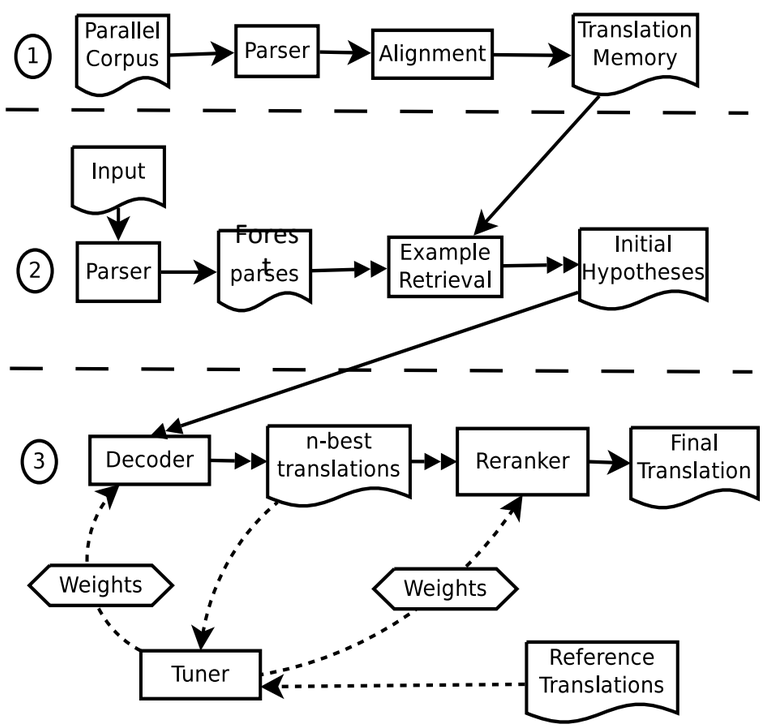
KyotoEBMT System Description for the 2nd Workshop on Asian Translation

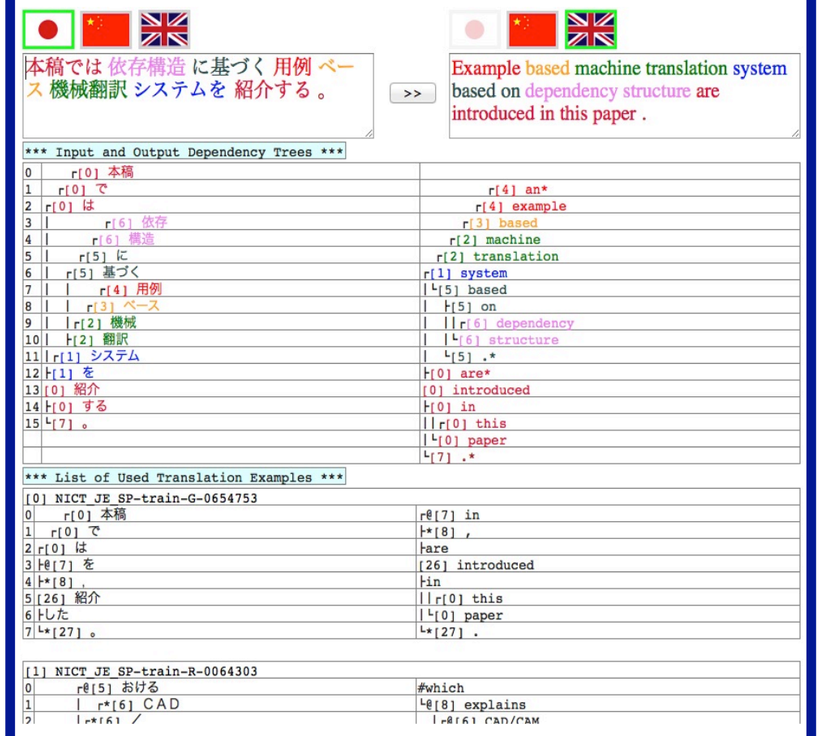
John Richardson john@nlp.ist.i.kyoto-u.ac.jp
Raj Dabre dabre@nlp.ist.i.kyoto-u.ac.jp
Chenhui Chu chu@pa.jst.jp
Fabien Cromières fabien@pa.jst.jp
Toshiaki Nakazawa nakazawa@pa.jst.jp
Sadao Kurohashi kuro@i.kyoto-u.ac.jp

Graduate School of Informatics, Kyoto University

KyotoEBMT System Pipeline



Web Interface of Translation



WAT2015 Official Results

Dependency Parsers

Ja: KNP [Kawahara and Kurohashi, 2006]
En: NLParser [Charniak and Johnson, 2005] with rules
Zh: SKP [Shen et al., 2012]

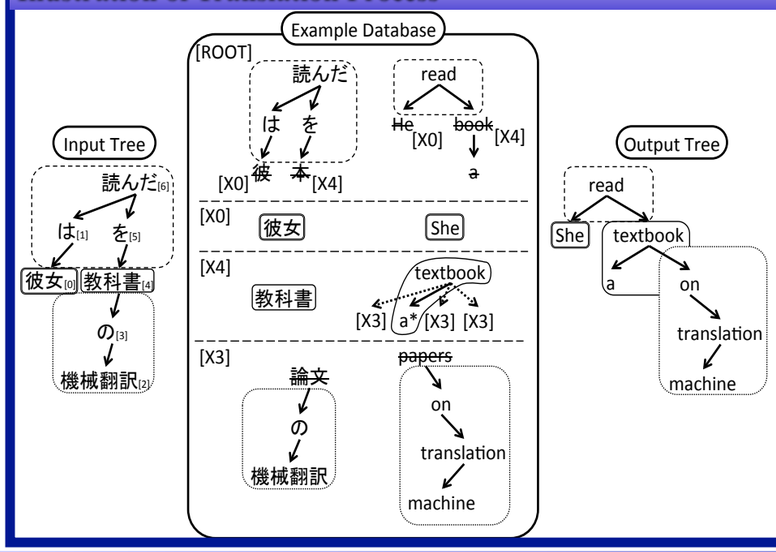
Reranking Features

7-gram language model with Modified Kneser-Ney smoothing
Recurrent Neural Network Language Model (hidden layer: 200) (Mikolov, 2011)
Bilingual RNN Language Model (Bahdanau et al., 2015)

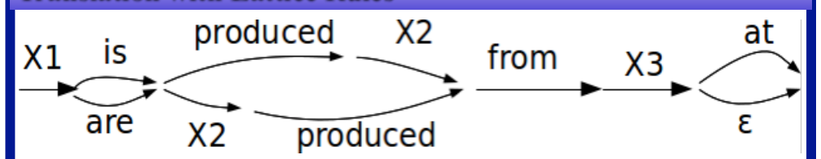
	Rerank	BLEU	RIBES	HUMAN
JE	NO	21.31 (+0.71)	70.65 (+0.53)	16.50
	YES	22.89 (+1.82)	72.46 (+2.56)	32.50
EJ	NO	30.69 (+0.92)	76.78 (+1.57)	40.50
	YES	33.06 (+1.97)	78.95 (+2.99)	51.00
JC	NO	29.99 (+2.78)	80.71 (+1.58)	16.00
	YES	31.40 (+3.83)	82.70 (+3.87)	12.50
CJ	NO	36.30 (+2.73)	81.97 (+1.87)	16.75
	YES	38.53 (+3.78)	84.07 (+3.81)	18.50

(Improvement over WAT2014 in parentheses)
Remark: For WAT2014, J->C was the only direction for which reranking was worsening BLEU and Human Evaluation. For WAT2015, J->C is still the only direction for which reranking worsens Human Evaluation (although it now does improve BLEU)

Illustration of Translation Process



Translation with Lattice Rules



[Cromieres and Kurohashi, EMNLP 2014]

Each path in this lattice corresponds to different choices of insertion position for X2, morphological forms of “be”, and the optional insertion of “at”.

- designed to handle an arbitrary number of non-terminals
- able to handle ambiguities of translation hypotheses
  - ➢ which target word is going to be used

Conclusion and Future Work

KyotoEBMT system

- source code available under a GPL license at http://nlp.ist.i.kyoto-u.ac.jp/kyotoebmt/ (version 1.0 just released!)
- uses both source and target dependency analysis
- online example retrieving
- availability of full translation examples at run time
- can use forest parses of input

Future work

- use a target-side tree language model
- online tuning of weights
- target-side structural features
- use of neural network language models in decoding