

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

星野 翔 宮尾 祐介 須藤 克仁 永田 昌明
総合研究大学院大学 国立情報学研究所 NTT コミュニケーション科学基礎研究所
{hoshino, yusuke}@nii.ac.jp {sudoh.katsuhito, nagata.masaaki}@lab.ntt.co.jp

1 はじめに

統計的機械翻訳においては、従来の並べ替えモデル (Yamada and Knight, 2001; Koehn et al., 2005; Chiang, 2007) に加えて、原言語と目的言語の語順を近づける事前並べ替え (Xia and McCord, 2004; Xu et al., 2009; Isozaki et al., 2010b; Wu et al., 2011; Neubig et al., 2012) や事後並べ替え (Sudoh et al., 2011; Goto et al., 2012) を行うことによって、日本語と英語のように語順が大きく異なる言語対でも翻訳結果を大幅に改善することができる。

本研究では、述語項構造解析と簡潔な並べ替え規則だけで局所と大域の両方の並べ替えを改善することのできる、日英事前並べ替え手法を提案する。提案手法では、まず述語項関係の解析を行い、日本語の述語、主語、目的語を同定する。次に文レベル・句レベルそれぞれの並べ替えを行い、大域と局所、両方での語順を改善する。最後に並べ替え後の中間表現から英語への翻訳を行う。

既存手法では全体の並べ替えが最適となるよう語順を決めていたのに対し、提案手法では文レベルと句レベルという2つの並べ替え問題に分割し、それぞれが最適となるよう並べ替える点が異なる。これにより、簡潔かつ少数の並べ替え規則を用いるだけで、統計的機械翻訳の精度が大幅に向上することを示す。また提案手法は少数の規則で構成されているため、容易に実装することができ、作成にコストのかかる教師データは必要ない。実験では、同じく規則を用いる既存手法 (Katz-Brown and Collins, 2008; Komachi et al., 2006) と比較を行い、提案手法の有効性を示す。

2 提案手法

提案手法では、入力文に対して係り受け・述語項構造解析を行い、文レベルでは述語、主語、目的語を、句レベルでは内容語と機能語を認識し、それを下に示す規則1～3により並べる。なお本論文では、原言語で係り受けを定義する単位を句と定義して、句自体の並べ替えを文レベル、句内部の並べ替えを句レベルと呼ぶ。日本語では文節間が文レベルに、文節内が句レベルに相当する。

• 規則1 (文レベル規則)

$S ::= S^* \text{ Subj Pred } S^* \mid S^* \text{ Pred Obj } S^* \mid S^* \text{ Pred } S \mid C$

• 規則2 (句レベル規則)

$C ::= \text{Function Content}$

• 規則3 (補助規則)

$S' ::= \{\text{Conjunct}^* S \text{ Comma} \mid \text{Conjunct}^* S\}^* S \text{ Period}$

Pred は述語、Subj は主語、Obj は目的語、S, S' は文 (単文または複文)、C は句、Conjunct は並列句、Function は機能語、Content は内容語、Comma は読点、Period は句点を意味する。

これらの並べ替え規則は並べ替え後の語順を示している。例えば入力文が「Subj C Pred」の場合、規則1のうち $S^* \text{ Subj Pred } S^*$ は、「Subj C Pred」を「Subj Pred C」と書き換えることを表している。ただし、述語、主語、目的語、並列句もまた句であり、他の句と同じく規則2が適用される。

規則1では、VをSの直後、Oの直前に移動し、SVOの語順にすることを目的として、述語を主語の直後、目的語の直前、または最後の文の直前に移動する。

規則2では、目的言語での順序に基づき句中の機能語と内容語を並べ替える。目的言語において内容語が先の場合には、 $C ::= \text{Content Function}$ と書き

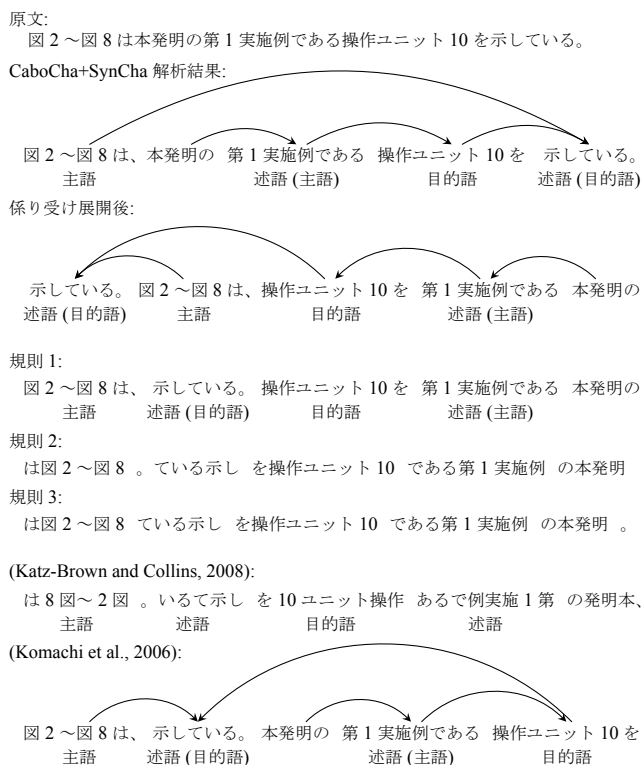


図1 事前並べ替え例

換える必要がある。

規則 3 では、句読点と並列句の位置を原文そのままにするため、S の句読点を移動した S' を出力する。また同時に S の句のうち並列句を先に出力し、「A and B」のような並列句が「B and A」と変化してしまうを防ぐ。この規則は、句読点や並列句の語順には言語による違いがあまりないため加えている。

図 1 に例として、原文、CaboCha と SynCha による解析結果、解析結果を係り受け関係に基づき深さ優先で展開した出力、提案手法の並べ替え結果、先行研究の並べ替え結果を順に示す。図中の上部の矢印は係り受け関係を、下部のラベルは述語項関係を示している。

提案手法では、展開後の出力を基として、まず規則 1 によって述語「示している」を目的語「操作ユニット 10 を」の直前に移動する。述語「第 1 実施例である」は主語が無いため移動しない。次に規則 2 によって各句の機能語を内容語の前に移動する。最後に規則 3 によって句点を文の最後に移動する。

先行研究のうち、(Katz-Brown and Collins, 2008) は提案手法と同じく展開後の出力を基として、述語「示している」を目的語「操作ユニット 10 を」の直前に移動し、また句の中の語順を全て逆転させる。(Komachi

et al., 2006) は解析結果を基として、述語「示している」を目的語を含む文「本発明の第 1 実施例である操作ユニット 10 を」の直前に移動する。

提案手法の規則は、原言語において述語、主語、目的語の 3 つが定義されていれば、SVO の語順で出力されるため、目的言語が SVO の言語対に一般的に適用することができる。今回の実験では、日本語の動詞、形容詞、名詞 + だ、さらに文全体の主辞を述語として、また述語項構造解析によるガ格を主語、二格とヲ格を目的語と定義した。

著者による提案手法の実装では、述語項構造解析を行った状態から規則 1 ～ 3 を適用するために、Python で 200 行程度の簡潔なプログラムで KNP と CaboCha+SynCha の 2 通りの解析結果を処理することができた。

3 実験設定

本研究では日本語から英語への翻訳実験を行い、提案手法と既存の並べ替え手法の比較を行った。また提案手法の規則の有用性を確かめるため、より少ない規則の組み合わせでの実験も行った。

実験に使用した統計的機械翻訳システムでは、言語モデル作成に SRILM 1.6.0 (Stolcke et al., 2011)、フレーズ対学習に MGIZA++ 0.7.1 (Gao and Vogel, 2008)、デコーダに Moses 0.91 (Koehn et al., 2007) を使用した。機械翻訳システムは事前並べ替え手法の違いによらず共通である。システムの設定として、言語モデルに 6-gram、Moses の並べ替えオプションに msd-bidirectional-fe、チューニングに MERT を使用した。distortion limit は事前実験で各手法ごとに最適の値を調べ、ベースラインと (Komachi et al., 2006) では 20、その他の手法では 10 に設定した。またベースラインでは、形態素解析器の JUMAN 7.0^{*1} と MeCab 0.994^{*2} を使用して、JUMAN と IPA の 2 種類の分かち書き基準で前処理した。

提案手法では、日本語解析で代表的である、係り受け・述語項構造解析器 KNP 4.01^{*3}、係り受け解析器 CaboCha 0.65 (Kudo and Matsumoto, 2002)^{*4}、述語項

^{*1} <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

^{*2} <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

^{*3} <http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

^{*4} <http://code.google.com/p/cabocha/>

	BLEU-4	RIBES
ベースライン 1 [JUMAN]	27.33	0.650
ベースライン 2 [McCab]	29.19	0.685
(Katz-Brown and Collins, 2008)	27.59	0.661
(Komachi et al., 2006)	29.58	0.691
提案手法 1 [KNP]	30.65	0.723
提案手法 2 [CaboCha+SynCha]	30.01	0.724

表1 既存研究との比較

構造解析器 SynCha 0.3 (Iida and Poesio, 2011)*⁵の3つを使用し、KNP のみの場合と、CaboCha と SynCha を組み合わせた場合の2通りで実験し、比較を行った。それぞれの解析器で解析に失敗した部分は訓練データから除いたが、開発データとテストデータでの解析の失敗は無く、全データを使用した。

実験データには NTCIR-9 特許機械翻訳テストコレクションの日英翻訳データを用いた。訓練データに約 300 万文対*⁶、テストデータには 2000 文対、開発データは 500 文対を用いた。

評価手法には BLEU-4 (Papineni et al., 2002) と RIBES (Isozaki et al., 2010a) を用いた。BLEU-4 の実験結果のうち、全ての組み合わせについて有意水準 5% で二項検定を行い、有意差が無かった結果を†で表した。

4 実験結果

4.1 既存研究との比較

表 1 に、ベースライン、既存手法 (Katz-Brown and Collins, 2008; Komachi et al., 2006)、提案手法の比較結果を示す。

提案手法の結果は、ベースラインと既存手法の全てを上回った。また KNP のみの提案手法 1 と CaboCha と SynCha を組み合わせた提案手法 2 に BLEU-4 で約 0.6 ポイントの違いがあった。解析器の違いが翻訳結果でも大きな差となることが分かり、将来の解析精度上昇による翻訳結果の改善が期待できる。

既存手法のうち (Komachi et al., 2006) の結果はベースライン 1・2 の双方を上回り、語順が改善されていることが確認できるが、一方で提案手法の結果を BLEU-4 と RIBES の両方で下回った。この手法では、文レベルの並べ替えは行うものの係り受け関係に基づいた解析結果の展開が無く、述語に対応する主語・目的語が無

規則 1	規則 2	規則 3	BLEU-4	RIBES
	✓	✓	28.84	0.704
✓		✓	30.41†	0.717
✓	✓		30.94	0.713
✓	✓	✓	30.65†	0.723

表2 3つの規則の組み合わせ結果

い場合の規則が無い場合、述語が少ない文や主語・目的語が無い文での述語の並べ替えが不十分だったと考えられる。

(Katz-Brown and Collins, 2008) の結果は、同じ分ち書き基準のベースライン 1 をやや上回ったが、ベースライン 2 を大きく下回った。この手法では、述語項関係を展開し、局所と大域の両方を並べ替えるものの、主語や目的語が無い場合に述語を子要素の左に移動する規則があるため、「A は B です」という文が「です B は A」と並べ替えられて訳文が「is B A」のように SVO の順序から遠ざかってしまう場合が多く、そのため規則が有効に働かなかったと考えられる。

4.2 規則の組み合わせ結果

表 2 に、規則 1～3 のそれぞれが無い場合と、3つの規則全てを用いた場合の比較結果を示す。3つの規則全てを使った場合の結果は規則 1・2 が無い場合を上回り、また規則 3 が無い場合と比べて、BLEU-4 は下回ったが RIBES では大きく上回る結果となり、規則の有用性を確認することができた。また規則 1 が無い場合のスコアが規則 2 が無い場合を大きく下回ったことから、評価手法のスコア上は、文レベルの並べ替え結果が句レベルの並べ替え結果以上に翻訳に影響することが分かった。

4.3 順位相関係数による並べ替え評価

図 2, 3 に、(Isozaki et al., 2010b) と同様に訓練データでの Kendall の τ の分布をベースライン 2 と提案手法で測定した結果を示す。^{*7} ベースライン 2 では τ が 0.8 以上の文の割合は 10.2% だったのに対し、提案手法 2 では 33.9% となり、原言語と目的言語の語順の違いが大幅に減っていることが確認できる。

5 おわりに

本研究では、述語項構造解析と 3つの規則で局所と大域の両方の並べ替えを改善することのできる、日英統計的機械翻訳のための事前並べ替え手法を提案した。

*⁵ <http://www.cl.cs.titech.ac.jp/~ryu-i/syncha/>

*⁶ 訓練データ約 320 万文対から 1 語以上 65 語以下の文のみ取り出して使用した。

*⁷ en-ja.A3.final ではなく ja-en.A3.final を用いた。

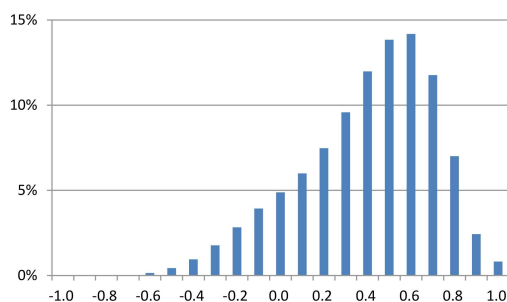


図2 ベースライン 2 の τ

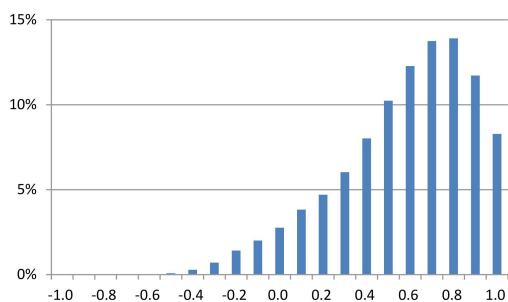


図3 提案手法 1 の τ

実験では、提案手法は既存の並べ替え手法を上回る結果となった。将来は学習ベースの並べ替えを行う既存研究との比較を行い、規則を用いた並べ替えとの違いを明らかにしたい。

参考文献

- Chiang, D. (2007). Hierarchical phrase-based translation. *Computational Linguistics*, 33(2).
- Gao, Q. and Vogel, S. (2008). Parallel implementations of word alignment tool. In *Software Engineering, Testing, and Quality Assurance for Natural Language Processing*, pages 49–57.
- Goto, I., Utiyama, M., and Sumita, E. (2012). Post-ordering by parsing for Japanese-English statistical machine translation. In *Proc. of ACL*, pages 311–316.
- Iida, R. and Poesio, M. (2011). A cross-lingual ILP solution to zero anaphora resolution. In *Proc. of ACL-HLT*, pages 804–813.
- Isozaki, H., Hirao, T., Duh, K., Sudoh, K., and Tsukada, H. (2010a). Automatic evaluation of translation quality for distant language pairs. In *Proc. of EMNLP*, pages 944–952.
- Isozaki, H., Sudoh, K., Tsukada, H., and Duh, K. (2010b). Head finalization: A simple reordering rule for SOV languages. In *Proc. of WMT-MetricsMATR*, pages 244–251.
- Katz-Brown, J. and Collins, M. (2008). Syntactic reordering in preprocessing for Japanese→English translation: MIT system description for NTCIR-7 patent translation task. In *Proc. of the NTCIR-7 Workshop Meeting*.
- Koehn, P., Axelrod, A., Mayne, R. B., Callison-burch, C., Os-

borne, M., and Talbot, D. (2005). Edinburgh system description for the 2005 IWSLT speech translation evaluation. In *Proc. of IWSLT*.

- Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R., Dyer, C., Bojar, O., Constantin, A., and Herbst, E. (2007). Moses: Open source toolkit for statistical machine translation. In *Proc. of ACL Demo and Poster Sessions*, pages 177–180.
- Komachi, M., Matsumoto, Y., and Nagata, M. (2006). Phrase reordering for statistical machine translation based on predicate-argument structure. In *Proc. of IWSLT*, pages 77–82.
- Kudo, T. and Matsumoto, Y. (2002). Japanese dependency analysis using cascaded chunking. In *Proc. of CoNLL*, pages 63–69.
- Neubig, G., Watanabe, T., and Mori, S. (2012). Inducing a discriminative parser to optimize machine translation reordering. In *Proc. of EMNLP-CoNLL*, pages 843–853.
- Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J. (2002). BLEU: A method for automatic evaluation of machine translation. In *Proc. of ACL*, pages 311–318.
- Stolcke, A., Zheng, J., Wang, W., and Abrash, V. (2011). SRILM at sixteen: Update and outlook. In *Proc. of IEEE Automatic Speech Recognition and Understanding Workshop*.
- Sudoh, K., Wu, X., Duh, K., Tsukada, H., and Nagata, M. (2011). Post-ordering in statistical machine translation. In *Proc. of the Machine Translation Summit XIII*.
- Wu, X., Sudoh, K., Duh, K., Tsukada, H., and Nagata, M. (2011). Extracting pre-ordering rules from predicate-argument structures. In *Proc. of IJCNLP*, pages 29–37.
- Xia, F. and McCord, M. (2004). Improving a statistical MT system with automatically learned rewrite patterns. In *Proc. of COLING*, pages 508–514.
- Xu, P., Kang, J., Ringgaard, M., and Och, F. (2009). Using a dependency parser to improve SMT for Subject-Object-Verb languages. In *Proc. of HLT-NAACL*, pages 245–253.
- Yamada, K. and Knight, K. (2001). A syntax-based statistical translation model. In *Proc. of ACL*, pages 523–530.