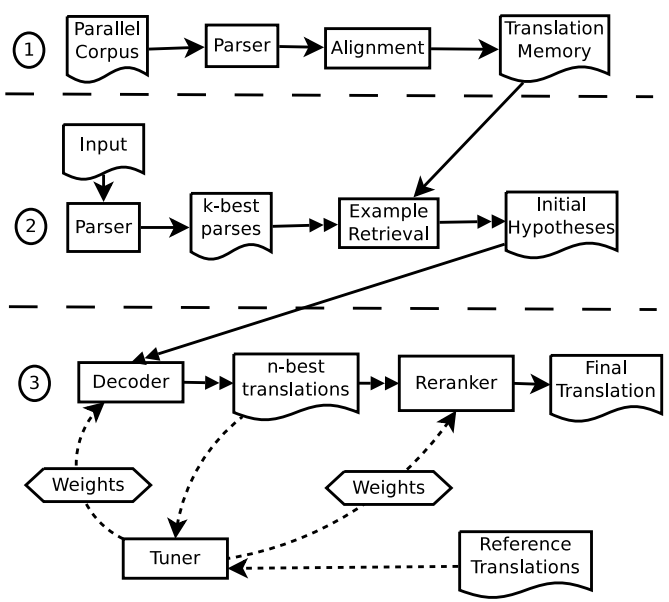
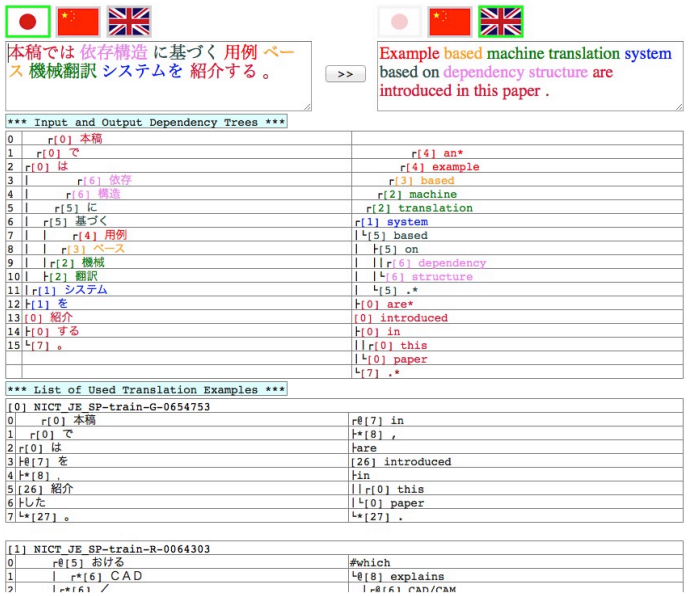
KyotoEBMT System Description for the 1st Workshop on Asian Translation

John Richardson john@nlp.ist.i.kyoto-u.ac.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



WAT2014 Official Results

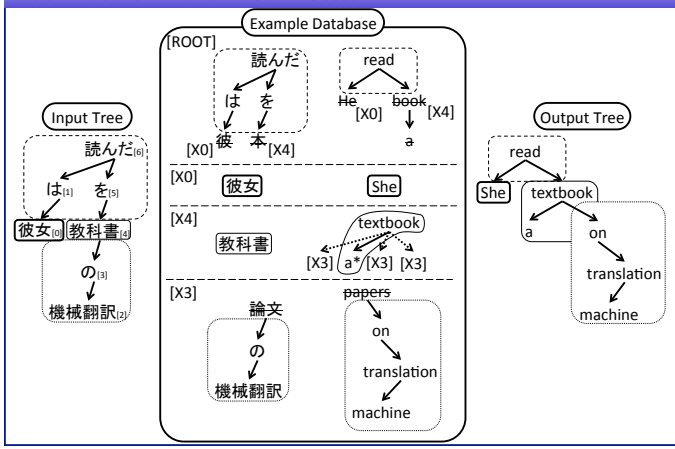
Dependency Parsers
- Ja: KNP [Kawahara and Kurohashi, 2006]
- En: NLParser [Charniak and Johnson, 2005] with rules
- Zh: SKP [Shen et al., 2012]

Reranking Features
- 20-best parses with its parsing scores
- 7-gram language model with Modified Kneser-Ney smoothing
- Recurrent Neural Network Language Model with hidden layer size 200

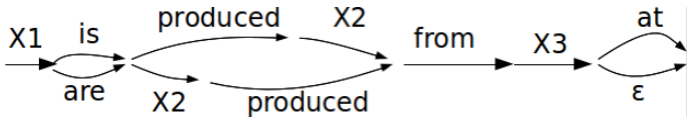
	Reranking	BLEU	RIBES	HUMAN
JE	NO	20.60	70.12	21.50
	YES	21.07	69.90	25.00
EJ	NO	29.76	75.21	33.75
	YES	31.09	75.96	38.00
JC	NO	27.21	79.13	-0.75
	YES	27.67	78.83	-8.75
CJ	NO	33.57	80.10	6.00
	YES	34.75	80.26	7.50

Findings
- Team-based rankings of HUMAN score were 2nd, 3rd, 5th and 4th
- Reranking works well for all the directions other than J->C
  - need more investigations

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
  - ➢ which will be the final position of each non-terminal

Conclusion and Future Perspective

KyotoEBMT system
- source code available under a GPL license at http://nlp.ist.i.kyoto-u.ac.jp/kyotoebmt/
- uses both source and target dependency analysis
- online example retrieving
- availability of full translation examples at run time

Future perspectives
- use of the input parse forest instead of k-best parses
- parsing the example parallel corpus as forests
- use a target-side tree language model as a decoding/ reranking feature
- online tuning of weights
- target-side structural features