

対訳句追加による NMT の翻訳

浅井 奏人 村上 仁一

鳥取大学 工学部

b18t2002m@edu.tottori-u.ac.jp

murakami@tottori-u.ac.jp

概要

近年、機械翻訳の分野においてニューラル機械翻訳 (Neural Machine Translation; NMT) がある。しかし、NMT の翻訳精度は人手による翻訳の精度には及ばない。翻訳精度の向上には大量の対訳学習文を用いる必要がある。しかし、学習に用いるデータが文のみの場合精度の向上に限界がある。本論文では、日英 NMT において対訳句を利用する手法を示す。次に、対訳句の正答率を変化した場合について述べる。最後に、初期値を変更した場合について述べる。

1 はじめに

近年、機械翻訳の分野においてニューラル機械翻訳 (Neural Machine Translation; NMT) がある。NMT は、従来の手法と比較して、より流暢性の高い翻訳を出力することができる。一方で、NMT の翻訳精度は人手による翻訳の精度には及ばない。翻訳精度の向上には大量の対訳学習文を用いる必要がある。これを行うには、コストがかかる。

そこで本研究では、日英 NMT において対訳句を利用する手法を提案する。日英 NMT の学習において、対訳学習文に、対訳学習文から抽出した対訳句を追加する。対訳句を追加することで、語句の対応情報を強化できるため、翻訳精度を向上すると仮定する。

2 対訳句の利用

要素合成法の問題を以下に示す。この問題は、影響の大小はあるが、すべての翻訳システムに当てはまる。

1. 文を翻訳するとき、入力文を単語ごとに翻訳しても、文の意味は間違った翻訳になることが多い。

2. 翻訳精度を向上させるには、文全体の意味を考慮して翻訳する必要がある。これを考慮するために、できるだけ長い句ごとに翻訳を行う必要がある。

例を以下に示す。

- 入力文: 彼は我を通した。
- 参照文: He had his own way.
- 出力文: He passed me.

一般的に、「通す」は"pass", 「我」は"me"と訳すことが多い。しかし、文全体の意味を考えると間違った翻訳になる。そこで、対訳学習文に対訳句を追加学習することで、文全体の意味を考慮した翻訳を行うことができる。具体的には、対訳学習文に、以下に示す対訳句を追加学習する。

- 日本語句: 我を通した
- 英語句: had his own way

対訳学習文に対訳句を追加することで、翻訳精度が向上すると考えられる。

3 提案手法

本研究では、対訳学習文から抽出した対訳句を、対訳学習文に追加する。対訳句追加の流れを図 1 に示し、手順を以下に示す。

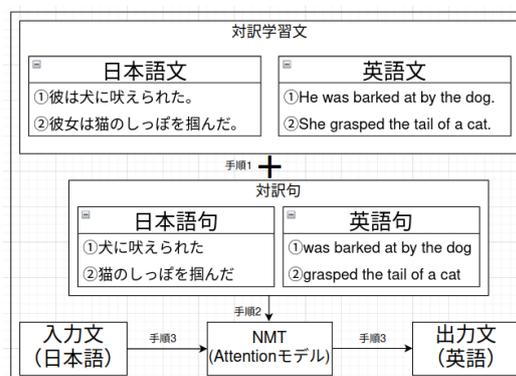


図 1. 対訳句追加の流れ

手順1対訳学習文と、対訳学習文から抽出した対訳句を組み合わせて、新たな学習データを作成する。

手順2手順1で作成した学習データを用いてNMTのモデルの学習を行う。

手順3手順2で学習したモデルを用いて日英ニューラル機械翻訳を行う。

4 実験

4.1 対訳学習データ

本研究では、実験データとして日英重文複文コーパス [1] から抽出した単文を用いる。対訳学習データの内訳を表1に、対訳学習文の例を表2に示す。

表1. 実験データの内訳

対訳学習文	163188 文
テスト文	16328 文

表2. 対訳学習文の例

入力文 1: 星印のノートを買った。	出力文 1:I bought a star brand notebook .
入力文 2: この動物園には、コアラがいます。	出力文 2:There are koala bears in this zoo .
入力文 3: 彼は、この会社の主要な人物である。	出力文 3:He is a central figure in this company .

4.2 対訳句

本研究の対訳句は、対訳学習文から自動抽出したものをを用いる。(森本 [2] 参照) 正答率は95%である。対訳句の例を表4に示す。

表3. 対訳句データの内訳

対訳句	41926 句
-----	---------

表4. 対訳句の例

日本語句	英語句
星印のノート	star brand notebook
コアラがいます	are koala bears
この会社の主要な人物	central figure in this company

4.3 評価実験

本研究では、NMTの学習データに対訳学習文のみを用いる手法(以下、ベースライン)と、対訳学習

文に対訳句を追加した学習データを用いる手法(以下、提案手法)を比較する。

評価方法として、自動評価と人手評価を行う。自動評価では実験で得られた16328文の出力文に対して、4つの自動評価指標(BLEU, METEOR, RIBES, TER)で評価を行う。人手対比較評価では、出力文16328文より無作為に抽出した100文に対して、正確性に基づいて対比較評価を行う。

5 実験結果

5.1 自動評価

出力文16328文において、提案手法とベースラインの自動評価の結果を表5に示す。

表5. 自動評価結果

手法	BLEU	METEOR	RIBES	TER
提案手法	0.2002	0.4686	0.7744	0.6060
ベースライン	0.1856	0.4564	0.7656	0.6215

表5より、自動評価の結果では、提案手法の翻訳精度が高いことが確認できる。

5.2 人手評価

出力文16328文から無作為に抽出した100文に対して人手評価を行った。人手評価は、提案手法とベースラインとの対比較評価である。評価者は5名である。結果を表6に示す。表6の結果は、5名が同じ100文を評価し、その結果をまとめた。また、評価の説明を以下に示す。

- 提案手法○
提案手法とベースラインを比較した際、対訳句追加時の出力文の方が入力文の意味に近い。
- ベースライン○
提案手法とベースラインの比較した際、ベースラインの出力文の方が入力文の意味に近い。
- 差なし
提案手法とベースラインを比較した際、両手法の出力文の評価には差がない。
- 同一
提案手法とベースラインを比較した際、両手法の出力文がほぼ同じである。

表 6. 人手対比較評価結果

提案手法○	ベースライン○	差なし	同一
27 文	14 文	44 文	15 文

表 5 より、提案手法とベースラインを比較して、提案手法の方が翻訳精度が高いことが確認できる。

5.3 人手評価例

ベースラインと提案手法との対比較評価において、提案手法○とした例を表 7 に、ベースライン○とした例を表 9 に、差なしの例を表 10 にそれぞれ示す。

5.3.1 提案手法○の例

表 7. 提案手法○の例

入力文	森の静寂の中でかすかな音が聞こえた。
参照文	A slight sound could be heard in the quietude of the woods .
ベースライン	I heard a faint sound in the woods .
提案手法	I heard a faint noise in the stillness of the woods .

表 6 の例において、入力文中の語句「静寂」に対して、ベースラインでは相当する翻訳が得られていない。一方、提案手法では比較的正しい出力「stillness」を得られている。表 7 で使用した対訳句を表 8 に示す。

表 8. 使用した対訳句

日本語句	英語句
森	the woods
静けさ	stillness
森の静寂の中	quietude of the woods
かすかな	a faint
音	noise
聞こえた	heard

5.3.2 ベースライン○の例

表 9. ベースライン○の例

入力文	強風で屋根が剥がされた。
参照文	The roof was torn off by the strong wind .
ベースライン	The roof was blown out by the strong wind .
提案手法	The roof was stripped of its roof .

表 9 の例において、提案手法では一部しか翻訳できていない。一方、ベースラインでは比較的正しい出力文を得られている。

5.3.3 差なしの例

表 10. 差なしの例

入力文	その会社は輸出入業を行なっている。
参照文	The company deals in importing and exporting .
ベースライン	The company administers a profitability trade .
提案手法	The company has an open business .

表 10 の例において、ベースラインおよび提案手法に入力文中の語句「輸出入業」に対する出力が得られていない。このことから、どちらも翻訳ができていないと判断して差なし判定とした。

6 考察

6.1 対訳句の正答率に対する翻訳精度の変化

本研究で用いた対訳句の正答率は、95%である。この精度が低下したときの、翻訳精度の変化を調査した。本実験の対訳句の正答率を 100%と仮定して、正答率を 50%、75%、90%に設定した。表 11 に自動評価の結果を示す。

表 11. 正答率変更後の自動評価結果

手法	BLEU	METEOR	RIBES	TER
ベースライン	0.1856	0.4564	0.7656	0.6215
100%	0.2002	0.4686	0.7744	0.6060
90%	0.1986	0.4686	0.7719	0.6076
75%	0.1855	0.4470	0.7558	0.6266
50%	0.1743	0.4357	0.7468	0.6345

表 11 より対訳句の正答率を下げた場合、翻訳精度が下がることを確認した。75%まで下げた場合、翻訳精度がベースラインより低下している。

6.2 初期値の変更

NMT は、初期値によって翻訳精度が変化する。本節では、提案手法と別の初期値を用いて翻訳した。結果を表 12 に示す。

表 12. 初期値変更後の自動評価結果

手法	BLEU	METEOR	RIBES	TER
ベースライン	0.1841	0.4542	0.7674	0.6219
100%	0.2022	0.4673	0.7735	0.6096
90%	0.1991	0.4695	0.7737	0.6225
75%	0.1946	0.4653	0.7716	0.6104
50%	0.1595	0.4095	0.7176	0.6649

表 12 より、初期値を変更すると、翻訳精度が変化していることがわかる。しかし、表 11 と表 12 を比較すると、ほぼ同じ傾向があることがわかる。つまり、対訳句の正答率が低いと翻訳精度は向上しない。

6.3 対訳句追加による効果

本実験の結果として、対訳句を追加することで翻訳精度が向上することが確認できた。しかし、対訳句の正解率を 75% もしくは 50% まで下げた場合、対訳句追加前の翻訳精度を下回ることを確認した。このことから翻訳精度の向上には、高精度な対訳句を追加する必要があると推察される。

6.4 関連研究

6.4.1 PBSMT に対訳句を追加

本論で述べた手法と似た手法を用いる機械翻訳に、句に基づく統計機械翻訳 (Phrase-Based Statistical Machine Translation; PBSMT) がある。(日野 [3]) 日野は、日英翻訳において PBSMT を用いて翻訳精度の向上に成功している。なお、日野の研究結果と本論の研究結果から、NMT の翻訳精度は PBSMT と比較して対訳句の精度による影響が大きいと推察される。

6.4.2 複文での調査

本研究の関連研究として、今仁 [4] の研究を挙げる。今仁は本論で述べた実験を、自動で作成した対訳句と人手で作成した対訳句で比較を行った。その結果として、高精度かつ多量の手対訳句を用いた翻訳が精度が高いことが確認できた。また、本論で述べた研究は単文で行ったが、今仁は複文で行っていた。複文で行った場合、人手評価が不安定になる問題点がある。

7 おわりに

NMT の翻訳精度は人手の翻訳には及ばないことが問題となっていた。そこで、本研究では、対訳学習文から作成した対訳句を、対訳学習文に追加する手法を提案した。実験の結果、自動評価および人手評価ともに、提案手法の有効性が確認できた。一方で、対訳句の精度が低下した場合、翻訳精度はベースラインより低くなることが確認できた。翻訳精度が向上した理由は、対訳句を追加することで、語句の対応情報が強化されたと考えられる。今後はより人手の翻訳精度に近づける手法を検討し、さらなる翻訳精度の向上を試みたい。

謝辞

人手評価には、以下 5 名の学生の協力を得ました。感謝いたします。(新田 玲輔, 矢野 貴大, 齋藤 永, 森 唯人, 柳原 弘哉)

参考文献

- [1] 村上仁一, 藤波進. 日本語と英語の対訳文対の収集と著作権の考察. 第一回コーパス日本語学ワークショップ, pp.119-130, 2012.
- [2] 森本世人, 村上仁一. 類似度を利用した b 変換テーブルの精度向上. 2021 年度鳥取大学卒業論文, 2021.
- [3] 日野聡子, 村上仁一, 徳久雅人, 村田真樹. 日英統計翻訳における対訳句コーパスの効果. 言語処理学会台 19 回年次大会, pp.556-559, 2013.
- [4] 今仁優希, 村上仁一. 日英ニューラル機械翻訳における対訳句の追加. 言語処理学会台 26 回年次大会, pp.169-172, 2020.