NATURAL LANGUAGE PROCESSING

LECTURE 6: RNNs with Attention

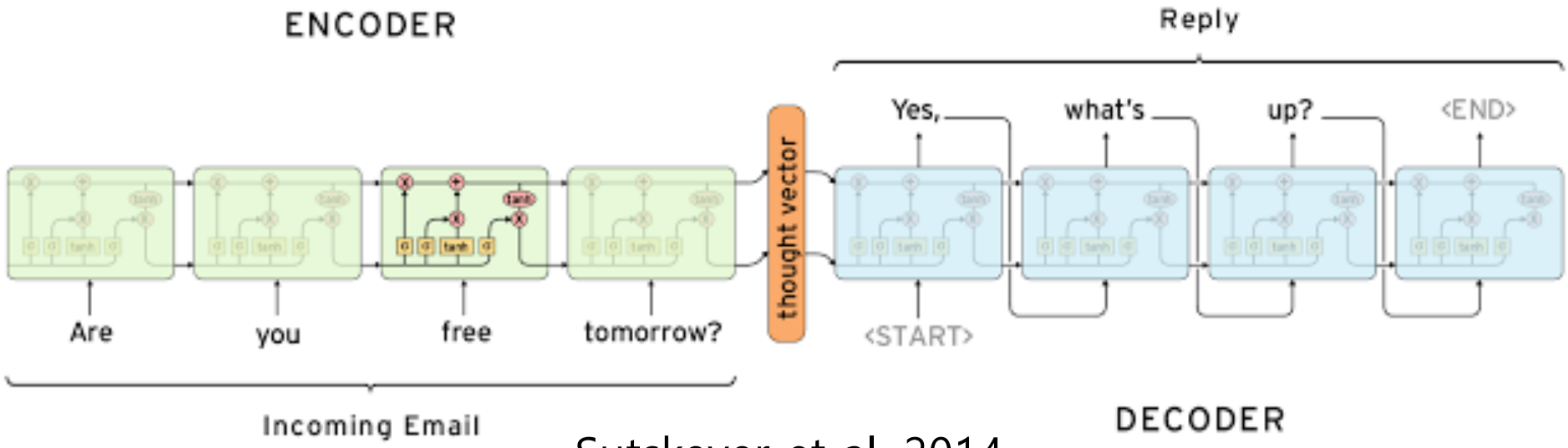
goorm

KAIST AI
Graduate School of AI

DAVIAN

Sequence to Sequence Model (seq2seq)

- 시퀀스를 입력으로 받아서, 시퀀스를 출력으로 생성
- 많은 NLP task 들에서 기본 모델로 활용됨: 챗봇, 기계번역 등



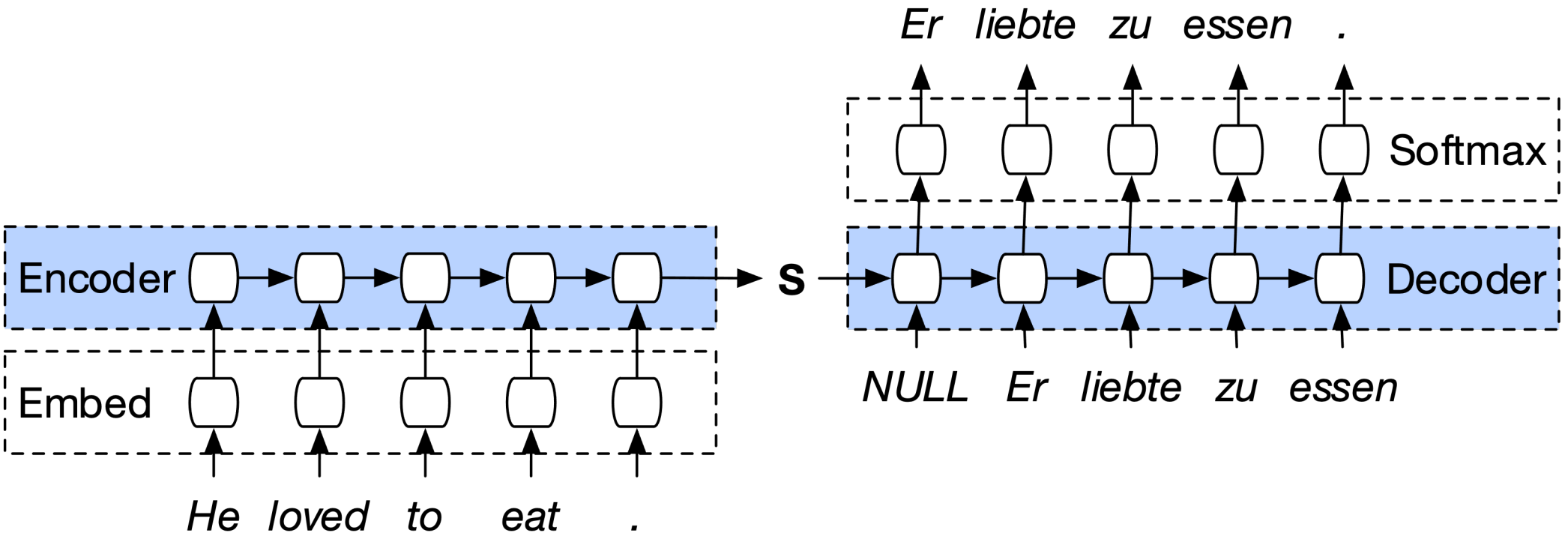
Sutskever et al. 2014

"***Sequence to Sequence Learning with Neural Networks***"

Encode source into fixed length vector, use it as
initial recurrent state for target decoder model

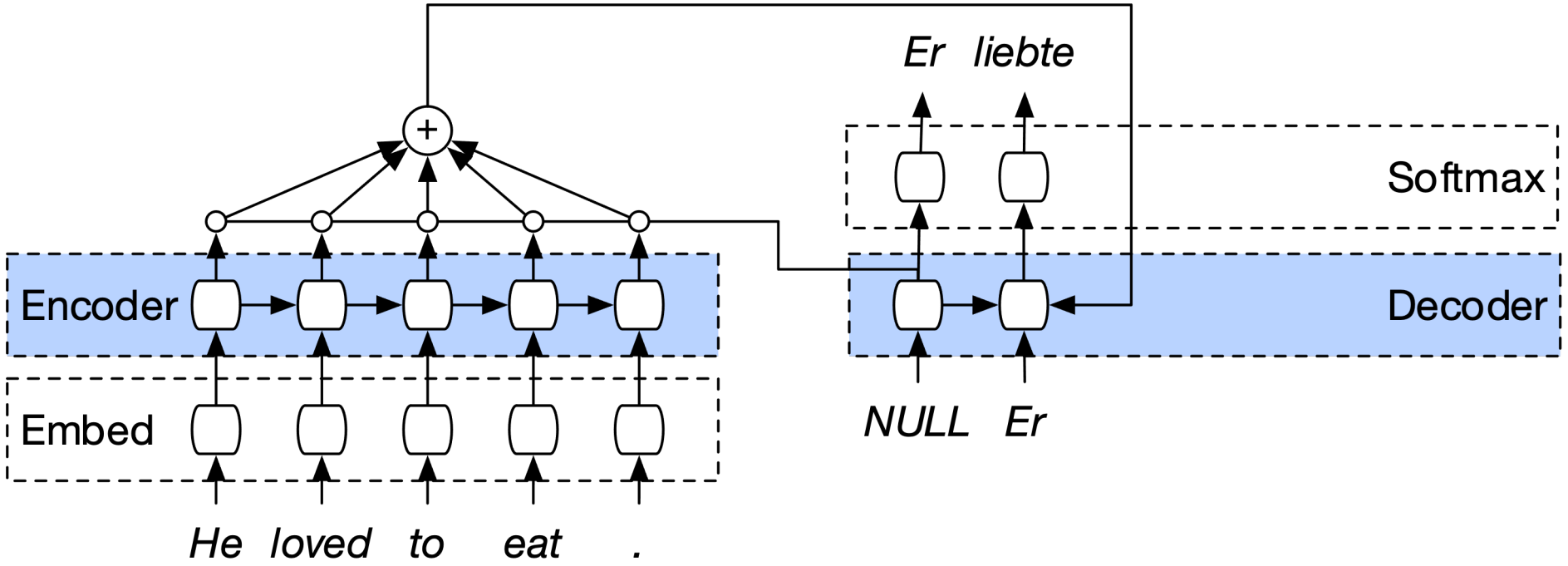
Seq2seq for Machine Translation

- 학습 데이터에서 입력 시퀀스-출력 시퀀스를 번역 데이터로 사용



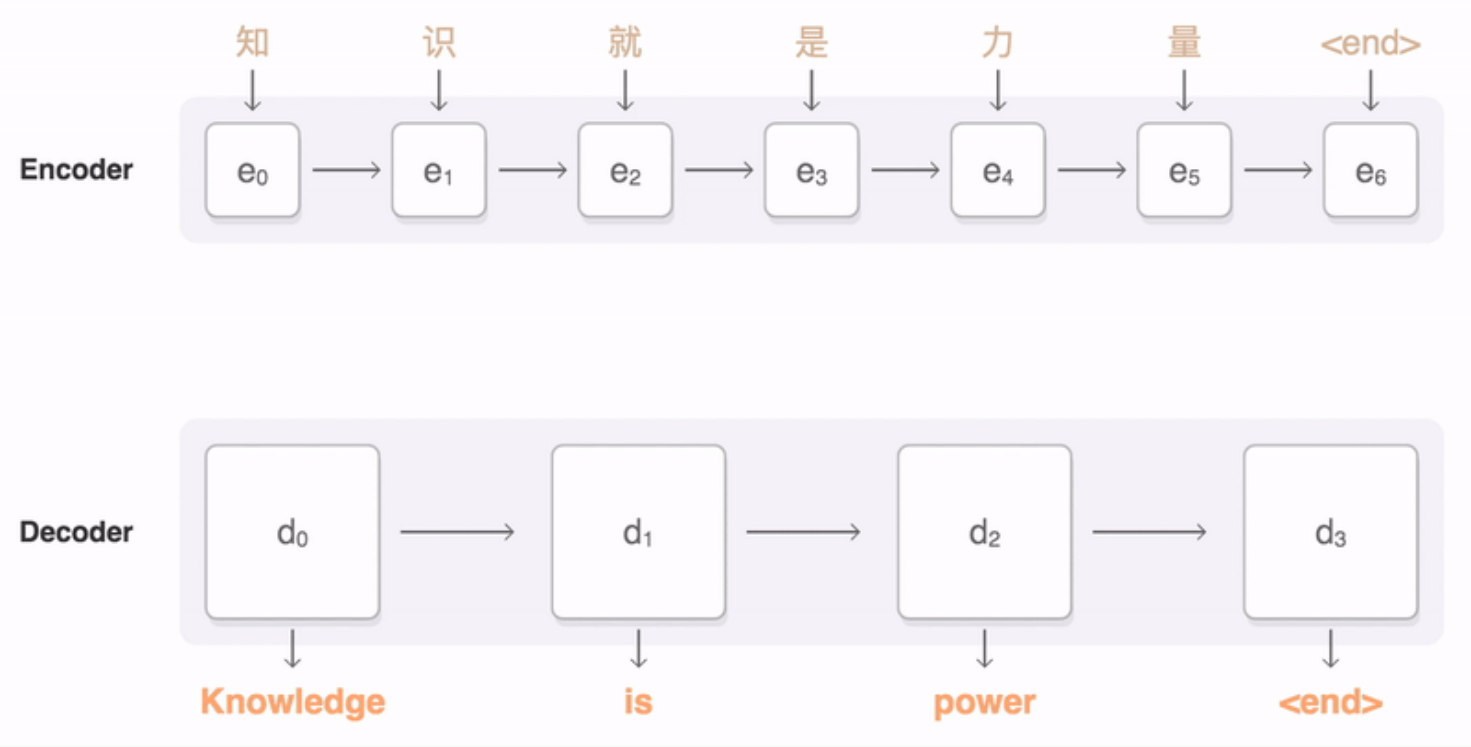
Seq2seq with Attention

- 입력 시퀀스의 마지막 시점의 벡터에 모든 정보를 다 담기가 버거우므로, 모든 입력 시퀀스의 정보를 조합하여 각 출력 단어를 생성
- 기계 번역 예



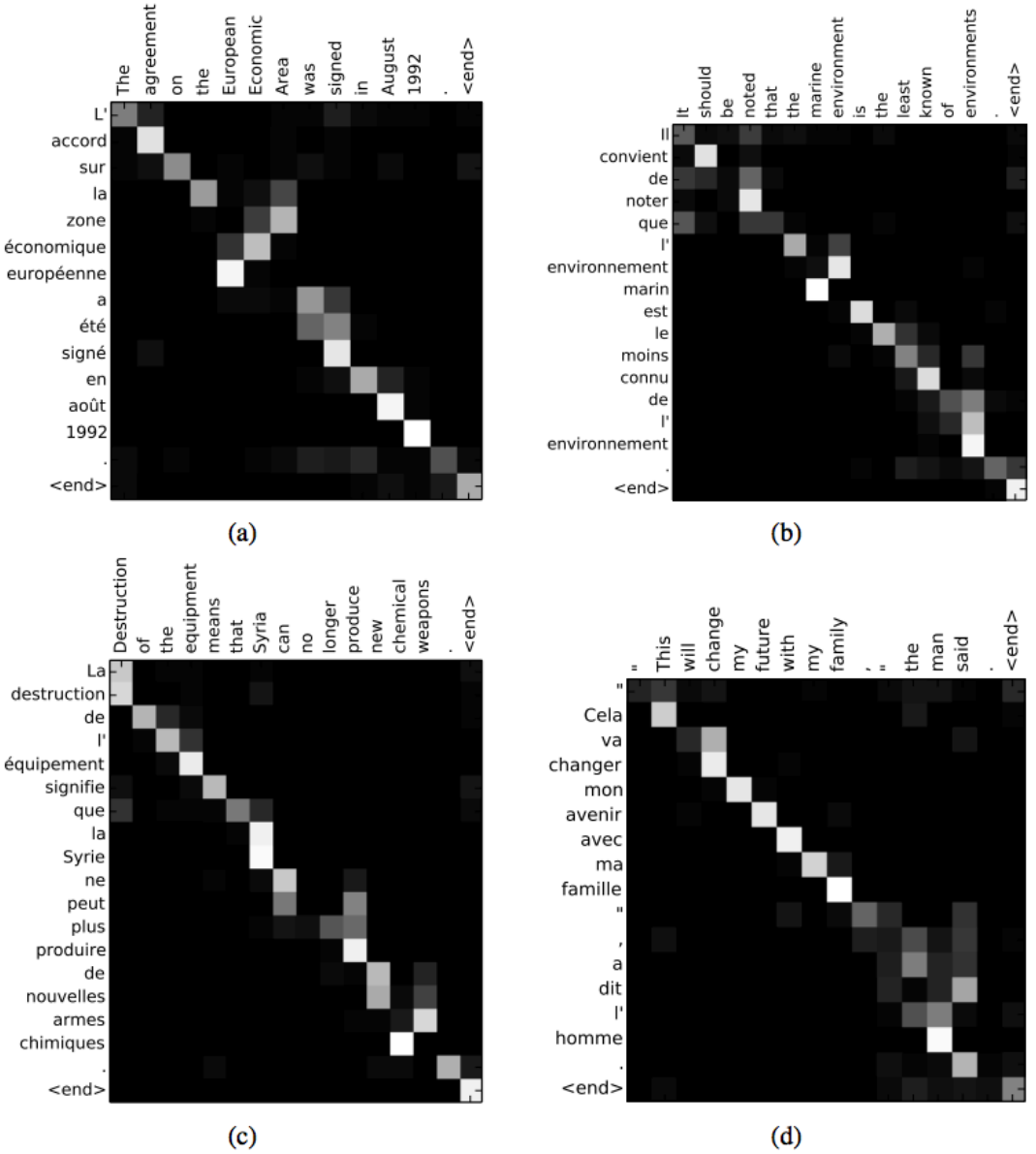
Seq2seq with Attention

- 입력 시퀀스의 마지막 시점의 벡터에 모든 정보를 다 담기가 버거우므로, 모든 입력 시퀀스의 정보를 조합하여 각 출력 단어를 생성
- 기계 번역 예: https://github.com/google/seq2seq



Attention Example in Machine Translation

- 다른 언어들 간의 어순을 학습함
- 관사 등의 필요없는 단어는 건너뜀

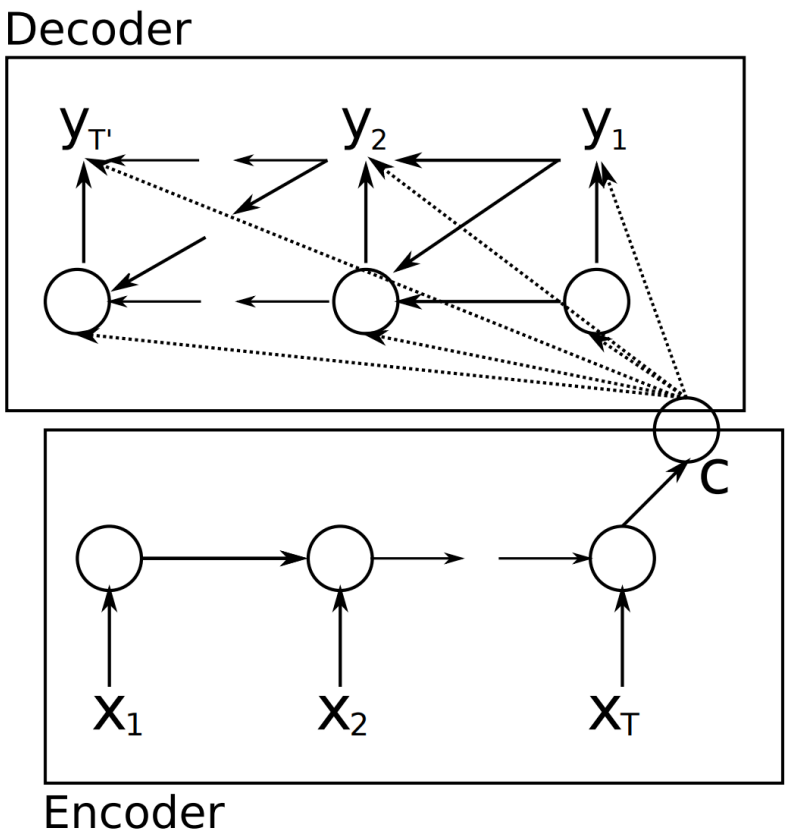


Trends of MT

- Machine translation 의 발전
- Statistical machine translation (SMT) (2015) learns probabilistic model of word-to-word transition to trai ning data.
- SMT requires advanced feature engineering.
- On the other hand, neural machine translation (NMT) based on neural network automatically extracts fe ature from data.
- NMT outperformed SMT in 2016.

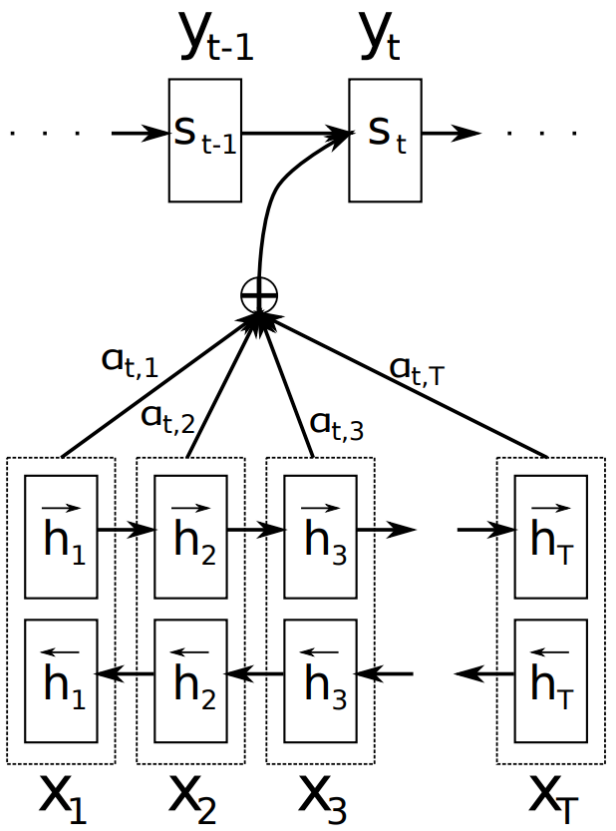
Trends of Attention

- Encoder-Decoder (seq2seq) (Cho et al., 2014)



Trends of Attention

- Encoder-Decoder (seq2seq) (Cho et al., 2014)
- Decoder with Attention (Bahdanau et al., 2015)



Trends of Attention

- Encoder-Decoder (seq2seq) (Cho et al., 2014)
- Decoder with Attention (Bahdanau et al., 2015)
- Encoder with Attention ?

MRC

- Machine reading comprehension
- Question answering

Article: Endangered Species Act
Paragraph: " . . . *Other legislation followed, including the Migratory Bird Conservation Act of 1929, a 1937 treaty prohibiting the hunting of right and gray whales, and the Bald Eagle Protection Act of 1940. These later laws had a low cost to society—the species were relatively rare—and little opposition was raised.*"

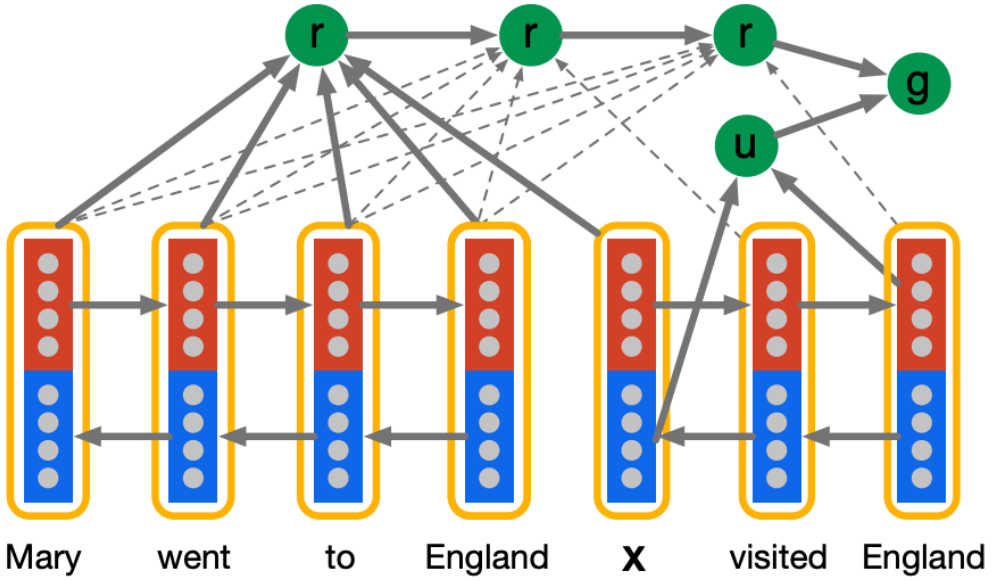
Question 1: "*Which laws faced significant opposition?*"
Plausible Answer: *later laws*

Question 2: "*What was the name of the 1937 treaty?*"
Plausible Answer: *Bald Eagle Protection Act*

Rajpurkar, P., Jia, R., & Liang, P. (2018). Know what you don't know: Unanswerable questions for SQuAD. *arXiv preprint arXiv:1806.03822*.

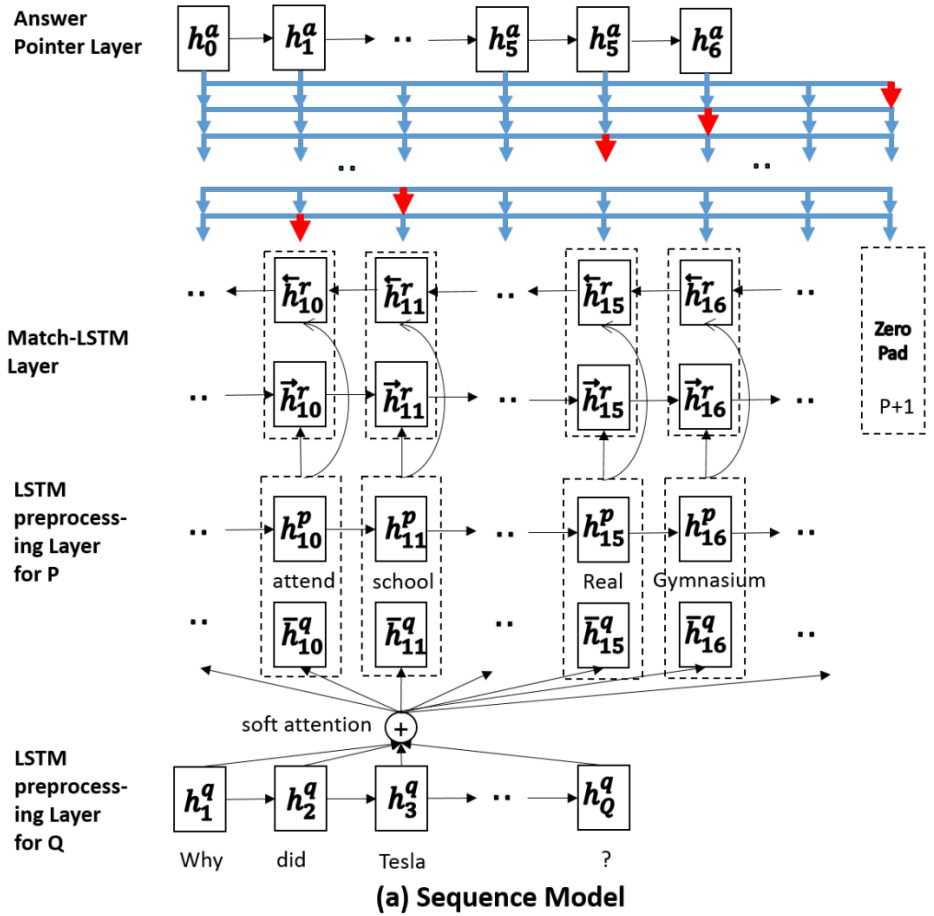
Trends of Attention in MRC

- Question-to-Context Attention (Hermann et al., 2015)



Trends of Attention in MRC

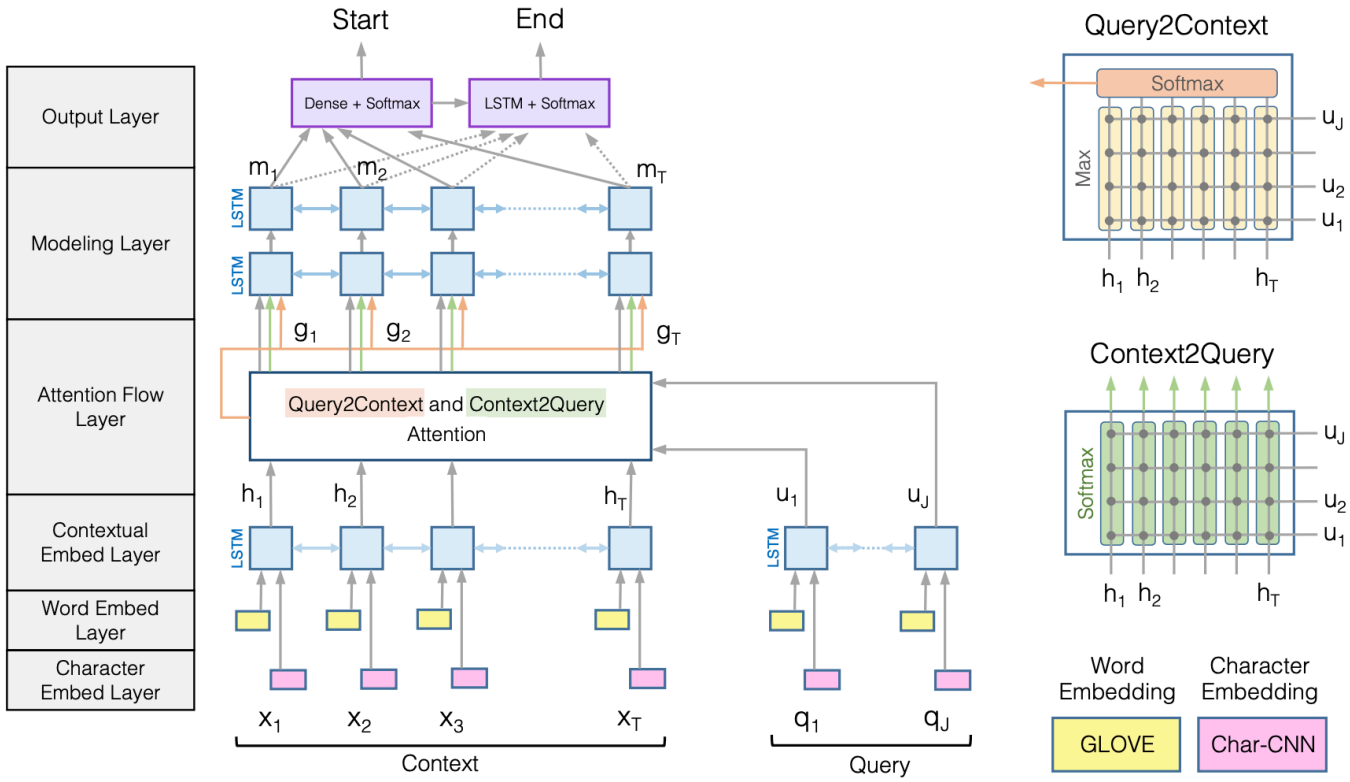
- Question-to-Context Attention (Hermann et al., 2015)
- Context-to-Question Attention (Wang & Jian, 2017)



(a) Sequence Model

Trends of Attention in MRC

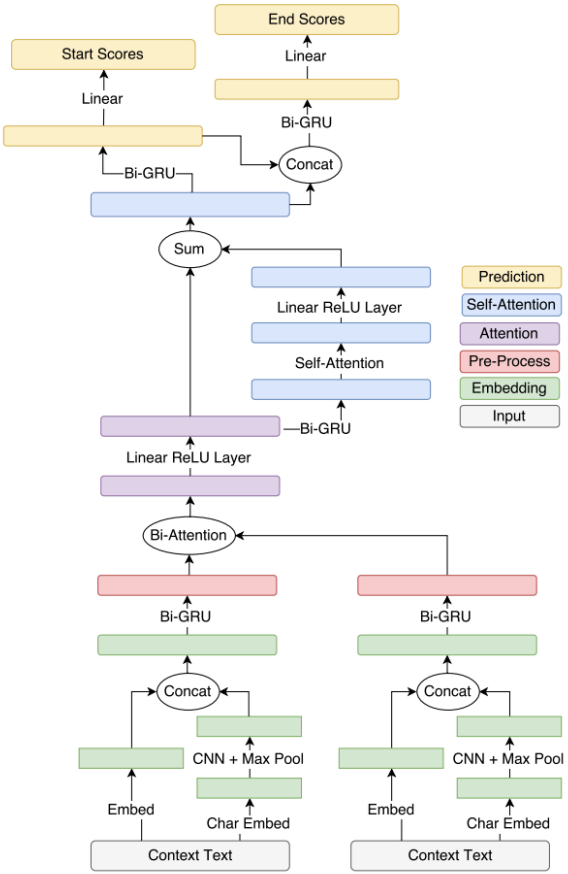
- Question-to-Context Attention (Hermann et al., 2015)
- Context-to-Question Attention (Wang & Jian, 2017)
- Bidirectional Attention (Seo et al., 2017)



Trends of Attention in MRC

- Question-to-Context Attention (Hermann et al., 2015)
- Context-to-Question Attention (Wang & Jian, 2017)
- Bidirectional Attention (Seo et al., 2017)
- Context-to-Context Attention (Self Attention)

  (Clark and Garnder, 2017)



References

Stanford University CS231n: Convolutional Neural Networks for Visual Recognition

Deep Learning Summer School, Montreal 2016 - VideoLectures.NET

Understanding LSTM Networks -- colah's blog