

# On the Properties of Neural Machine Translation: Encoder-Decoder Approaches

SSST-8 (ACL 2014)

Tag: Neural Machine Translation, RNN, grConv

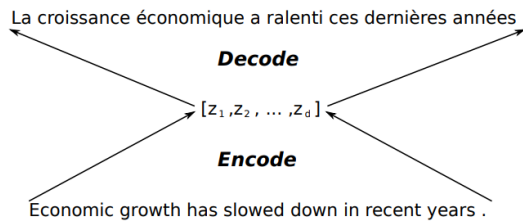
## 0. Summary

- 통계적 모델인 SMT(Statistical Machine Translation)와 비교하여 **Encoder-Decoder 구조인 RNN Encoder-Decoder 모델과 grConv 구조**에서 NMT 속성을 분석함

## 1. Introduction

### > NMT (Neural Machine Translation)

#### - Encoder-Decoder Approach



#### - Encoder

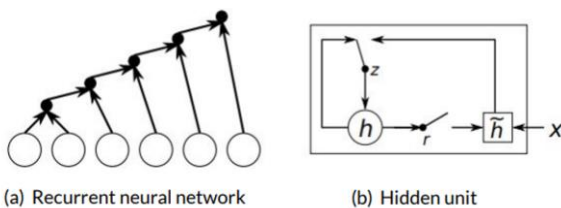
Input: variable-length sentence  
Output: fixed-length vector representation

#### - Decoder

Input: fixed-length vector representation  
Output: variable-length target translation

## 2. Neural networks for variable-length sequences

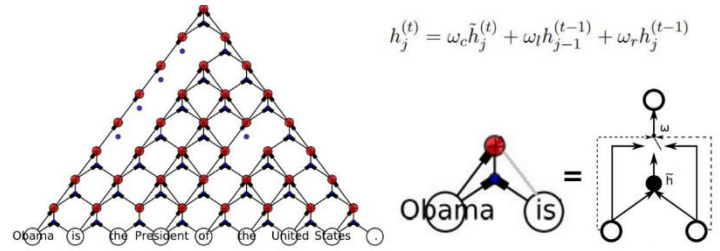
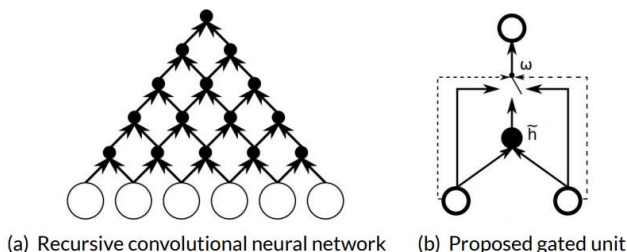
### > RNN with gated hidden neurons



$$\mathbf{h}^{(t)} = f(\mathbf{h}^{(t-1)}, \mathbf{x}_t)$$

- 각 timestep t마다 hidden state  $\mathbf{h}(t)$ 가 업데이트 됨
- Next input에 대한 분포를 학습하여 Variable-length sequence를 잘 학습함

### > grConv (gated recursive convolutional neural network)



- 논문에서 제안한 **Encoder 모델**, Decoder는 RNN 사용
- Fixed-length vector를 출력할 때까지 binary convolution 신경망을 적용

## 3. Experiments

### > Experiments Setting

- **Dataset:** English-to-French translation task
- **Model:** RNNenc (RNN Encoder-Decoder), grConv  
비교: SMT system(Moses)
- BLEU score로 성능 측정, Beam-Search 사용

### > Quantitative case

#### 1) sentence 길이에 따른 성능 변화

- 짧은 문장에서 RNNenc, grConv 성능 좋음
- 문장 길이 길어질수록 성능 저하

#### 2) unknown words 포함 여부

- UNK 많아질수록 RNNenc, grConv의 성능 저하됨

### > Qualitative Analysis

- BLEU score과 무관하게 짧은 문장에서는 모두 좋은 번역을 보임
- 긴 문장에서는 신경망 모델의 성능 저하 나타남  
→ fixed-length vector에는 긴 문장 encoding할 용량 X
- grConv는 문법 구조를 알아서 학습하는 특징을 보임

끝!!

질문 있으시면 정민지에게 DM 주세요!!