

Japanese to English Machine Translation using Preordering and Compositional Distributed Semantics

Sho Hoshino, Hubert Soyer, Yusuke Miyao, Akiko Aizawa
National Institute of Informatics (NII), Japan

Summary

- Applied **preordering** and **RNNLM reranking** (both worked well in the NTCIR)
- Employed **neural network based tuning** in the Ja-En task for the first time

Contributions / Findings

- ASPEC corpus ($\tau=0.29$) is much more difficult than NTCIR corpus ($\tau=0.39$)
- Obtaining better word alignment ($\tau=0.37$) did not improve BLEU and RIBES

Three Components (Preordering, Tuning with Embeddings, and Reranking with RNNLM) and Intrinsic Evaluation

Preordering rules (Hoshino+ 2014)

Rule 1 (chunking):

merge **coordinated chunks** into one
ex: 表1と、図2は → 表1と図2、は

Rule 2 (inter-chunk reordering):

reverse chunks before/after topic-marker

$S \{は,が\} O V \rightarrow S \{は,が\} V O$
ex: 表1と図2は 次のものを表している
表1と図2は 表しているものを次の

Rule 3 (intra-chunk reordering):

postpositional phrase → prepositional

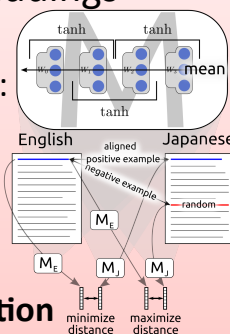
ex: 表している → いる 表して

Our System	BLEU	RIBES	Kendall's tau
Baseline	19.16	63.41	0.2990
+Preordering	18.55	61.44	0.3712

Tuning with Embeddings

Modified "BI" model by Hermann and Blunsom (2014):

- vector = 128 dimensions
- sentence = **mean** vector
- min. positive alignment
- max. negative alignment



Improved preordered translation

Preordering	Embeddings	BLEU	RIBES
	None	19.16	63.41
	Feature Function	18.95	63.48
	Phrase Table (ext.)	18.82	63.23
✓	None	18.55	61.44
✓	Feature Function	18.92	61.76
	Phrase Table (1 col.)	14.53	60.59

Reranking 100-best with RNNLM

- #classes = 220 (sqrt of vocabulary size)
- backpropagation-through-time steps = 5

Improved RIBES by trading off BLEU

Preordering	Configuration	BLEU	RIBES
	h = 300, d=150	18.37	63.90
	h = 300, d=100	18.40	63.74
	h = 200, d=100	18.61	63.72
	h = 200, d=150	18.30	63.66
	None (Baseline)	19.16	63.41
✓	h = 200, d=100	18.36	62.06
✓	h = 200, d=150	18.29	62.03
✓	h = 300, d=150	18.34	61.99
✓	h = 300, d=100	18.48	61.98
✓	None (+Preordering)	18.55	61.44

Extrinsic Evaluation

Our System	BLEU	RIBES	Human
Baseline	17.47	63.08	-5.750
+Preordering	17.01	61.08	-14.250

- negative: 197
- tie: 63
- positive: 139

Our intuition "better word alignment leads to translation improvement" did not hold in the official human evaluation.

Negative Effect (judged as -1 -1 -1)

ex: 論じた ついてに 義肢 装具 材料の 非 金属 製の プラスチック 以外。
out: The non - metallic materials of a prosthesis apparatus is made of plastics.

Positive Effect (judged as 1 1 1)

ex: 開発したを インライン 流量計 新しいの 標記。
out: **We have developed** a new in-line flowmeter.

Future Work

- Error Analysis
- Application for Embeddings
- Improvement

Contact: Sho Hoshino / National Institute of Informatics
Email : hoshino@nii.ac.jp