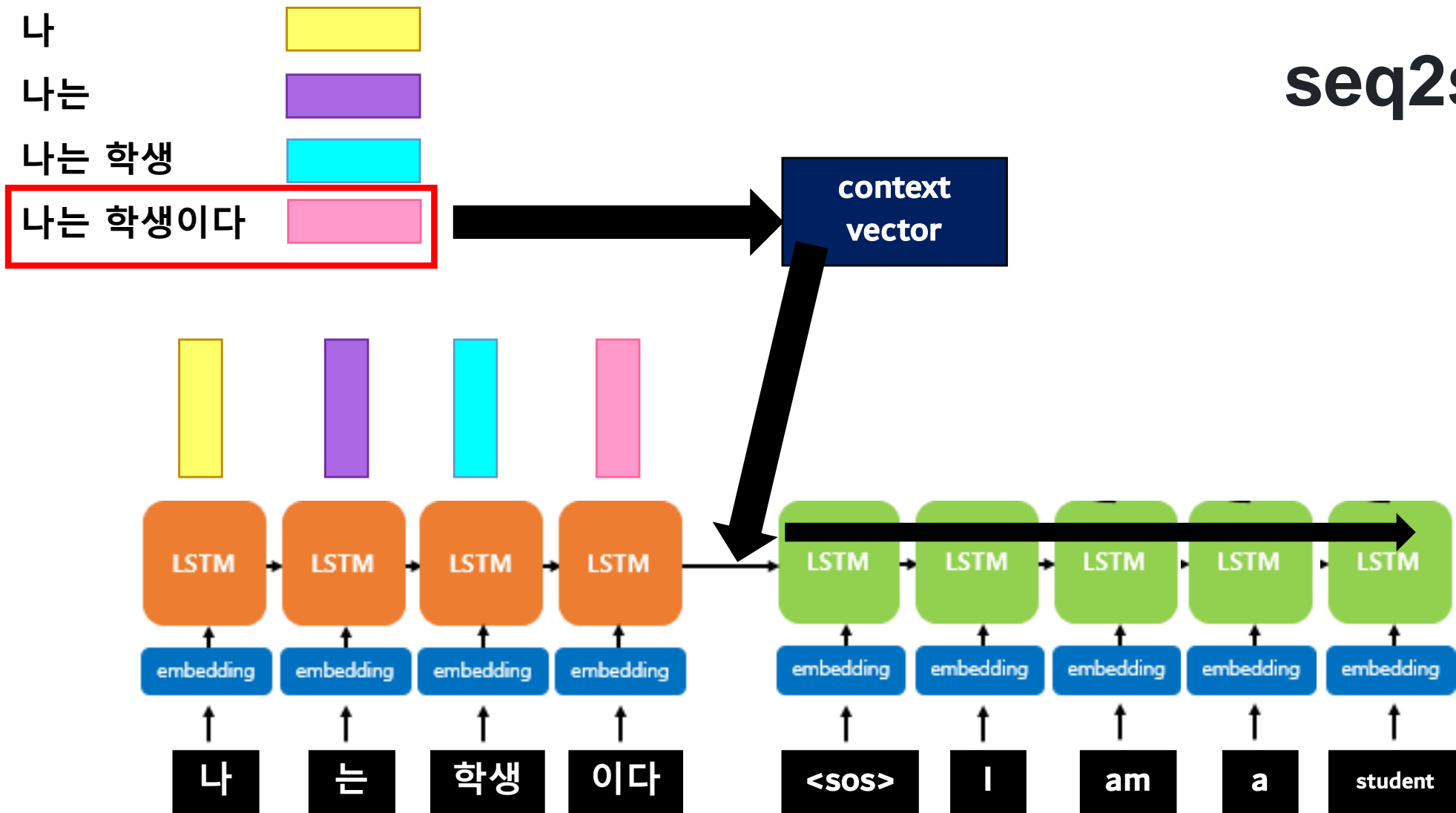




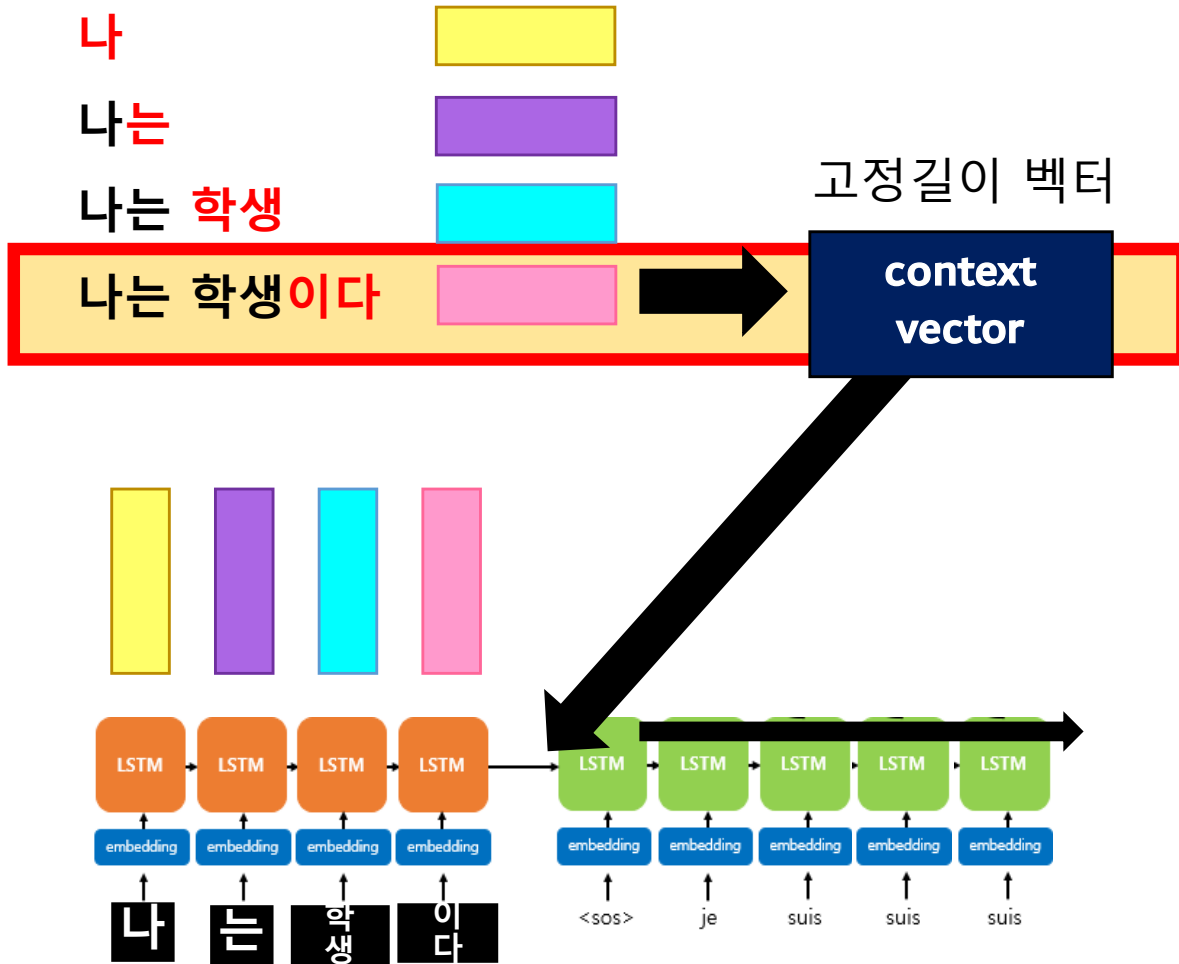
Attention

23.07.26 유하영

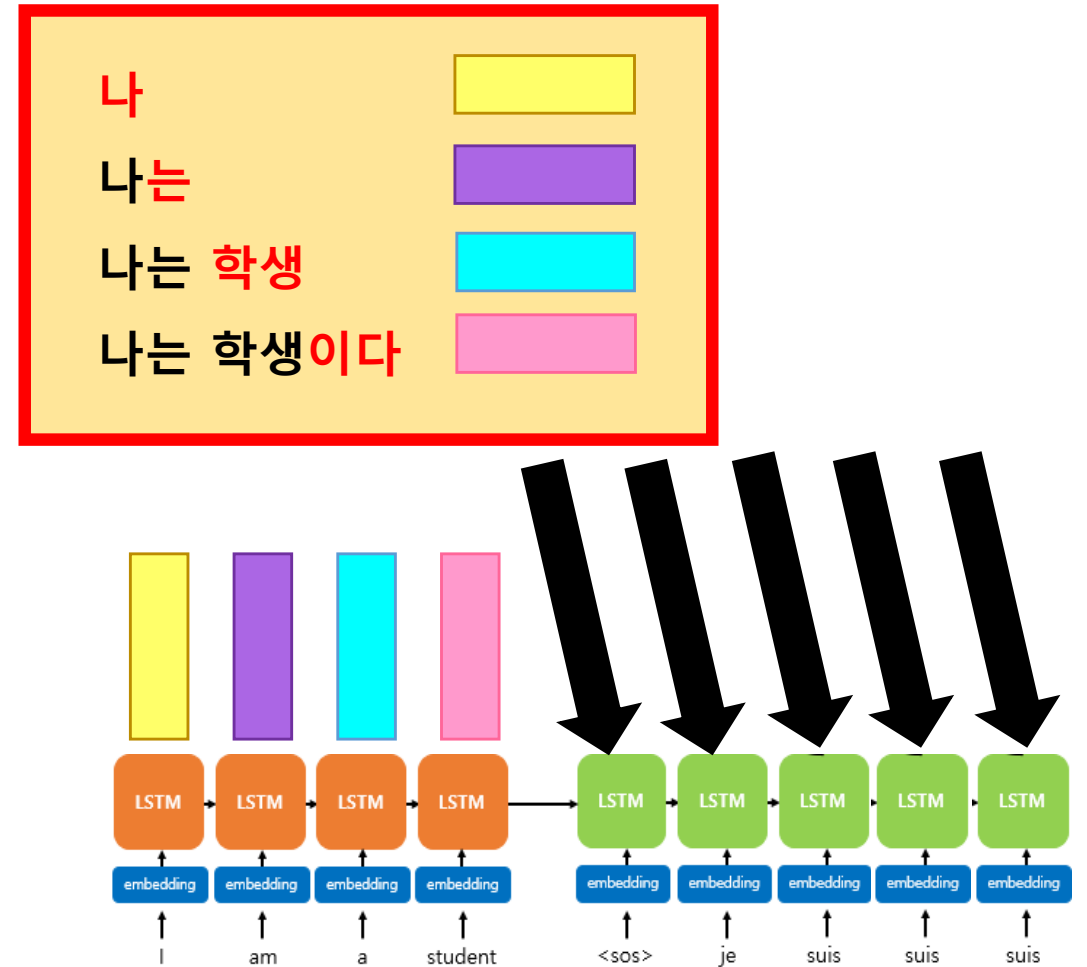
seq2seq



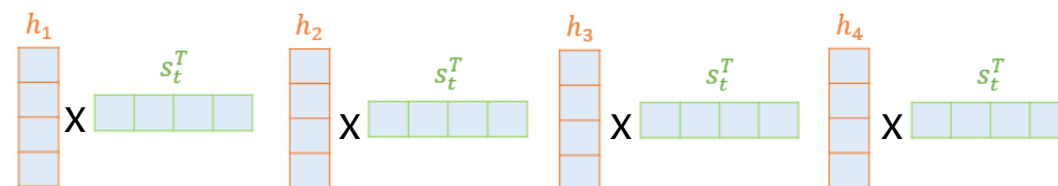
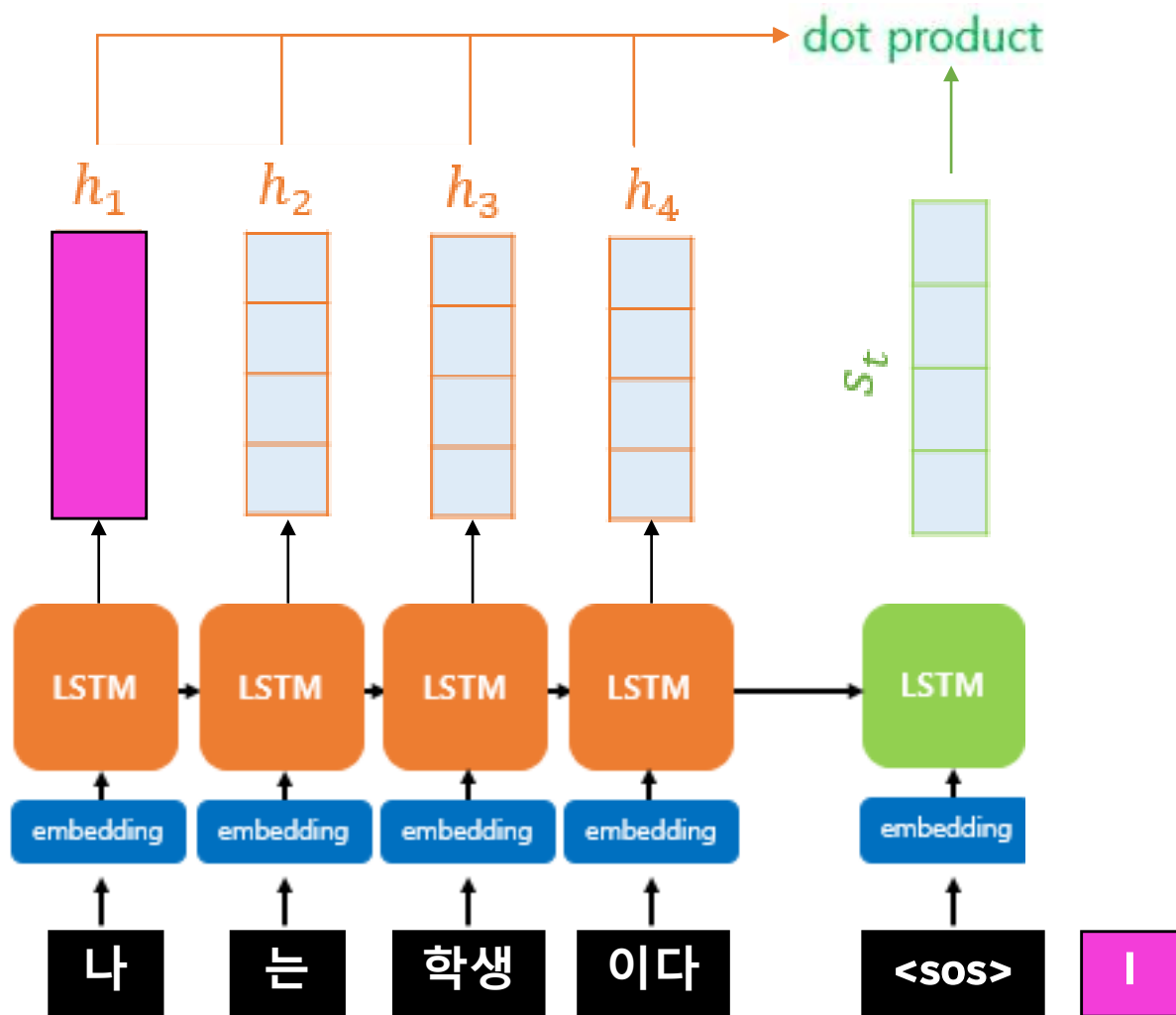
seq2seq



Attention이 적용된 seq2seq



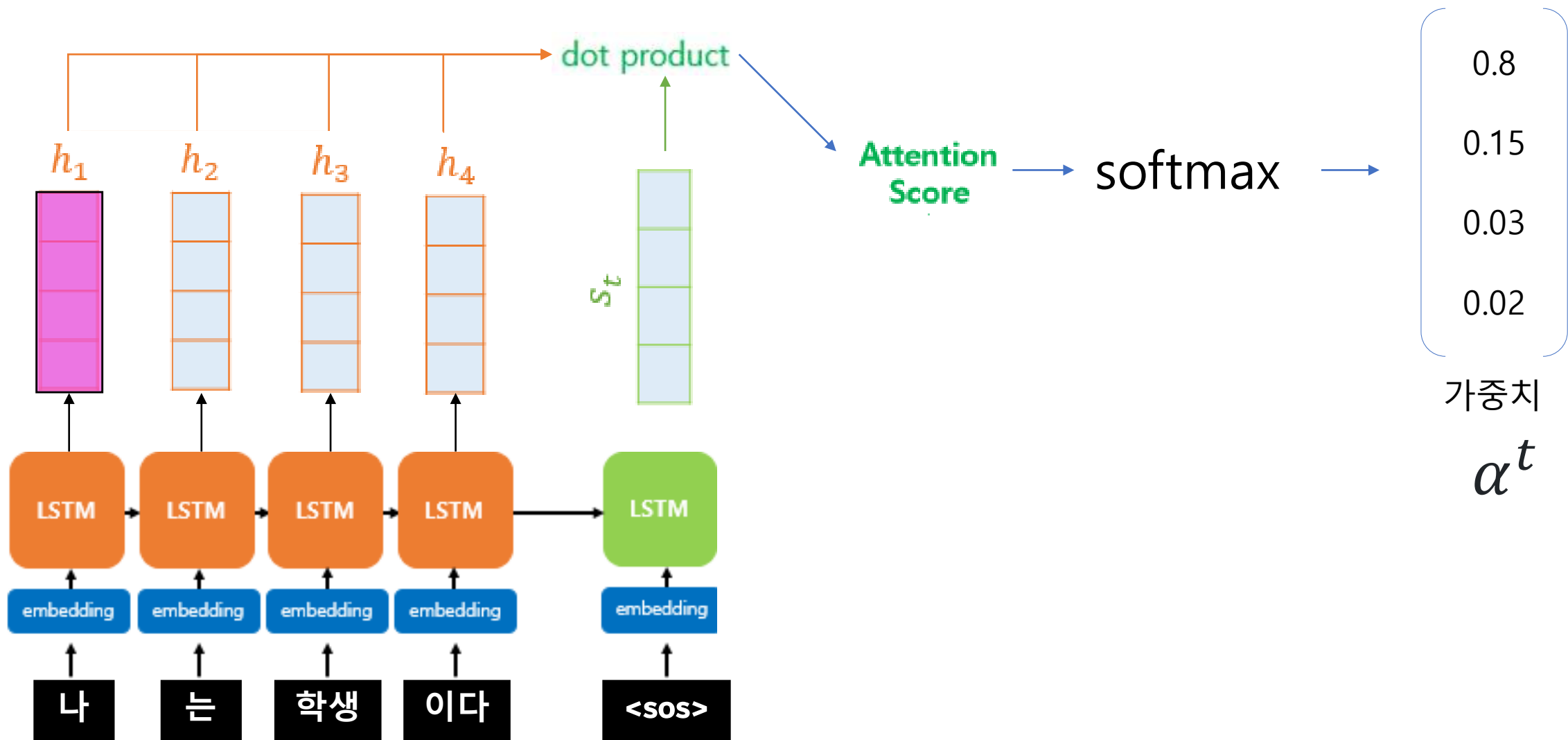
서로 대응되는 단어에 큰 가중치를 줄 수 있도록
어텐션 계층을 추가하는 것

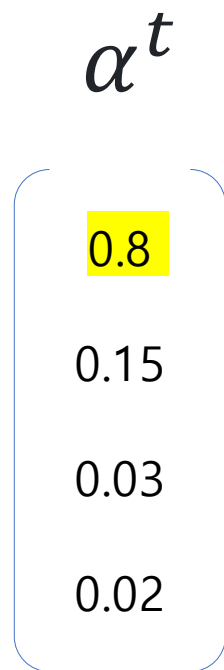
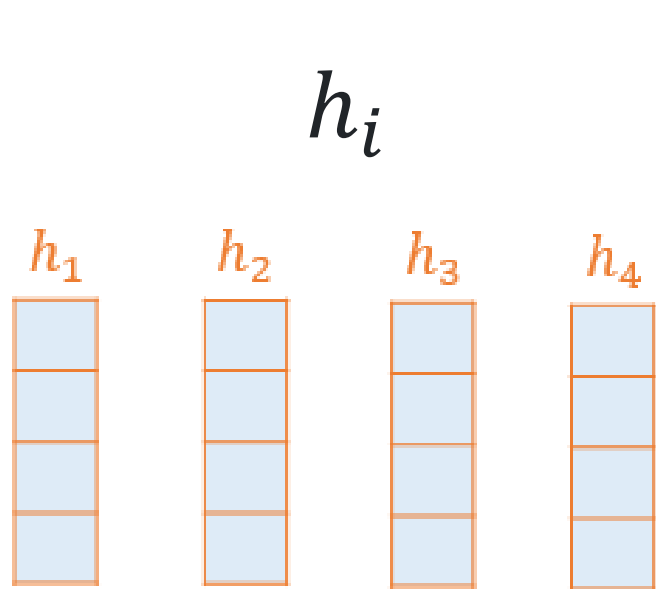


$$\text{score}(s_t, h_i) = s_t^T h_i$$

Attention Score $e^t = [s_t^T h_1, \dots, s_t^T h_N]$

인코더-디코더 각 단어들 간의
관련성 계산

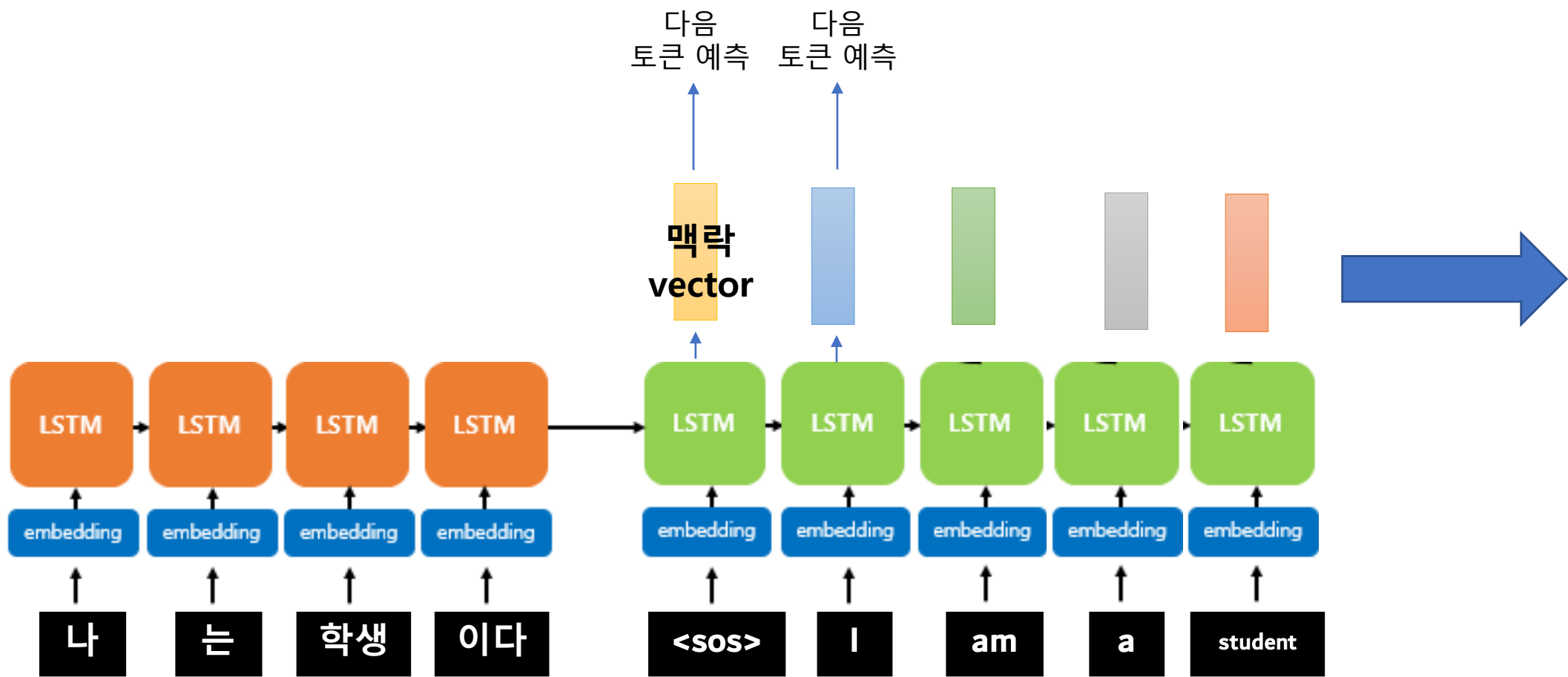




$$a_t = \sum_{i=1}^N \alpha_i^t h_i$$

맥락 vector
= **Attention value**

디코더가 인코더의 정보를
얼마나 사용해야하는지



더욱 집중적인 어텐션을 적용하는 어텐션 최적화 기법

self-attention mechanism