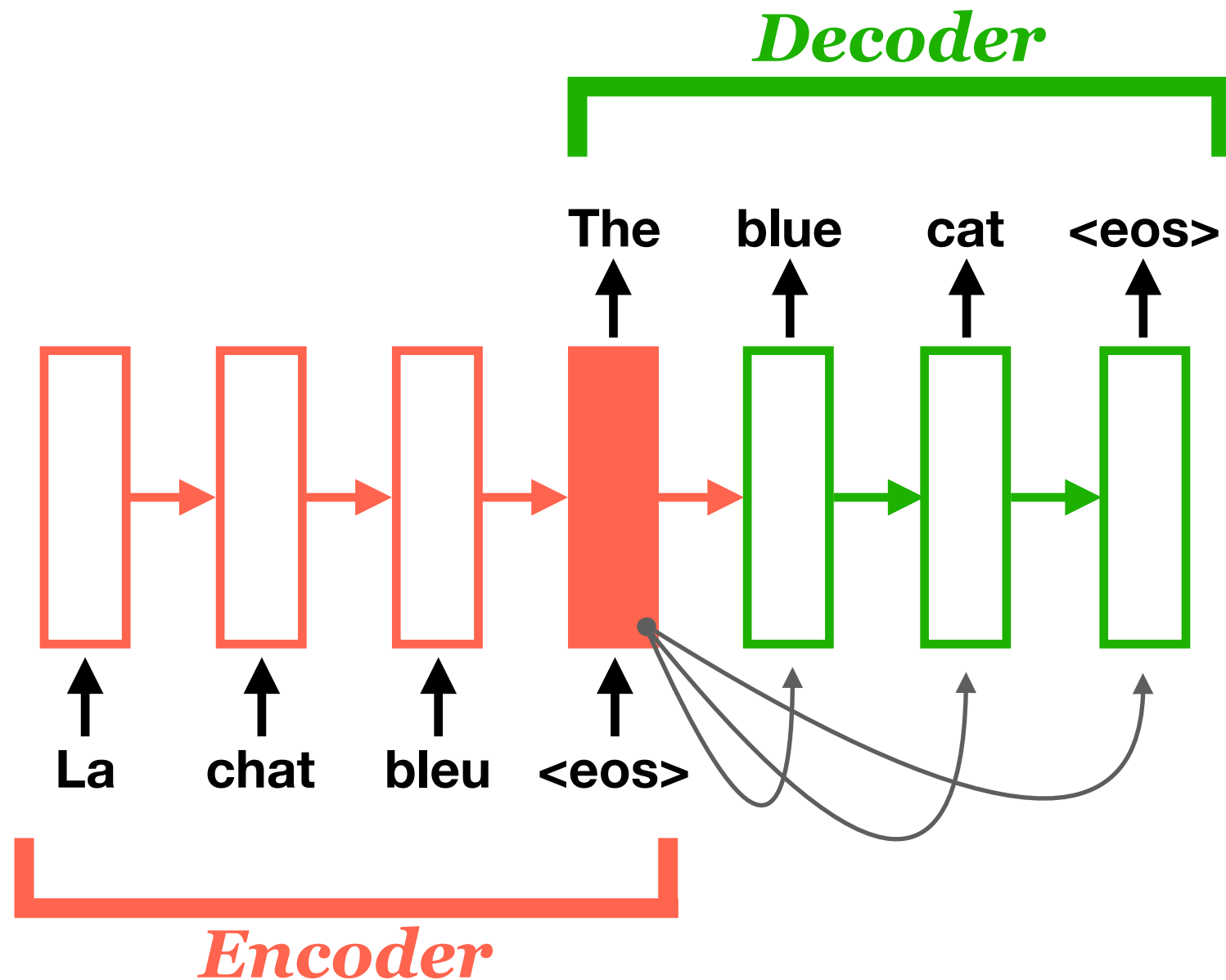
Sequence to sequence



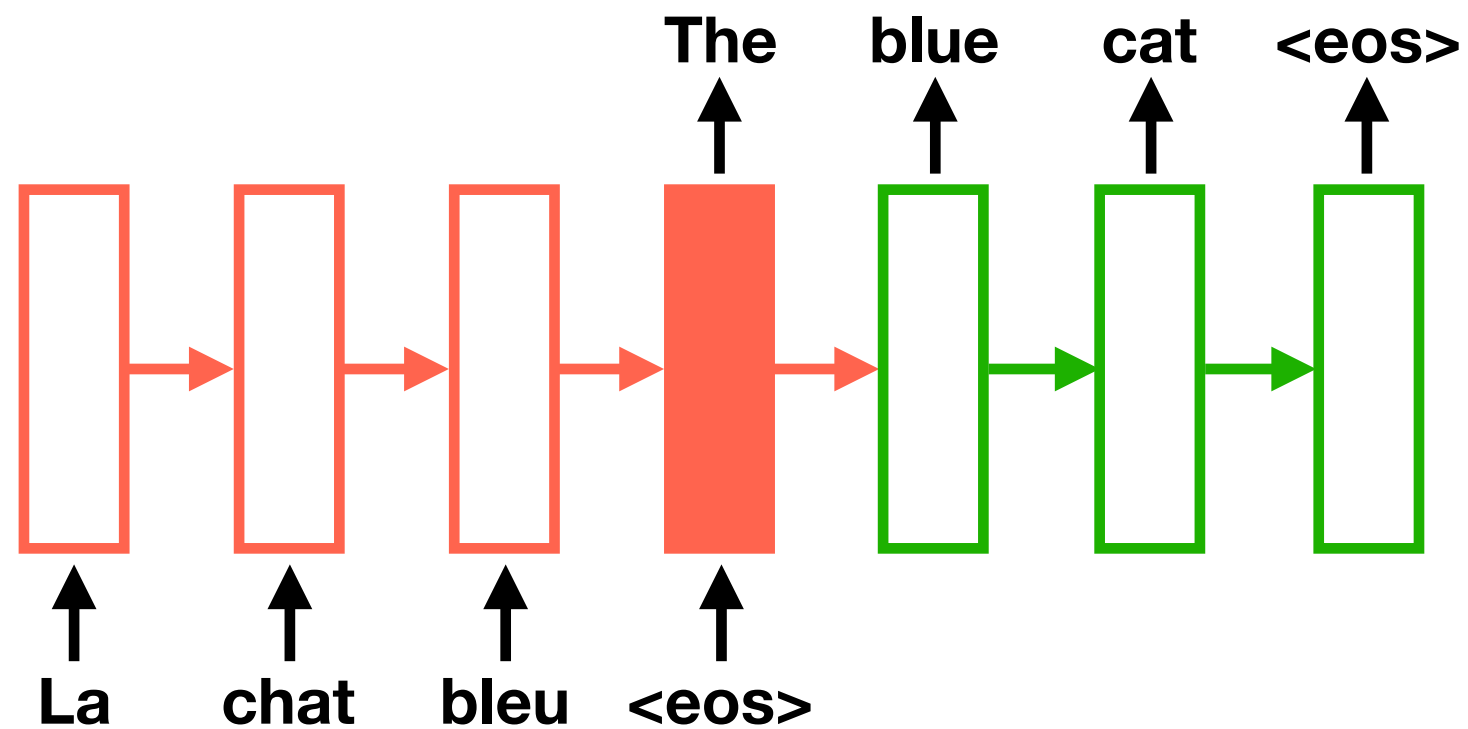
Sequence to sequence



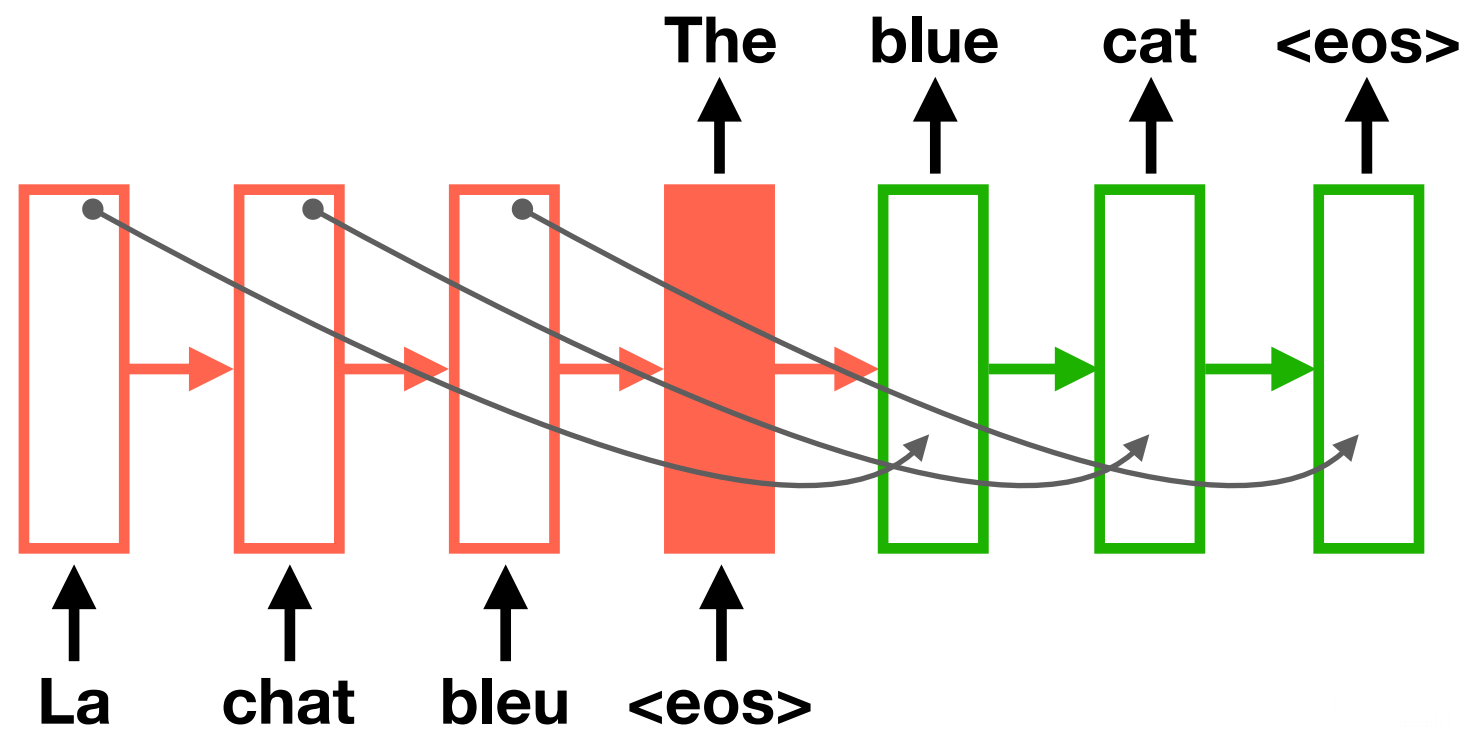
Sequence to sequence



Sequence to sequence

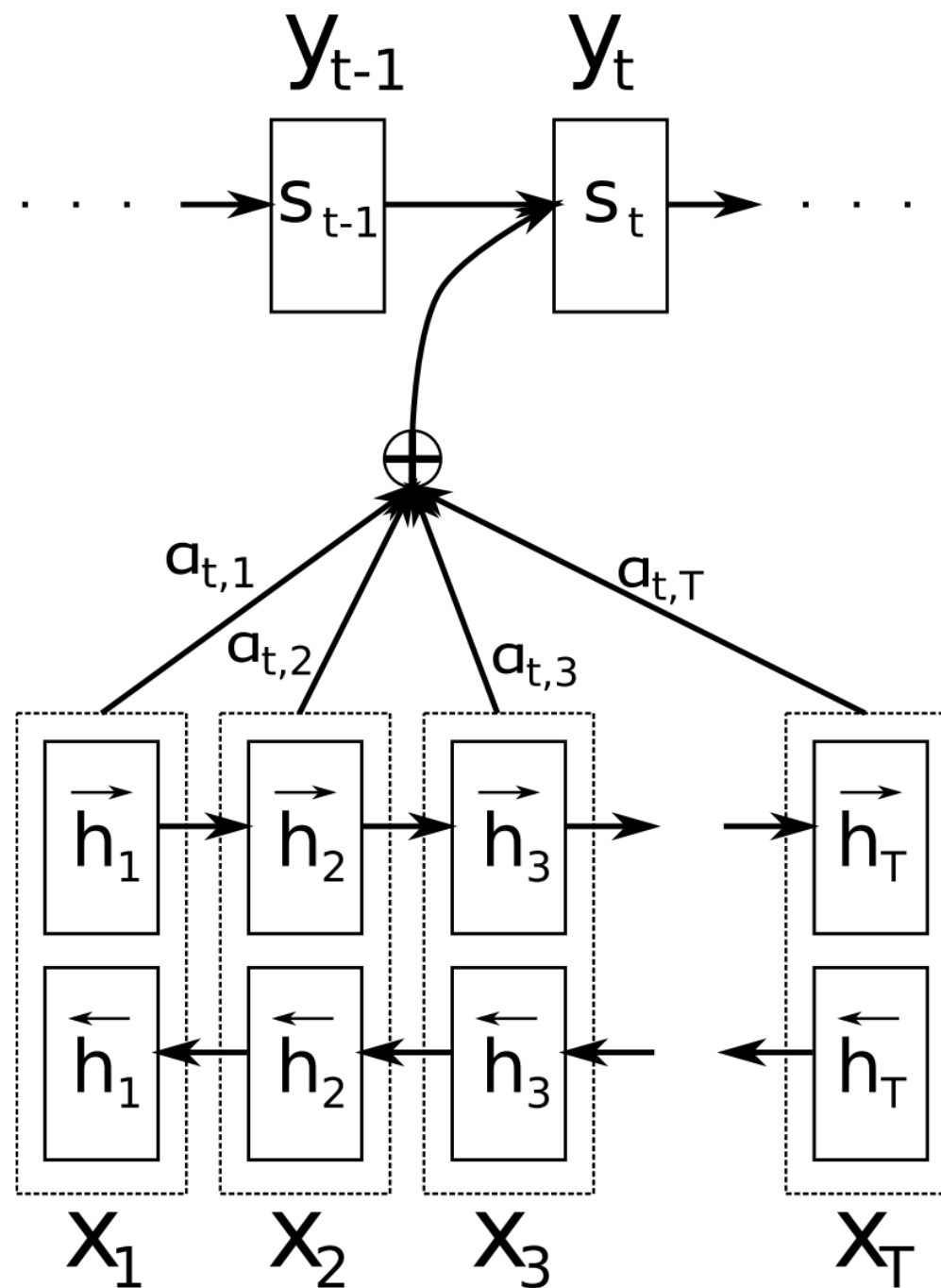


Sequence to sequence

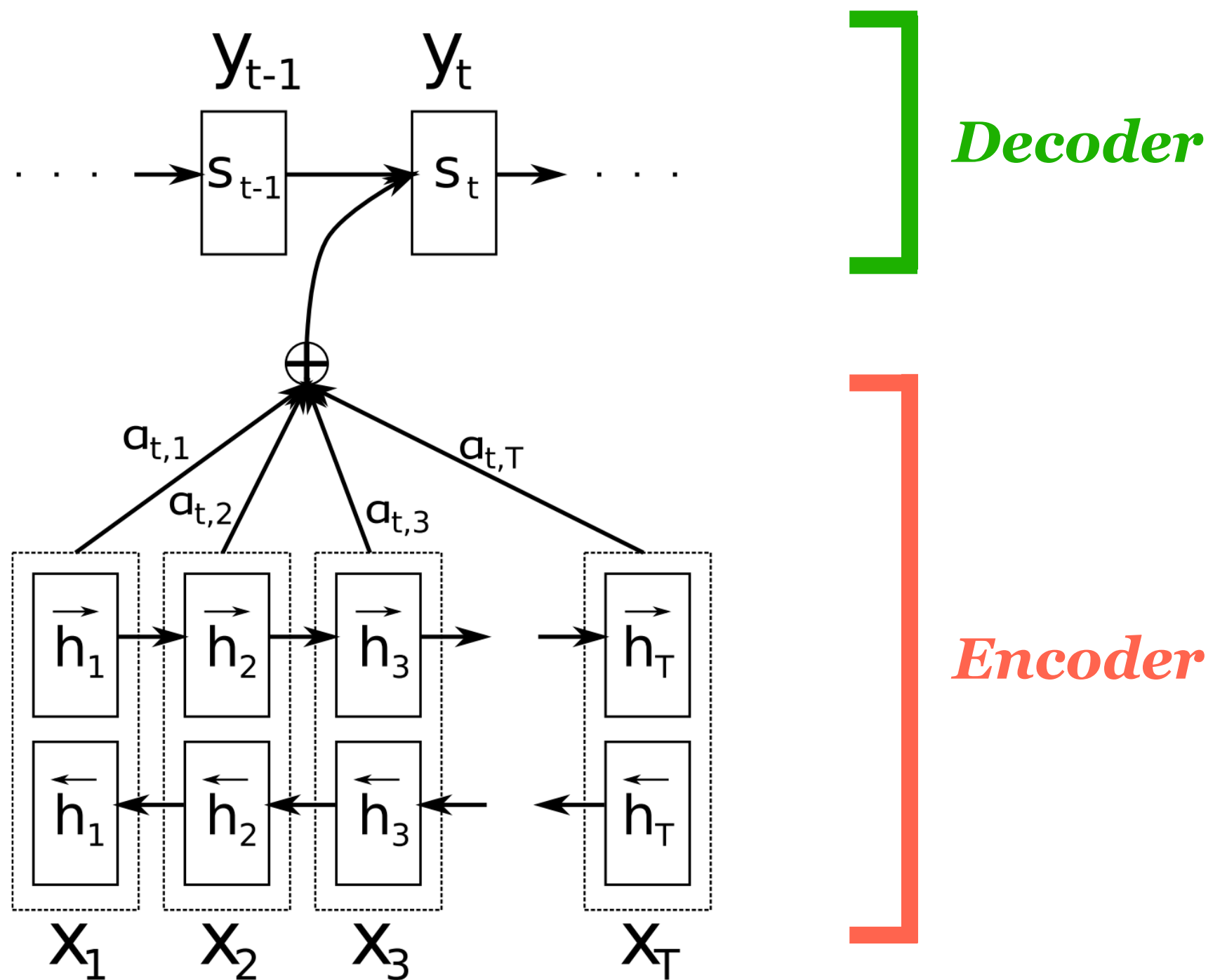


Attention Is All You Need

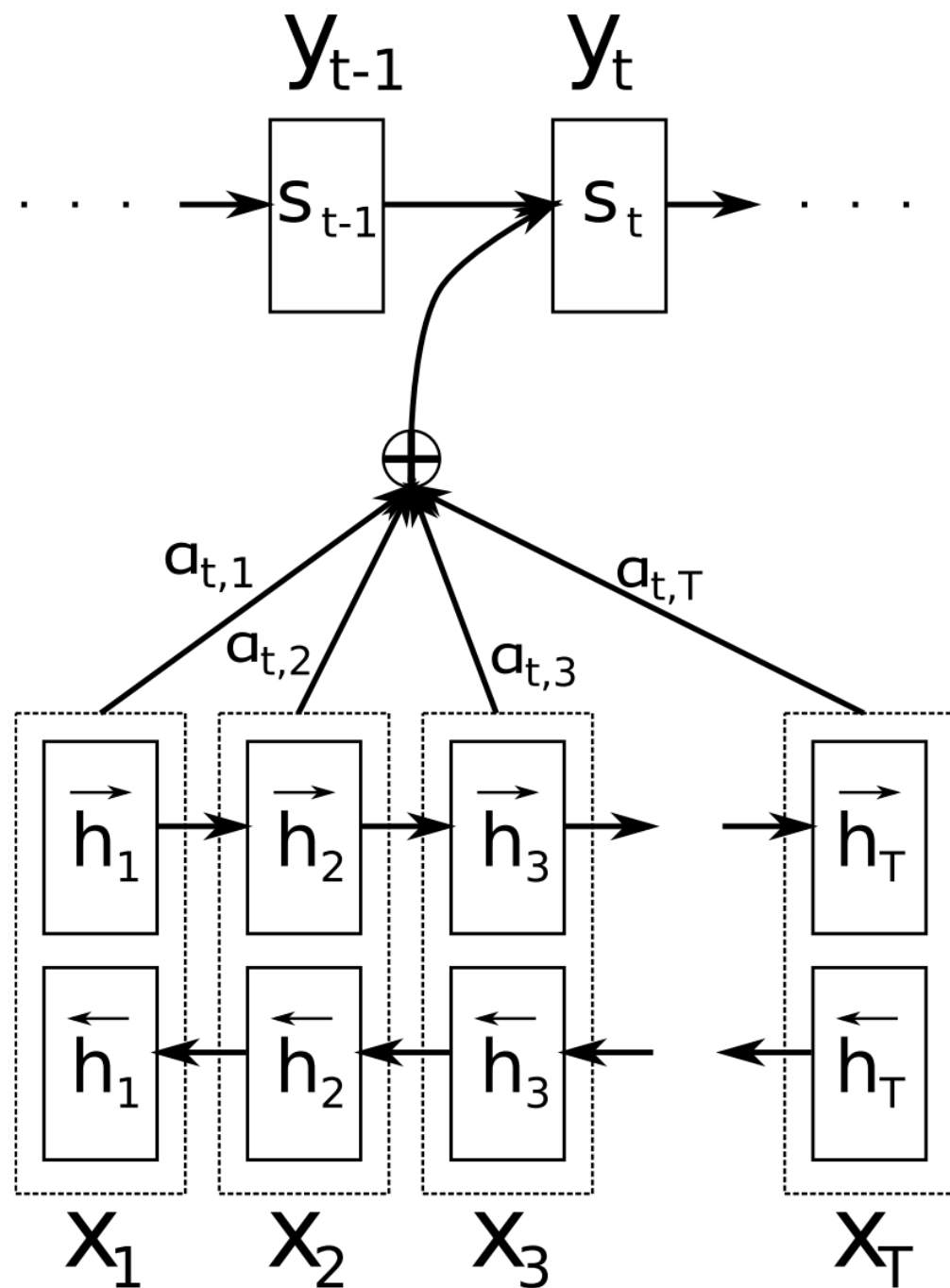
Attention



Attention



Attention



Скрытый слой RNN

$$s_i = f(s_{i-1}, y_{i-1}, c_i).$$

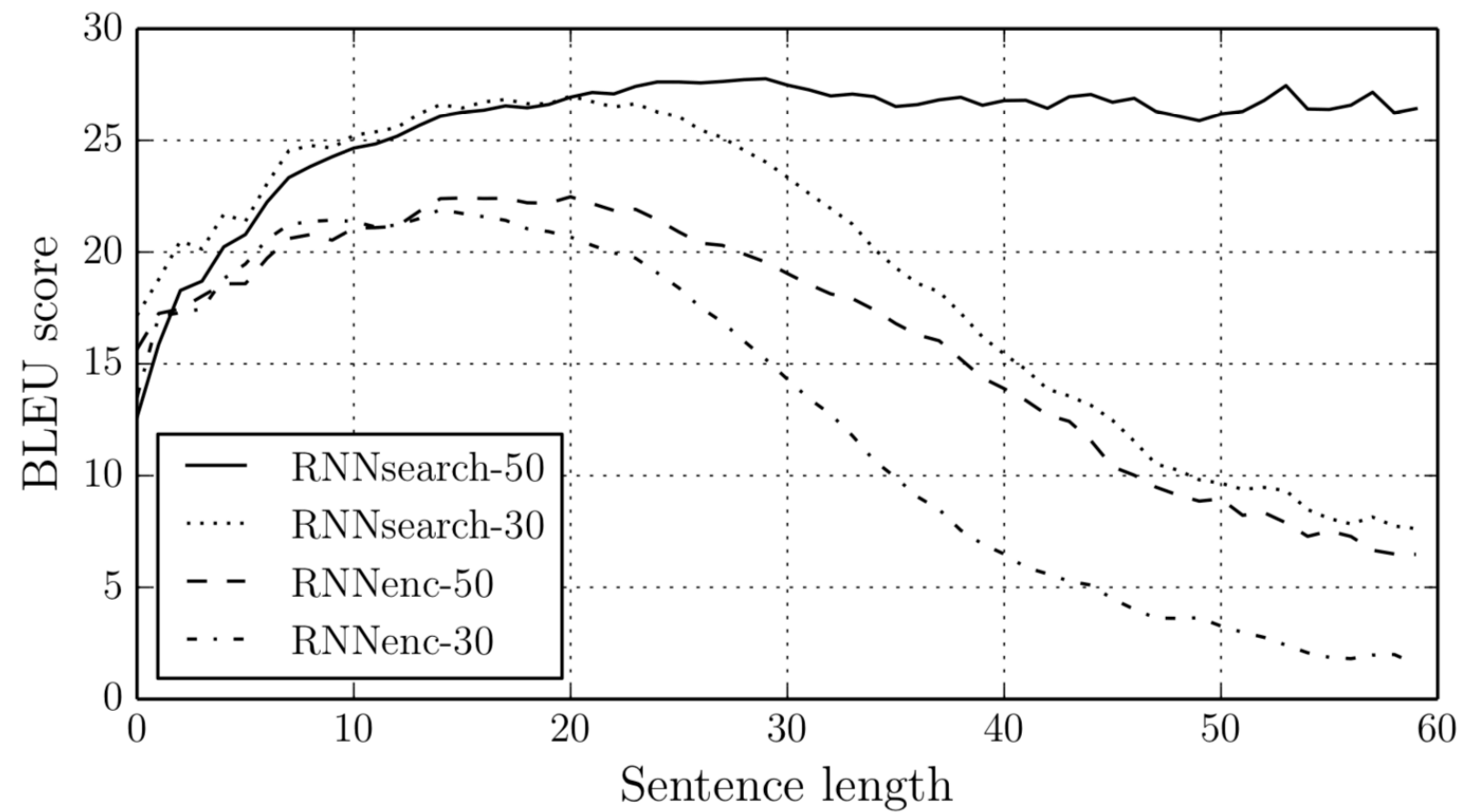
Взвешенная сумма

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j.$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})}$$

$$e_{ij} = a(s_{i-1}, h_j)$$

Attention



Attention

