

# 統計的後編集による英日・中日・韓日特許翻訳の精度向上

Improvement of English/Chinese/Korean-Japanese patent translation by Statistical Post-Editing.

株式会社東芝 研究開発センター **園尾 聡**

**PROFILE** 2009年九州工業大学大学院生命体工学研究科博士課程修了。博士（工学）。同年株式会社東芝入社。AAMT/Japio 特許翻訳研究会 拡大評価部会メンバー。自然言語処理の研究に従事。

✉ satoshi.sonoo@toshiba.co.jp

株式会社東芝 研究開発センター **木下 聡**

**PROFILE** 1985年東京工業大学大学院理工学研究科修士課程修了。同年株式会社東芝入社。自然言語処理の研究に従事。

## 1 はじめに

外国語で書かれた大量の特許文書を検索・調査するために機械翻訳技術が広く用いられている。機械翻訳技術には、大きく2つの枠組みがある。1つは、長年研究開発が続けられてきた、辞書や訳し分け知識を記述した規則に基づくRBMT（規則ベース機械翻訳）である。もう一つは、対訳コーパスを利用したSMT（統計的機械翻訳）であり、近年盛んに研究が行われている。SMTの一手法であるフレーズベースSMTでは、対訳コーパスから対訳となるフレーズを統計的に抽出して翻訳モデルを構築し、翻訳を行う。特許翻訳においては、パテントファミリーから大規模な対訳特許コーパスを比較的容易に構築することが可能であり、英日特許翻訳タスクにおいて、SMTがRBMTに比べて高い翻訳精度を示している[1]。

一方、RBMTの翻訳結果に対してSMTの枠組みを用いて後編集を実現するSPE（統計的後編集）を適用することで、特許翻訳や科学技術論文の翻訳タスクにおいて自動評価スコアが大幅に改善することが報告されている[2,3]。

本稿では、RBMTにSPEを適用するメリットについて紹介し、特許翻訳の自動評価および人手評価による分析結果について報告する。

## 2 統計的後編集

SPEによる翻訳方式の概要を図1に示す。はじめに、対訳コーパス中の原言語文をRBMTによって翻訳し、目的言語による翻訳結果文を得る。次に、翻訳結果と参照訳文を対訳コーパスとみなして、翻訳結果から参照訳文への変換を行う翻訳モデルおよび目的言語による言語モデルを学習によって構築する。学習は、従来のSMTと同じ枠組みを用いることができる。翻訳実行時には、RBMTによる翻訳結果に対して、あらかじめ構築したモデルを適用することによって、後編集された翻訳結果を得ることができる。図1の例では、RBMTによって中国語から日本語に翻訳し、SMTによって（翻訳された）日本語から日本語への翻訳（後編集）を行っている。

ある分野において十分な対訳コーパスが利用可能であれば、RBMTに対してSPEを適用することで以下のような効果が得られる。

### (1) 未知語・誤訳の修正

RBMTにおいて未知語や訳し分け誤りであったとしても、該当する訳語を参照訳文に対応した訳語に修正し、翻訳精度が向上する。例えば、図1のRBMT翻訳結果の「持っている」という訳語を、特許文書により適した表現である「有する」へと修正することが出来る。

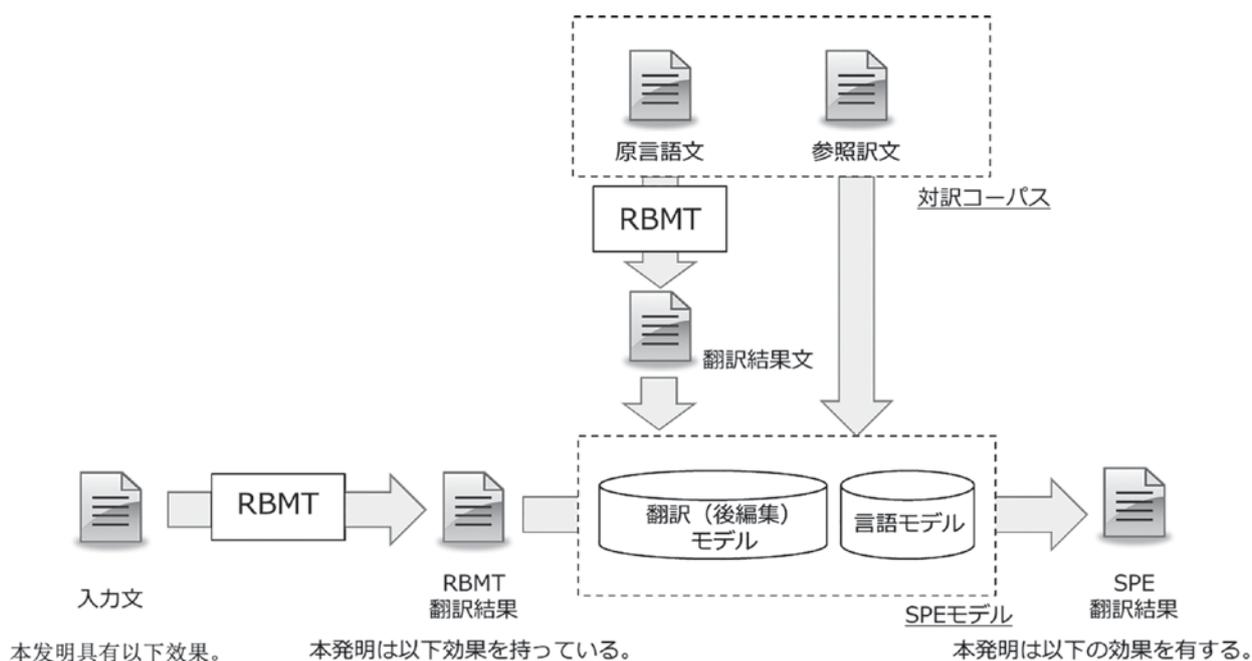


図1 SPE (統計的後編集) の概要

## (2) 翻訳結果の自然性の向上

助詞等の追加・削除や、特定の言い回し（フレーズ）をより自然な表現に変換することで、RBMT が課題としている翻訳結果の流暢さが改善される。図1の例では、「以下効果」に対して助詞の追加を行い、「以下の効果」と修正され、より自然な訳文となっている。

## (3) 翻訳処理失敗文の補正

例外として、規則の不備や誤字・脱字等の原文の誤りが原因でRBMTの翻訳処理が失敗したとしても、同様の失敗例が学習データ中に存在すれば、翻訳文のある程度の補正が可能となる。

スのモデルを構築した。テスト文は、各翻訳方向ともに2000文を学習用とは別に用意し、自動評価には2000文全て、人手評価には100文を用いた。

## 3.2 自動評価

機械翻訳の自動評価において一般的に用いられているBLEU[9]による評価結果を表1に示す。英日、中日ではSPEが最も高いスコアを示し、韓日ではSMTおよびSPEが高いスコアを示している。BLEUスコアは、参照訳文との一致度を示す指標であり、各翻訳方向において、SPEによってRBMTの翻訳結果が参照訳文により近い翻訳結果へと後編集されたことが分かる。

表1 自動評価結果

翻訳方向	システム	BLEU(%)
英日	RBMT	28.60
	SMT	37.18
	SPE	46.62
中日	RBMT	27.19
	SMT	38.60
	SPE	39.95
韓日	RBMT	51.40
	SMT	70.57
	SPE	68.71

## 3 特許翻訳におけるSPEの効果

### 3.1 実験環境

英日・中日・韓日特許翻訳において、RBMT、SMT、SPEの翻訳精度の比較を行った。RBMTは商用システム[4]と同等のものを使用し、中日方向においては220万語の専門用語辞書[5]を追加した。SMTおよびSPEの学習には、英日は200万文対([6]の一部)、中日および韓日は100万文対([7])の対訳コーパスから、Moses-toolkit[8]を用いてフレーズベ



### 3.3 人手評価

人手評価は、Adequacy および Acceptability[1] によって評価した。英日、中日および韓日の人手評価結果をそれぞれ図2、3、4に示す。Adequacy は、原文の内容をどの程度正確に伝達しているかの指標で、5段階評価で数値が高いほど正確となる。Acceptability は、フローチャートに従い、文法的で内容が理解可能な訳文 (AA,A)、非文法的であるが内容が理解可能な訳文 (B,C)、内容が理解できない (F) として、評価ラベルを付与する。

英日方向では、Adequacy の平均値が RBMT3.53、SMT2.12、SPE3.93 と、SPE が他のシステムより優位であることを確認した。なお、図中の \* は、ウィルコクソンの符号順位検定による有意差 ( $p<0.05$ ) を示す。また、Acceptability において、内容が理解できるレベル (C 以上) の割合が RBMT81.8%、SMT13.1%、SPE87.9% と、SPE の内容理解度が向上していることが分かる。

中日方向においても同様に、Adequacy の平均値が RBMT2.95、SMT3.02、SPE3.34、内容理解度が、RBMT39.8%、SMT38.8%、SPE43.9% と、SPE

の有効性を確認した。

韓日方向では、言語が類似していることから、各システムともに高い評価結果となっているが、SPE によって RBMT の評価結果が改善していることを確認した。

また、Acceptability の分布に注目すると、各翻訳方向において、SPE は RBMT と比べて自然な訳文 (AA) の割合が高くなっており、RBMT が課題としている翻訳結果の流暢さが改善していることが分かる。

### 3.4 翻訳例

表2に RBMT および SPE の翻訳結果の一例を示す。EJ-1 では、「展示スペース」がより適切な「表示領域」に修正されている。EJ-2 では、不自然な訳出が修正され、内容理解度および可読性の高い翻訳結果となっている。また EJ-3 の RBMT 訳では、解析失敗を示す「……」が訳出され、内容を理解することが困難であるが、SPE 訳ではある程度内容が理解できるレベルまで改善していることが分かる。

中日翻訳においても同様に、誤訳の修正「胸骨→前駆体」や、訳出フレーズの自然性の向上を確認することが出来る。

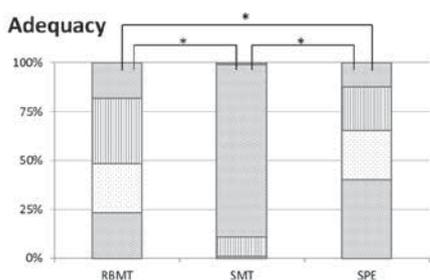


図2 人手評価結果 (英日)

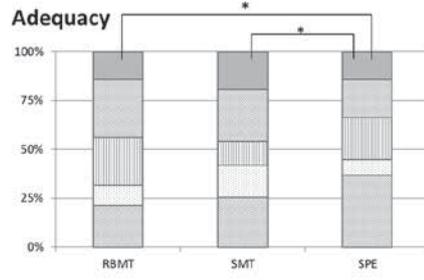


図3 人手評価結果 (中日)

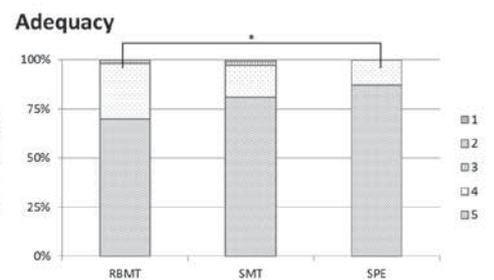


図4 人手評価結果 (韓日)

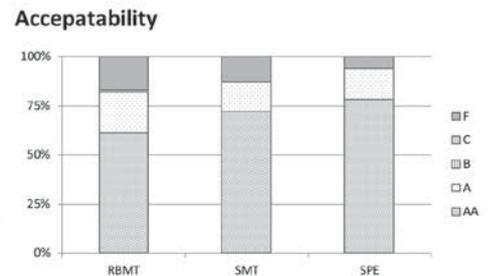
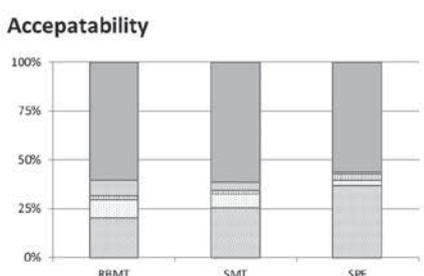
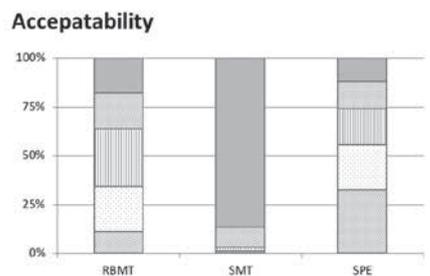


表 2 翻訳結果例

EJ-1	原文	If the mouse is in display space 213, the following occurs:
	参照訳文	マウスが表示空間 2 1 3 にある場合、つぎようになる。
	RBMT訳	マウスが <b>展示スペース</b> 213にある場合、下記が生じる:
	SPE訳	マウスが <b>表示領域</b> 2 1 3 にあるとき、以下が発生する。
EJ-2	原文	Therefore, the stored word models 19 can be modified or adapted by the input speech of a different user.
	参照訳文	したがって、記憶されたワードモデル 1 9 は、別のユーザの入力音声によって、修正されたり適応させられたりする。
	RBMT訳	したがって、記憶式ワード・モデル19は、異なるユーザの入力音声によって <b>修正するか適しているかもしれない</b> 。
	SPE訳	従って、格納されたワードモデル 1 9 は、異なるユーザの入力音声により <b>修正または適合させることができる</b> 。
EJ-3	原文	First circular buffer 140 a has an input that receives an output from first state machine 145 a .
	参照訳文	第 1 循環バッファ 1 4 0 a は、第 1 ステートマシン 1 4 5 a からの出力を受信する入力を持している。
	RBMT訳	第1の円環状バッファ140...1つの...has...第1の状態機械145 aから出力を受け取る入力。
	SPE訳	第 1 循環バッファ 1 4 0 a は、第 1 ステートマシン 1 4 5 a からの出力を受け取る入力。
CJ-1	原文	也可采用各种前体的混合物。
	参照訳文	異なる前駆体の混合物も使用することができる。
	RBMT訳	各種の <b>胸骨</b> の混合物も採用することができる。
	SPE訳	各種 <b>前駆体</b> の混合物を使用することができる。
CJ-2	原文	该结晶通过后述的试验例4的示差扫描热量分析显示出与图7实质上相同的光谱, 确认到该结晶是本案化合物1的B型结晶。
	参照訳文	当該結晶は、後述する試験例 4 の示差走査熱量分析により図 7 と実質的に同様なスペクトルを示し、本件化合物 1 の B 形結晶であることが確認された。
	RBMT訳	この結晶は、後述する試験例 4 の示差走査熱量分析を通して図 7 と <b>実質上が同じスペクトルを現れて</b> 、この結晶がこの事件化合物 1 である B 型結晶を確認した。
	SPE訳	この結晶は、後述する試験例 4 の示差走査熱量分析によって図 7 と <b>実質的に同じスペクトルを示し</b> 、この結晶は、本化合物 1 の B 型結晶を確認した。

## 4 おわりに

本稿では、RBMT（規則ベース機械翻訳）に SPE（統計的後編集）を適用し、英日・中日・韓日特許翻訳における有効性を確認した。SPE は、自動評価においてフレーズベース SMT と同等以上のスコアを実現すると同時に、人手評価においても正確さと内容理解度の高い翻訳結果となっていることを確認した。

今後は、RBMT の特徴であるきめ細かいカスタマイズ機能と、SPE による翻訳精度の向上を両立させることにより、特許翻訳における実用性を高めていきたい。

### 参考文献

- [1] Goto, Isao, et al., "Overview of the patent machine translation task at the NTCIR-10 workshop." Proceedings of the 10th NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access (NTCIR-10) , pp.260-282, 2013.
- [2] 江原 暉将, "規則方式機械翻訳と統計的後編集を組

み合わせた特許文の日英機械翻訳", Japio YEAR BOOK 2010, pp.280-282, 2010.

- [3] Satoshi Sonoh, et al., "Toshiba MT System Description for the WAT2014 Workshop." Proceedings of the 1st Workshop on Asian Translation (WAT2014) , pp.26-32, 2014.
- [4] The 翻訳エンタープライズ TM, <http://mt-server.toshiba-sol.co.jp/>
- [5] JPO 中日辞書, <https://alaginrc.nict.go.jp/resources/jpo-info/jpo-outline.html#jpo-dic-zh>
- [6] JPO・NICT 英日対訳コーパス, <https://alaginrc.nict.go.jp/resources/jpo-info/jpo-outline.html#jpo-en-ja>
- [7] JPO Patent Corpus, <http://orchid.kuee.kyoto-u.ac.jp/WAT/patent/>
- [8] Moses, <http://www.statmt.org/moses/>
- [9] Papineni, Kishore, et al. "BLEU: a method for automatic evaluation of machine translation." Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics, pp.311-318, 2002.