

# KyotoEBMT System Description for the 1st Workshop on Asian Translation

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

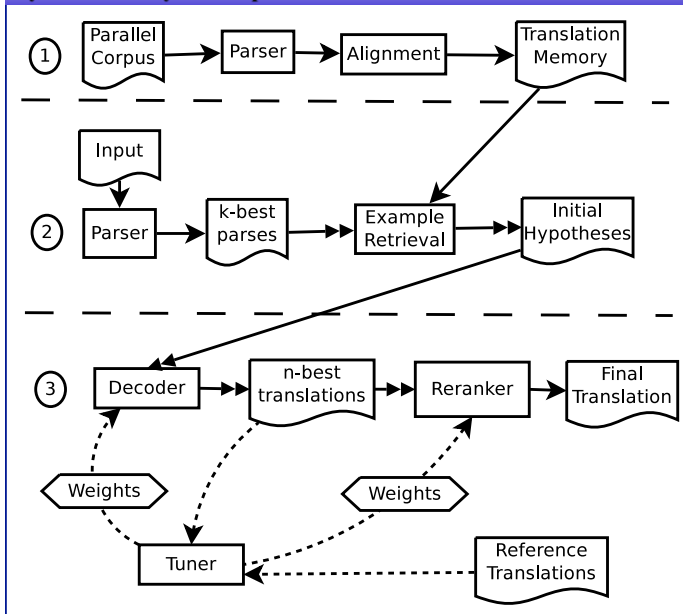
Fabien Cromières  
fabien@pa.jst.jp

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

Example based machine translation system based on dependency structure are introduced in this paper.

```

*** Input and Output Dependency Trees ***
0 | r[0] 本稿
1 | r[0] で
2 | r[0] は
3 | | r[6] 依存
4 | | r[6] 構造
5 | | r[5] に
6 | | r[5] 基づく
7 | | r[4] 用例
8 | | r[3] ベース
9 | | r[2] 機械
10 | | r[2] 翻訳
11 | | r[1] システム
12 | | r[1] を
13 | | r[0] 紹介
14 | | r[0] する
15 | | r[7] 。
    | | r[4] an*
    | | r[4] example
    | | r[3] based
    | | r[3] machine
    | | r[2] translation
    | | r[1] system
    | | r[5] based
    | | r[6] dependency
    | | r[6] structure
    | | r[5] *
    | | r[0] are*
    | | r[0] introduced
    | | r[0] in
    | | r[0] this
    | | r[0] paper
    | | r[7] .

*** List of Used Translation Examples ***
[0] NICT_JE_SP-train-G-0654753
0 | r[0] 本稿
1 | r[0] で
2 | r[0] は
3 | r[7] を
4 | r[8] .
5 | r[26] 紹介
6 | r[0] する
7 | r[27] 。
    | r[7] in
    | r[8] ,
    | here
    | r[26] introduced
    | r[8] in
    | r[0] this
    | r[0] paper
    | r[27] .

[1] NICT_JE_SP-train-R-0064303
0 | r[5] おける
1 | r[6] CAD
2 | r[6] /
    | r[8] explains
    | r[6] can/cam
    
```

## WAT2014 Official Results

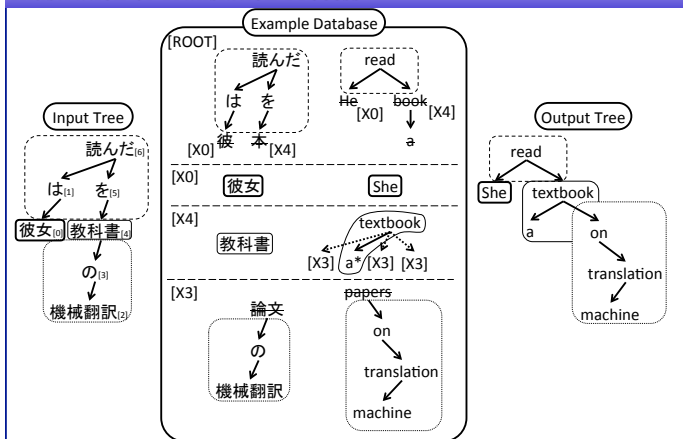
Dependency Parsers  
 Ja: KNP [Kawahara and Kurohashi, 2006]  
 En: NLParse [Charniak and Johnson, 2005] with rules  
 Zh: SKP [Shen et al., 2012]

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 2<sup>nd</sup>, 3<sup>rd</sup>, 5<sup>th</sup> and 4<sup>th</sup>  
 Reranking works well for all the directions other than J->C  
 need more investigations

## Illustration of Translation Process



## Translation with Lattice Rules

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