XMU Neural Machine Translation Systems for WMT 17

Zhixing Tan, Boli Wang, Jinming Hu, Yidong Chen and Xiaodong Shi
School of Information Science and Engineering, Xiamen University, China

MODELS

- **MiniNMT**
  - Encoder: IBE / SBE
    - **I**nterleaved **B**idirectional **E**ncoder (4 layers x 2 columns)

$$\overrightarrow{\mathbf{x}}_t^i = \mathrm{LSTM}_i^f(\overrightarrow{\mathbf{x}}_t^{i-1}, \overrightarrow{\mathbf{s}}_{t+(-1)^i}^i) \quad (1)$$

$$\overleftarrow{\mathbf{x}}_t^i = \mathrm{LSTM}_i^b(\overleftarrow{\mathbf{x}}_t^{i-1}, \overleftarrow{\mathbf{s}}_{t+(-1)^{i+1}}^i) \quad (2)$$

    - **S**tacked **B**idirectional **E**ncoder (8 layers)

$$\overrightarrow{\mathbf{x}}^i = \mathrm{LSTM}_i^f(\mathbf{x}_t^{i-1}, \overrightarrow{\mathbf{s}}_{t-1}^i) \quad (3)$$

$$\overleftarrow{\mathbf{x}}^i = \mathrm{LSTM}_i^b(\mathbf{x}_t^{i-1}, \overleftarrow{\mathbf{s}}_{t+1}^i) \quad (4)$$

$$\mathbf{x}^i = [\overrightarrow{\mathbf{x}}^{iT}; \overleftarrow{\mathbf{x}}^{iT}]^T \quad (5)$$

  - Decoder (8 layers)

$$\mathbf{a}_t = \sum_{i=1}^{S} \alpha_{t,i} \cdot \mathbf{x}_i \quad (6)$$

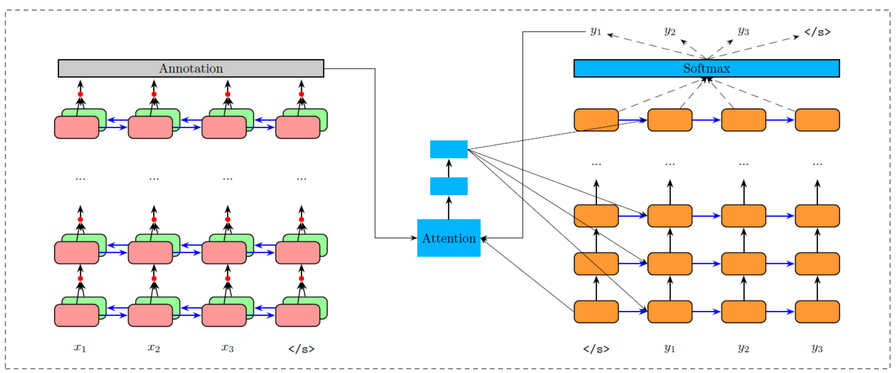
$$\mathbf{h}_t = [\mathbf{y}_{t-1}^{0\,T}; \mathbf{y}_{t-1}^{1\,T}]^T \quad (7)$$

$$e_{t,i} = \mathbf{v}_a^T \tanh(\mathbf{W}_a \mathbf{h}_t + \mathbf{U}_a \mathbf{x}_i) \quad (8)$$

$$\alpha_{t,i} = \frac{\exp(e_{t,i})}{\sum_{j=1}^{S} \exp(e_{t,j})} \quad (9)$$

$$p(y_t|\boldsymbol{x}, \boldsymbol{y}_{<t}) = \mathrm{softmax}(y_t, \mathbf{y}_t^{L_{\mathrm{dec}}}) \quad (10)$$

- **DL4MT** (Our reimplementation)
  - Use modified version of AmuNMT decoder



▲ The architecture of **MiniNMT**

EXPERIMENTAL FEATURES

- **Data Filtering**
  - According to the word alignment score (fast-align)
- **Segmentation Approaches**
  - BPE (Sennrich et al., 2016c)
  - Mixed word/character model (Wu et al., 2016)
- **Synthetic Training Data**
  - Back-translation (Sennrich et al., 2016b)
  - Mix synthetic bilingual pairs with sampled ones
- **Target-bidirectional Translation**
  - Rerank by L2R score + R2L score

RESULTS

- **English-German**
  - BPE
  - **MiniNMT**-IBE
  - Synthetic data
  - Ensemble
- **Chinese-English**
  - BPE
  - **MiniNMT**-SBE
  - Synthetic data
  - Ensemble & Reranking
- **English-Chinese**
  - BPE+Mix W/C
  - DL4MT
  - Synthetic data
  - Ensemble

	System	Test (BLEU)
EN-DE	Baseline	25.7
	+Synthetic	26.1
	+Ensemble	**26.7**
CH-EN	Baseline	23.1
	+Synthetic	23.7
	+Ensemble	25.3
	+R2L reranking	**26.0**
EN-CH	Baseline	30.4
	+Synthetic	34.3
	+Ensemble	**35.8**

MORE ABOUT US

- **Online NMT Demo**
- **Open-Source NMT System**
- **Contact Us**

XIAMEN UNIVERSITY
SCHOOL OF INFORMATION SCIENCE AND ENGINEERING

自强不息 止于至善
Pursure Excellence, Strive for Perfection.