

ニューラル機械翻訳における事前並び替えの影響分析

瓦 祐希[†] Chenhui Chu[‡] 荒瀬 由紀[†]

[†] 大阪大学大学院情報科学研究科 [‡] 大阪大学データビリティフロンティア機構

{kawara.yuki, arase}@ist.osaka-u.ac.jp, chu@ids.osaka-u.ac.jp

1 はじめに

統計的機械翻訳 (SMT) において、翻訳言語間の語順の相違は翻訳精度に大きな影響を与えることが知られている [7]。この問題を解決する手法の一つとして、翻訳器に入力する前に原言語側の語順を目的言語の語順に近づくように並び替える事前並び替えが提案されている [5, 7]。先行研究により、特に英日対のように語順の大きく異なる言語対において、事前並び替えは SMT で大幅に精度を改善することが示されてきた。

しかしニューラル機械翻訳 (NMT) では、事前並び替えを用いて翻訳を行うと並び替えをせずそのまま翻訳を行うよりも精度が低下することが報告されている [5, 10]。Sudoh and Nagata [10] は中日翻訳実験において、事前並び替えにより原言語側の言語の構造に一貫性がなくなってしまうことが翻訳精度低下の一因と考察している。本研究では、英日対のような語順の大きく異なる言語対における NMT での並び替えの効果を検証する。

ASPEC コーパス [8] を用いた英日・日英言語対での実験の結果：(1) 参照翻訳を利用したオラクルな事前並び替えを用いて翻訳を行うと NMT においても精度が大幅に向上すること；(2) 既存の事前並び替え手法として Head Finalization (HFE) [4]、Bracketing Transduction Grammar (BTG) [7]、Recursive Neural Network (RvNN) [5]、注意機構付きエンコーダ・デコーダモデル (AttnEncDec-re) [6] を用いた事前並び替えを行う場合、NMT では並び替えなしの場合に比べて翻訳精度が低下すること；(3) オラクルな並び替えに近いよりも、並び替えに一貫性があるかどうか NMT に事前並び替えを適用する際には重要であることが明らかとなった。

2 関連研究

Edunov ら [2] は目的言語の単言語コーパスによる back translation を用いた訓練において単語を並び替

えた文を用いることで翻訳精度を向上させている。生成した back translation において単語の削除や別の単語への置き換え、3 単語以内の単語との交換を行い訓練することで、エンコーダのロバスト性を高める効果があると考えられる。この手法は WMT2018 の英独翻訳タスクで 1 位を達成している [1] が、長距離の語順の変更は行っていない。

Huang ら [3] は、フレーズベースの NMT において局所的な並び替えを考慮する層を導入している。この手法では word embedding の重み付き和を計算することで並び替えを考慮した翻訳を行なっている。Zhang ら [12] はエンコーダとデコーダの間に並び替えを行うモデルを挿入し、Inversion Transduction Grammars に基づいた 2 分木の構文解析および並び替えを同時に学習している。実験の結果、Text Normalization タスクにおいて、注意機構付きエンコーダ・デコーダモデルよりも学習が速く精度も高いと報告している。しかし、並び替えを 10 単語以内の単語のみに制限しており、長距離の並び替えにおける提案手法の効果の検証は行っていない。

これらの先行研究から、語順は NMT においても重要なファクターであり、語順に関する学習を行うことで NMT の精度向上に効果があることが示されている。しかし、事前並び替えによる NMT の精度向上に対する効果はまだ示されていない。

3 事前並び替え手法

本稿で比較を行う事前並び替え手法について説明する。

3.1 オラクルな事前並び替え

本来テスト時には参照できない参照翻訳を用いることで、事前並び替え手法のなかで最も高精度であると期待できるオラクルな事前並び替えをシミュレーショ

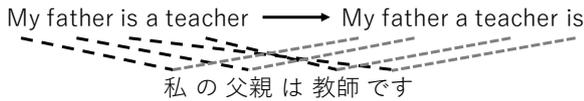


図 1: オラクルな並び替えの例

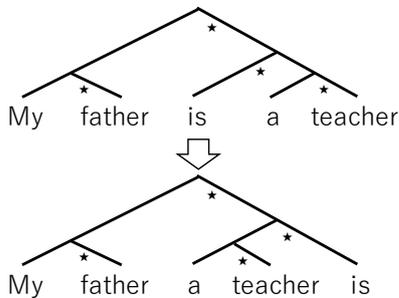


図 2: HFE での並び替えの例 (★ のついたノードは主辞を表す)

ンする。まずはじめに原言語と目的言語間の単語アライメントを計算する。計算された単語アライメントに基づき、アライメントの交差がなくなるように単語の位置を交換する。対応した単語がない原言語の単語は、Neubig らのルール [9] に従い、主辞前置型言語か主辞後置型言語かに応じて前または後ろの単語と同時に移動する。例えば、英語は主辞前置型言語であるため、単語アライメントのない単語は後ろの単語と同時に並び替える。日本語を並び替える場合は主辞後置型言語のため、前の単語と同時に並び替える。これにより、オラクルな事前並び替えではアライメントが存在しない単語は一貫したルールで並び替えられる。図 1 に英語文の並び替えの例を示す。ここでは単語アライメントを持たない “a” は “teacher” と同時に並び替えられ、“a teacher is” という語順となっている。

3.2 ルールベースの事前並び替え

Head Finalization (HFE) [4] は英語のような SVO 型言語から日本語のような SOV 型言語への翻訳における、ルールベースの並び替え手法である。与えられた構文木に対して、各節ノードで主辞であるノードが他のノードの後ろに来るように並び替えを行う。図 2 に並び替えの例を示す。この例では “is a teacher” に対応したノードにおいて、“is” が主辞のため、“a teacher” の後ろに “is” のノードを並び替える。

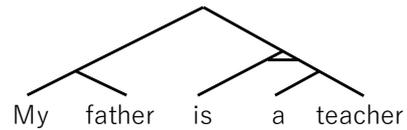


図 3: BTG による並び替えの例 (横線は子ノードを並び替えることを表す)

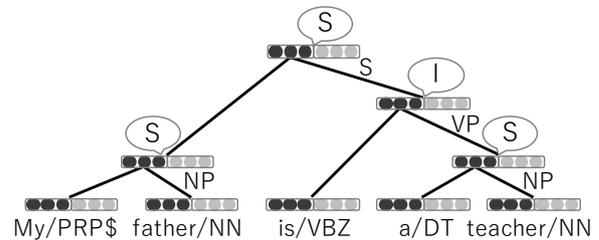


図 4: RvNN による並び替えの例 (I は子ノードを並び替えることを、S はそのままの順序であることを表す)

3.3 機械学習による事前並び替え

本節では、機械学習を用いた事前並び替え手法について述べる。Nakagawa [7] は Bracketing Transduction Grammar (BTG) を用いた事前並び替え手法を提案している。この手法では原言語文と単語アライメントからトップダウンに 2 分木を構築しつつ、各ノードで並び替えをするか否かを学習する。2 単語以上のスパンのうち、どこで区切るか、区切った部分の前後を並び替えるかを再帰的に計算していくことで、最終的に構文木および各節ノードでのラベルが決定される。SMT に適用することで、英日・日英機械翻訳において最高性能を達成している。図 3 に BTG による並び替えの例を示す。例えば “is a teacher” に対応したノードでは、“is” と “a teacher” にまず分割され、“is” と “a teacher” を並び替えること示すラベルが付与される。

Kawara ら [5] は Recursive Neural Network (RvNN) を用いた事前並び替え手法を提案し、英日・日英 SMT に適用することで BTG と同等の翻訳性能を達成した。RvNN を用いた事前並び替え手法では、入力された構文木に沿ってニューラルネットワークを構築する。それから、各節ノードに付与された並び替えラベルの学習を行う。図 4 に RvNN を用いた並び替えの例を示す。例えば “is a teacher” のノードでは、“is” のノードと “a teacher” のノードから自身のベクトルを計算し、そのベクトルを用いて並び替えラベルの予測を行う。

さらに本稿では、事前並び替えを「原言語文から並び替えた文を生成するタスク」とみなし、注意機構付きエンコーダ・デコーダモデル [6] での並び替えを行

う (AttnEncDec-re)。語彙は頻度の上位 50,000 単語を用い、それ以外の単語は “<UNK>” タグに置き換え訓練を行なった。本研究ではデコーダが出力する単語に制限をかけておらず、出力される文に元の文にない単語が含まれる場合や元の文に含まれていた単語が出力されない場合がある。今後、出力される単語を元の文に含まれていた単語のみになるよう出力に制限をかけて並び替えを行い、翻訳実験を行う予定である。

4 翻訳実験

4.1 実験設定

ASPEC コーパス [8] で英日、日英方向の翻訳実験を行なった。ASPEC コーパスに含まれる訓練データは約 3,000,000 文対、開発データは 1,790 文対、テストデータは 1,812 文対である。英語文は Moses の前処理スクリプトである tokenizer.perl¹ で単語分割を行い、日本語文は MeCab² で形態素解析を行なった。単語アライメントは MGIZA³ を用いた。オラクルの並び替えでは、英日翻訳では英語から日本語への、日英翻訳では日本語から英語への片方向の単語アライメントを用いた⁴。HFE、RvNN の構文解析には Enju⁵ を用いた。BTG と RvNN の単語アライメントは、英日、日英方向のアライメントを計算した後、Intersection ルールにしたがって最終的なアライメントを計算した。AttnEncDec-re の学習には OpenNMT-py⁶ を用いた。

BTG、RvNN での並び替えの学習には訓練データからサンプリングした文を用いた。ASPEC コーパスは Utiyama and Isahara [11] の類似度に基づいて対訳文がランク付けされている。本研究では上位 500,000 文対からランダムに 100,000 文対をサンプリングして BTG、RvNN での並び替えの訓練データとした。RvNN では単語ベクトル、品詞・構文タグベクトル、および節ノードのベクトルの次元数は 200 とした。バッチサイズは 500 とし、5 エポック学習を行い開発データでロスが最も低い時点でのモデルを用いて並び替えを行なった。AttnEncDec-re の並び替えの学習には訓練データ全文を用いた。AttnEncDec-re ではデフォルトの設定に従い、エンコーダ、デコーダともに 2 層の

¹<http://github.com/moses-smt/mosesdecoder/blob/master/scripts/tokenizer/tokenizer.perl>

²<http://taku910.github.io/mecab/>

³<https://github.com/moses-smt/mgiza>

⁴両方向のアライメントを用いるとアライメントの交差を 0 に出来ない場合があるため、本研究では片方向の単語アライメントを用いた。

⁵<https://github.com/mynlp/enju>

⁶<https://github.com/OpenNMT/OpenNMT-py>

表 1: 事前並び替えを用いた翻訳結果

	En-Ja		Ja-En	
	BLEU	RIBES	BLEU	RIBES
w/o preordering	31.51	81.26	22.88	74.49
w/ oracle	37.69	87.05	30.76	84.85
HFE	29.69	79.64	-	-
BTG	28.91	79.58	20.11	71.07
RvNN	29.01	79.63	19.92	71.74
AttnEncDec-re	27.48	79.15	20.37	72.15

LSTM を用いた。語彙数は頻度の上位 50,000 語を用い、単語ベクトルの次元数は 500、隠れ層のベクトルの次元数も 500 とした。バッチサイズは 64、200,000 ステップの学習を行った。10,000 ステップごとに開発データのパープレキシティを測り一番小さいモデルでテストデータの翻訳を行なった。

翻訳器の学習には上位 200 万文対のうち、原言語、目的言語ともに 50 単語以下の文対を用い、単語数の比が 9 以上のものをフィルタリングした 1,814,494 文対を用いた。翻訳器には、注意機構付きエンコーダ・デコーダモデル (AttnEncDec) として OpenNMT-py⁷ を用いた。語彙数、単語ベクトル、隠れ層のベクトルの次元数、バッチサイズ、学習ステップは AttnEncDec-re と同様に設定した。

4.2 評価指標

翻訳性能を BLEU および RIBES 値を用いて評価する。またオラクルな並び替えに対する自動事前並び替え手法の性能を評価するため、オラクルな並び替えを参照翻訳とみなして BLEU および RIBES 値を計測する。これらの値が高いことはよりオラクルな並び替えに近い並び替えを実現できていることを示す。

4.3 実験結果

表 1 に翻訳結果の BLEU 値、RIBES 値を示す。実験結果から、オラクルな並び替えにより翻訳精度が大きく向上しており、英日対で BLEU 値が 6.18 ポイント、RIBES 値が 5.79 ポイント、日英対では BLEU 値が 7.88 ポイント、RIBES 値が 10.36 ポイント向上した。しかし、自動事前並び替えを用いた場合、並び替えなしの場合に比べて BLEU 値、RIBES 値共に低下する結果となった。

⁷<https://github.com/OpenNMT/OpenNMT-py>

表 2: オラクルな並び替えとの BLEU 値、RIBES 値

	En-oracle		Ja-oracle	
	BLEU	RIBES	BLEU	RIBES
w/o reordering	33.61	62.95	43.19	67.64
HFE	35.13	72.31	-	-
BTG	38.93	76.31	39.84	75.22
RvNN	34.88	72.70	39.22	74.56
AttnEncDec-re	49.25	81.40	49.45	77.79

表 2 にオラクルな事前並び替えと比較した際の、自動事前並び替え結果の BLEU 値、RIBES 値を示す。英語文（英日対）では事前並び替えの BLEU 値、RIBES 値が全ての手法で並び替えなしと比較して向上しており、オラクルな並び替えに近づくように並び替えが出来ていることが分かる。日本語文（日英対）では BTG、RvNN を用いた事前並び替えによって BLEU 値は下がったが RIBES 値は向上し、AttnEncDec-re では BLEU 値および RIBES 値が向上した。

これらの結果を翻訳精度に照らし合わせると、オラクルの事前並び替えに対する BLEU 値、RIBES 値と翻訳精度との関連は薄いことが分かる。英日方向において、オラクルな並び替えに最も近い並び替えが出来ている AttnEncDec-re では翻訳精度が最も低くなっている一方、ルールベースの並び替えである HFE は、オラクルの並び替えに対する BLEU 値は低いが、翻訳精度は AttnEncDec-re と比べると高い。このことから、ルールベースのように一貫性のある並び替えを行うことが、NMT の精度を向上するためには重要と考えられる。オラクルな事前並び替えでは、単語アライメントが存在しない単語はルールに基づいた並び替えを行なっている。これらアライメントが存在しない単語は機械学習に基づく並び替え手法においてランダムな位置に並び替えられる可能性が高く、一貫性を大きく損なう要因となっていると考えられる。

5 まとめ

本稿では事前並び替えが NMT に及ぼす影響について調査を行なった。実験の結果、並び替えにおいて一貫性を保持することが、NMT での翻訳精度を向上する上で重要であることが示された。今後、並び替えと翻訳を同時に学習することで、一貫性を保持した並び替えを行いながら NMT での翻訳精度を向上する手法を検討する予定である。

謝辞

本研究は、日本電信電話株式会社 コミュニケーション科学基礎研究所および科研費#17H06822 の助成を受けたものである。

参考文献

- [1] Ondej Bojar, Christian Federmann, Mark Fishel, Yvette Graham, Barry Haddow, Matthias Huck, Philipp Koehn, and Christof Monz. Findings of the 2018 conference on machine translation (wmt18). In *Proc. of WMT*, pp. 272–307, Belgium, Brussels, October 2018.
- [2] Sergey Edunov, Myle Ott, Michael Auli, and David Grangier. Understanding back-translation at scale. In *Proc. of EMNLP*, pp. 489–500, 2018.
- [3] Po-Sen Huang, Chong Wang, Sitao Huang, Dengyong Zhou, and Li Deng. Towards neural phrase-based machine translation. In *Proc. of ICLR*, 2018.
- [4] Hideki Isozaki, Katsuhito Sudoh, Hajime Tsukada, and Kevin Duh. Head Finalization: A Simple Reordering Rule for SOV Languages. In *Proc. of WMTMetricsMATR*, pp. 244–251, July 2010.
- [5] Yuki Kawara, Chenhui Chu, and Yuki Arase. Recursive neural network based reordering for english-to-japanese machine translation. In *Proc. of ACL, Student Research Workshop*, pp. 21–27. Association for Computational Linguistics, July 2018.
- [6] Thang Luong, Hieu Pham, and Christopher D. Manning. Effective approaches to attention-based neural machine translation. In *Proc. of EMNLP*, pp. 1412–1421, 2015.
- [7] Tetsuji Nakagawa. Efficient Top-Down BTG Parsing for Machine Translation Preordering. In *Proc. of ACL-IJCNLP*, pp. 208–218, July 2015.
- [8] Toshiaki Nakazawa, Manabu Yaguchi, Kiyotaka Uchimoto, Masao Utiyama, Eiichiro Sumita, Sadao Kurohashi, and Hitoshi Isahara. ASPEC: Asian Scientific Paper Excerpt Corpus. In *Proc. of LREC*, pp. 2204–2208, May 2016.
- [9] Graham Neubig, Taro Watanabe, and Shinsuke Mori. Inducing a Discriminative Parser to Optimize Machine Translation Reordering. In *Proc. of EMNLP-CoNLLs*, pp. 843–853, July 2012.
- [10] Katsuhito Sudoh and Masaaki Nagata. Chinese-to-Japanese Patent Machine Translation based on Syntactic Pre-ordering for WAT 2016. In *Proc. of WAT*, pp. 211–215, December 2016.
- [11] Masao Utiyama and Hitoshi Isahara. A Japanese-English Patent Parallel Corpus. In *Proc. of the Machine Translation Summit XI*, September 2007.
- [12] Hao Zhang, Axel Ng, and Richard Sproat. Fast and accurate reordering with its transition rnn. In *Proc. of COLING*, pp. 1454–1463. Association for Computational Linguistics, 2018.