

2021年度(令和3年度) 卒業論文

対訳句の追加によるNMTの翻訳

指導教員

村上仁一
村田真樹

鳥取大学工学部 電気情報系学科

自然言語処理研究室

B18T2002M 浅井 奏人

概要

近年，機械翻訳の分野において Neural Machine Translation[1] 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．翻訳精度の向上の手法の一つに，大量の対訳学習文を用いる手法がある．しかし，学習に用いるデータが対訳学習文のみの場合，要素合成法の問題から精度の向上に限界がある．

要素合成法とは、入力文を単語ごとに翻訳し，各翻訳結果を組み合わせることで出力文を形成する手法である．要素合成法の問題として，翻訳結果の文全体の意味が通じなくなることもある．そこで，入力文を単語ごとではなく句ごとの翻訳で，文の意味を保ったまま翻訳が可能であると推察する．

本論文では，日英 NMT において対訳句を利用する手法 (以下，提案手法) を示す．対訳句は対訳学習文から自動抽出したものをを用いる．対訳句を追加で，入力文を句ごとに翻訳し，文全体の意味が通じる出力文が形成される．

評価は BLEU 等の自動評価値による自動評価と対比較で人手評価を行った．自動評価においては，提案手法がベースラインより自動評価値が良好であった．人手評価においては，提案手法がベースラインより良好だと評価された出力文が 173 文に対して，ベースラインが提案手法より良好だと評価された出力文が 112 文となった．両評価結果より，ベースラインと比較して提案手法では翻訳精度が向上することが確認された．

考察として，対訳句の精度を変更した場合の実験を行った．対訳句の精度を 50 %，75 %，90 % にそれぞれ変更し，学習を行った．結果として，対訳句の精度が 75 % の時点で翻訳精度はベースラインより低下していることが確認された．この結果から，提案手法において高い対訳句の精度が必要であると推察される．本研究で対訳句の追加をする手法で翻訳精度が向上した理由は，対訳句の追加で入力文を句ごとに翻訳したからだと考えている．句ごとの翻訳で入力文の意味が通じる出力文を得ることができる．対訳句は対訳学習文から自動で抽出したものをを用いており，外部の情報を持っていない．

本実験では，間違えてクローズドデータを用いて実験を行った．本論で述べる考え方を参考にしてほしい．

目次

第1章	はじめに	1
第2章	NMTについて	2
2.1	NMTの概要	2
2.2	学習と翻訳	2
2.3	Encoder-Decoderモデル	3
2.4	Attentionモデル	3
第3章	要素合成法	4
3.1	要素合成法の問題	4
3.2	要素合成法の問題の解決方法	5
第4章	提案手法	6
第5章	実験	7
5.1	対訳学習データ	7
5.2	対訳句	8
5.3	評価実験	9
第6章	実験結果	10
6.1	自動評価	10
6.2	人手評価	11
6.2.1	人手評価例	11
6.2.2	提案手法の例	12
6.2.3	ベースラインの例	13
6.2.4	差なしの例	14

第7章	考察	15
7.1	対訳句の精度と翻訳精度	15
7.2	初期値の変更	16
7.3	対訳句追加による効果	16
7.4	要素合成法の問題の緩和	16
第8章	関連研究	18
8.1	PBSMT に対訳句を追加	18
8.2	対訳学習文から作成した対訳句の追加	18
8.3	複文での調査	18
第9章	おわりに	19

目 次

2.1 NMT の仕組み	3
4.1 対訳句追加の流れ	6

表 目 次

5.1.1 実験データの内訳	7
5.1.2 対訳学習文の例	7
5.2.1 対訳句データの内訳	8
5.2.2 対訳句の例	8
6.1.1 自動評価結果	10
6.2.1 人手対比較評価結果	11
6.2.2 提案手法 の例	12
6.2.3 入力文 1 で使用した対訳句	12
6.2.4 入力文 2 で使用した対訳句	12
6.2.5 ベースライン の例	13
6.2.6 差なしの例	14
7.1.1 精度変更後の自動評価結果	15
7.1.2 精度低下時の翻訳結果	15
7.2.1 初期値変更後の自動評価結果	16

第1章 はじめに

機械翻訳の方式の一つに、Neural Machine Translation(以下、NMT)[1]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。しかし、対訳学習文だけでは要素合成法の観点から翻訳精度の向上に限界がある。

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。対訳句を追加する理由は、語句のつながりを強化できるからであると推察される。対訳句を追加で語句の対応情報を強化し、日英NMTの翻訳精度向上を試みる。

第2章 NMTについて

2.1 NMTの概要

NMTは、近年提案された機械翻訳の手法である。NMTは学習と翻訳の2つの過程を経て出力文を形成する。NMTの学習には、日本語文と英語文の対訳文を用いる。NMTの手法には、Encoder-Decoderモデルとそれを拡張したAttentionモデルが提案されている。Encoder-Decoderモデルは入力系列を固定長のベクトルに符号化(encode)し、固定長のベクトルより出力系列を復号化(decode)するモデルである。入出力の系統は単語のベクトル表現となる。AttentionモデルはEncoder-Decoderモデルの出力系統を生成する際に、出力系列と入力系列の参照を行う機構(Attention)を用いた手法である。Encoder-Decoderモデルと比較して、長い入力文に対する妥当な出力文を形成できる。

2.2 学習と翻訳

NMTは、学習の過程で対訳学習文から日本語と英語の語句の対応と語句間のベクトル関係を学習し、モデルを作成する。

(例:吠えられた=barked, 掴んだ=grasped など)

翻訳の過程では、NMTが学習の過程で作成したモデルを用いて、入力文を翻訳、出力文を形成する。NMTの全体的な仕組みを図2.2に示す。

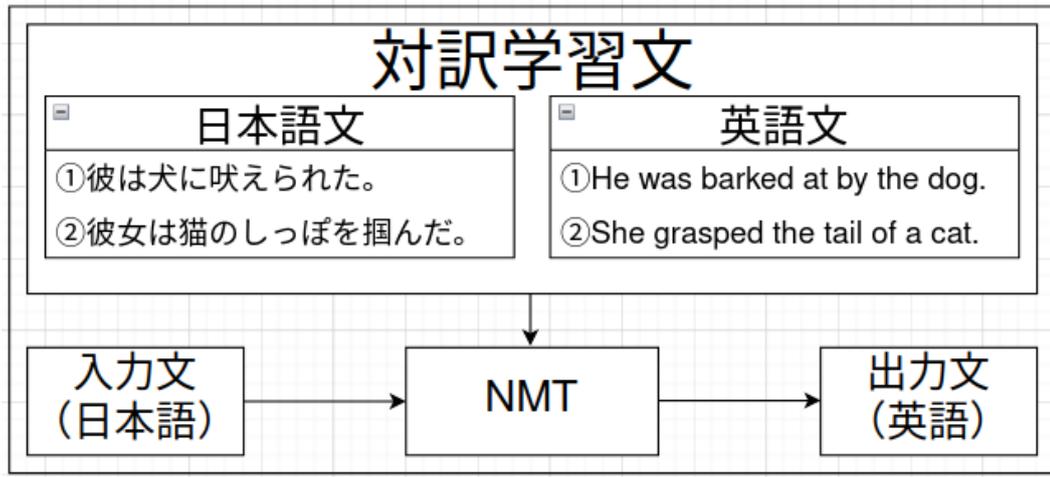


図 2.1: NMT の仕組み

2.3 Encoder-Decoder モデル

Encoder-Decoder モデルとは、2つのリカレントニューラルネットワーク (RNN) により構成されるニューラル機械翻訳のモデルである。1つの RNN が入力系列を1つの固定長のベクトルに符号化 (encode) し、もう1つの RNN により固定長のベクトル符号を出力系列へと復号化 (decode) する。

2.4 Attention モデル

Encoder-Decoder モデルの問題は、学習に用いた対訳学習文以上の長さの入力文を翻訳した場合に、翻訳精度が極端に低下する点である。

Attention モデルは、この問題を解決するために提案された NMT のモデルである。Attention モデルの Encoder は、入力単語を前後両方向から RNN に渡す手法を用いる。Encoder は、入力文を前から読み込んだノードと後ろから読み込んだノードを組み合わせることで各単語を符号化する。Decoder では、モデルが出力文中にある単語を生成する際に、生成した単語のベクトルと対訳学習文から次の単語を予測して出力文を出力する。Attention モデルは、Encoder-Decoder モデルと異なり、入力文の情報を出力文中の単語生成時に参照することで、より長い入力文の翻訳精度の向上を実現している。

第3章 要素合成法

3.1 要素合成法の問題

要素合成法とは、入力文を単語ごとに翻訳し、各翻訳結果を組み合わせることで翻訳を行う手法である。

要素合成法の問題は、ルールベース機械翻訳 (RMT) の時代から存在していた。しかし、未だに明確な解決方法が発見されていない。NMT においても同様の問題が存在すると推察されている。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。

- (a) 文を翻訳するとき、入力文を単語ごとに翻訳しても文全体の意味としては間違った翻訳になることが多い。
- (b) 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ために、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 1:彼は我を通した。
- 参照文 1:He had his own way.
- 出力文 1:He passed me.
- 入力文 2:彼女は公園で道草を食った。
- 参照文 2:She wasted time in the park.
- 出力文 2:She ate the road grass in the park.

要素合成法の翻訳方法は、学習した内容で一番頻度が高い単語を利用して翻訳する。そのため、文全体の意味を考えると間違った翻訳になる場合がある。

3.2 要素合成法の問題の解決方法

対訳学習文に対訳句を追加学習する．対訳句の追加で，NMT が語句の対応情報を強化し，句ごとに翻訳を行う可能性が高くなる．上記の例に，以下に示す対訳句を追加学習する．

- 日本語句 1:我を通した
- 英語句 1:had his own way
- 日本語句 2:道草を食う
- 英語句 2:wasted time

上記の例から，対訳句を追加学習で翻訳精度が向上すると考えられる．

第4章 提案手法

本章では、提案手法について記述する。本研究では、日英 NMT の学習データに対訳学習文に対訳句を追加したデータを用いる。日英ニューラル機械翻訳における対訳句追加の流れを図 4.1 に示し、手順を以下に示す。

手順 1 対訳学習文と対訳句を組み合わせる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。

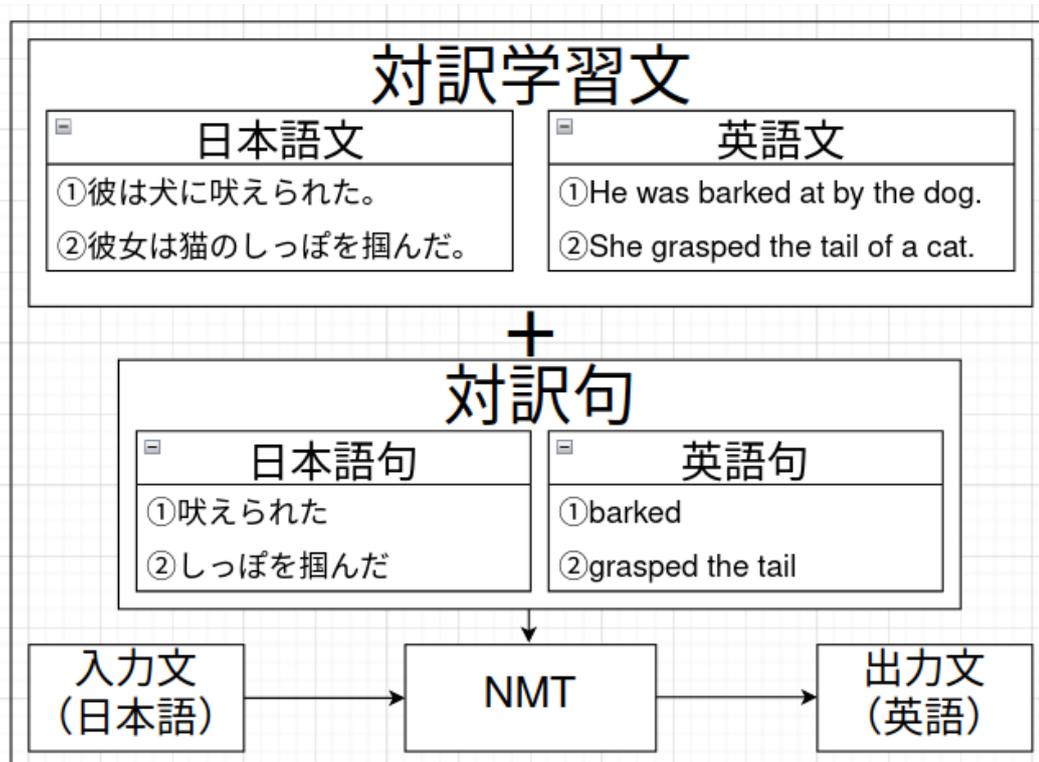


図 4.1: 対訳句追加の流れ

第5章 実験

5.1 対訳学習データ

本研究では、実験データとして電子辞書 [2] から抽出した単文を用いる。対訳学習データの内訳を表 5.1.1 に、対訳学習文の例を表 5.2.2 に示す。

表 5.1.1: 実験データの内訳

対訳学習文	163188 文
テスト文	16328 文

表 5.1.2: 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 .

5.2 対訳句

本研究の対訳句は, 森本 [3] が対訳学習文から自動作成したものをを用いる . 精度は 95 % である .

森本は , 変換テーブルの「A=B なら C=D である」の関係に基づき , 対訳学習文対から対訳句を自動作成した .

対訳句データの内訳を表 5.2.1 に , 対訳句の例を表 5.2.2 に示す .

表 5.2.1: 対訳句データの内訳
対訳句 41926 句

表 5.2.2: 対訳句の例

日本語句	英語句
星印のノート	star brabd 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

森本は , 対訳学習文対から「A が B なら C は D である」という変換テーブルを用いて自動で作成している . 作成方法を示す . 例として , 「犬」と「猫」を生成する .

手順 1 対訳文 (日本語文 1 と英語文 1) と対訳単語確率を用いて対訳単語を作成する . 対訳単語は変換テーブルの A と B に相当する .

- 日本語文 1: 私は犬が好きだ。
- 英語文 1: I like a dog.

手順 2 手順 1 で作成した対訳単語に相当する部分を変数に置き換えてパターンを作成する .

- パターン (日本語): 私は X が好きだ。
- パターン (英語): I like a X.

手順3 対訳文(日本語文2と英語文2)とパターンを照合する。パターンにおいて変数が相当する部分は変換テーブルのCとDに相当する。

- 日本語文2 2:私は猫が好きだ。
- 英語文2:I like a cat.

5.3 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6人がそれぞれ対比較評価を行い、合計600文の集計を取った。

第6章 実験結果

6.1 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 6.1.1 に示す．

表 6.1.1: 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 6.1.1 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．

6.2 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。評価方法は、同じ 100 文 6 人がそれぞれ評価し、集計を取った。結果を表 6.2.1 に示す。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く、両出力文が不一致である。

- 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 6.2.1: 人手対比較評価結果

提案手法	ベースライン	差なし	同一
173 文	112 文	219 文	96 文

表 6.2.1 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。

6.2.1 人手評価例

ベースラインと提案手法との対比較評価の結果を示す。

6.2.2 提案手法 の例

提案手法 としての例を表 6.2.2 に示す .

表 6.2.2: 提案手法 の例

入力文 1	私はそのグループに何か違和感を覚えた。
参照文 1	I felt somewhat awkward with that group .
ベースライン 1	I felt something affected in the group .
提案手法 1	I felt somewhat awkward with the group .
入力文 2	森の静寂の中でかすかな音が聞こえた。
参照文 2	A slight sound could be heard in the quietude of the woods .
ベースライン 2	I heard a faint sound in the woods .
提案手法 2	I heard a faint noise in the stillness of the woods .
入力文 3	かばんにあの本が全部収まりました。
参照文 3	I have put all those books in the bag .
ベースライン 3	All that book sank into the shape of the book .
提案手法 3	I've got all that book in my bag .

表 6.2.2 で使用した対訳句を表 6.2.3 と表 6.2.4 に示す .

表 6.2.3: 入力文 1 で使用した対訳句

日本語句	英語句
そのグループに何か違和感を覚えた	felt somewhat awkward with that group

表 6.2.4: 入力文 2 で使用した対訳句

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

入力文 1 では、「そのグループに何か違和感を覚えた」と”felt somewhat awkward with that group”の対訳句を学習したため、翻訳精度が向上したと考えられる .

入力文 2 では、提案手法がベースラインより良好だと評価した . しかし、「森の静寂の中」と”quietude of the woods”の対訳句を学習したにもかかわらず「静寂」に対応する語句

として”stillnes”が出力されている．NMT が学習した対訳句と違う翻訳結果を出力した理由は，「森の静寂の中」とは別に学習した対訳句”stillness”が乱数によって「静寂」の翻訳として出力されたからだと考えられる．

入力文3では，提案手法がベースラインより良好と評価したが，学習した対訳句が存在しなかった．対訳句の学習無しで提案手法の翻訳結果が良好であった理由は，乱数によって偶然提案手法がベースラインより良好になったからだと考えられる．

6.2.3 ベースライン の例

ベースライン とした例を表 6.2.5 に示す．

表 6.2.5: ベースライン の例

入力文 4	強風 で 屋根 が 剥がされた 。
参照文 4	The roof was torn off by the strong wind .
ベースライン 4	The roof was blown out by the strong wind .
提案手法 4	The roof was stripped of its roof .
入力文 5	彼の 豪邸 は 町中の 評判 になった 。
参照文 5	His palatial residence became the talk of the town .
ベースライン 5	His mansion attracted a reputation from the town .
提案手法 5	His high health has become a whole town .

入力文4は，ベースラインでは「強風」に対する出力”the strong wind”が得られている．一方で，提案手法では「強風」に対する出力が得られていない．そのため，ベースラインが提案手法より良好だと評価した．

入力文5では，ベースラインでは「豪邸」に対する出力”mansion”が得られている．一方で，提案手法では「豪邸」に対する出力が”high health”で，適切な出力ではない．そのため，ベースラインが提案手法より良好だと評価した．

人手評価で行った 100 文において，ベースライン と評価したときに出力された提案手法の出力文では対訳句が使用されていないことを確認した．

6.2.4 差なしの例

差なしの例を表 6.2.6 に示す。

表 6.2.6: 差なしの例

入力文 6	その会社は輸出入業を行なっている。
参照文 6	The company deals in importing and exporting .
ベースライン 6	The company administers a profitability trade .
提案手法 6	The company has an open business .
入力文 7	その理由を上げなさい。
参照文 7	Cive the reasons for it .
ベースライン 7	Be sure of the reason .
提案手法 7	Put up the whys .
入力文 8	今年の夏はたぶん暑い日が続くだろう。
参照文 8	We expect hot days this summer .
ベースライン 8	It is likely to stay hot this summer .
提案手法 8	We will have hot weather this summer .

入力文 6 において、両手法で出力文が「輸出入業」に対する出力を得られていないため、差なしと評価した。

入力文 7 と入力文 8 では、両手法で出力文が入力文全体の意味を捉えているため、差なしと評価した。

第7章 考察

7.1 対訳句の精度と翻訳精度

6章の実験結果から，対訳句の学習で翻訳精度が向上すると推察される．そこで，対訳句の精度が低下した場合の考察する．

本研究で用いた対訳句の精度は，95%である．対訳句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7.1.1に自動評価の結果を示す．

表 7.1.1: 精度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
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7.1.1より対訳句の精度を下げた場合，75%まで下げたとき翻訳精度がベースラインより低下している．例を表7.1.2に示す．

表 7.1.2: 精度低下時の翻訳結果

入力文:あの チーム には とても 歯 が 立た ない。

参照文:We're no match for that team . .

90 %:That team is very powerful .

75 %: That team has no toothache .

50 %:That team doesn't have much teeth .

表7.1.2より，対訳句の精度が下がるにつれて翻訳精度が低下している．90%は、「歯が立たない」を「相手チームがとても強い」と言い換えているため，入力文の意味を概

ね捉えた翻訳となっている。一方で、75%と50%ではそれぞれ「あのチームには歯痛がない」、「あのチームにはあまり歯がない」と翻訳している。この結果から、75%と50%は文全体の意味を考慮した翻訳ができないと考えられる。

7.2 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7.2.1に示す。

表 7.2.1: 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
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7.2.1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7.1.1と表7.2.1を比較すると、ほぼ同じ傾向があることがわかる。

7.3 対訳句追加による効果

本実験の結果として、対訳句の追加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。このことから提案手法における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。

7.4 要素合成法の問題の緩和

本実験では、要素合成法の問題の観点からNMTの翻訳精度の向上を試みた。私は、対訳句の追加で要素合成法の問題が緩和できたと考えている。両評価方法において提案手法がベースラインより良好と評価されたからである。しかし、1つの対訳句が複数の出力結果に影響する問題がある。例を示す。

(a) 「私は我を通した。」を翻訳する．翻訳精度向上のために対訳句の追加学習を行う．

- 日本語句:我を通した。
- 英語句:had my own way.

結果として，正しい出力文が得られる．

- 英語文 1:I had my own way.

(b) しかし、「彼は私を通してくれた。」(参照文:He passd me.)を翻訳したとき，学習した対訳句の影響で”He had his own way.”と出力される場合がある．

これは，提案手法において避けることができない問題である．しかし，この問題に対する明確な解法はない．そこで，私が考えた解法を示す．

私が考えた解法は，NMTの学習に電子辞書などの外部から得た対訳句を追加する手法である．本実験はクローズドデータで行ったため，外部の情報を有していない．外部情報の追加で多種多様な入力文に対応する可能性があるかと推察される．

第8章 関連研究

本章では関連研究について挙げる．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．PBSMT の関連研究は多く存在した．しかし，日英 NMT において対訳句を追加する手法は今仁の研究だけである．

8.1 PBSMT に対訳句を追加

日野 [4] 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．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．

8.2 対訳学習文から作成した対訳句の追加

本研究と似た手法を用いた研究として、池淵 [5] の研究を挙げる．池淵らは，日英 PBSMT において対訳学習文から作成した対訳句を学習データに追加し，翻訳精度の向上が得られた．池淵らは，対訳句の追加で語句の対応が強化されたためであると考察している．

8.3 複文での調査

本研究の関連研究として，今仁 [6] 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．

第9章 おわりに

NMTの翻訳精度は人手の翻訳精度に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、対訳句を追加で語句の対応情報が強化されたと考えられる。本実験はクローズドデータで行ったが、オープンデータで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] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. CoRR, abs/1409.0473, 2014.
- [2] 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.
- [3] 森本世人, 村上仁一. 類似度を利用した変換テーブルの精度向上. 2021 年度鳥取大学卒業論文, 2021.
- [4] 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.
- [5] 池淵堅斗. 対訳分から抽出した対訳句を用いた統計翻訳. 2015 年度修士論文, 2015.
- [6] 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. 言語処理学会第 26 回年次大会, pp.169-172, 2020.