

「私、行く、に京都」の方が訳しやすいんです！

～日本語の述語項構造(SVO)を利用した日英・日中翻訳～

どんな研究

英語や中国語は日本語と語順が異なるため、翻訳するときに正しい語順になるように並べ替える必要があります。この研究は、日本語を英語や中国語に翻訳するときに、日本語の語順を英語や中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ためのものです。

どこが凄い

日本語は主語・目的語・述語(SOV)の語順、英語や中国語は主語・述語・目的語の語順なので、**日本語の述語を主語と目的語の間に移動**します。そのために、述語と主語・目的語の関係を見つける「**述語項構造解析**」を高精度に行うことで、翻訳精度を高めることができました。

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】

私は 京都に 行く
長い 尻尾の 猫を見た

機械翻訳

*I to Kyoto go
Long tail of cat saw*

よく語順を間違える！

係り受け解析・ ①述語項構造解析

私は 京都に 行く
主語 述語

(私は) 長い 尻尾の 猫を 見た
主語(省略) 目的語 述語

②英語or中国語の 語順に並べ替え

私 行く に 京都
私 見た を 猫 の 長い 尻尾

③翻訳 (ほぼ)逐語訳

I go to Kyoto
I saw a cat with a long tail

正しい語順で訳せる！

技術のポイント①: 述語項構造解析 [1]

- 述語に対応する主語や目的語を見つける
- 主語等の省略に対応した高速高精度な方式を考案

技術のポイント②: 二段階の単語並べ替え [2]

- SOV語順を英・中のSVO語順に(文節単位)
- 助詞等を文節先頭に移動 or 削除(単語単位)

技術のポイント③: 統計的機械翻訳 [3]

- 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！

* この研究は国立情報学研究所の
宮尾祐介教授らと共同で行ったものです

手法	RIBES	BLEU
従来手法	68.48	29.19
本手法	72.26	30.65

語順に注目した
翻訳評価尺度
(NTTが考案)

語句の翻訳に
注目した
翻訳評価尺度

関連文献

- [1] 平博順, 永田昌明, “述語項構造解析を伴った日本語省略解析の検討,” 言語処理学会第19回年次大会, 2013.
[2] 星野翔, 宮尾祐介, 須藤克仁, 永田昌明, “日英統計的機械翻訳のための述語項構造に基づく事前並べ替え,” 言語処理学会第19回年次大会, 2013.
[3] 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 ({at}の部分をも@に置き換えてください)