

# ニューラル日英翻訳における適切な出力文数の推定

伊藤均 衣川和堯 美野秀弥 後藤功雄 山田一郎

NHK

{itou.h-ce, kinugawa.k-jg, mino.h-gq, goto.i-es, yamada.i-hy} @ nhk.or.jp

## 1 はじめに

NHK で放送される英語ニュース用の原稿は、日本語ニュース用の原稿を基に人手で翻訳・編集し制作している。我々はこの英語ニュース用の原稿制作を支援するためのツールとして日英機械翻訳の研究を進めている。

NHK のニュース原稿は、日本語ニュース用と英語ニュース用で異なる特徴を有する。日本語ニュース用の原稿は一文が長いことが特徴であるのに対し、英語ニュース用の原稿は一文を短く区切る特徴がある。

これまでに我々は、この特徴の異なりを考慮した日英機械翻訳システム実現のため、人手で指定した文数に翻訳結果を制御する手法を提案した [1]。

本稿では、日英機械翻訳処理における日本語一文に対する適切な英語出力文数の推定手法と、出力文数推定から制御までの一連の処理について提案する。本手法は推定モデルと制御モデルの2つのモデルを用いる。日本語一文に対する翻訳結果が、推定モデルが適切であると予想した文数で出力されるよう制御モデルで出力を制御し、所望の翻訳結果を得る。実験結果から、提案手法は出力文数を推定・制御しない従来の日英ニューラル機械翻訳モデルの出力と比較して、人手による翻訳結果の文数との一致率が高いことを確認した。

## 2 出力文数の推定・制御タスク

本稿は、機械翻訳における出力文数を以下の2つの処理によって推定、制御する。

- (a) 日本語各文をいくつかの英文で翻訳するべきかを推定する。
- (b) 指定した文数で翻訳するよう、機械翻訳の出力を制御する。

我々はこれまで、(b)の課題に対する手法を提案して

いる [1]。本稿では (a) の課題に取り組み、(b) の課題も含む枠組み全体についても評価する。

関連研究として、目的言語における読者の語学力に併せて翻訳文の読解難易度を制御することを目的とした NMT の手法 [2] や、翻訳品質の向上を目的として NMT の訓練データ中の長文を分割する手法 [3] があるが、これらは訳文の文数を制御する手法ではない。

## 3 提案手法

提案手法は推定モデルと制御モデルの2つのモデルを用いて適切な文数へ翻訳結果を制御する。具体的には、以下の手順で処理を進める。

- (1) 日本語各文に対する適切な出力英文数を推定する推定モデルと、指定した文数で翻訳結果を出力する制御モデルを、ニューラル機械翻訳モデルを利用し学習する。
- (2) 推定モデルに翻訳したい日本語一文を入力し、適切な出力英文数を推定する。
- (3) 制御モデルに (2) の推定結果を入力し、指定した文数で日英翻訳する。

本研究では、更なる推定性能向上のため制御モデルへ文数を入力する前に推定モデルの出力を補正する前処理を追加することも想定しており、上記のような2段階の処理を導入している。

本節ではこのために必要な出力文数の推定手法と、この推定結果を利用した出力文数制御手法について説明する。

### 3.1 適切な出力文数の推定

日本語各文に対する適切な出力英文数を推定するため、日本語一文を入力するとその日本語を日英翻訳する際に適切な英文数を表すタグを出力する推定モデルを構築する。

日本語	全国一の生産量を誇る広島特産のカキを夏にも観光客らに楽しんでもらおうと、広島県は新たに生食用のカキを開発し、16日、関係者を集めて試食会を開きました。
英語	<3> Hiroshima is the number one producer of oysters in the country. The prefecture has newly developed oysters for eating raw so that tourists can enjoy oysters, a specialty of Hiroshima, also in summer. Officials held a tasting event on July 16, inviting people concerned.

表 1. 推定モデルの学習データへのタグの付与例

まず、日本語一文に対する翻訳英文のペアを用意する。これは、日本語一文に対応する英文は必ずしも一文ずつではなく、一対多のデータである。次に、この一対多のデータの各出力の最初の英文の文頭に、その対訳英文が何文であることを表すタグ(文数タグ)を付与し、日本語一文とこの文数タグ付き対訳英文とのペアを推定モデルの学習データとしてニューラル機械翻訳モデルを学習する。

表 1. に推定モデルの学習に用いる学習データの例を示す。こちらの例では、日本語一文に対して英語 3 文の対訳データであるため、ターゲット言語文の冒頭に” <3>” と、英文数の情報を示す文数タグを付与している。文数タグは、出力文数が一文であれば” <1>”、2 文であれば” <2>” と、翻訳英文中の数値と区別するために” <>” 内の数字で文数情報を与える。

この学習データから構築したモデルのシステムに任意の日本語一文を入力すると、翻訳結果の先頭に文数タグが出力される。このタグの示す文数を推定モデルにおける適切な英文数の推定結果とし、後述する制御モデルの入力に用いる。

### 3.2 出力文数の制御

我々はこれまでに、日本語一文と共に適切な出力英文数を同時に指定することで、指定した文数で日英翻訳結果を出力する制御モデルを提案した[1]。この制御モデルでは手動での出力文数指定も可能であるが、上述の推定モデルの出力によって得られた適切な出力文数を引用することで自動的に出力文数を制御することが可能となる。

制御モデルでは、日英対訳データの学習データに

日本語	<3> 財務のプロとして知られ、アメリカの大手会計事務所や中国のオンラインゲーム会社を経てアリババの経営陣に加わり、4年前からCEOを務めています。
英語	He is known to be an expert in financial affairs. After working at a major accounting firm in the United States and a Chinese online gaming company, he joined Alibaba's management. He has served as the company's CEO since 4 years ago.

表 2. 制御モデルの学習データへのタグの付与例

において、文数タグを入力である日本語各文の文頭に配置する(表 2)。学習データ作成時にターゲット言語の出力文数を数え、ソース言語の冒頭に文数タグとして付与する。

この学習データをニューラル機械翻訳モデルに与えて学習することにより、翻訳文数の制御モデルを構築する。

この制御モデルを用いて翻訳結果を出力する際には、翻訳したい日本語各文の冒頭に、推定モデルによって得られた文数タグを付与し、制御モデルへ入力することで最終的な翻訳結果を取得する。

## 4 実験

### 4.1 実験条件

NHK の日本語ニュースを人手により英訳し、約 10 万文対の実験データを構築した。このうち学習データ、開発データにはそれぞれ約 96,000 文対、5,000 文対を用いた。評価データには日本語文 772 文を用意し、評価データのみ翻訳者 2 名(翻訳者 A, 翻訳者 B とする)がそれぞれ日英翻訳したマルチリファレンスデータとした。

学習モデルは NMT ツールキット ” Sockeye [4]” を使用した。パラメータについては表 3 の項目について表記の値に変更し、その他のパラメータについてはデフォルト値を使用した。出力ユニット単位の決定には Byte Pair Encoding [5] でサブワードに分割した。

推定モデル・制御モデルの学習にはそれぞれ学習データの出力英文数を文数タグとして用い、出力英文数はサブワード” .” の数とした。

項目	パラメータ
max-seq-len	300
weight-tying-type	trg_softmax
label-smoothing	0.3
embed-dropout	0.5
transformer-dropout-attention	0.5
transformer-dropout-act	0.2
transformer-dropout-prepost	0.3

表 3. Sockeye で用いたパラメータ

	一致率[%]
Baseline	81.8
Assume_Tag	87.0
Tag(refA)	97.8
Tag(refB)	99.1
推定モデル	87.0

表 4. 出力文数の一致率の比較

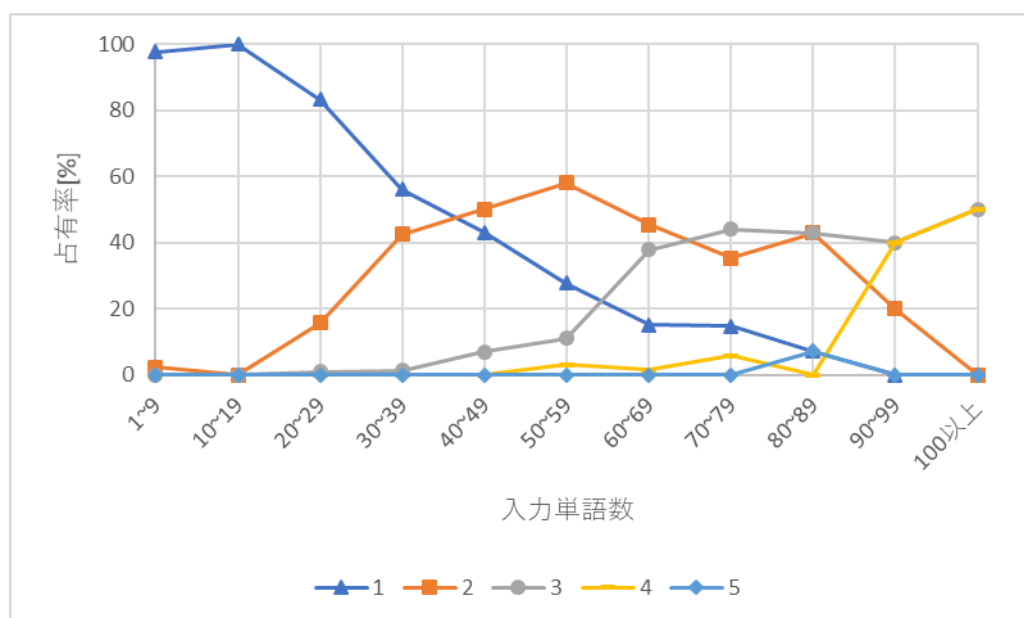


図 1. 入力単語数ごとの人手翻訳英文数の占有率

## 4.2 翻訳手法

4.1 節の条件下で推定モデルを学習した上で、以下の翻訳手法から得られる翻訳結果を比較した。

- (a) **Baseline**: 上記のデータを用いて学習データをそのまま使用し学習、翻訳したもの
- (b) **Tag**: 制御モデルにおいて評価データの出力英文数を入力の文数タグとして与えて翻訳したもの。それぞれ制作者 A の出力文数を用いたものを **Tag(refA)**、制作者 B の出力文数を用いたものを **Tag(refB)** とする。

- (c) **Assume\_Tag**: 推定モデルによって推定した出力文数を制御モデルの入力の文数タグとして与えて翻訳したもの。

## 4.3 実験結果

### 4.3.1 出力文数の一致率の比較

まず、評価データの日本語各文に対する出力英文数と評価データの英文数との一致率を比較した (表 4)。ここで一致率とは、翻訳者 A, B いずれかの出力英文数と一致していれば一致とみなして算出したも

	BLEU
Baseline	31.23
Assume_Tag	30.26
Tag(refA)	32.07
Tag(refB)	31.98

表 5. BLEU 値の比較

のを指す。文数を制御せずに翻訳した場合 (Baseline) と比較して、推定モデルを用いた手法の一致率が 81.8% から 87.0% に向上し、学習データに出力文数を明示的に与えることの効果を観測することが出来た。また、推定モデルと制御モデルの両方を用いた手法 (Assume\_Tag) でも 87.0% の一致率を得た。

制御モデルだけに着目すると、評価データの出力文数を文数タグとして与えた場合も推定モデルの推定結果を文数タグとして与えた場合も、文献 [1] で示したように制御モデルによる一致率の低下なく文数を制御することが出来た。Tag(refA) の一致率は 97.8%、Tag(refB) は 99.1% と、出力文数を指定すれば、ほぼその通りの文数で翻訳できることが分かる。

出力文数の推定タスクの難しさを調べるために、入力文長と出力文数との関係を調査した。図 1 に翻訳者 A の翻訳結果における、入力文の単語数ごとの出力文数の占有率のグラフを示す。こちらのグラフは各入力文を入力単語数で 10 単語単位に分類した際の人手翻訳英文数の占有率を示している。グラフから、日本語一文に対する入力単語数と人手翻訳英文数には一定の相関があるが、重なりのある部分も多く、難しいことが分かる。

また、評価データにおける翻訳者 2 者の平均英文数は翻訳者 A が 1.73 文、翻訳者 B が 1.47 文で、両者の文数一致率は 65.8% と両者の一致度は高くはなかった。推定プロセスを経ることで文数一致率が向上していることから入力文の内容と翻訳英文数との間には何らかの相関があると推測できるが、更なる推定性能向上のため翻訳結果を複数文に分割する要因の分析については今後の課題とする。

#### 4.3.2 翻訳性能の比較

次に、各手法の BLEU 値について比較した (表 5)。BLEU 値は翻訳者 A・B 両者の翻訳結果をマルチリフ

ァレンスとして測定した。推定モデルの結果を反映させた手法 (Assume\_Tag) では Baseline と比較して BLEU 値の低下が観測された。人手翻訳時の出力文数を与えて翻訳した手法 (Tag\_refA, Tag\_refB) では Baseline と比較して BLEU 値が向上していることから、出力文数の一致率と BLEU スコアには相関があることが分かる。

一方で、出力文数を制御出来たとしても BLEU スコアの改善には繋がらなかった。これは出力文数が一致することとリファレンスとの翻訳表現が一致することは同義ではないためと考察する。文数推定モデルにより文数一致率が向上したことから、人手翻訳において適切な出力文数は一定の基準に従って定まることは推測出来るが、翻訳表現については翻訳者ごとに個性があり、出力文数が一致したとしても出力結果は必ずしも一致しない。更なる原稿制作支援のため、今後入力文の情報が翻訳時に各文にどのような分割されるか分析を進めたい。

## 5 おわりに

本稿では、NHK のニュースを対象とした日英翻訳において、日本語一文に対する翻訳として適切な英文数を推定し、推定結果に基づいた文数で翻訳するためのニューラル機械翻訳モデルを提案した。提案手法は適切な文数を推定する推定モデルと、指定した文数で翻訳する制御モデルの 2 つのモデルで構成され、推定・制御を経由しない従来の機械翻訳モデルの出力と比較して人手翻訳結果の出力文数との一致率を 6.4% 向上させることが出来た。今後は更なる翻訳性能と制作支援効果向上に向けた手法を検討する。

## 参考文献

1. ニューラル機械翻訳における出力文数の制御. 伊藤ほか: 人工知能学会, 2020
2. Controlling Text Complexity in Neural Machine Translation. S. Agrawal et al. : EMNLP, 2019
3. Automatic Long Sentence Segmentation for Neural Machine Translation. S.Kuang et al. : ICCPOL, 2016
4. Sockeye: A Toolkit for Neural Machine Translation. F. Hieber et al. : arxiv, 2017
5. Neural Machine Translation of Rare Words with Subword Units. R. Sennrich et al. : ACL, 2016.