

3.14 翻訳

Graham Neubig (奈良先端科学技術大学院大学)

工藤 拓 (Google (株)) 赤部 晃一 (奈良先端科学技術大学院大学)

機械翻訳の誤り分析

日本語から英語へと自動的に翻訳する機械翻訳では、最先端の翻訳機は何ができるのか？ 何ができないのか？ 本稿は、この疑問に答えるべく、翻訳の誤り分析を行った結果を報告する。

機械翻訳には、さまざまな手法がある。たとえば、人手で翻訳規則を記述するルールベース翻訳 (RBMT) やデータから自動構築する統計翻訳 (SMT) が挙げられる。SMT では単語列を置き換えて並べ替えるフレーズベース翻訳 (PBMT) や文の構造を解析して翻訳を行う統語ベース翻訳 (SBMT) が存在する。本調査では、特定のシステムに分析結果が偏らないよう、下記の通り3つの商用システム、3つのオープンソースソフト (OSS) システムを評価した：

RBMT：商用のルールベース翻訳システム

SMT1, SMT2：商用の統計翻訳システム2種類

Moses：OSS のフレーズベース翻訳システム

KyotoEBMT, Travatar：OSS の統語ベース翻訳システム2種類

分析の対象として、ブログやQA サイト、白書、法律文書など、さまざまな分野をカバーするデータを利用し、OSS のシステムの学習には約 300 万文を用いた。日本語が原文であり、上記の翻訳システムで英語へと翻訳した。また、比較のために文を人間の翻訳者 (Human) にも翻訳してもらった。

この翻訳結果に対して、まず 0～6 の間の7段階評価で翻訳の正確さの総合的な数量評価を行った。次に、全システムの数量評価が低かった文から、図-1 に示す翻訳誤り分類体系を用いて誤りの分類を行い、各種類の翻訳誤りの割合を定量化した。

各システムの総合評価

まず、各システムの総合評価を、評価値 4～6 (良い)、2～3 (理解可能)、0～1 (理解不可能) の内訳として図-2 に示す。この結果から、RBMT が最も精度が高く、商用システムの SMT1 と文の構造を考慮する OSS システム Travatar が次ぐ。しかし、いずれのシステムでも、3割以上の文は理解不能となっており、課題が残ることも分かる。

一方、人手による翻訳はすべての機械翻訳システムを大幅に上回っていることが分かるが、0か1と評価された文も一部あり、人手翻訳でも厳しい評価がされている。低い評価値になった文を分析したところ、原因は主に1) 人手翻訳は文脈を用いて行っているが、翻訳の評価は文脈を考慮しておらず、翻

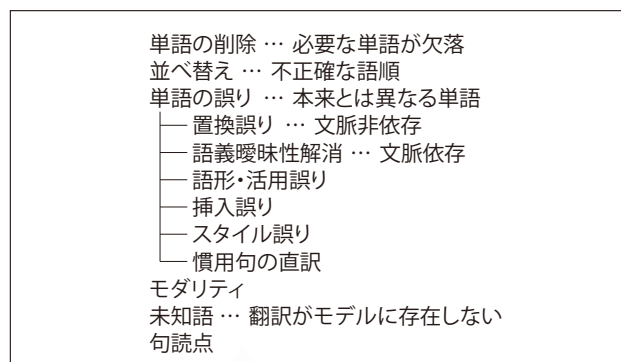


図-1 分析対象の誤りの種類

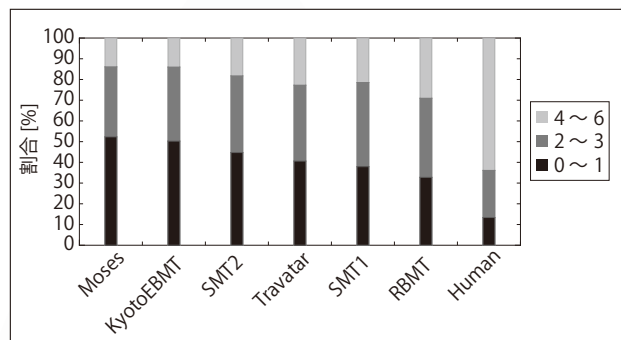


図-2 各システムの評価値の内訳

訳時と評価時に差が生じたこと、2) 単純な翻訳誤り、もしくは3) 原文がそもそも曖昧で、直訳することが難しいことに起因した。

各システムの誤り傾向

次に、誤りの種類の内訳を図-3に示す。この中で、「単語の誤り」が最も多く、「並べ替え誤り」「単語の削除」も多く存在した。

削除誤りに目を向けると、Moses, KyotoEBMTとSMT2, Travatar, SMT1, RBMTの順に減っていくことが分かる。これは人手評価の順と同等であり、単語の削除はシステムの手評価に比例する。特に内容語の削除によって文の意味が損なわれるため、直感に合った結果であるといえる。

並べ替えに関して、日本語から英語への翻訳では文法の構造が異なるため、一般的に難しいとされている。しかし、その割には今回の結果で並べ替えの誤りが比較的少ない。今回の分析対象の文が短いこともあるが、最先端の翻訳システムは文の構造を考慮することで以前に比べて並べ替えに頑健であるともいえる。

次に、図-3(b)に、単語に関する誤りの詳細な分析結果を示す。中では、文脈に依存する語義曖昧性誤り(例:「米」は「米国」の「U.S.」か食物の「rice」に翻訳できる中、誤った意味を選択した誤り)や文脈に依存しない置換誤り(学習の失敗により「米」を「the」などの関係ない単語へ翻訳する誤り)が最も多かった。文脈に依存しない置換誤りに着目すると、OSSのシステムに比べて商用システムは置換誤りが少ないことが分かる。これは、OSSのシステムに比べて、商用システムが大規模な学習データを利用しており、誤った翻訳ルールを学習する比率が少ないためであると考えられる。また、RBMTはほぼ文脈依存の置換誤りを起こさず、安定した翻訳ルールを用いていることが分かる。その一方、全システムで語義曖昧性による誤りが多く

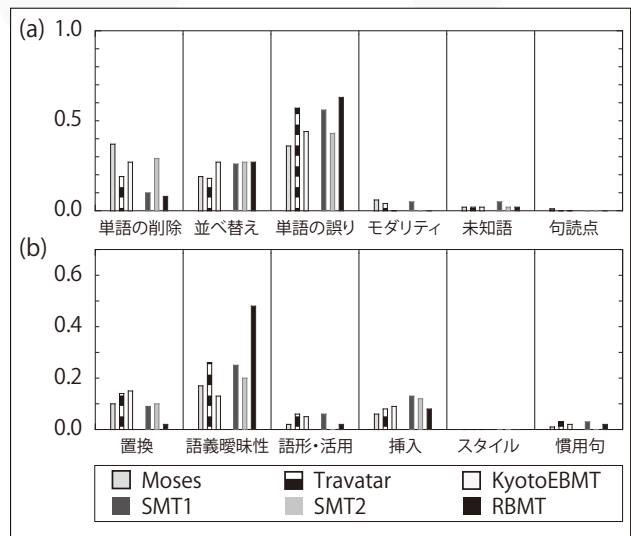


図-3 (a) 全誤りの傾向, (b) 単語誤りの傾向

見受けられた。RBMTの誤りは特に多く、SMTシステムは統計情報を用いて周りの文脈で曖昧性を解消しているのに対して、RBMTシステムはこのような統計情報を取り入れていないことが原因として考えられる。

機械翻訳の目指すべき方向

今回の分析では、機械翻訳にはさまざまな課題があるが、曖昧な語彙の選択が特に大きな課題であることが明らかとなった。今後、既存の語義曖昧性解消技術の適応、もしくは機械翻訳独自の曖昧性解消法で解決していくことが重要であろう。

(2015年9月8日受付)

Graham Neubig (正会員) neubig@is.naist.jp

イリノイ大学工学部計算機科学専攻卒業。京都大学大学院情報学研究科博士課程修了。奈良先端科学技術大学院大学助教。自然言語処理に関する研究に従事。

工藤拓 taku@google.com

Google (株) ソフトウェアエンジニア。京都大学工学部卒業、奈良先端科学技術大学院大学博士後期課程修了、NTTコミュニケーション科学基礎研究所リサーチアソシエイトを経て現職。専門は統計的自然言語処理、機械学習、データマイニング。

赤部 晃一 akabe.koichi.zx8@is.naist.jp

2015年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。同大学院博士後期課程在学中。機械翻訳、自然言語処理に関する研究に従事。