

## 12

## 日本語を詳細に読み解けばもっと上手に訳せる！

～深い構文解析による事前並べ替え機械翻訳～

## どんな研究

英語や中国語は日本語と語順が異なります。語順は文法で決まるため、各言語の文法に基づいた構文解析が語順の正しい翻訳のためにとても重要です。この研究は、日本語を翻訳する時に、日本語の構文解析を利用して語順を外国語に近くなるようにあらかじめ並べ替えて、機械翻訳しやすくする方法に関するものです。

## どこが凄い

日本語と英語や中国語は主語・述語・目的語や修飾語等の順番が異なります。私たちは、日本語の構文解析で各単語間の係り受けとその役割を同時に求め、係り受けの役割に応じて語順を並べ替えるようにしました。これにより、役割の区別がない文節単位の係り受けを用いた場合より柔軟な並べ替えを可能にしました。

## 目指す未来

日本語と外国語の間の双方向の機械翻訳を高精度に行うことで、言語の壁を越えて情報収集、情報発信、コミュニケーションを円滑にします。今後はより多くの対象分野、言語で高精度の翻訳を行うことができるように、言語の解析と翻訳の研究を進めていきます。

## 【日本語】

彼は長い尻尾の猫を見た

昨年の方式

彼は 長い 尻尾の 猫を 見た  
主語 目的語 述語

一般的な方式

語順を探索して並べ替え

He is a long tail of cat saw

よく語順を間違える！

最新の方式

問題点① 係り受け役割が不明  
(別処理で主語・目的語を推定)  
問題点② 文節単位の処理  
(文節を崩す並べ替えは複雑に)

は 彼 見た を 猫 の 長い 尻尾  
主語 述語 目的語

## ①単語間の係り受けと係り受け役割の解析

主語  
連体 連体 目的語 助動詞  
彼は 長い 尻尾の 猫を 見た

## ②外国語の語順に並べ替え

主語 助動詞 目的語 連体 連体  
は 彼 見た 猫 の 長い 尻尾

## ③翻訳(ほぼ逐語訳)

He saw a cat with a long tail

正しい語順で訳せる！

技術のポイント①: 係り受け解析<sup>[1]</sup>

- 単語間の係り受け関係と係り受け役割を求める
- 従来日本語では役割を考慮しない文節単位処理
- 英語中国語等と同様の情報を高精度に解析可能

## 技術のポイント②: 単語「事前」並べ替え

- SOV語順を英・中のSVO語順に
- 形容詞・助詞・助動詞などの語順調整

技術のポイント③: 統計的機械翻訳<sup>[2]</sup>

- 大量の対訳データを利用して翻訳規則を学習
- 並べ替えが(ほぼ)不要なので楽に解ける！

## 関連文献

- [1] 田中貴秋, 永田昌明, “日本語のラベル付き依存構造解析の検討,” 言語処理学会第21回年次大会発表論文集, 2015.  
[2] 永田昌明, 須藤克仁, 鈴木潤, 秋葉泰弘, 平尾努, 塚田元, “革新的発展期を迎えた統計翻訳,” NTT技術ジャーナル, Vol. 25, No. 9, 2013.

## 連絡先

須藤克仁 (Katsuhito Sudoh) 協創情報研究部 言語知能研究グループ  
E-mail: sudoh.katsuhito(at)lab.ntt.co.jp

