

英文和訳評価における 類義語変換を組み入れた評価値算出方式

小林 健太[†] 坂部 司[†] 秋吉 政徳[†]

神奈川大学[†]

1 はじめに

近年 e-learning などの英語学習のシステム化が進んでいる中、学習者の英語和訳の採点には教員が必須であるという状況が続いている。

一方、機械翻訳自動評価法の一つである RIBES [1]では、値を算出する過程で参照文(正しい翻訳文)と機械による翻訳文の単語の出現順序を考慮した一致率を計算に用いているが、単語そのものの一致が前提となっているために同義語が使われていると RIBES 値が低くなるという問題が存在する。

そこで、類語辞書や日本語コーパスによる単語検索によって、翻訳文の置き換えをすることで RIBES 値を向上させることで、この問題を解決しようと試みた。

本稿では、学習者の英語和訳をよりよく評価できるように、翻訳文の類義語置き換えによる評価指標の算出方式を提案する。

2 提案方式

2.1 アプローチ

RIBES(Rank-based In-tuitive Bilingual Evaluation Score)は語順相関に基づく評価法を採用した自動評価システムである。算出方法は参照文を単語ごとに分け、先頭から順番に番号をつける。その後、翻訳文のなかで参照文と同じ単語に同じ番号を割り当て、それぞれの文の単語の出現順序の近さを Kendall の順位相関係数から算出している。RIBES 評価値の算出例を表1に示す。

表1: 同じ意味の単語が使われる例

種別	例文	RIBES 値
参照文	綺麗な/あの娘/は/ 木/と/比べて/非常に/低い	-
翻訳文	美しい/あの娘/は/ 樹/と/比較して/とても/低い	0.78

A Calculation Method with Synonyms Conversion for English-Japanese Translation Evaluation

[†] Kenta Kobayashi, Tukasa Sakabe, Masanori Akiyoshi
Kanagawa University

表1内の例に示すように、「非常に」と「とても」は同じ意味の単語であるが、単語自体は一致していないために、RIBES 値も低くなる。そこで、単語は違っても同じ意味なら同じ評価となるように、類義語辞書による類義語検索を利用し翻訳文の単語の置き換えを行う。

2.2 方式の構成

図1に提案方式の構成を示す。

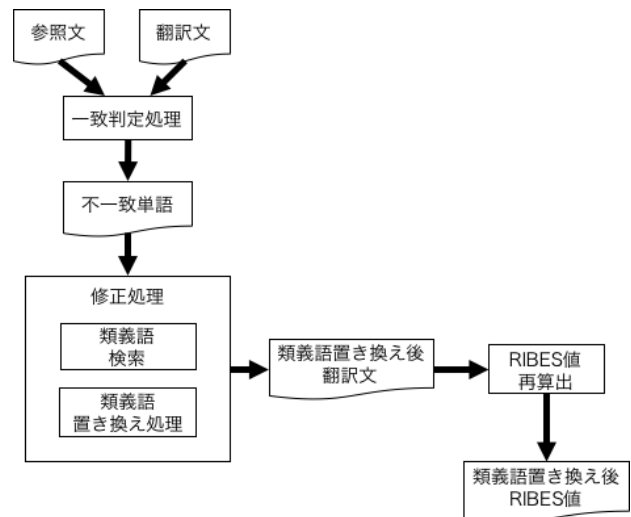


図1: 提案方式の構成

参照文と翻訳文の一致判定処理を行い、不一致単語を日本語 WordNet[2]で類義語検索し、類義語であることが確認できた場合に、翻訳文の単語を置き換えて新たに RIBES 値を算出する。

3 実験と考察

3.1 実験方法

300 組の英文とその日本語訳文(参照文)を集め、擬似的に学習者の英語和訳として用いるために、英文を三種類の翻訳サイト(google 翻訳, @nifty 翻訳, bing 翻訳)を用いて翻訳する。それらの翻訳文をどの翻訳が文として一番意味が近いかという観点で、予め人手で順位付けする。その後、RIBES 値と類義語置き換え RIBES 値を算出し、それぞれの RIBES 値の順位を前出の順位と比較し評価が向上できたかを調べる。

3.2 実験結果

類義語置き換えを使い RIBES 値の算出を試みた。

結果を表2と表3に示す。

表2: RIBES 値と類義語置き換え RIBES 値の比較

	文の数
翻訳文の数	900
評価できた翻訳数	499
評価値が上がった翻訳文	54
評価値が下がった翻訳文	8

表3: RIBES 値による順位の人手による順位との比較

	文の数
RIBES 値の変動があった参照文	47
人手による順位に近づいた参照文	9
人手による順位から遠ざかった参照文	4

表2に示すように三種類の翻訳文の合計である900文のうちプログラムが評価できた数は約半数の499文あり、RIBES 値が上がった文の数は10%程度である。また RIBES 値が下がった文の数は2%以下である。

表3は参照文とそれぞれの翻訳文の人手順位と RIBES 値の順位を比較したデータである。3つの翻訳文のうち一つでも RIBES 値が変動していた参照文の数は300のうち約15%の47あり、その中で、人手順位に近づいた文の数は20%弱存在し、また人手順位から遠ざかった文の数は10%弱である。

3.3 考察

評価できた翻訳文の数に対して評価値が変動している翻訳文の割合が少ない。これは用言の活用形の単語や丁寧語など類義語辞書に載っていない単語の変換ができていないことが考えられる。この結果、評価値の変動している翻訳文が少なくなったため、順位の変動の数も少なくなっている。

3.4 方式の改良

類義語辞書に載っていない単語の変換については、コーパスを利用することを考える。コーパスによる単語の結果を類義語辞書にはない新しい擬似的な類義語として判断できれば、置き換えを行うことで RIBES 値の向上が期待できる。

さらに、類義語検索処理で使われ方の違う単語を正しく置き換えることができない場合がある。例えば、「すごく」と「ひどく」は類義語だが、その後に「可愛い」という単語が来る場合、「すごく可愛い」は使われるが「ひどく可愛い」は使われない。このように類義語でも日本語として使われない単語のつながりが正しく置き換えられない。そこで、一致判定処理で一致した単語の前後に不一致の単語がある場合、まず一致した単語を日本語の書き言葉均衡コーパス[3]で検索する。その検索結果の中で、翻訳文の不一致単語があるならば、類義語辞書による置き換えに処理を進める。

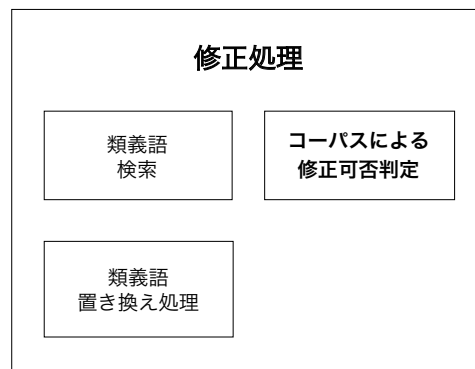


図2: 修正可否判定の追加

このように単なる類義語辞書による置き換えでない処理を組み入れることで、評価値の精度の向上が図られると考える。

4 まとめ

本稿では、学習者の英語和訳をよりよく評価できるように、翻訳文の類義語置き換えによる評価指標の算出方法を提案した。現状の手法のままでは類義語辞書が活用形や丁寧語に対応していないなどの理由で、評価値の向上が制限されてしまうことがわかった。今後の検討課題として、手法としてコーパスの検索による単語置き換えが RIBES 値にどのような影響を与えるかを考慮しながら、コーパスをうまく活用できる手法を検討し、文章の評価値の精度を上げることで人手による順位に近づけていく予定である。

参考文献

- [1] 平尾努, 磯崎秀樹, 須藤克仁, DuhKevin, 塚田元, 永田昌明, “語順の相関に基づく機械翻訳の自動評価法”, 自然言語処理, Vol.21 No.3, pp.421-444 (2014)
- [2] 国立研究開発法人情報通信研究機構 (NICT), NICT 情報分析研究室, 日本語 WordNet, <http://nlpwww.nict.go.jp/wn-ja/>
- [3] 国立国語研究所, コーパス開発センター, 現代日本語書き言葉コーパス, http://pj.ninjal.ac.jp/corpus_center/bccwj/