

Translation: Final Frontier for Artificial Intelligence?

永田昌明

(NTT コミュニケーション科学基礎研究所)

コンピュータを用いて人間の言葉のある言語から別の言語へ変換する機械翻訳の研究は 1950 年代から今日まで約 70 年の歴史がある。1950 年代から 2005 年頃までは、専門家が翻訳規則を作成する「ルールベース翻訳」が主流であった。しかし、ルールベース翻訳は人手作業による精度の限界に到達し、2000 年代に入ると、対訳データから翻訳規則に相当する翻訳モデルを自動的に学習する「統計的機械翻訳」(SMT, Statistical Machine Translation) が主流となった。統計的機械翻訳は、英語とフランス語のような語順が近い言語間では早くから実用レベルの精度が得られたが、日本語と英語のような語順が大きく違う言語間では翻訳の精度が低いという問題があった。しかし 2010 年頃から翻訳元の言語の文を翻訳先の言語の語順に変換してから統計翻訳を行う「事前並べ替え」の研究が進み、日本語と英語の翻訳精度も改善が進んでいた。ところが 2014 年に「ニューラル機械翻訳」(Neural Machine Translation) と呼ばれる、ニューラルネットを用いた従来とは全く異なる機械翻訳の手法が考案され、2015 年には英語とフランス語および英語とドイツ語の翻訳で統計翻訳の精度を超え、2016 年には日本語と英語の翻訳でもニューラル翻訳の精度が従来の統計翻訳の精度を大きく上回るようになった。

今回のシンポジウムの講演では、Mark Jewel 先生(早稲田大学)が課題となる日本語文(81 文)とその Google 翻訳による英語訳を提示し(Google は 2016 年 8 月時点では統計翻訳を使用)、小田先生(毎日新聞)が人手による英語訳(理想訳)を作成した。これに対して筆者は、ニュース記事や書籍から収集した約 100 万文の日英対訳データを用いて独自にニューラル翻訳システムを作成し、統計翻訳(Google 翻訳)とニューラル翻訳の精度の比較を行った。

一般に、ニューラル翻訳は母語話者に匹敵する流暢な文を生成するが、必ずしも翻訳元の言語の文の内容と忠実に反映しない。同じ内容を何度も繰り返したり、原文の一部を無視して訳出しないことがある。ときには原文とは全く関係ない内容の訳文を生成することもある。これに対して統計翻訳は文法的に正しくない文を生成することが多いが、だいたい原文の内容を反映した訳文を生成する。一言でまとめると、統計翻訳は「不器用だが努力を感じる逐語訳」、ニューラル翻訳は「流暢だが怪しい意識」を生成する。

現在、機械翻訳技術は技術革新の真っ最中であり、研究レベルでは既にニューラル翻訳が主流となっている。やがて実用レベルでもニューラル翻訳が主流になり、皆さんにお使い頂けるようになると思う(Google は 2016 年 11 月にニューラル翻訳のサービスを開始した)。しかし、人間並みの翻訳が実現できるのはまだまだ先の話であり、機械翻訳が人間の翻訳者の仕事を奪うことはない。例えば、ニューラル翻訳の翻訳結果を人間が後編集するなど、技術の変化について行けば、人間の仕事は決してなくなるならない。