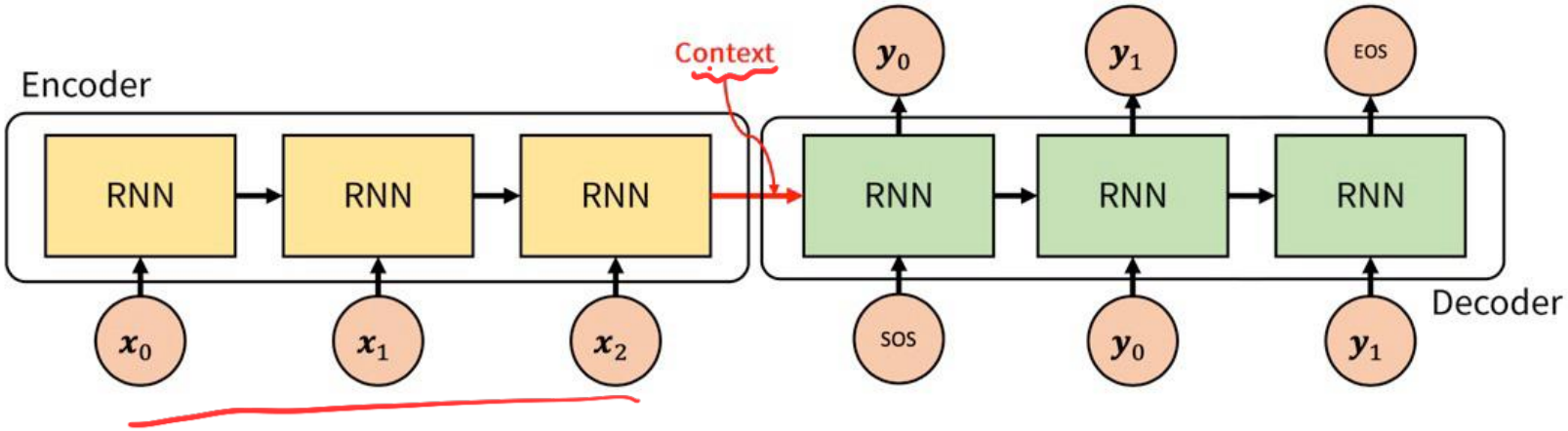
NEURAL MACHINE TRANSLATION BY JOINTLY LEARNING TO ALIGN AND TRANSLATE

2021.04.08

BOAZ 16기 NLP 1팀
김윤기

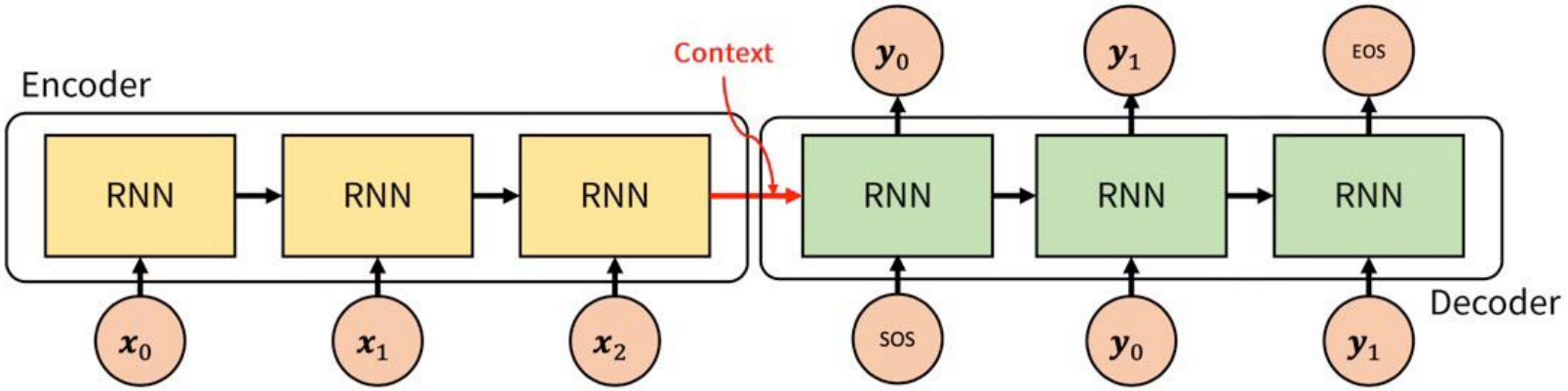
Problem Definition

- Machine Translation
  - 하나의 문장을 받아서, 다른 언어로 번역하는 작업



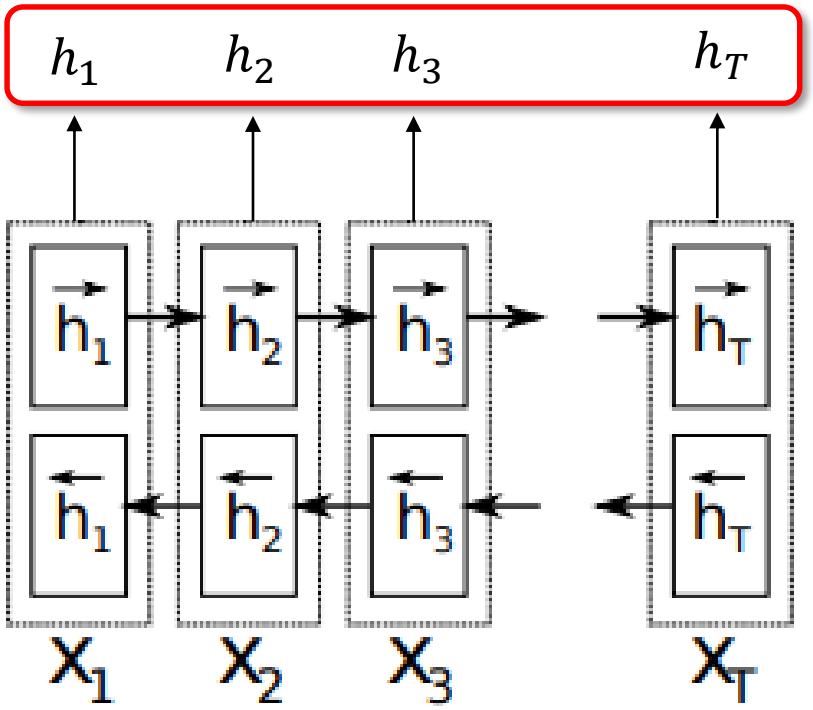
Motivation

- 장기 의존성 문제
  - 입력 문장의 모든 정보를 Fixed-length context vector에 나타내야 하기 때문에 발생
  - 입력 문장의 길이가 길어질 수록 두드러짐



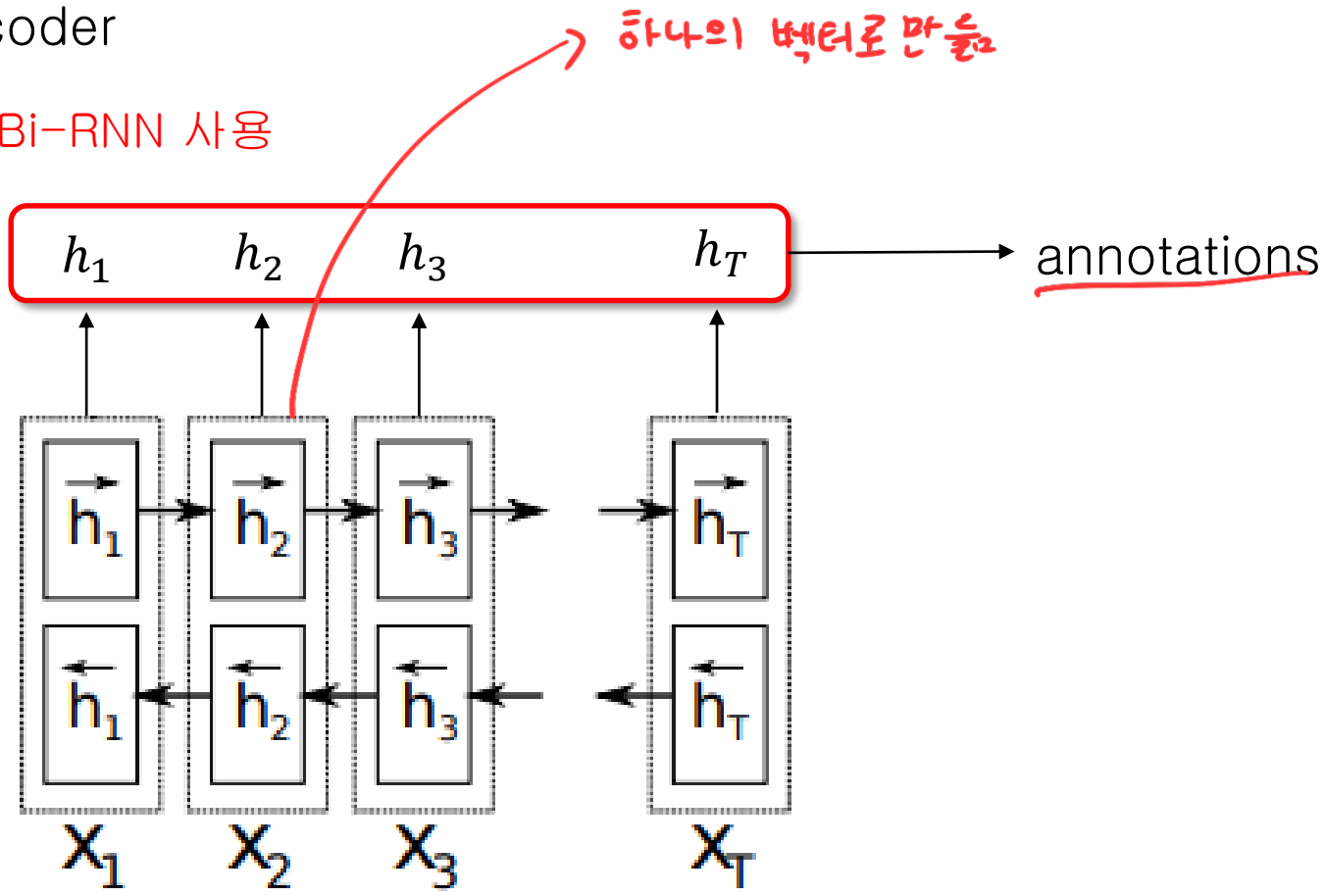
Idea

- 입력 문장을 vectors의 sequence로 나타내고 이를 decoder에서 활용하고자 함



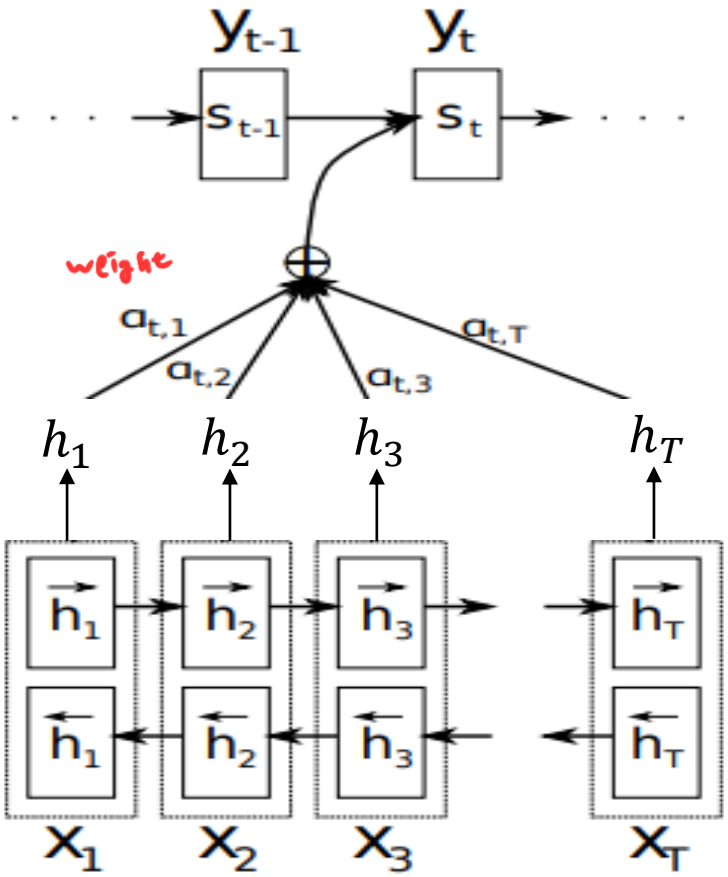
Proposed Method (1)

- Encoder
  - Bi-RNN 사용



Proposed Method (2)

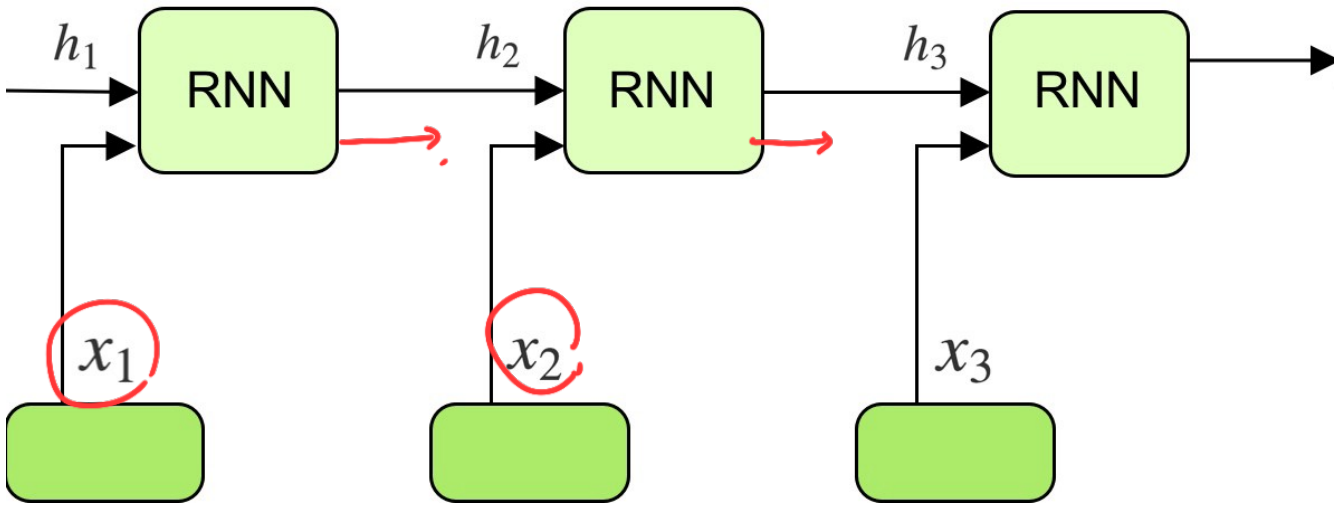
- Decoder



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

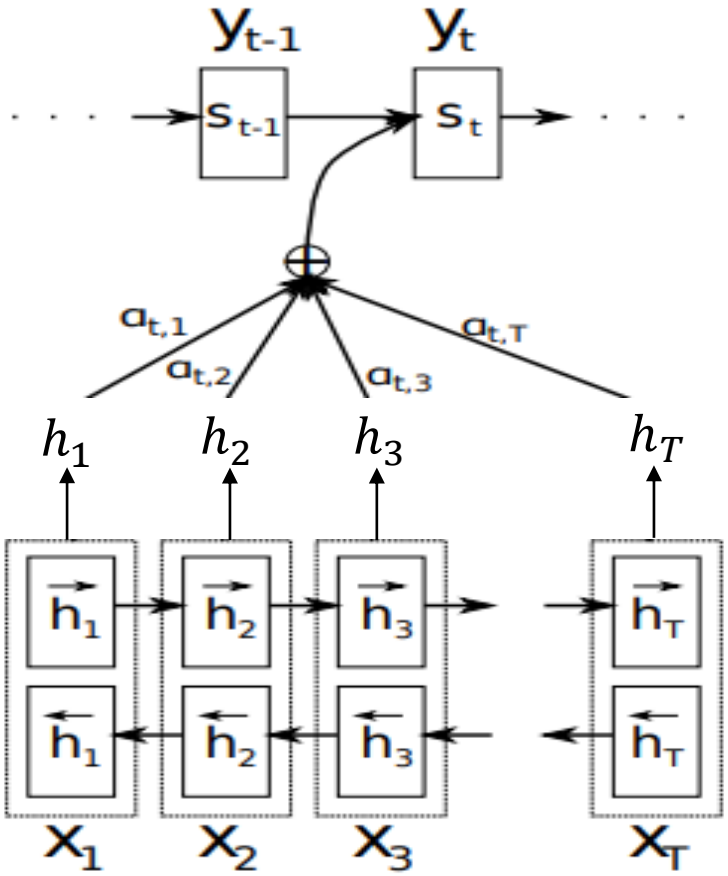
Proposed Method (3)

- Hidden state meaning
  - 해당 시점까지의 "문맥 정보"



Proposed Method (4)

- Decoder



유사도

align

=

디코더 정보

입력 문장 단어

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

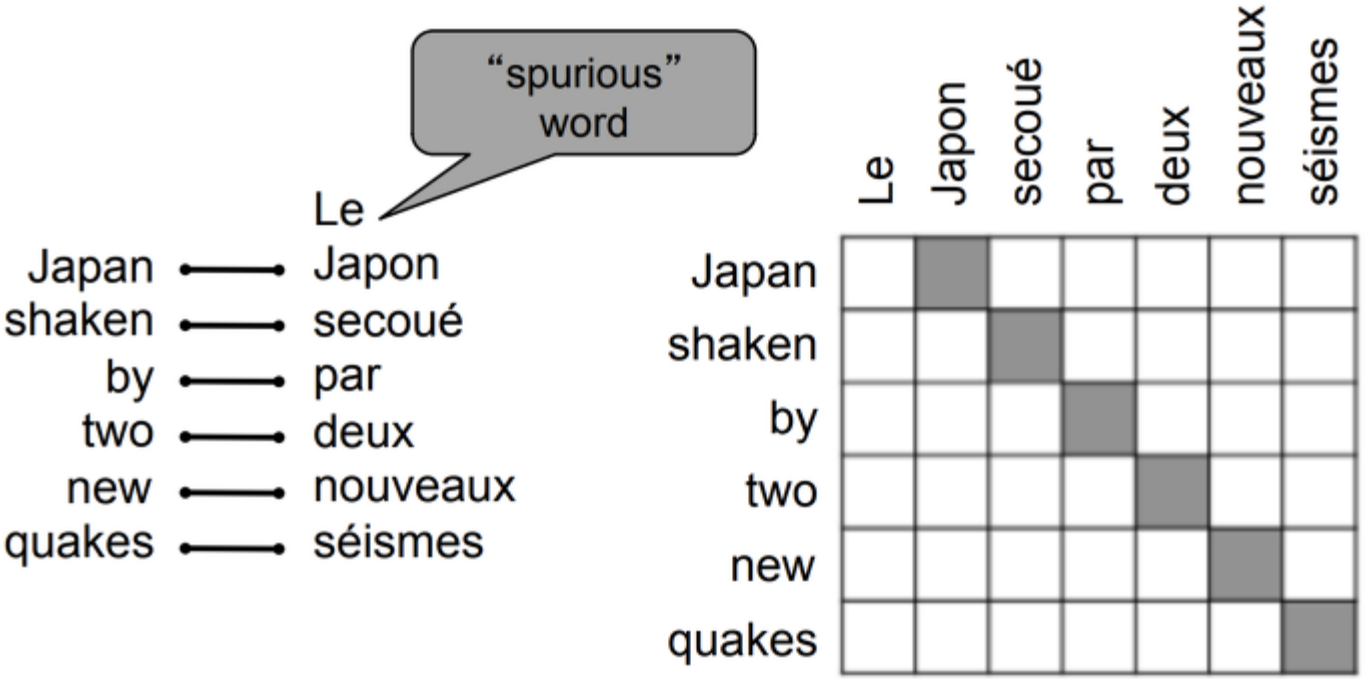
decoder hidden

input 단어

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

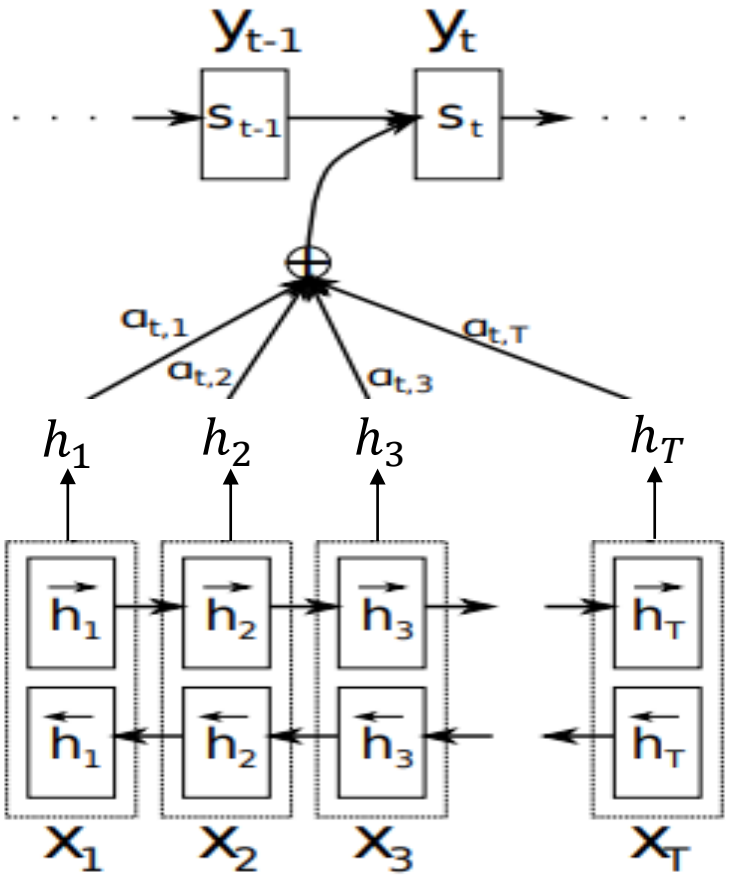
Proposed Method (5)

- Alignment (soft) => Attention
  - 두 언어간의 어순이 다른 경우, 그것을 자동으로 match 시켜준다는 의미



Proposed Method (6)

- Decoder



Experiment

- Experiment Setting
  - Data : WMT'14 (English-to-French)
  - Metric : BLEU

Experiment : RQ1

- 제안 방법이 장기 의존성 문제를 해결했는가 ?