DFKI Character-based Neural MT

Georg Heigold, **Jon Dehdari**, Josef van Genabith

DFKI

March 14, 2017

Quality Translation 21 (QT21) has received funding from the EUs Horizon 2020 research and innovation programme under grant no. 645452. META-QT has received funding from the EUs Horizon 2020 research and innovation programme through the contract CRACKER (grant agreement no.: 645357). Formerly co-funded by FP7 and ICT PSP through the contracts T4ME (grant agreement no.: 249119), CESAR (grant agreement no.: 271022), METANET4U (grant agreement no.: 270893) and META-NORD (grant agreement no.: 270899).

Using Sub-word Units in Neural MT

Morph-based:

- Morphs are induced morphemes
- BPE (Sennrich et al, 2016)

Character-composition-based:

- Ling et al (2015)
- Lee et al (2016)

Overlapping CNN+RNN-based:

- Kim et al (2016)
- Costa-jussà & Fonollosa (2016)

Dilated CNN-based

- Kalchbrenner et al (2016)

Character-based NMT at DFKI

- Replace lookup table with two-layer LSTM
- With tokenization: one such LSTM for each word
- Without tokenization: use sliding windows of size 20 and with stride 5 (approx. average word length)

WMT Results

Newstest2015, BLEU	**DE-EN**	**EN-DE**
Nematus (BPE, single, no backtrans)	26.4	22.4
Stanford (word, single, no backtrans)	24.9	
DFKI (BPE→BPE, single, no backtrans)	27.0	24.8
DFKI (char→BPE, single, no backtrans)	26.1	
DFKI (notok+char→BPE, single, no backtrans)	25.6	

WMT Results

Newstest2015, BLEU	**DE-EN**	**EN-DE**
Nematus (BPE, single, no backtrans)	26.4	22.4
Stanford (word, single, no backtrans)	24.9	
DFKI (BPE→BPE, single, no backtrans)	27.0	24.8
DFKI (char→BPE, single, no backtrans)	26.1	
DFKI (notok+char→BPE, single, no backtrans)	25.6	

Newstest2016, BLEU	**DE-EN**	**EN-DE**
Nematus (BPE, single, no backtrans)	28.5	26.8
DFKI (BPE→BPE, single, no backtrans)	31.6	29.8
DFKI (char→BPE, single, no backtrans)	30.7	
DFKI (notok+char→BPE, single, no backtrans)	29.9	

References I

Costa-jussà, M. R. and Fonollosa, J. A. R. (2016).

Character-based neural machine translation.
In *Proceedings of the ACL-2016*, pages 357–361, Berlin, Germany. Association for Computational Linguistics.

Kalchbrenner, N., Espeholt, L., Simonyan, K., van den Oord, A., Graves, A., and Kavukcuoglu, K. (2016).

Neural machine translation in linear time.
Unpublished manuscript.

Kim, Y., Jernite, Y., Sontag, D., and Rush, A. M. (2016).

Character-aware neural language models.
In *Proceedings of the Association for the Advancement of Artificial Intelligence (AAAI)*, Phoenix, AZ, USA. AAAI.

Lee, J., Cho, K., and Hofmann, T. (2016).

Fully character-level neural machine translation without explicit segmentation.
Unpublished manuscript.

Ling, W., Trancoso, I., Dyer, C., and Black, A. W. (2015).

Character-based neural machine translation.
Unpublished manuscript.

Sennrich, R., Haddow, B., and Birch, A. (2016).

Neural machine translation of rare words with subword units.
In *Proceedings of the ACL-2016*, pages 1715–1725, Berlin, Germany. Association for Computational Linguistics.

More Results: TR-EN

Newstest2016	**TR-EN BLEU**
Best WMT-2016 (additional data)	15.6
DFKI (BPE→BPE)	12.5
DFKI (char→word)	14.4
DFKI (+ensemble)	16.9