Sampling-based Alignment and Hierarchical Sub-sentential Alignment in Chinese–Japanese Translation of Patents

Wei Yang, Zhongwen Zhao, Baosong Yang and Yves Lepage

EBMT/NLP Lab of IPS, Waseda University (WASUIPS)

October 16, 2015
SMT Experiments

- Experimental results of SMT

| s→t       | Moses | Aligner | BLEU  | RIBES | Training time |
|-----------|-------|---------|-------|-------|---------------|
| zh→ja     | 2.1.1 | MGIZA   | 37.70 | 0.783000 | 5:34:28       |
|           | 2.1.1 | GIZA++  | 37.46 | 0.778914 | 4:43:56       |

**Table**: Evaluation results by using different aligner (GIZA++ and MGIZA) based on the data of JPC.

| Language | Moses | Anymalign + Cutnalign | BLEU | Training time |
|----------|-------|-----------------------|------|---------------|
| zh-ja    | 3.0   | 1200                  | 2 (c)| 36.11         | 1:2:8         |
|          | 3.0   | 5400                  | 2 (c)| 36.07         | 2:9:29        |
| zh-ja    | 2.1.1 | 1200                  | 2 (c)| 35.95         | 0:57:1        |
| zh-ja    | 2.1.1 | 1200                  | 2 (python)| 35.93 | 1:1:16        |

**Table**: Evaluation results by using the alignment method of combining sampling-based alignment and bilingual hierarchical sub-sentential alignment methods.
Thank you for listening.
ご清聴ありがとうございました。
谢谢大家。