

学修番号 16890528

修士論文

目的言語の言い換えによる 日英ニューラル機械翻訳の改善

関沢 祐樹

2018年2月23日

首都大学東京大学院
システムデザイン研究科 情報通信システム学域

関沢 祐樹

審査委員：

小町 守 准教授 (主指導教員)

山口 亨 教授 (副指導教員)

高間 康史 教授 (副指導教員)

目的言語の言い換えによる 日英ニューラル機械翻訳の改善*

関沢 祐樹

修論要旨

近年、自然言語処理においてある言語の文を異なる言語の文へと機械が自動で書き換える、機械翻訳の研究が盛んに行なわれている。機械翻訳はある言語で書かれた文を他の言語の文に自動かつ高速に翻訳することができる。機械翻訳は、使用者が言語の知識を持たずとも文を自動で翻訳できることから重要な技術であると言える。

従来の機械翻訳は、翻訳前の言語（原言語）の句に対して翻訳後の言語（目的言語）の句が与えられている統計的機械翻訳が高精度であった。統計的機械翻訳はフレーズ翻訳のスコアや言語モデルスコアなどの様々なスコアを学習し、これらのスコアを組み合わせた結果最も適した翻訳規則を適用することで文を翻訳する。しかし、翻訳された文は流暢性に欠けており、人間が読むには不自然な文が多いという課題が存在する。

一方、ニューラルネットワークを用いた機械翻訳手法である、ニューラル機械翻訳が提案された。ニューラル機械翻訳は統計的機械翻訳よりも自然な文を出力でき、実用的な機械翻訳手法であると言える。その一方で、ニューラル機械翻訳はsoftmaxを出力の語彙サイズで取るため、トレーニングをする際に時間が掛かるという問題がある。したがって、すべての単語を用いることは現実的でない。ニューラル機械翻訳では通常、使用する語彙を制限し、計算時間を削減する。語彙を制限する際、学習に用いる単語の出現頻度を用い、高頻度な単語のみを使用し、それ以外の単語である低頻度語は語彙から外れる（Out-Of-Vocabulary; OOV）。このOOVは、まとめて1つの特殊記号“<unk>”で出力され、意味を持たないため、翻訳前後の意味の保持ができなくなる。

*首都大学東京大学院 システムデザイン研究科 情報通信システム学域 修士論文, 学修番号 16890528, 2018年2月23日.

意味を考慮しつつ OOV の削減を試みた研究として、トレーニングコーパスの OOV を高頻度な類義語に置換する前処理手法がある。この手法は出力文の OOV を減少させる一方、同義語でない類義語に置換することがあり、意味が異なる可能性がある。また、この手法ではトレーニングコーパスにおいて、対応する単語が存在しない OOV を消去するため、翻訳前後で内容の損失が発生する。

そこで、本研究では翻訳前後の意味を保持しつつ、OOV の出力を抑制する前処理手法を提案する。本手法では、トレーニングコーパスのうち、目的言語において、OOV となる単語を高頻度な同義語にあらかじめ言い換えてから翻訳の学習を行う。提案手法では、言い換えの際に同義語の言い換えを収録している言い換え辞書を使用し、OOV を高頻度語に言い換える言い換え対を使用し言い換えラティスを作成する。言い換えの際 OOV が全て OOV に言い換えられる場合、さらに言い換えを行うことで高頻度語への言い換えを探索することが可能である。

言い換えを選択する際、言い換えラティスに対して言い換え辞書に付随するスコアと言語モデルスコアを与え、動的計画法を用いて最も高いスコアとなる言い換えを選択する。2つのスコアを組み合わせることで言い換え後の文を自然にしつつ言い換え前後の意味を保持することができる。したがって、本手法は文の意味を変化させることなく出力の OOV を削減することができる。本研究では、日英翻訳での実験を行い評価する。

本論文の主要な貢献を以下に示す。

1. 本論文では OOV を考慮しつつ、日英ニューラル翻訳を改善する言い換えに基づく前処理手法を提案した。提案手法は言い換え辞書を用いてトレーニングコーパスに存在する OOV を同義な高頻度語へと言い換えることができるため、意味を変化させずに OOV を削減することが期待できる。
2. 提案する手法が従来手法と比較して翻訳文の精度を向上しつつ OOV の出現率を減少させた。言い換えを行わない手法と比較すると、翻訳の質を向上しつつ出力文の OOV を減らすことができた。また OOV を類義語に置換する手法と比べても、OOV をより多く削減することができた。
3. 単語の完全一致だけではなく同義語も正解であると考慮する評価尺度を用いると、提案手法がより良い翻訳を出力していることが明らかとなった。提案手法は OOV の単語に対してその同義語を出力するため、この評価尺度を

用いることで提案手法が意味を考慮した前処理手法であることが明らかとなった。

本論文の構成は以下の通りである。第 1 章では本研究の概要，貢献を述べる。第 2 章ではニューラル機械翻訳における OOV の解決を試みた関連研究について述べる。第 3 章ではトレーニングコーパスを言い換えることでニューラル機械翻訳の OOV を削減する提案手法について述べる。第 4 章では提案手法と比較手法を用いた日英翻訳の結果を述べる。第 5 章では実験結果に基づき考察する。最後に第 6 章では本研究のまとめ及び課題について述べる。

Paraphrasing the Target Language Corpus to Improve Japanese-to-English Neural Machine Translation*

Yuki Sekizawa

Abstract

Recently, machine translation which translates a sentence to another sentence in other language is actively researched in natural language processing. Machine translation can translate a sentence automatically and fast. Machine translation is an important technique since it helps people communicate with each other using a non-native language.

Previously, statistical machine translation (SMT) has been mainly researched. SMT translates sentences according to the score of automatically extracted translation rules. However, SMT has a problem that the translated sentences are not fluent.

On the other hand, neural machine translation (NMT) was proposed few years ago. NMT produces sentences that are more fluent than those produced by SMT. However, NMT requires a very high computational cost for training. Since NMT calculates softmax using the vocabulary size of output, it is not realistic to use all of words. Generally, NMT restricts the size of the vocabulary to reduce the computational cost. When restricting vocabulary, NMT uses only frequent words according to the frequency of words in training corpus, which results in infrequent words being treated as out-of-vocabulary (OOV). The infrequent words are output with a special symbol “<unk>” and it degrades the performance of the translated sentence.

*Master’s Thesis, Department of Information and Communication Systems, Graduate School of System Design, Tokyo Metropolitan University, Student ID 16890528, February 23, 2018.

In order to reduce OOV while considering meaning, a simple but effective preprocess method was proposed. This method reduces OOV in output with frequent similar words in training corpus whereas it might replace OOV words with similar but non-synonymous words. In addition, this method deletes OOV words from the training corpus if they aligned to null, which leads to a loss of sentence meaning.

In this research, I propose a preprocess method to suppress outputting OOV while keeping the sentence meaning. My method that paraphrases infrequent words or phrases expressed as OOV with frequent synonyms from the translated language (target language) side of the training corpus before training. Proposed method uses a paraphrase lexicon recording of paraphrasing with synonyms and makes a paraphrase lattice. This method can search paraphrasing for frequent words by additional paraphrasing if a OOV paraphrases with other OOV words.

Proposed method gives paraphrase lexicon score and language model score to nodes and edges for a paraphrasing lattice and chooses maximum score paraphrases using dynamic programming. This method can keep meaning and be fluent for paraphrased sentences because of combination of these two scores. Therefore, this method can reduce OOV in output while not changing sentence meaning. Since I use a database collecting paraphrases of synonyms, I can reduce OOV in output keeping the meaning. In this research, I evaluate my method with Japanese-to-English translation.

The contributions of this thesis are as follows.

1. I propose a paraphrasing-based preprocessing method for Japanese-to-English NMT to improve translation accuracy with regard to OOV words. I expect to reduce OOV while keeping sentence meaning since proposed method can paraphrase OOV in training corpus using a paraphrase lexicon.
2. Proposed method improves translated sentences accuracy reducing OOV compared with previous works. Comparing with non-paraphrase method, proposed method reduces OOV in output and improves accuracy of trans-

lation. Also, comparing the previous method replacing OOV with a similar word, proposed method reduces more OOV.

3. Using an evaluation method considering synonyms, proposed method clearly outputs better translation.

The structure of this thesis is shown below. In Section 1, I show the abstract and contribution of this thesis. In Section 2, I explain related works attempting to reduce OOV in machine translation and improve machine translation by paraphrasing corpora. In Section 3, I present the proposed method reducing OOV of NMT by paraphrasing the training corpus. In Section 4, I describe the result of Japanese-to-English translation of related method and proposed method. In Section 5, I consider according to the result. Finally, In Section 6, I conclude this research with future works.

目次

図目次	viii
第 1 章 はじめに	1
第 2 章 関連研究	4
2.1 ニューラル機械翻訳の OOV の削減を試みる関連研究	4
2.2 機械翻訳の前処理としてコーパスを書き換える関連研究	6
第 3 章 OOV を同義語に言い換える提案手法	8
3.1 言い換えラティスの構築	8
3.2 言い換えの選択	11
第 4 章 日英翻訳実験	13
4.1 実験設定	13
4.2 実験結果	15
第 5 章 考察	19
第 6 章 おわりに	21
謝辞	23
参考文献	24
発表リスト	26

目次

2.1	Luong らの手法	4
2.2	Sennrich らの手法	5
3.1	言い換えラティスを作成する Python の実装	9
3.2	言い換えラティスの例	10
3.3	複数回言い換えの例	10
3.4	言い換えを選択する Python の実装	11
4.1	提案手法の日英翻訳の BLEU スコア	15
4.2	提案手法の日英翻訳の OOV の数	16

第 1 章 はじめに

近年，自然言語処理においてある言語の文を異なる言語へと機械が自動で書き換える，機械翻訳の研究が盛んに行なわれている．機械翻訳はある言語で書かれた文を他の言語の文に自動かつ高速に翻訳することができる．機械翻訳は，使用者が言語の知識を持たずとも文を自動で翻訳できることから重要な技術であると言える．

従来の機械翻訳は，翻訳前の言語（原言語）の句に対して翻訳後の言語（目的言語）の句が与えられている統計的機械翻訳が高精度であった．統計的機械翻訳はフレーズ翻訳のスコアや言語モデルスコアなどの様々なスコアを学習し，これらのスコアを組み合わせた結果最も適した翻訳規則を適用することで文を翻訳する．しかし，翻訳された文は流暢性に欠けており，人間が読むには不自然な文が多いという課題が存在する．

一方，自然言語処理における多くのタスクにおいてニューラルネットワークに基づく手法が大きな成果を上げている．機械翻訳の分野では，従来研究されてきた統計的機械翻訳と比較してより自然な文を出力できるという利点から，ニューラル機械翻訳 [1] が盛んに研究されている．ニューラル機械翻訳は翻訳文を出力するために単語を 1 つずつ生成する．しかし，ニューラル機械翻訳は語彙次元の分類問題を順番に解いていく生成タスクであり，出力層が高次元となる．ニューラル機械翻訳は softmax を出力の語彙サイズで取るため計算量が多いという課題がある．そのため，ニューラル機械翻訳では通常，使用する語彙を制限し，計算時間を削減する．語彙を制限する際，学習に用いる単語の出現頻度を用い，高頻度な単語のみを使用し，それ以外の単語である低頻度語は語彙から外れる（Out-Of-Vocabulary; OOV）．この OOV は，まとめて 1 つの特殊記号 “<unk>” で出力され，意味を持たないため，翻訳前後の意味の保持ができなくなる．

ニューラル機械翻訳の OOV の削減を試みる先行研究として，Mi ら [2] はトレーニングに使用する語彙を文ごとに選択することで，トレーニングの計算量を減少させ，全体の語彙を拡張した．しかし，この手法では翻訳のトレーニング方法を変更する必要がある．Luong ら [3] は OOV との対応関係にある原言語の単語を翻訳辞書を用いて直接翻訳する後処理を提案した．この手法では，トレーニングデータを用いて原言語と目的言語の単語アライメントを取る必要がある．Jean ら [4] は

OOV をその単語が持つ attention が最も大きい原言語の単語を翻訳辞書によって翻訳する後処理手法を提案した．この手法は単語アライメントを必要としない一方，目的言語での意味を考慮できていない．さらに，Sennrich ら [5] は，系列に対するデータ圧縮手法である Byte Pair Encoding (BPE) を文字列に適用し，単語を頻出する部分文字列の系列に分解して学習することで OOV を削減した．この手法では，意味を考慮せずに単語を部分文字列に分解する．

一方，我々の手法のように機械翻訳の前処理段階においてコーパスを言い換え，原言語および/あるいは目的言語の文の複雑さを減少させる手法が存在する．Sanja ら [6] は機械翻訳の前処理として原言語文の語彙を簡単な文法を用いて言い換えた．本研究では入力文を簡単にせず，出力文の OOV を割合を減少させることで翻訳の質の向上を試みる．さらに，Li ら [7] は前処理の段階でトレーニングコーパスや入力文の OOV を使用する語彙に含まれる類義語に置換する手法を提案した．彼らは OOV を高頻度語に置換する際，単語同士の類似度や，置換後の文が自然であるかを判断するために言語モデルを使用した．単語の意味が似ているかどうかを判定するために，彼らは似た意味の単語はその文脈も似ているという分布仮説に基づいた分散表現を用いてコサイン類似度を計算した．また，彼らはトレーニングコーパスにおいて OOV の単語に対応する異なる言語の単語のアライメントを用い，OOV がどの単語にも対応していない場合その単語を削除した．しかし，OOV の削除は文の意味の損失につながり，翻訳後の文の情報が欠落する．加えて，彼らは分布類似度を用いるため OOV を同義でない類義語に置換する可能性がある．例えば，彼らは “surfing” を “snowboard” に置換するため，“internet surfing” を “internet snowboard” と書き換えてしまうため意味の変化が発生する．本研究では分布類似度ではなくあらかじめ計算された言い換えスコアを用いる．そのため，本手法では OOV の不適切な表現への言い換えを抑制できる．前述の例では，“surfing” を “browser” へと言い換えるため元の意味をある程度残すことができる．

本研究では，翻訳前後の意味を保持しつつ，OOV の出力を抑制する前処理手法を提案する．本手法では，トレーニングコーパスのうち，目的言語において，OOV となる単語を高頻度な同義語にあらかじめ言い換えてから翻訳の学習を行う．提案手法では，言い換えの際に同義語の言い換えを収録している言い換え辞書を使用し，OOV を高頻度語に言い換える言い換え対を使用し言い換えラティスを作成す

る．言い換えの際 OOV が全て OOV に言い換えられる場合，さらに言い換えを行うことで高頻度語への言い換えを探索することが可能である．

言い換えを選択する際，言い換えラティスに対して言い換え辞書に付随するスコアと言語モデルスコアを与え，動的計画法を用いて最も高いスコアとなる言い換えを選択する．2つのスコアを組み合わせることで言い換え後の文を自然にしつつ言い換え前後の意味を保持することができる．したがって，本手法は文の意味を変化させることなく出力の OOV を削減することができる．本研究では，日英翻訳での実験を行い評価する．

本論文の主要な貢献を以下に示す．

1. 本論文では OOV を考慮しつつ，日英ニューラル翻訳を改善する言い換えに基づく前処理手法を提案した．提案手法は言い換え辞書を用いてトレーニングコーパスに存在する OOV を同義な高頻度語へと言い換えることができるため，意味を変化させずに OOV を削減することが期待できる．
2. 提案する手法が従来手法と比較して翻訳文の精度を向上しつつ OOV の出現率を減少させた．言い換えを行わない手法と比較すると，翻訳の質を向上しつつ出力文の OOV を減らすことができた．また OOV を類義語に置換する手法と比べても，OOV をより多く削減することができた．
3. 単語の完全一致だけではなく同義語も正解であると考慮する評価尺度を用いると，提案手法がより良い翻訳を出力していることが明らかとなった．提案手法は OOV の単語に対してその同義語を出力するため，この評価尺度を用いることで提案手法が意味を考慮した前処理手法であることが明らかとなった．

本論文の構成は以下の通りである．第 1 章では本研究の概要，貢献を述べる．第 2 章ではニューラル機械翻訳における OOV の解決を試みた関連研究について述べる．第 3 章ではトレーニングコーパスを言い換えることでニューラル機械翻訳の OOV を削減する提案手法について述べる．第 4 章では提案手法と比較手法を用いた日英翻訳の結果を述べる．第 5 章では実験結果から考えられる考察を述べる．最後に第 6 章では本研究のまとめ及び課題について述べる．

樋 補修 時間 が 短縮 し た 。
the <unk_1> repair hour was shortened .



図 2.1 Luong らの手法

第 2 章 関連研究

この章では，ニューラル機械翻訳の OOV の削減を試みる関連研究や機械翻訳の前処理としてコーパスを言い換える関連研究について述べる．

2.1 ニューラル機械翻訳の OOV の削減を試みる関連研究

ニューラル機械翻訳のトレーニング方法の変更によって OOV の削減を試み，翻訳の精度を向上させる先行研究が存在する．Jean ら [8] は，トレーニングにおいて対訳コーパスを分割し，分割された対訳コーパスを用いたトレーニングにおいて，使用する語彙を目的言語側の語彙からサンプリングし，得られた一部分の語彙を用いてトレーニングを行うことでトレーニングの計算量を減少させ，全体の語彙を広く取ることで OOV の削減を試みた．Mi ら [2] はトレーニングに使用する語彙を文ごとに選択することで，トレーニングの計算量を減少させ，全体の語彙を拡張した．Luong ら [9] は文字ベースの学習によって OOV を減少させた．これらの手法はトレーニング方法を変更する必要がある．本研究では，トレーニング方法を変更せず，トレーニングデータにおける目的言語の語彙的言い換えによって前処理のみで翻訳結果の OOV を削減する．

一方，ニューラル機械翻訳のトレーニング方法は変更せずに OOV の削減を試みる研究も存在する．Luong ら [3] は OOV との対応関係にある原言語の単語を翻訳辞書を用いて直接翻訳する後処理を提案した．OOV を後処理によって翻訳するためには OOV が原言語側の文のどの単語に対応しているかを出力する必要がある．そのため，Luong らは OOV を出力する際，OOV に対応する原言語側の単語との

初めに、クリニカルパスのメリットおよびデ@@メリットについて述べた。

merit and de@@ merit of the clinical pass are described .

図 2.2 Sennrich らの手法

相対距離を出力するトレーニングを行うことによって直接翻訳を可能とした。対応関係を利用する例を図 2.1 に示す。この例では原言語側の 0 番目の単語“樋”に対応する目的言語側の 1 番目の単語が OOV である。OOV は対応する原言語の単語よりも 1 単語だけ後ろに存在するので、相対距離 1 を出力することで単語対応を獲得する。この手法では、トレーニングデータを用いて原言語と目的言語の単語アライメントを取る必要がある。

Jean ら [4] は OOV をその単語が持つアテンションの確率が最も大きい原言語の単語を翻訳辞書によって翻訳する後処理手法を提案した。ニューラル機械翻訳では単語を出力する際、入力文のどの単語へと注目するか（アテンション）という情報を用いる。そのアテンションの確率が最も高い入力単語に対応している単語であるとみなし、Luong らと同様に翻訳辞書を用いて直接翻訳する。この手法は単語アライメントを必要としない一方、目的言語での意味を考慮できていない。

さらに、Sennrich ら [5] は、系列に対するデータ圧縮手法である Byte Pair Encoding (BPE) を文字列に適用し、単語を頻出する部分文字列の系列に分解して学習することで OOV を削減した。この手法では部分文字列単位での翻訳を行い単語にするために部分文字列を結合する。図 2.2 は BPE を用いた分割例である。“@@”は本来結合されて単語になる部分を表している。日本語側では OOV である“デメリット”が“デ”と“メリット”という 2 つの高頻度な部分文字列に分割され、英語側においても同様に“demerit”が“de”と“merit”に分割される。翻訳後において“@@”がその後続く部分文字列と結合されることで 1 つの単語を生成するため、この例では OOV “demerit”を出力する。

この手法は、頻度に基づいた部分文字列への分割を行う貪欲なアルゴリズムである。そのため、意味を考慮せずに単語を部分文字列に分解する。例えば、日本語での OOV “ガリ”は“ガ”と“リ”に分けられるが、この 2 つの文字それぞれは意味を持っていない。そのため、翻訳結果では“g”と“asses”が出力され“gasses”（ガ

ス)となるため，翻訳前後で意味が変化する．また，結合によって生成される単語は実際には存在しない単語である可能性がある．日本語の例では“ピ@@ピ@@ネ@@メチル@@アミン”という出力から“ピピネメチルアミン”という単語が生成されるがこの単語は存在しない．この現象は生成された単語が実際に存在するかどうか分からないために起こり，それを確かめるためには人間が確認を行う必要がある．本研究では単語の生成は行わないため，出力される単語はすべて存在する単語である．

2.2 機械翻訳の前処理としてコーパスを書き換える関連研究

本研究のように機械翻訳の前処理段階においてコーパスを言い換え，原言語および/あるいは目的言語の文の複雑さを減少させる手法が存在する．Štajner ら [6] は機械翻訳の前処理として入力文の語彙と文法を平易にするテキスト平易化を適用した．本研究では，入力文のテキスト平易化を用いず，語彙の言い換えのみを用いて OOV の削減を試みる．

さらに，Li ら [7] は前処理の段階でトレーニングコーパスや入力文の OOV を使用する語彙に含まれる類義語に置換する手法を提案した．彼らは OOV を高頻度語に置換する際，単語同士の類似度や，置換後の文が自然であるかを判断するために言語モデルを使用した．言語モデルは文がどれほど自然であるかを判定するモデルであり，大規模なコーパスから作成される．文のスコアは一定区間の単語列がコーパス中でどれほどの頻度であるかに基づくため，文法が間違っているようなコーパス中に存在しにくい単語列の場合は不自然と判断される．単語の意味が似ているかどうかを判断するために，彼らは似た意味の単語はその文脈も似ているという分布仮説に基づいた分散表現を用いてコサイン類似度を計算した．また，彼らはトレーニングコーパスにおいて OOV の単語に対応する原言語，あるいは目的言語の単語アライメントを用い，OOV がどの単語にも対応していない場合その単語を削除した．

しかし，OOV の削除は文の意味の損失につながり，翻訳後の文の情報が欠落する．加えて，彼らはコサイン類似度を用いるため OOV を同義でない類義語に置換する可能性がある．例えば，彼らは“surfing”を“snowboard”に置換するため，

“internet surfing”を“internet snowboard”と書き換えてしまうため意味の変化が発生する．本研究では分布類似度ではなく，言い換えそれぞれに登録されている言い換え辞書のスコアを用いる．そのため，本手法では OOV の不適切な表現への言い換いを抑制できる．前述の例では，言い換え辞書に登録されている言い換えによって，“surfing”を“browser”へと言い換えるため元の意味をある程度残すことができる．

第 3 章 OOV を同義語に言い換える提案手法

本研究では、元の文の意味を保持しつつニューラル機械翻訳の OOV を減らすために、トレーニングデータの目的言語文に存在する OOV を高頻度語に言い換えてから翻訳する手法を提案する。我々は言い換え対および言い換えスコアが登録されている言い換え辞書を用いて OOV を高頻度語に言い換える。我々は 3 つのスコアを使用する：(1) 言い換え辞書スコア，(2) 言語モデルスコア，(3) これらのスコアを結合したスコア。言い換え辞書スコアは翻訳前後の意味の保持を考慮し，言語モデルスコアは文の自然さを考慮する。我々は言い換えスコアと言語モデルスコアを以下のように線形補間によって結合する：*

$$\begin{aligned} \text{言い換えスコア} = \\ \lambda(\text{言い換え辞書スコア}) + (1 - \lambda)(\text{言語モデルスコア}) \end{aligned} \quad (3.01)$$

3.1 言い換えラティスの構築

言い換えラティスを作成する Python の実装を図 3.1 に示す。言い換えの際、OOV を高頻度語に言い換える辞書 (OOV2in-vocabulary) と OOV を OOV に言い換える辞書 (OOV2OOV) を用いる。言い換える対象となる区間を指定し、高頻度語への言い換えがある場合に言い換えスコアを計算しもっと高いスコアである時に best_score にそのスコアを、best_edge にその言い換えを区間 begin2end の情報とともに記録する。

図 3.2 は言い換えラティスの例である。原文 “they assert defending the rights.” において、“defending” が OOV である。“defending” は高頻度語である “guaranteeing” への言い換えが可能であり，“defending the rights” は全て高頻度語である “the protection of the rights” への言い換えが可能である。この例では，“assert guaranteeing the rights.” の 2-gram 言語モデル確率，“assert the”，“rights.” の 2-gram 言語モデル確率を計算する。この手法ではフレーズの言い換えの際、フレーズ外の言語モデル確率は計算するが、フレーズ内の言語モデル確率は計

* 予備実験において、これらのスコアの正規化は性能の改善とはならなかった。

```

def make_lattice(list sentence, dictionary OOV2in-vocabulary, dictionary phrase2score):
    """
    sentence is a list of words: [w1, w2, ..., wn]
    OOV2in-vocabulary paraphrases phrases to in-vocabulary phrases
    Phrase2 score gives PPDB score
    """
    n = length(sentence)
    best_score = dictionary{}
    best_score = dictionary{}
    end2phrases = dictionary{}
    end2phrases[0] = "BOS"
    for begin in range(0, n):
        for end in range(begin+1, n+1):
            original_phrase = sentence[begin:end]
            phrases = sentence[begin:end]
            while (1):
                paraphrases = list()
                for phrase in phrases:
                    paraphrases.append(OOV2in-vocabulary[phrase])
                if paraphrases exist:
                    for paraphrase in paraphrases:
                        for before_phrase in end2phrases[begin-1]:
                            m = length(paraphrase)
                            PPDB_score = phrase2score[paraphrase] #PPDB: ParaPhrase DataBase
                            LM_score = 2-gram(wbegin-1, paraphrase1) + 2-gram(paraphrasem, wend+1) #LM: Language Model
                            score = best_score[str(begin-1) + " ||| " + before_phrase] + λ * PPDB_score + (1-λ) * LM_score
                            if score is the best:
                                begin2phrase = str(end) + " ||| " + paraphrase
                                best_score[begin2phrase] = score
                                best_edge[begin2phrase] = str(begin-1) + " ||| " + before_phrase
                        end2words[end].append(paraphrases)
                    break
            else:
                OOV_paraphrases = get_OOV2OOV(phrases)
                if OOV_paraphrases exist:
                    phrases = OOV_paraphrases
                else:
                    break
    return best_edge

```

図 3.1 言い換えラティスを作成する Python の実装

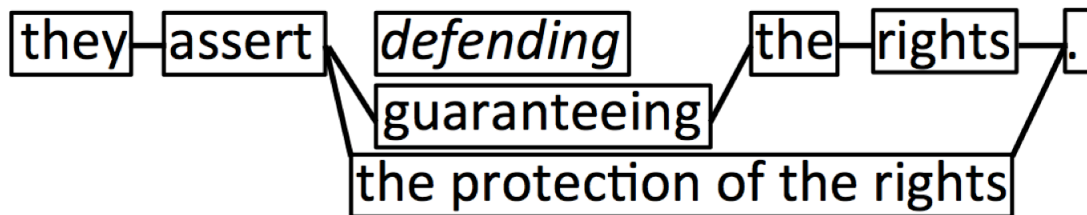


図 3.2 言い換えラティスの例

original: the *pedagogues* had *quarrels*.
 paraphrase, first round: the educators had discussions.
 paraphrase, second round: the teachers had discussions.

図 3.3 複数回言い換えの例

算しない。入力文を $X(x_0, x_1, \dots, x_m)$ 、低頻度語を x_i 、言い換えられる高頻度な 1 つのフレーズを $P(p_1, p_2, \dots, p_n)$ とすると、言語モデルスコアが計算されるのは、2-gram (x_{i-1}, p_1) および 2-gram (p_n, x_{i+1}) である。図 3.2 の例では “assert guaranteeing” や “assert the” などの言語モデル確率を計算し、フレーズ “the protection of the rights” の言語モデル確率は計算しない[†]。また、動名詞句 “defending the rights” に対しての言い換え “the protection of the rights” は名詞句であり、文法上の変化があるが文法は考慮しない。

さらに、我々の手法は OOV を高頻度語になるまで繰り返し言い換えることができる。複数回言い換えの例を図 3.3 に示す。図 3.3 の言い換え例では “pedagogues” 及び “quarrels” が OOV である。原文の OOV “quarrels” は 1 回目の言い換えで高頻度語 “discussions” へと言い換えられる。一方、OOV “pedagogues” は 1 回目の言い換えで OOV “educators” へと言い換えられる。我々はさらに “educators” を言い換えることが可能であり、2 回目の言い換えで高頻度語 “teachers” へと言い換えられる。OOV “pedagogues” は言い換えを 1 度のみ行う場合、“teachers” へは言い換えられない。なぜなら、言い換え辞書には “pedagogues” を “teachers” へ

[†] フレーズ内の言語モデルスコアの計算はニューラル機械翻訳を改善しなかった。

```

def select_paraphrase(dictionary best_edge, int sentence_length):
    words = list()
    current_index = sentence_length
    next_edge = best_edge[str(current_index) + " ||| " + "EOS"]
    while (current_index) > 0:
        current_word = next_edge.split(" ||| ")[1]
        words.append(current_word)
        next_edge = best_edge[next_edge]
        current_index = next_edge.split(" ||| ")[0]
    reverse words
    return words

```

図 3.4 言い換えを選択する Python の実装

と言い換える登録がないためである。

複数回の言い換えの場合も図 3.2 のような言い換えラティスを作成する。まず最初に言い換え辞書に登録されている言い換えすべてを用いてラティスを作成し、OOV でない高頻度語が存在する場合にその中からスコア最大の言い換えを選択する。一方、言い換えすべてが OOV の場合は OOV それぞれに対してさらなる言い換えを行い、2 回目の言い換えを生成する。さらなる言い換えによって OOV でない単語が出現した場合はスコアに基づいて言い換えを選択し、OOV のみの場合は再び言い換えを行う。この言い換えは OOV でない高頻度語への言い換えが行われるまで実行され、高頻度語への言い換えが行われずに他の OOV への言い換えができなくなるまで実行された場合は原文の単語のままにする。これは言い換えによって意味がわずかに変化するため、OOV から OOV への言い換えは出力の OOV を削減できずに文の意味を変化させるだけとなるからである。

3.2 言い換えの選択

言い換えラティスから言い換えを選択する Python の実装を図 3.4 に示す。前節で作成したラティスを用いて最もスコアが高くなる言い換えを選択し言い換え後の文を作成する。まず、それぞれの言い換え候補において、最も高いスコアとなる経路とそのスコアを動的計画法によって先頭から計算する。その後、文末までの経路

を作成し最も高いスコアの言い換えを文末からたどることで最もスコアが高くなる言い換えを選択する．図 3.2 の例では，“they assert the protection of the rights.”が生成される．

第 4 章 日英翻訳実験

4.1 実験設定

本研究では，アジア学術論文抜粋コーパス (ASPEC)[10] 日英対訳データを使用した．トレーニングにはアライメント確度によって分割されているトレーニングコーパスから最も高い文 100 万文すべてを採用し，そのうち，文長 40 単語以下の文 827,503 文対を使用し，チューニングには 1,790 文対すべてを，テストには 1,812 文対すべてを使用した．これらの文対はすべて重複していない．開発データセットを用いて最も精度の高いモデルを選択し，そのモデルによってテストデータセットを翻訳し評価した．コーパスの単語分割のために，日本語では MeCab[11] (IPAdic) を，英語では Moses [12] に付随するスクリプトを使用した．言語モデルには KenLM*を用いて，ASPEC の文全てを使用し 2-gram 言語モデルを構築した．単語アライメントの獲得には GIZA++[†]を使用した．言い換え辞書には英語では PPDB2.0 [13] の最大サイズ XXX-L[‡]を，日本語では PPDB:Japanese [14] を使用した．これらの辞書は ASPEC の内容を含まずに構築されたものである．PPDB2.0 は様々な分野のコーパスから作成されている一方，翻訳に使用している ASPEC コーパスは科学技術論文の内容のみであるため，言い換え辞書がコーパスに適応していないと考えられる．ASPEC コーパスと同じ分野の言い換え辞書を作成するため，PPDB を ASPEC コーパスで作成した．作成する際，公開されているコード[§]を使用し，閾値を 0.01 とした．本実験ではトレーニングコーパスの原言語側，目的言語側のどちらか，あるいは両方を言い換えた．原言語側の言い換えを行った場合，チューニング，テストの文も言い換えた．言い換えスコアの計算に用いる線形補間係数は $\lambda = 0.0, 0.25, 0.50, 0.75, 1.0$ とした．

ニューラル機械翻訳は，OpenNMT [15] を用いた．モデル構築のためのパラメータは以下の通りである．双方向リカレントネットワークを使用し，バッチサイズは 64，トレーニングのエポック数は 20，単語ベクトルの次元数は 500，原言語の

*<http://kheafield.com/code/kenlm/>

[†]<https://github.com/moses-smt/giza-pp>

[‡]gz 圧縮ファイルで 824MB

[§]<https://github.com/masahiro-mi/Paraphrasing>

表 4.1 日英翻訳の結果． * はブートストラップリサンプリングを用いて提案手法が統計的有意差 $p < 0.05$ で上回っていることを示す．

手法	BLEU	METEOR	OOV の数
ベースライン	26.13*	31.61	286
Luong ら	26.21*	31.30	406
Sennrich ら	27.32	32.72	0
Li ら	26.48	31.67	302
提案手法 (multi . (word))	26.73	31.93	275

語彙数，目的言語の語彙数は共に 30,000，ドロップアウトの割合は 0.3，最適化には SGD を用い，学習率は 1.0，エンコーダ側のリカレントネットワークは 2 層の LSTM，サイズは 500 とした．ベースラインは上記の設定でコーパスの言い換えを一切しないものとした．また，本論文で紹介した従来手法 [3, 7, 5] を上記の設定で再実験した．また，Li らの手法と比較するため，ベースラインと Li らの手法と提案手法では出力に OOV が出現する際，アテンションを用いて最も対応している確率の高い原言語の単語を翻訳辞書を用いて翻訳した．翻訳の評価には BLEU [16]，METEOR [17] を用いた．また，翻訳後に現れる OOV の数の変化による評価を行った．

PPDB2.0 の最大サイズに登録されている言い換えは質が悪い場合がある．PPDB2.0 は言い換えスコアに閾値をつけることで，質の高い言い換えのみが登録されているより小さい PPDB を提供している．PPDB の質が翻訳の質に関係するかを確かめるため，PPDB のサイズを小さくし質の高い言い換えのみを利用する実験を行った[¶]．また，目的言語の言い換えが効果的であるか，英語の言い換えが効果的であるかを明らかにするために，英日翻訳を行った．英日翻訳についても言い換え辞書に登録されている言い換えのうち，スコアの高いもののみを使用した翻訳を行った．

[¶]本研究では S-size と L-size を使用し，それぞれの容量は圧縮ファイルで 26.5MB，105MB である．

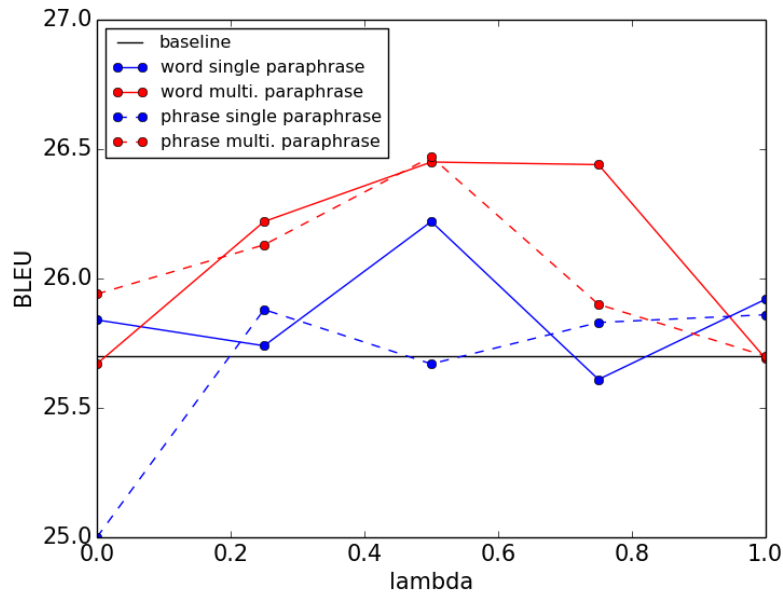


図 4.1 提案手法の日英翻訳の BLEU スコア

4.2 実験結果

先行研究と比較した日英翻訳の実験結果を表 4.1 に示す．提案手法は高頻度語への言い換えが見つかるまで単語の言い換えを行った．目的言語の OOV を言い換える提案手法はベースラインおよび Luong らの手法に対して，ブートストラップリサンプリングを用いた BLEU の評価において統計的有意差 ($p < 0.05$) を持って上回った．提案手法はベースラインと比較して BLEU スコアが 0.60，METEOR スコアが 0.31 上昇し OOV を 3% 減少させた．

線形補間係数と言い換え回数による提案手法の BLEU スコアを図 4.1 に示す．出力の OOV の数を図 4.2 に示す．図 4.1, 4.2 において，“word” は単語のみの言い換えを行い “phrase” は単語とフレーズの言い換えを行うことを示す．また，“single” は OOV に対して言い換えを 1 回まで行い “multi.” は高頻度語への言い換えが見つかるまで言い換えを行うことを示す．これらの図の結果において，アテンションを用いた OOV の翻訳は行っていない．提案手法は言い換え回数や線形補間係数に関係なく OOV の数を削減でき，特に複数回言い換えて言い換え辞書スコ

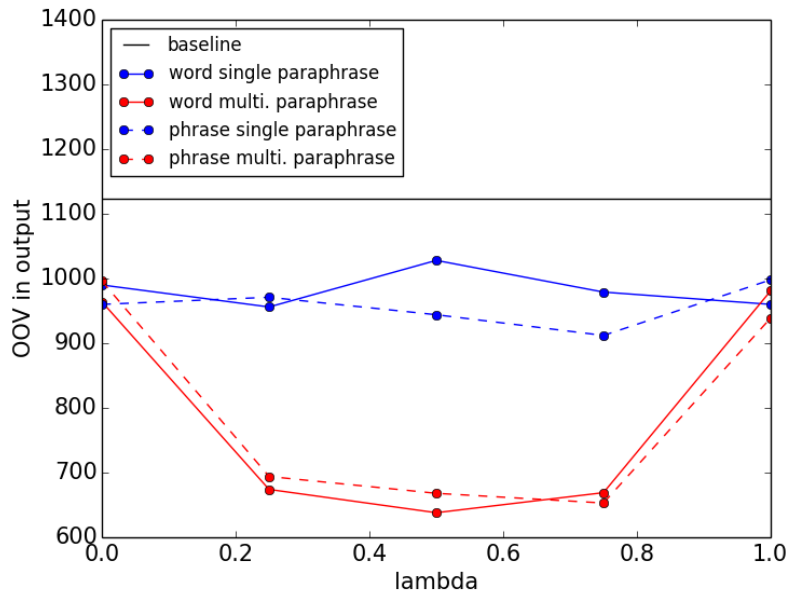


図 4.2 提案手法の日英翻訳の OOV の数

表 4.2 英日翻訳の結果． * はブートストラップリサンプリングを用いて閾値ありの提案手法が統計的有意差 $p < 0.05$ で上回っていることを示す．

手法	BLEU	OOV の数
ベースライン *	33.91	589
Luong ら	34.42	534
Sennrich ら	35.88	0
Li ら	34.22	544
提案手法 (multi . word)	34.07	573
提案手法 (multi . word , 閾値あり)	34.37	573

アと言語モデルスコアを組み合わせた時に、OOV をより多く削減できた。また、 $\lambda = 0.50$ の時に最も BLEU スコアを達成し、これは言い換えスコアは言い換え辞書スコアと言語モデルスコアをバランスよく足し合わせることが効果的であることを意味する。

英日翻訳の結果の BLEU スコアを表 4.2 に示す。ベ - スラインと比較すると提

表 4.3 言い換え辞書のサイズを変更した場合の日英翻訳の結果

手法	BLEU	METEOR	OOV の数	トレーニング中の OOV
言い換えなし	26.13	31.61	286	474,468
s-size (single)	26.13	31.57	279	453,158
s-size (multi.)	26.60	31.85	281	449,547
l-size (single)	25.90	31.53	281	441,112
l-size (multi.)	26.64	31.71	283	436,460
xxxl-size (single)	26.72	31.85	289	383,715
xxxl-size (multi.)	26.73	31.93	275	377,018

表 4.4 言い換え辞書を同分野にした場合の日英翻訳の結果

言い換え辞書	BLEU	METEOR	OOV	OOV (training)
未使用	26.13	31.61	286	474,468
PPDB2.0 (multi. word)	26.73	31.93	275	383,715
PPDB-ASPEC (multi. word)	26.41	31.73	289	385,173

表 4.5 言い換え対象を変更した場合の日英翻訳の結果

言い換え対象	BLEU	METEOR	OOV の数
原言語	26.69	31.82	295
目的言語	26.73	31.93	302
両方	26.63	31.78	281

案手法は BLEU スコアが 0.16 向上し、質のいい言い換えのみを使用すると 0.46 向上した。一方、翻訳後に存在する OOV の数は言い換え数が多いほど減少しているわけではない。これは OOV が出力される時にアテンションを用いて直接翻訳しているためだと考えられる。

PPDB2.0 のサイズを変更した日英翻訳の提案手法の結果を表 4.3 に示す。PPDB のサイズに関係なく 1 回のみ言い換えよりも複数回言い換えた方が翻訳精度が良かった。

PPDB-ASPEC を用いて言い換えた場合の翻訳結果を表 4.4 に示す。PPDB-ASPEC で言い換えを行った場合、言い換えをしないベースラインよりも翻訳精度

表 4.6 トレーニング文に対する人手評価

手法	翻訳の正しさ	言い換えの正しさ	文の自然さ
Li ら	1.63	1.46	2.41
提案手法	1.57	1.47	2.69

表 4.7 翻訳された文に対する人手評価

手法	翻訳の正しさ	文の自然さ
ベースライン	2.22	2.57
Sennrich ら	2.39	2.70
提案手法	2.18	2.61

が向上した一方 PPDB2.0 を用いた場合よりも悪い結果となった。これより、分野に対応した言い換え辞書を用いるよりも一般的な言い換え辞書を用いるほうが提案手法をよくすることが明らかとなった。この結果から分野に対応して専門用語を言い換えるよりも、一般的で低頻度な動詞や形容詞などを言い換えることが有効であると考えられる。

原言語と目的言語どちらか、あるいは両方を言い換えた日英翻訳の結果を表 4.5 に示す。原言語側のみを言い換える、両方を言い換える場合はベースラインを上回る一方、提案手法を改善しなかった。

トレーニング文に対する人手評価の結果および翻訳された文に対する人手評価の結果を表 4.6 , 4.7 に示す。この人手評価はトレーニング文対テスト文対それぞれ 200 文対に対して筆者、情報系大学院生の 2 人で 1~3 の 3 段階の評価を付与した。表内の数字は 2 人の評価の平均である。トレーニングの言い換えにおいて Li らの手法は高頻度語すべてを対象にしているため、文法の考慮があまりされていないことから文の自然さが提案手法よりも失われている。文をより自然に言い換えることで翻訳が改善されたと考えられる。また、テスト文の自然さの評価においてベースラインと提案手法が Sennrich らの手法に対して大きく下回っている。これは、Sennrich らの手法が OOV を一切出力しない一方ベースラインと提案手法は OOV を出力しており、OOV が出現した場合評価が最低の 1 であるためであると考えられる。

第 5 章 考察

図 4.1 および 4.2 から複数回の言い換えが 1 回のみ言い換えよりも BLEU スコアを高くし、出力に存在する OOV の数を削減することがわかる。OOV の数は BLEU スコアとは負の相関があり、本論文の仮説が正しいことを示している。一方、複数回の言い換えの場合、言い換えが成功していても元の意味を保持できるとは限らない。これは言い換えをすることによって、少しずつ意味のズレが発生することに起因する。したがって、複数回の言い換えによって OOV を高頻度語に言い換えるほど BLEU スコアの上昇幅が小さくなる傾向がある。

英日翻訳では言い換え辞書のスコアに閾値を適用することでベースラインよりも統計的有意差のある改善が得られた。日本語の言い換え辞書は不適切な言い換えが多く、それらを取り除くために閾値を用いることで適切な言い換えができた。その結果、提案手法は言語に関係なく目的言語の OOV の言い換えによってニューラル機械翻訳を改善できると考えられる。

表 4.3 から、PPDB のサイズや質に関係なく 1 回のみ言い換えよりも複数回言い換えた方が翻訳精度が良くなることがわかる。このことから、言い換え辞書に記載されていない言い換えを獲得し、より多く言い換えることによって学習すべき単語がより多く出現し、うまく学習が行われてモデルがより良くなると考えられる。PPDB のサイズが大きいほどより良い翻訳結果になる傾向がある一方、言い換えの質は翻訳精度にあまり関係していない傾向がある。言い換えを行う際スコアが高い言い換えを選択するため、スコアが高い言い換えがある場合は PPDB の大きさに関係なく同じ言い換えが選択され、これらの言い換えが翻訳の質を高めていると考えられる。

表 4.5 から、原言語のみの言い換えや両方の言い換えはさらなる改善とはならなかった。原言語側のみの言い換えは出力の OOV の減少に関係せず、言い換えをしないベースラインよりも翻訳精度が良くなる一方 OOV の数が増加している。両方を言い換えた場合、OOV を減少させているが、BLEU スコア、METEOR スコアは提案手法の中では最も悪くなっている。言い換えを行うことによって意味がわずかに異なり、原言語側の学習と目的言語側の学習それぞれが少し変化することにより、入力の実現と出力の実現が変わることが予想される。したがって、翻訳前後の

表 5.1 日英翻訳の例

method	translation
source	ロックインアンプを使用すれば，ノイズを著しく減少できることを期待できる。
reference	with the lock in amplifier used , significant reduction of the noise is expected .
baseline	it is expected that the noise can be reduced remarkably , if the <unk> is used .
multi. (word)	it is expected that the noise can be remarkably decreased , if the amplifier is used .
multi. (phrase)	it is expected that the noise can be remarkably reduced by using the lock-in amplifier .

表 5.2 ASPEC 特有の単語の言い換え例（複数回の言い換え）

言い換え前の OOV	言い換え後
megahertz	mhz
deflagration	combustion
cone-shaped	conical
revalued	examined
titrated	measured
teleportation	transport

意味が変化しやすくなることで翻訳精度が悪化したと考えられる。

翻訳例を表 5.1 に示す。ベースラインでは“amplifier”の代わりに“<unk>”を出力している。対して，言い換えを行う提案手法は“amplifier”に対応する単語が“amplifier”に言い換えられることで“amplifier”を出力できる。結果として，提案手法が正しく“amplifier”を出力できる。

ASPEC 特有の単語を複数の言い換えによって言い換えた例を表 5.2 に示す。提案手法は分野に特有な単語を言い換えることができ，これらの言い換えが翻訳を改善できると考えられる。表の中段は意味を保持できる言い換えである一方，下段は元の意味を少し失っている言い換えである。

第 6 章 おわりに

ある言語の文を異なる言語の文へと機械が自動で書き換える，機械翻訳の研究が盛んに行なわれており，近年，ニューラルネットワークを用いた機械翻訳手法である，ニューラル機械翻訳が提案された．ニューラル機械翻訳は softmax を出力の語彙サイズで取るため，トレーニングをする際に時間が掛かるという問題がある．そのため，すべての単語を用いることは現実的でなく，ニューラル機械翻訳では通常，使用する語彙を制限し，計算時間を削減する．その際，学習に用いる単語の出現頻度を用い高頻度な単語のみを使用し，それ以外の単語である低頻度語は語彙から外れる (Out-Of-Vocabulary; OOV)．この OOV は，まとめて 1 つの特殊記号 “<unk>” で出力され，意味を持たないため，翻訳前後の意味の保持ができなくなる．

先行研究では，トレーニングコーパスの OOV を高頻度な類義語に置換する前処理が提案されたが，この手法は出力文の OOV を減少させる一方，同義語でない類義語に置換することがある．また，この手法ではトレーニングコーパスにおいて，対応する単語が存在しない OOV を消去するため，翻訳前後で内容の損失が発生する．

本研究では，翻訳前後の意味を保持しつつ，OOV の出力を抑制する前処理手法を提案した．本手法ではトレーニングコーパスのうち，翻訳後の言語において OOV となる単語を高頻度な同義語にあらかじめ言い換えてから翻訳の学習を行った．言い換えの際には同義語の言い換えを収録しているデータベースを使用するため，言い換え前後での文の意味を変化させることなく出力の OOV を削減することができた．日英翻訳の実験の結果，言い換えを行わない翻訳よりも提案手法が翻訳結果の一致による BLEU スコアにおいて 0.60 ポイント向上し，同義語を正解とする METEOR スコアにおいて 0.32 ポイント向上した．

今後の課題として，言い換えの際文法を考慮することが挙げられる．本研究では言い換え辞書のエントリー全てを使用しており，言い換え前後の文法が変化していることがある．例えば，動名詞が普通名詞に言い換えられることが挙げられる．このような言い換えは文の意味を変化させるため文法を考慮する必要がある．また，OOV の言い換えの際，入力となる言語と出力となる言語両方を言い換える実験を

行ったが、翻訳を改善することができなかった。これはそれぞれを独立に言い換えているためである可能性があるため、両方を考慮した言い換えをすることで意味のズレを抑制できると考えられる。

謝辞

研究室配属から3年間研究の指導をしてくださった指導教員の小町 守先生，研究室の先輩である梶原 智之さん，PPDB:Japanese や PPDB を作成するコードを公開し，使用方法をお教えくださった水上 雅博さん，本研究の副査を務めてくださった山口 亨先生，高間 康史先生に多大なる感謝をしています．そして，秘書である馬縹 亜紀子さん，相談に乗ってくださった同期の金子 正弘さん，小平 知範さん，塩田 健人さん，研究室の先輩方後輩方の皆さんに感謝の意を申し上げます．

参考文献

- [1] D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” 3th International Conference on Learning Representations, 2015.
- [2] H. Mi, Z. Wang, and A. Ittycheriah, “Vocabulary manipulation for neural machine translation,” Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pp.124–129, 2016. <http://www.aclweb.org/anthology/P16-2021>
- [3] M.-T. Luong, I. Sutskever, Q. Le, O. Vinyals, and W. Zaremba, “Addressing the rare word problem in neural machine translation,” Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, pp.11–19, 2015. <http://www.aclweb.org/anthology/P15-1002>
- [4] S. Jean, O. Firat, K. Cho, R. Memisevic, and Y. Bengio, “Montreal neural machine translation systems for WMT’15,” Proceedings of the Tenth Workshop on Statistical Machine Translation, pp.134–140, 2015. <http://www.aclweb.org/anthology/W15-3014>
- [5] R. Sennrich, B. Haddow, and A. Birch, “Neural machine translation of rare words with subword units,” Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pp.1715–1725, 2016. <http://www.aclweb.org/anthology/P16-1162>
- [6] S. Štajner and M. Popovic, “Can text simplification help machine translation?,” *Baltic Journal of Modern Computing*, vol.4, no.2, pp.230–242, 2016.
- [7] X. Li, J. Zhang, and C. Zong, “Towards zero unknown word in neural machine translation,” Proceedings of the 25th International Joint Conference on Artificial Intelligence, pp.2852–2858, 2016. <http://www.ijcai.org/Proceedings/16/Papers/405.pdf>
- [8] S. Jean, K. Cho, R. Memisevic, and Y. Bengio, “On using very large target vocabulary for neural machine translation,” Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, pp.1–10, 2015. <http://www.aclweb.org/anthology/P15-1001>
- [9] M.-T. Luong and C.D. Manning, “Achieving open vocabulary neural machine translation with hybrid word-character models,” Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pp.1054–1063, 2016. <http://www.aclweb.org/anthology/P16-1100>

- [10] T. Nakazawa, M. Yaguchi, K. Uchimoto, M. Utiyama, E. Sumita, S. Kurohashi, and H. Isahara, “ASPEC: Asian scientific paper excerpt corpus,” Proceedings of the 10th edition of the Language Resources and Evaluation Conference, pp.2204–2208, 2016. http://www.lrec-conf.org/proceedings/lrec2016/pdf/621_Paper.pdf
- [11] T. Kudo, K. Yamamoto, and Y. Matsumoto, “Applying conditional random fields to Japanese morphological analysis,” Proceedings of the Empirical Methods in Natural Language Processing, pp.230–237, 2004. <http://www.aclweb.org/anthology/W/W04/W04-3230.pdf>
- [12] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. Dyer, O. Bojar, A. Constantin, and E. Herbst, “Moses: Open source toolkit for statistical machine translation,” Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions, pp.177–180, 2007. <http://www.aclweb.org/anthology/P07-2045>
- [13] E. Pavlick, P. Rastogi, J. Ganitkevitch, B. Van Durme, and C. Callison-Burch, “PPDB 2.0: Better paraphrase ranking, fine-grained entailment relations, word embeddings, and style classification,” Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics, pp.425–430, 2015. <http://www.aclweb.org/anthology/P15-2070>
- [14] M. Mizukami, G. Neubig, S. Sakti, T. Toda, and S. Nakamura, “Building a free, general-domain paraphrase database for Japanese,” Proceedings of the 17th International Oriental Chapter of the International Committee for the Co-ordination and Standardization of Speech Databases and Assessment Techniques, pp.1–4, 2014. <http://www.phontron.com/paper/mizukami14cocosda.pdf>
- [15] G. Klein, Y. Kim, Y. Deng, J. Senellart, and A.M. Rush, “OpenNMT: Open-source toolkit for neural machine translation,” arXiv preprint arXiv:, vol.1701.02810, pp.1–6, 2017.
- [16] P. Kishore, R. Salim, W. Todd, and Z. Wei-Jing, “BLEU: a method for automatic evaluation of machine translation,” Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, pp.311–318, 2002. <http://aclweb.org/anthology/P/P02/P02-1040.pdf>
- [17] A. Lavie and A. Agarwal, “METEOR: An automatic metric for MT evaluation with high levels of correlation with human judgments,” Proceedings of the Second Workshop on Statistical Machine Translation, pp.228–231, 2007. <http://www.cs.cmu.edu/~alavie/METEOR/pdf/Banerjee-Lavie-2005-METEOR.pdf>

発表リスト

査読なし

・ 関沢祐樹, 梶原智之, 小町守. 語構成情報と言い換えパターンを用いた二字漢字の句への言い換え. 言語処理学会第 22 回年次大会. pp.725-728. 2016 年. http://www.anlp.jp/proceedings/annual_meeting/2016/pdf_dir/B4-4.pdf

・ 関沢祐樹, 梶原智之, 小町守. 目的言語の低頻度語の高頻度語への言い換えによるニューラル機械翻訳の改善. 言語処理学会第 23 回年次大会. pp.982-985. 2017 年. http://www.anlp.jp/proceedings/annual_meeting/2017/pdf_dir/P17-1.pdf

査読あり

・ Yuuki Sekizawa, Tomoyuki, Kajiwara, Mamoru Komachi. Improving Japanese-to-English Neural Machine Translation by Paraphrasing the Target Language. The 4th Workshop on Asian Translation (WAT2017). pp.64-69. 2017. <http://www.aclweb.org/anthology/W/W17/W17-5703.pdf>