

Reordering Japanese for better translation

~ Translation using Japanese predicate argument structure ~

Abstract

Word order differences between Japanese and other languages cause serious translation errors by machine translation. We propose a method for reordering Japanese words into English or Chinese word order prior to machine translation, to make the translation easy. Our key idea is reordering Japanese “subject-object-verb” order into “subject-verb-object” one. We use our predicate argument structure analyzer to find subjects and objects for verbs, and use this information for re-ordering Japanese words. We can translate the re-ordered Japanese sentences into English/Chinese more easily and accurately than before. This kind of machine translation technologies will break language barriers in global communication.

<Japanese>

私は 京都に 行く
長い 尻尾の 猫を見た

Machine
Translation

*I to Kyoto go
Long tail of cat saw*

Error in word order

**Dependency &
① PAS analysis**

私は 京都に 行く
Subject Predicate

(私は) 長い 尻尾の 猫を見た
Subject Object Predicate

**② Reordering into
English(Chinese)
order**

私 行く に 京都
私 見た を 猫 の 長い 尻尾

**③
Translation**

I go to Kyoto

I saw a cat with a long tail

Correct word order

Key Technology①: PAS analysis [1]

- Find subjects and objects for predicates (verbs)
- High speed & accuracy with ellipsis resolution

Key Technology②: 2-stage reordering [2]

- Chunk-based reordering from SOV to SVO
- Word-based reordering for prepositions

Key Technology③: Statistical MT [3]

- Learn translation patterns from translation data
- Easily solvable with (almost) monotone translation

* This is a collaborative work with Dr. Y. Miyao at National Institute of Informatics (NII)

Method	RIBES	BLEU
Baseline	68.48	29.19
This work	72.26	30.65

Order-sensitive
Evaluation Metric
(developed in NTT)

Standard
Evaluation
Metric

Related work

- [1] H. Taira and M. Nagata, “A study of Japanese ellipsis resolution with predicate argument structure analysis,” in *Proc. ANLP*, 2013 (In Japanese)
- [2] S. Hoshino, Y. Miyao, K. Sudoh, M. Nagata, “Two-stage pre-ordering for Japanese-to-English statistical machine translation,” in *Proc. IJCNLP*, pp. 1062-1066, 2013
- [3] M. Nagata, K. Sudoh, J. Suzuki, Y. Akiba, T. Hirao, H. Tsukada, “Recent innovations in NTT’s statistical machine translation,” *NTT Technical Review*, Vol. 11, No. 12, 2013

Contact

Katsuhito Sudoh Linguistic Intelligence Research Group, Innovative Communication Laboratory
E-mail : sudoh.katsuhito[at]lab.ntt.co.jp (Please replace {at} with @)