

## 概要

近年、機械翻訳においてニューラル機械翻訳 (Neural Machine Translation; NMT)[1] が主流となっている。NMTは Attention という仕組みに基づき入力文と出力文における語句の対応情報を学習する手法であり、従来の句に基づく統計翻訳 (Phrase-based SMT; PBSMT) などの手法と比較して流暢性の高い翻訳文を出力することが可能である。一方で、入力文中の語句に対する出力文中の語句の対応 (Attention) の学習に誤りがある場合には、翻訳精度が低下するという問題がある。

従来の PBSMT において、語句の対応 (アライメント) を学習する目的で用いられる翻訳モデルに関しても同様の問題が指摘されており、アライメントの精度を向上させる手法が提案されている。Popovic ら [2] は、セルビア語英語間およびスペイン語英語間の PBSMT において、翻訳モデルの学習データとして、対訳文に対訳文外の対訳句データを追加し、翻訳精度の向上を実現している。日野ら [3] も同様の方法を用いて日本語英語間における PBSMT の翻訳精度を向上させている。また、池淵ら [4] は外部の対訳句データを利用せず、対訳文より作成した対訳句データを日英 PBSMT の翻訳モデルの学習データに追加し、翻訳精度の向上が得られたとしている。池淵らはこの理由として、対訳句を追加することで正しい句の対応が強調されたためであると考察している。

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 の学習データにおいて、対訳文に対訳文より作成した対訳句を追加することで、句レベルの対応情報を強調する。これにより、語句の対応情報をモデル化する Attention を強化し、日英 NMT の翻訳精度の向上を目指す。

また、本研究の提案手法において利用する対訳句には、対訳文から作成した対訳句のみを用い、外部の情報を有する対訳句データは利用しない。これは、対訳文外の対訳句コーパスを用いる場合、句レベルの対応情報が強調されることに加えて、それ以外の情報 (語彙情報など) が拡大することにより翻訳精度に影響を与える可能性があるためである。したがって、本研究では、対訳文から作成した対訳句のみを追加する手法を用いて日英 NMT の翻訳実験を行い、語彙などの情報が拡大しない条件で、句レベルの対応情報を強調することのみにより翻訳精度が向上することを検証する。

結果として、人手で作成した対訳句を用いた手法の評価では、ベースラインの方が正確性が高いと評価された文が出力文 100 文中 14 文であるのに対して、人手で作成した対訳句を追加する手法の方が正確性が高いと評価された文が 44 文となった。また、自動で作成した対訳句を用いた手法では、ベースラインの方が正確であると評価された文が出

力文 100 文中 10 文であるのに対して，自動で作成した対訳句を追加する手法の方が正確であると評価された文が 29 文となった．これより，ベースラインと比較して対訳句を追加する 2 つの手法では翻訳精度が向上することが確認された．本研究で用いた 2 種類の対訳句はいずれも対訳文から作成され，対訳文外の語彙などの情報を有しないものである．したがって，対訳句を追加する提案手法により，NMT において語句の対応をモデル化する Attention が強化され，翻訳精度が向上したと考えられる．

# 目次

|          |                          |           |
|----------|--------------------------|-----------|
| <b>1</b> | <b>はじめに</b>              | <b>1</b>  |
| <b>2</b> | <b>ニューラル機械翻訳</b>         | <b>3</b>  |
| 2.1      | ニューラル機械翻訳の概要             | 3         |
| 2.2      | Encoder-Decoder モデル      | 4         |
| 2.2.1    | Encoder-decoder モデルの概要   | 4         |
| 2.2.2    | Encoder-decoder モデルの枠組み  | 4         |
| 2.3      | Attention モデル            | 6         |
| 2.3.1    | Attention モデルの概要         | 6         |
| 2.3.2    | Attention モデルの枠組み        | 7         |
| <b>3</b> | <b>関連研究</b>              | <b>9</b>  |
| 3.1      | ニューラル機械翻訳において対訳句を扱う研究    | 9         |
| 3.2      | 関連研究と本研究との相違点            | 10        |
| <b>4</b> | <b>対訳句を利用したニューラル機械翻訳</b> | <b>12</b> |
| 4.1      | 対訳句作成システム                | 14        |
| 4.1.1    | 人手対訳句の作成                 | 14        |
| 4.1.2    | 自動対訳句の作成                 | 15        |
| 4.2      | 提案手法の設定                  | 18        |
| <b>5</b> | <b>実験環境</b>              | <b>19</b> |
| 5.1      | 実験データ                    | 19        |
| 5.2      | ニューラル機械翻訳の実験設定           | 20        |
| 5.3      | 評価実験                     | 20        |
| <b>6</b> | <b>対訳句の抽出</b>            | <b>21</b> |
| 6.1      | 人手対訳句                    | 21        |
| 6.2      | 自動対訳句                    | 22        |
| <b>7</b> | <b>翻訳実験</b>              | <b>24</b> |
| 7.1      | 自動評価結果                   | 24        |
| 7.2      | 人手対比較評価結果                | 24        |

|           |  |           |
|-----------|--|-----------|
| 7.3       | 人手対比較評価例 . . . . .                     | 25        |
| 7.3.1     | ベースラインと+人手対訳句の人手対比較評価例 . . . . .       | 25        |
| 7.3.2     | ベースラインと+自動対訳句の人手対比較評価例 . . . . .       | 27        |
| <b>8</b>  | <b>考察</b>                              | <b>29</b> |
| 8.1       | 対訳句利用の効果 . . . . .                     | 29        |
| 8.2       | 翻訳結果と対訳句の分析 . . . . .                  | 29        |
| 8.2.1     | 翻訳精度が向上した結果 . . . . .                  | 29        |
| 8.2.2     | 翻訳精度が低下した結果 . . . . .                  | 30        |
| 8.3       | 人手対訳句と自動対訳句の比較 . . . . .               | 30        |
| <b>9</b>  | <b>追加実験：外部辞書データより抽出した対訳句</b>           | <b>31</b> |
| 9.1       | 外部辞書データより抽出した対訳句の利用 . . . . .          | 31        |
| 9.2       | 実験設定 . . . . .                         | 32        |
| 9.3       | 英辞郎の精度 . . . . .                       | 32        |
| 9.4       | NMT 学習時の語彙数の制限 . . . . .               | 33        |
| 9.5       | 自動評価結果 . . . . .                       | 34        |
| 9.6       | 人手対比較評価結果 . . . . .                    | 35        |
| 9.7       | 人手対比較評価例 . . . . .                     | 35        |
| 9.8       | 追加した外部辞書と翻訳結果の分析 . . . . .             | 37        |
| <b>10</b> | <b>追加実験：1変数対訳文パターンを用いた対訳句</b>          | <b>38</b> |
| 10.1      | 1変数対訳文パターンを用いた対訳句の抽出手法 . . . . .       | 38        |
| 10.2      | 実験設定 . . . . .                         | 39        |
| 10.3      | 多変数対訳句および1変数対訳句の作成結果 . . . . .         | 40        |
| 10.4      | 多変数対訳句および1変数対訳句の精度 . . . . .           | 40        |
| 10.5      | 自動評価結果 . . . . .                       | 42        |
| 10.6      | 人手対比較評価結果 . . . . .                    | 42        |
| 10.7      | 人手対比較評価例 . . . . .                     | 43        |
| 10.8      | 1変数対訳句を用いた繰り返し処理 . . . . .             | 45        |
| 10.8.1    | 1変数対訳文パターンと繰り返し処理による対訳句の作成結果 . . . . . | 45        |
| 10.8.2    | 1変数対訳文パターンと繰り返し処理による対訳句の精度 . . . . .   | 46        |
| 10.8.3    | 自動評価結果 . . . . .                       | 47        |

|           |                                  |           |
|-----------|----------------------------------|-----------|
| 10.8.4    | 人手対比較評価結果 . . . . .              | 47        |
| 10.8.5    | 人手対比較評価例 . . . . .               | 48        |
| 10.9      | 1 変数対訳文パターンを用いた対訳句のまとめ . . . . . | 50        |
| <b>11</b> | <b>おわりに</b>                      | <b>51</b> |

## 目 次

|   |                                   |    |
|---|-----------------------------------|----|
| 1 | Encoder-Decoder モデルの枠組み . . . . . | 5  |
| 2 | Attention モデルの枠組み . . . . .       | 8  |
| 3 | 対訳句を利用した日英ニューラル機械翻訳の流れ . . . . .  | 13 |
| 4 | 対訳単語作成の例 . . . . .                | 16 |
| 5 | 対訳文パターン作成の例 . . . . .             | 17 |
| 6 | 対訳句作成の例 . . . . .                 | 18 |

## 目 次

|     |  |    |
|-----|--|----|
| 4.1 | 対訳データの例                                  | 12 |
| 4.2 | 対訳文の例                                    | 14 |
| 4.3 | 対訳文パターンの例                                | 14 |
| 4.4 | 対訳文と対訳文パターンの比較により作成された対訳句の例              | 15 |
| 5.1 | 実験データの内訳                                 | 19 |
| 5.2 | 対訳文の例                                    | 19 |
| 6.1 | 人手対訳句の内訳                                 | 21 |
| 6.2 | 人手対訳句の例                                  | 21 |
| 6.3 | 人手対訳句の評価結果                               | 22 |
| 6.4 | 人手対訳句の評価例                                | 22 |
| 6.5 | 自動対訳句の内訳                                 | 23 |
| 6.6 | 自動対訳句の評価結果                               | 23 |
| 6.7 | 自動対訳句の評価例                                | 23 |
| 7.1 | ベースライン, +人手対訳句, +自動対訳句の自動評価結果 (10,000 文) | 24 |
| 7.2 | ベースラインと+人手対訳句の人手対比較評価結果 (100 文)          | 25 |
| 7.3 | ベースラインと+自動対訳句の人手対比較評価結果 (100 文)          | 25 |
| 7.4 | +人手対訳句○                                  | 26 |
| 7.5 | +ベースライン○                                 | 26 |
| 7.6 | 差なし                                      | 27 |
| 7.7 | +自動対訳句○                                  | 27 |
| 7.8 | +ベースライン○                                 | 28 |
| 7.9 | +差なし                                     | 28 |
| 9.1 | 英辞郎の内訳                                   | 32 |
| 9.2 | 英辞郎の例                                    | 32 |
| 9.3 | 英辞郎の評価結果                                 | 32 |
| 9.4 | 英辞郎の評価例                                  | 33 |
| 9.5 | 語彙数を制限した学習データの例                          | 34 |
| 9.6 | ベースライン, +英辞郎の自動評価結果 (10,000 文)           | 34 |
| 9.7 | ベースラインと+英辞郎の人手対比較評価結果 (100 文)            | 35 |
| 9.8 | +英辞郎○                                    | 35 |

|       |   |    |
|-------|---|----|
| 9.9   | ベースライン○                                     | 36 |
| 9.10  | 差なし   | 36 |
| 9.11  | 「信用度」を含む対訳学習文中の対                            | 37 |
| 9.12  | 「信用度」を含む英辞郎中の対訳句                            | 37 |
| 10.1  | 実験データの内訳                                    | 39 |
| 10.2  | 多変数対訳句の内訳                                   | 40 |
| 10.3  | 1変数対訳句の内訳                                   | 40 |
| 10.4  | 多変数対訳句の評価結果                                 | 40 |
| 10.5  | 1変数対訳句の評価結果                                 | 40 |
| 10.6  | 多変数対訳句の評価例                                  | 41 |
| 10.7  | 1変数対訳句の評価例                                  | 41 |
| 10.8  | ベースライン, +多変数対訳句, +1変数対訳句の自動評価結果 (10,000文)   | 42 |
| 10.9  | ベースラインと+多変数対訳句の人手対比較評価結果 (100文)             | 42 |
| 10.10 | ベースラインと+1変数対訳句の人手対比較評価結果 (100文)             | 42 |
| 10.11 | +1変数対訳句○                                    | 43 |
| 10.12 | ベースライン○                                     | 43 |
| 10.13 | 差なし   | 44 |
| 10.14 | 1変数対訳文パターンと繰り返し処理により作成した対訳句の内訳              | 45 |
| 10.15 | ループ1対訳句の評価結果                                | 46 |
| 10.16 | ループ2対訳句の評価結果                                | 46 |
| 10.17 | ループ1対訳句の評価例                                 | 46 |
| 10.18 | ループ2対訳句の評価例                                 | 47 |
| 10.19 | ベースライン, +ループ1対訳句, +ループ2対訳句の自動評価結果 (10,000文) | 47 |
| 10.20 | ベースラインと+ループ1対訳句の人手対比較評価結果 (100文)            | 48 |
| 10.21 | ベースラインと+ループ2対訳句の人手対比較評価結果 (100文)            | 48 |
| 10.22 | +ループ1対訳句○                                   | 48 |
| 10.23 | ベースライン○                                     | 48 |
| 10.24 | 差なし   | 49 |
| 10.25 | +ループ2対訳句○                                   | 49 |
| 10.26 | ベースライン○                                     | 49 |
| 10.27 | 差なし   | 49 |

# 1 はじめに

近年、機械翻訳においてニューラル機械翻訳 (Neural Machine Translation; NMT)[1] が主流となっている。NMT は Attention という仕組みに基づき入力文と出力文における語句の対応情報を学習する手法であり、従来の句に基づく統計翻訳 (Phrase-based SMT; PBSMT) などの手法と比較して流暢性の高い翻訳文を出力することが可能である。一方で、入力文中の語句に対する出力文中の語句の対応 (Attention) の学習に誤りがある場合には、翻訳精度が低下するという問題がある。

従来の PBSMT において、語句の対応 (アライメント) を学習する目的で用いられる翻訳モデルに関しても同様の問題が指摘されており、アライメントの精度を向上させる手法が提案されている。Popovic ら [2] は、セルビア語英語間およびスペイン語英語間の PBSMT において、翻訳モデルの学習データとして、対訳文に対訳文外の対訳句データを追加し、翻訳精度の向上を実現している。日野ら [3] も同様の方法を用いて日本語英語間における PBSMT の翻訳精度を向上させている。また、池淵ら [4] は外部の対訳句データを利用せず、対訳文より作成した対訳句データを日英 PBSMT の翻訳モデルの学習データに追加し、翻訳精度の向上が得られたとしている。池淵らはこの理由として、対訳句を追加することで正しい句の対応が強調されたためであると考察している。

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 の学習データにおいて、対訳文に対訳文より作成した対訳句を追加することで、句レベルの対応情報を強調する。これにより、語句の対応情報をモデル化する Attention を強化し、日英 NMT の翻訳精度の向上を目指す。

また、本研究の提案手法において利用する対訳句には、対訳文から作成した対訳句のみを用い、外部の情報を有する対訳句データは利用しない。これは、対訳文外の対訳句コーパスを用いる場合、句レベルの対応情報が強調されることに加えて、それ以外の情報 (語彙情報など) が拡大することにより翻訳精度に影響を与える可能性があるためである。したがって、本研究では、対訳文から作成した対訳句のみを追加する手法を用いて日英 NMT の翻訳実験を行い、語彙などの情報が拡大しない条件で、句レベルの対応情報を強調することのみにより翻訳精度が向上することを検証する。

さらに、追加する対訳句の作成手法として、人手で作成する手法と自動で作成する手法の 2 通りの手法を提案し、それぞれについて対訳句の精度と、対訳句を日英 NMT に利用する際の翻訳精度を調査する。

本論文の構成は以下の通りである。まず、2 章でニューラル機械翻訳システムの概要を

示し，各モデルの構成について述べる．3章では，本研究に関連する研究について述べる．4章では，本研究の提案手法について述べる．5章では，実験に用いるデータやツールといった実験環境について述べる．6章では，対訳句の抽出結果を示す．7章では，提案手法の結果を示す．8章では，考察を行う．9章および10章では，それまでの議論を踏まえて，追加実験を行う．最後に，11章で結論を述べ，まとめる．

## 2 ニューラル機械翻訳

### 2.1 ニューラル機械翻訳の概要

ニューラル機械翻訳 (Neural Machine Translation; NMT) とは、近年提案された機械翻訳の手法である。多くの細かい構成要素によって成立する従来の統計翻訳のシステムとは異なり、NMT は巨大なニューラルネットワークを一つ (もしくは少数) 用いてシステムを構築する。NMT の手法には、Encoder-Decoder モデルおよびそれを拡張した Attention モデルが提案されている。Encoder-Decoder モデルは入力の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の系列を復号化 (decode) するモデルである。機械翻訳のタスクにおいて、Encoder-Decoder モデルの入出力系列の要素は単語のベクトル表現となる。Attention モデルは Encoder-Decoder モデルにおいて出力系列を生成する際に、出力系列と入力系列の参照を行う機構 (Attention) を用いた手法であり、Encoder-Decoder モデルと比較して、長い入力に対するより妥当な出力を得られるとされている。

## 2.2 Encoder-Decoder モデル

### 2.2.1 Encoder-decoder モデルの概要

Encoder-Decoder モデル [5] 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 は入力系列を一つの固定長のベクトルに符号化 (encode) し、もう一方の RNN により固定長のベクトル符号を出力の系列へと復号化 (decode) する。Encoder-Decoder モデルでは、同時に2つの RNN の訓練を行い、入力の原言語文に対して出力の目的言語文を得る条件付き確率を最大化する。

### 2.2.2 Encoder-decoder モデルの枠組み

図1に Encoder-Decoder モデルの枠組みを示す。Encoder-Decoder モデルでは、Encoder は入力文をベクトル系列  $\mathbf{x} = (x_1, \dots, x_{T_x})$  として読み込み、ベクトル  $c$  に符号化する。一般的な Encoder-Decoder モデルは、1式および2式で表される RNN を用いて構成される。

$$h_t = f(x_t, h_{t-1}) \quad (1)$$

$$c = q(\{h_1, \dots, h_{T_x}\}) \quad (2)$$

ここで  $h_t$  は各時刻  $t$  の隠れ層の状態であり、 $c$  は隠れ層を用いて生成されたベクトルである。 $f$  および  $q$  は活性化関数であり、基本的な Encoder-Decoder モデルにおいて、活性化関数  $f$  には LSTM が用いられ、 $q(\{h_1, \dots, h_{T_x}\}) = h_{T_x}$  として定義されている [6]。Decoder は文脈ベクトル  $c$  と既に生成された単語列  $\{y_1, \dots, y_{t-1}\}$  が与えられた際の次の単語  $y_t$  を予測するように訓練され、結合確率を3式に示す条件式に分解することで翻訳文  $\mathbf{y} = \{y_1, \dots, y_{T_y}\}$  を得る条件付き確率を定義している。

$$p(\mathbf{y}) = \prod_{t=1}^{T_y} p(y_t | \{y_1, \dots, y_{t-1}\}, c) \quad (3)$$

1式および2式で表される RNN を用いて、それぞれの条件付き確率は4式によりモデル化される。ここで、 $g$  は  $y_t$  の確率を生成する非線形が多層関数であり、 $s_t$  は RNN の隠れ層の状態となる。

$$p(y_t | \{y_1, \dots, y_{t-1}\}, c) = g(y_{t-1}, s_t, c) \quad (4)$$

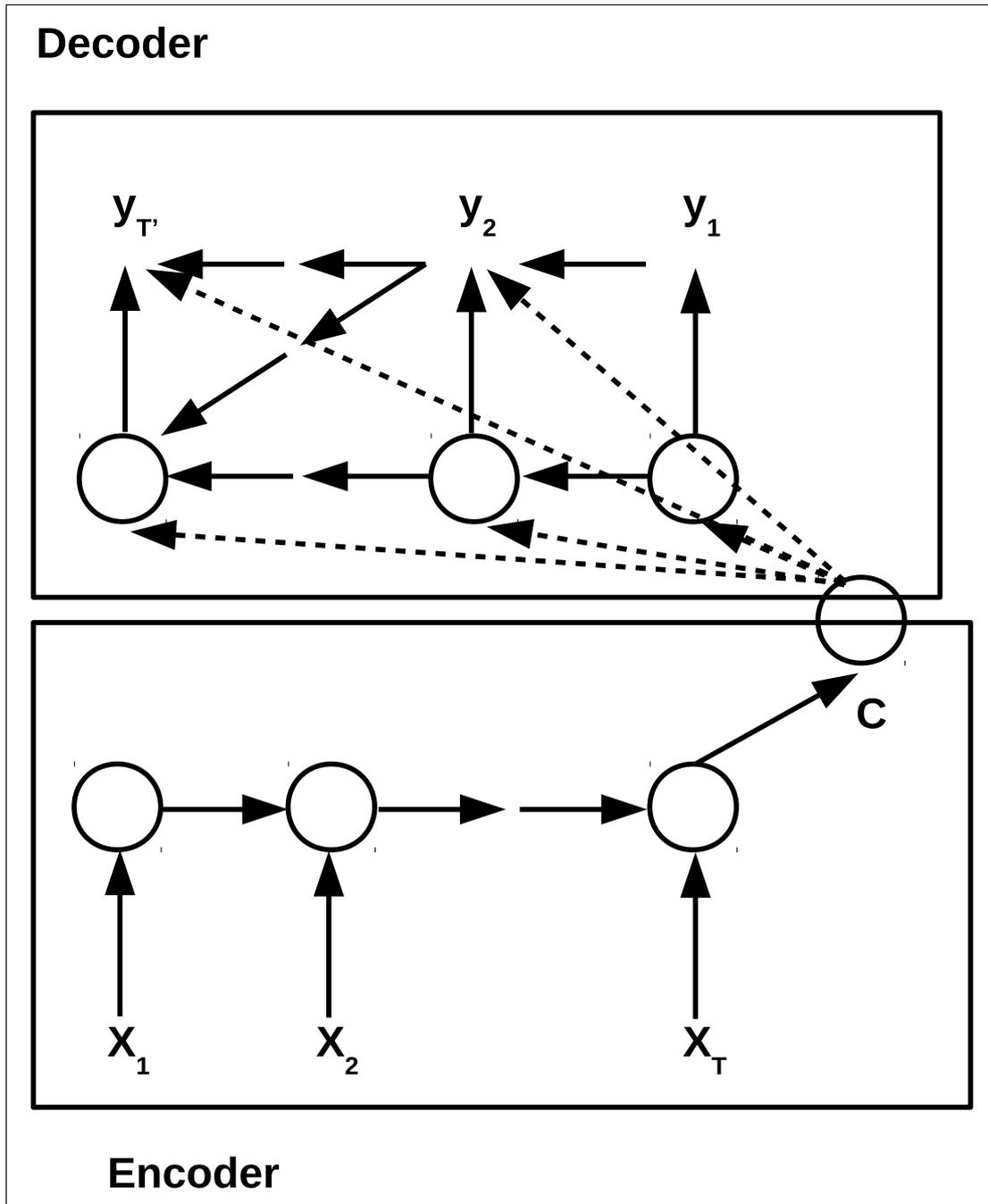


図 1: Encoder-Decoder モデルの枠組み

## 2.3 Attention モデル

### 2.3.1 Attention モデルの概要

Encoder-Decoder モデルの問題は、原言語文中の全ての情報を一つの固定長のベクトルに圧縮する点である。Encoder-Decoder モデルでは対訳学習文中で用いられている文よりも長い原言語文が入力された場合に、極端に精度が低下することが報告されている。これは、長い原言語文の全情報が一つの固定長のベクトルに圧縮されるために、目的言語文を生成する際に必要な情報が損失することが原因である。

Attention モデル [1] はこの問題を改善するために提案された NMT の手法である。Attention モデルの Encoder では入力単語を前後両方向から RNN に渡す手法 (bidirectional RNN) を用いている。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各単語を符号化 (encode) する。Decoder では、モデルが出力文中のある単語を生成する際に、その単語が最も相関する原言語文中の単語に相当する符号を探索する。その後、探索により得られた原言語単語の文中の位置情報を有した文脈ベクトル、および既に生成された全ての目的言語単語を参照し、次の目的言語単語を予測する。Attention モデルは Encoder-Decoder モデルと異なり、原言語文の情報を一つの固定長のベクトルに圧縮せずに目的言語文中の単語生成時に参照することで、より長い文における精度の向上を実現している。

### 2.3.2 Attention モデルの枠組み

図2に Attention モデルの枠組みを示す. Encoder-Decoder モデルにおいて3式で定義されていた条件付き確率に対して, Attention モデルでは時刻  $i$  の時の条件付き確率を5式により定義する. Encoder-Decoder モデルにおける条件付き確率  $p$  は, 既に出力された単語列と全ての入力系列  $\mathbf{x}$  を変換した文脈ベクトル  $c$  により定義されていたが, Attention モデルにおいては入力系列を一つの文脈ベクトルのみに圧縮することなく, 全ての入力系列  $\mathbf{x}$  より導かれる.

$$p(y_i|y_1, \dots, y_{i-1}, \mathbf{x}) = g(y_{i-1}, s_i, c_i) \quad (5)$$

また,  $s_i$  は時刻  $i$  での隠れ層の状態であり, 6式により計算される.

$$s_i = f(s_{i-1}, y_{i-1}, c_i) \quad (6)$$

ここで, 3式の Encoder-Decoder モデルと異なり, 各目的言語単語  $y_i$  は文脈ベクトル  $c_i$  により状態が与えられる. 文脈ベクトル  $c_i$  はアノテーション系列  $(h_i, \dots, h_{T_x})$  に依存し, Encoder により入力文と対応付けられる. 入力文の前後両方向より得られる各アノテーション  $h_i$  は全入力文の情報を保持しており, 特に入力文中の  $i$  番目の単語付近の情報を強く対応付けている. また, 7式に示す文脈ベクトル  $c_i$  はアノテーション  $h_i$  の重み付き和により計算される.

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j \quad (7)$$

各アノテーション  $h_j$  の重み  $\alpha_{ij}$  は

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})} \quad (8)$$

により計算される. ここで

$$e_{ij} = a(s_{i-1}, h_j) \quad (9)$$

は  $j$  番目付近の入力の対応および  $i$  番目の出力の適合を示す値であり, アライメントモデルと呼ばれる. この値は  $y_i$  を出力する直前の RNN の隠れ層の状態  $s_{i-1}$  および入力文中の  $j$  番目のアノテーション  $h_j$  に基づいている.

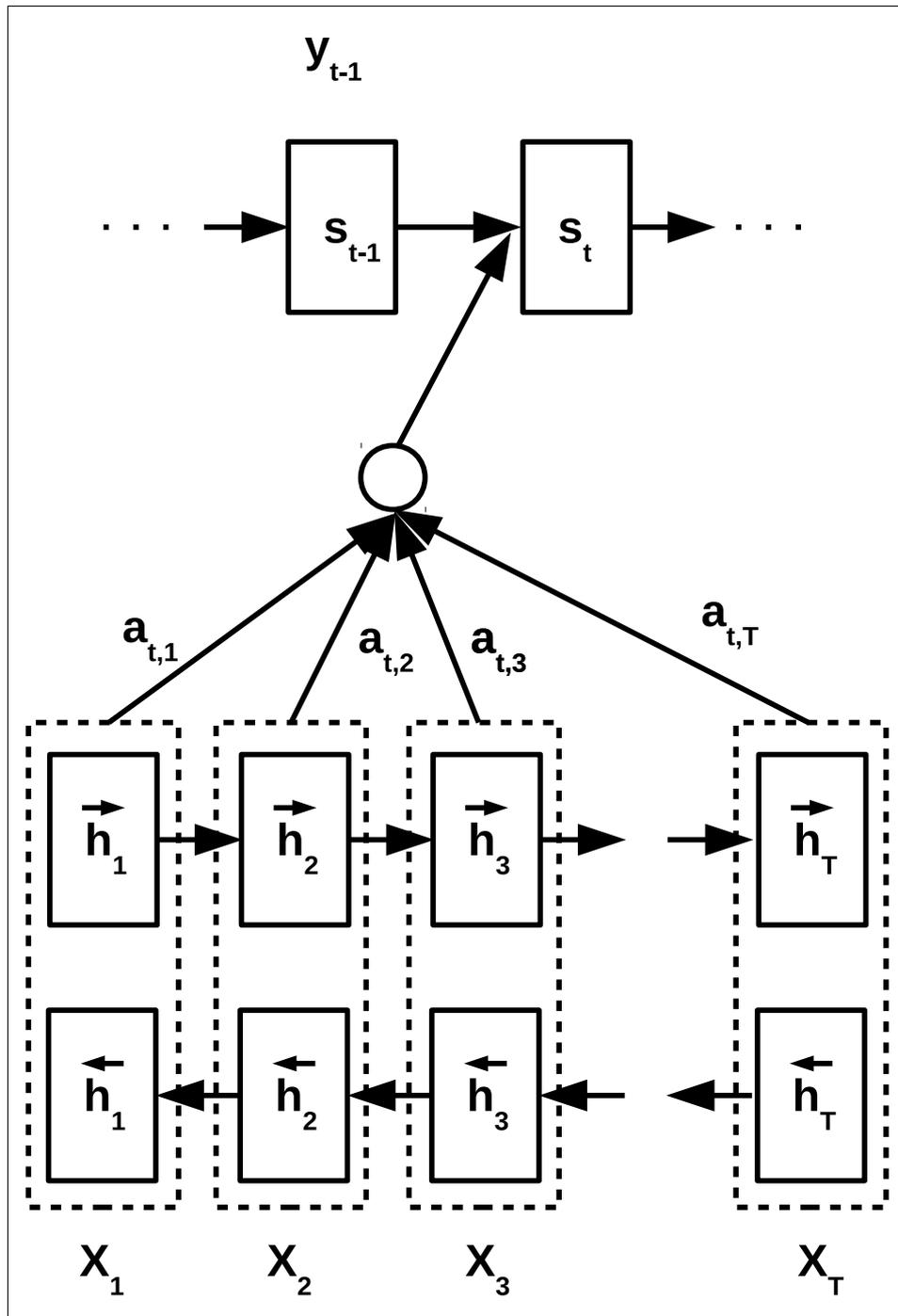


図 2: Attention モデルの枠組み

## 3 関連研究

従来の統計翻訳において、句に基づく手法は主流となっていた。しかし、Encoder-Decoder および Attention に基づくニューラル機械翻訳が提案されて以後、句レベルのアプローチをとる NMT 手法の提案は少なくなっている。本章ではニューラル機械翻訳において対訳句を扱う研究について記述し、それら関連研究と本研究との相違点について述べる。

### 3.1 ニューラル機械翻訳において対訳句を扱う研究

- Neural Machine Translation with External Phrase Memory[7]

外部句辞書を用いて入力に対応する出力フレーズを決定する phraseNet を提案している。中国語英語間の翻訳結果では、語彙数を頻度上位 16K 単語のみに制限する一般的な NMT モデルと比較して、語彙数を制限したモデルに、phraseNet を用いて出力文中の未知語句を置き換える提案手法では、BLEU 値が平均 3.45 ポイント上昇したと報告している。

- Sequence Modeling via Segmentations[8]

文を構成する系列を segment 化する仕組み「SWAN」を提案する。SWAN により、文をフレーズ単位に segment 化することが可能となる。

- Towards Neural Phrase-based Machine Translation[9]

目的言語側のネットワークにおいて SWAN を用いた明確な句レベルに基づく NMT 「Neural Phrase-based Machine Translation(NPMT)」を実装している。NPMT では SWAN の入力文を reordering する soft layer を用い、decoder では attention を用いることなく、SWAN による対訳句を直接出力する。英語ドイツ語間の翻訳結果では、Attention モデルの NMT と比較して BLEU 値が greedy 手法で 2.46、Beam Search 手法で 2.49 上昇し、英語ベトナム語間の翻訳では同様の比較で BLEU 値が greedy 手法で 1.41、Beam Search 手法で 1.59 上昇したと報告している。

- Neural Phrase-to-Phrase Machine Translation[10]

先行研究の NPMT において、目的言語側のみが明確な句ベースであったが、NP<sup>2</sup>MT では原言語側においても明確な句に基づく方法を用いている。

NP<sup>2</sup>MT において、低頻度語を 1 つの記号 (UNK) に置換する手法を用いた学習を行うことで、出力文中に入力文中の句を未知語としてそのまま生成することが可能である。出力された入力文中の未知語に対し、句辞書を用いて未知語を翻訳する方法により精度を向上させている。NP<sup>2</sup>MT における句辞書については、Moses [11] を用いて対訳学習文のみから作成した in-domain dictionary (ドイツ語英語:97,399 対, 英語ベトナム語:84817 対), および外部ドメインのコーパスから作成した辞書 (out-of-domain dictionary, ドイツ語英語:1,243,722 対) を利用している。また、同研究では、UNK とする低頻度語を下位 3 ~100 頻度で設定する実験を行い、低頻度語として妥当な順位についても考察している。

ドイツ語英語間および英語ベトナム語間において、Transformer モデル [12] の NMT をベースラインとした翻訳実験の結果として、in-domain dictionary を利用した条件の比較では、ベースラインと同程度の性能となった。また、ドイツ語英語間における out-of-domain dictionary を利用した実験においても、ベースラインと比較した時の BLEU 値の向上は小幅となっている。

- Phrase-Based Attentions [13]

Bahdanau ら [1], Luong ら [14] などの Attention 手法と比較して、高い性能を有すると報告される Transformer モデル [12] を 1-gram および 2-gram ベースに実装する手法を提案している。本章において記述している他の手法と異なり、アライメントベースの対訳句辞書は利用していない。Transformer モデルをベースラインとした比較実験の結果では、BLEU 値が最大で約 0.5 向上したと示している。

## 3.2 関連研究と本研究との相違点

先行研究の手法と、本研究の手法との相違として、まず、いずれの先行研究においても日本語英語間の翻訳実験が行われていない点が挙げられる。本研究では、対訳句を利用した日英ニューラル機械翻訳を提案し、語順の大きく異なる日英翻訳間において対訳句を利用する手法が有効であるかの検証を行う。また、本研究において、人手対訳句として対訳文から作成した大規模かつ高精度の対訳句を利用する実験を行っているが、同様の実験は先行研究において行われていない。さらに、先行研究において、対訳句は未

知語の置換処理に用いる手法や，入・中間・出力層でのネットワークの計算に用いる手法がとられているが，学習データに追加する手法は提案されていないという点が異なっている．本研究の提案手法では，対訳句を学習データに直接追加することで句レベルの対応を強化させ，Attention の精度を向上させることを目的としている．提案手法の利点としては，対訳句を学習データに直接追加する手法であるため，モデルの拡張が容易である点があげられる．提案手法は，Attention を利用し，より翻訳精度の高いと報告される Transformer モデル [12] や，句に基づくアプローチをとる他のニューラル機械翻訳手法と提案手法を組み合わせることが容易であるという特徴がある．

## 4 対訳句を利用したニューラル機械翻訳

本研究では Attention に基づく日英ニューラル機械翻訳 [1] における学習データとして、日英の対訳文に対訳文から作成した対訳句を追加する。本研究における対訳句を利用した日英ニューラル機械翻訳の流れを図 3 に、手順を以下に示す。

手順 1 対訳文から対訳句を人手または自動で作成する。

手順 2 対訳文に作成した対訳句を追加し、対訳データを作成する。

手順 3 手順 2 で作成した対訳データを学習データとして日英ニューラル機械翻訳の学習を行う。

手順 4 手順 3 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。

また、表 4.1 に対訳文に作成した対訳句を追加した対訳データの例を示す。

表 4.1: 対訳データの例

| 日本語側                     | 英語側                               |
|--------------------------|-----------------------------------|
| 彼女はにっこり笑ってぼくに<br>あいさつした。 | She greeted me with a big smile . |
| 彼女                       | She                               |
| にっこり笑っ                   | a big smile                       |
| ぼくにあいさつし                 | greeted me                        |

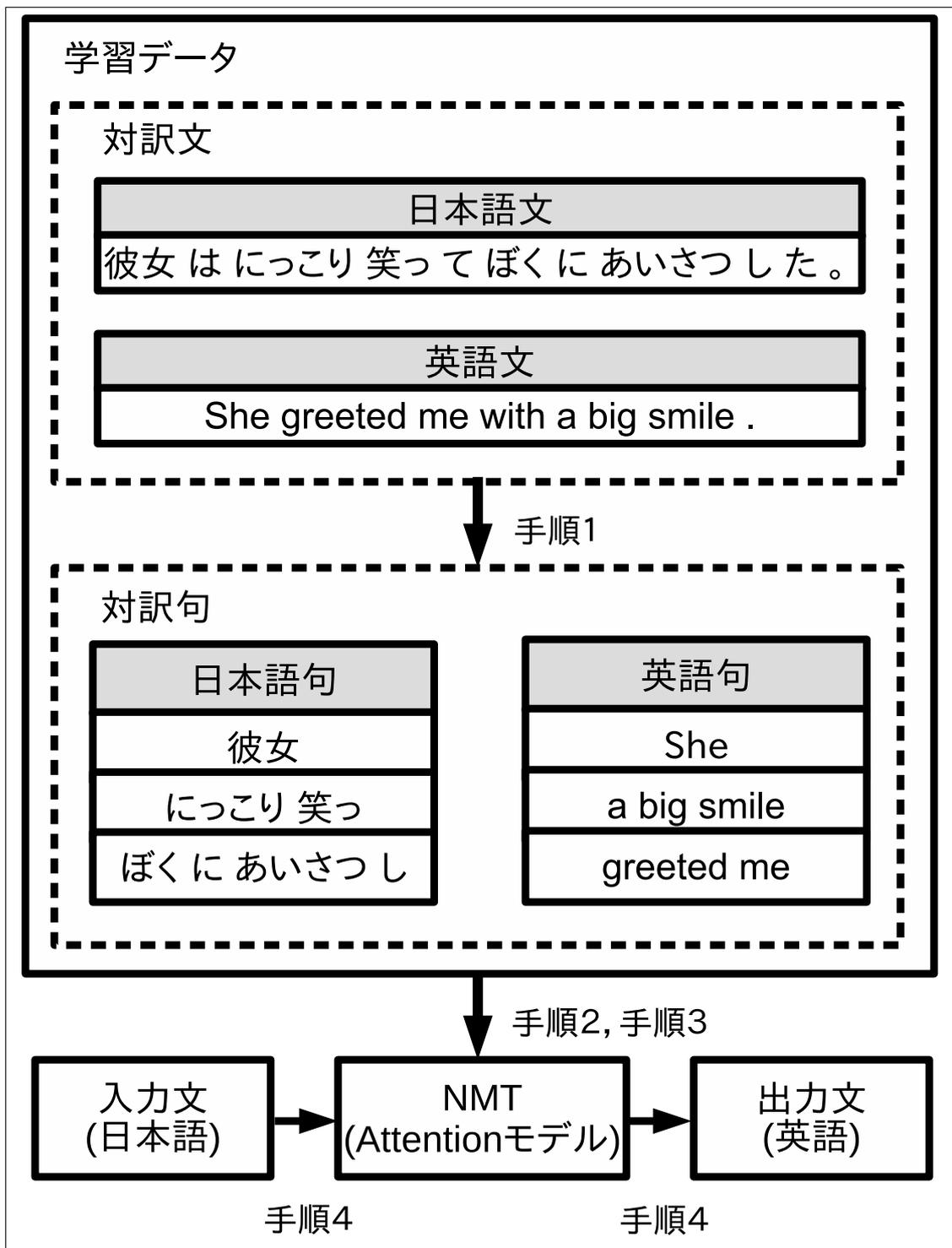


図 3: 対訳句を利用した日英ニューラル機械翻訳の流れ

## 4.1 対訳句作成システム

提案手法において、対訳文から対訳句を作成する必要がある。対訳文から対訳句を作成する手法には人手で作成する手法と自動で作成する手法の2通りある。本章では、この2通りの手法の説明を行う。なお、以下、人手で作成した対訳句を人手対訳句、自動で作成した対訳句を自動対訳句と呼ぶ。

### 4.1.1 人手対訳句の作成

対訳句を人手で作成する手法として、対訳文とその対訳文から人手で作成した対訳文パターンを用いる手法がある。具体的には、対訳文とその対訳文から人手で作成した対訳文パターンを比較した変数部から品詞が *AJ*(形容詞), *V*(動詞), *N*(名詞), *ADV*(副詞), *AJV*(形容動詞), *VP*(動詞句), *AJVP*(形容動詞句), *ADJP*(形容詞句), *NP*(名詞句) の対訳句を抽出し作成する。また、対訳句は基本的に句単位で作成しているが、単語単位や節単位の形で作成されたものも含まれる。なお、詳しい対訳文パターン作成方法は池原らの論文 [15] を参照のこと。

対訳文の例を表 4.2, 対訳文パターンの例を表 4.3, 対訳文と対訳文パターンの比較により作成された対訳句の例を表 4.4 に示す。この例では、表 4.2 の対訳文と表 4.3 の対訳文パターンを比較し、対訳文パターンにおける、「*NP*」や「*AJ*」などの変数に当たる対訳文中の語句を抽出した結果、表 4.4 に示す「彼の, his」「若い, young」などの対訳句が作成されている。

表 4.2: 対訳文の例

---

彼のお母さんがああ若いとは思わなかった。  
I never expected his mother to be so young .

---

表 4.3: 対訳文パターンの例

---

*NP N* が ああ *AJ* とは *V* なかった。  
I never *V NP N* to be so *AJ* .

---

表 4.4: 対訳文と対訳文パターンの比較により作成された対訳句の例

| 日本語句    | 英語句        |
|---------|------------|
| 彼の      | his        |
| お母さん    | mother     |
| 彼の お母さん | his mother |
| 若い      | young      |
| 思わ      | expected   |

#### 4.1.2 自動対訳句の作成

対訳句を自動で作成する手法として、対訳文とその対訳文から自動で作成した対訳文パターンを用いる手法がある。具体的には、まず、GIZA++[16]を用いて、対訳単語を作成する。次に、その対訳単語と対訳文を用いて対訳文パターンを作成する。そして、対訳文と対訳パターンを用いて対訳句を作成する。

以下、自動対訳句の作成過程における各手順について記述する。

##### 4.1.2.1 対訳単語の作成

GIZA++を用いて対訳単語を作成する。対訳単語作成の手順を以下に、例を図4に示す。

手順1 GIZA++を用いて対訳文から日英方向と英日方向の単語対応を得る。

手順2 単語対応より対訳単語を得る。

手順3 日英方向と英日方向の単語の翻訳確率を掛け合わせ、対訳単語の翻訳確率(以下、対訳単語翻訳確率と呼ぶ)を得る。

手順4 対訳単語翻訳確率が一定の閾値( $\alpha$ )以上である対訳単語を抽出する。

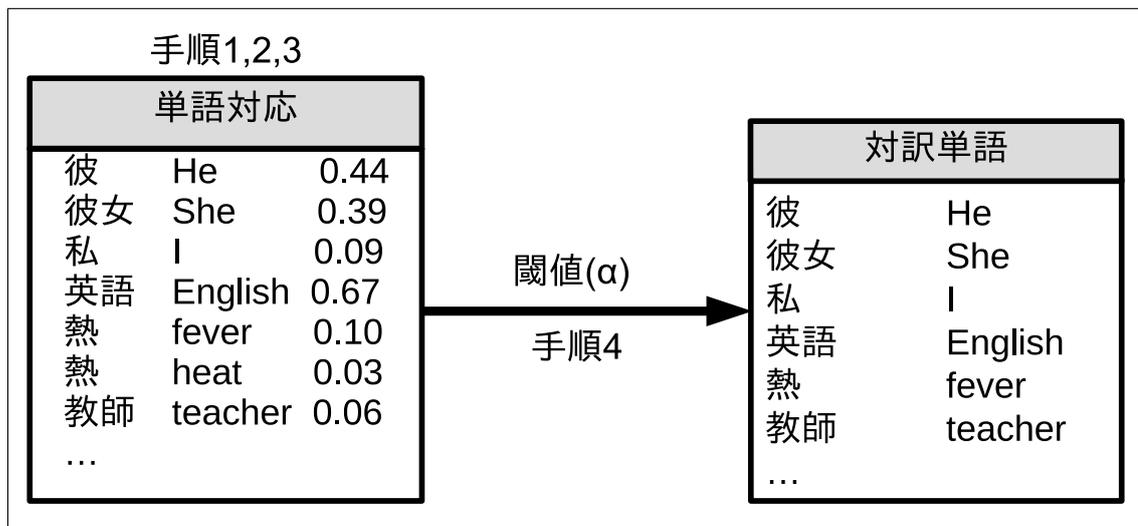


図 4: 対訳単語作成の例

#### 4.1.2.2 対訳文パターンの作成

対訳単語と対訳文を用いて対訳文パターンを作成する。対訳文パターンの作成の手順を以下に、例を図5に示す。

**手順1** 4.1.2.1節で抽出した対訳単語が対訳文中で適合した場合、変数化を行い、対訳文パターンを得る。

**手順2** 対訳文パターンの英語側において、変数の直前に冠詞がある場合、冠詞を除去する。

なお、変数が連続しない対訳文パターンのみを本研究で用いる対訳文パターンとする。

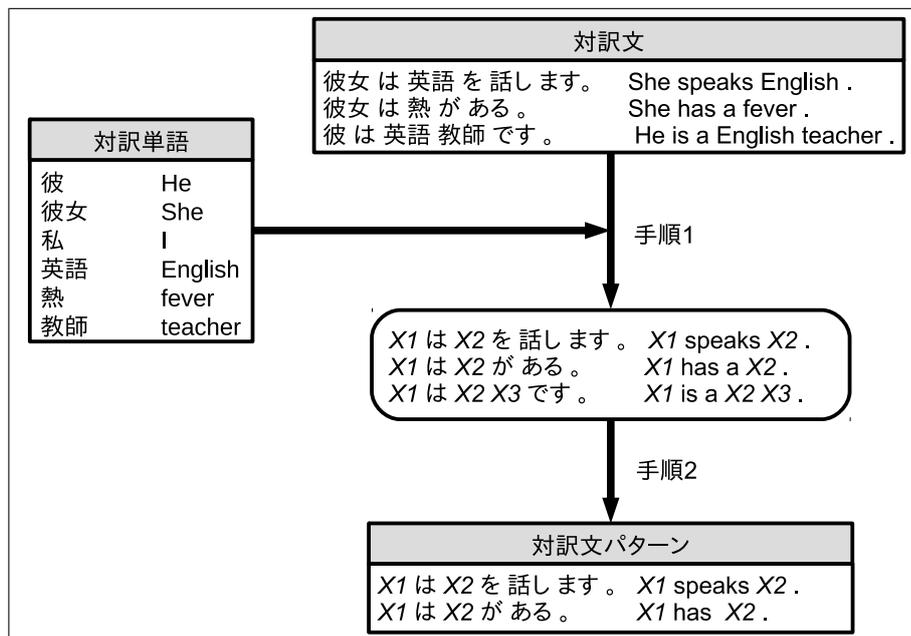


図 5: 対訳文パターン作成の例

#### 4.1.2.3 対訳句の作成

対訳文と対訳文パターンを用いて対訳句を作成する。対訳句の作成の手順を以下に、例を図6に示す。

**手順1** 対訳文と対訳文パターンを照合する。

**手順2** 対訳文が対訳文パターンに適合した場合、対訳文パターンの変数部に対応する対訳文を対訳句として抽出する。

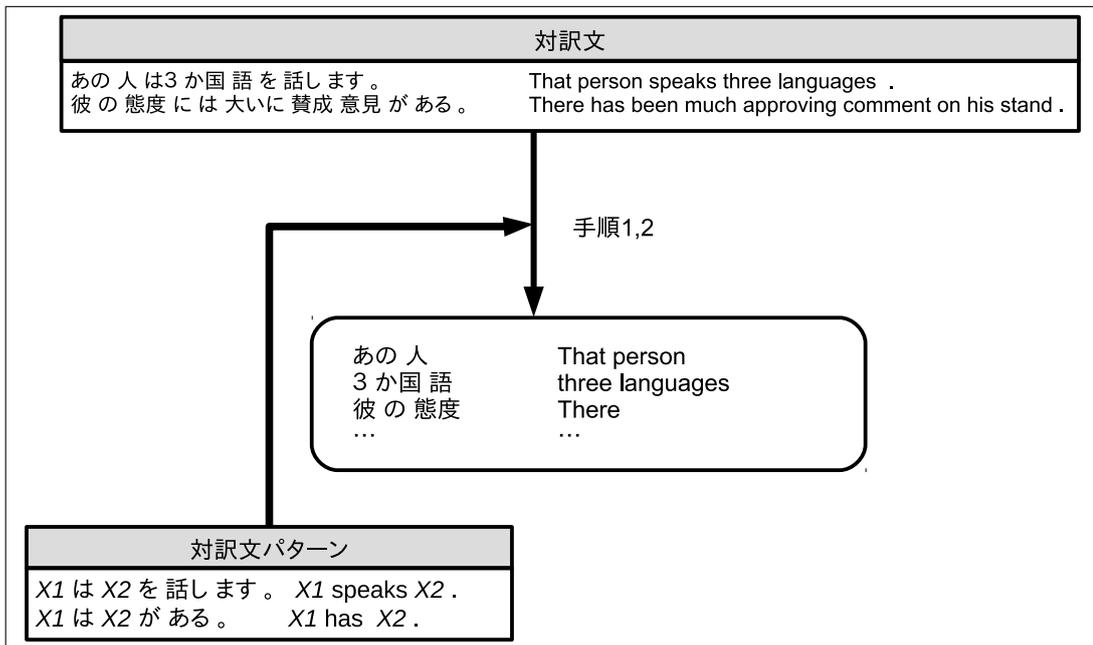


図 6: 対訳句作成の例

## 4.2 提案手法の設定

本研究の提案手法として、対訳文に対訳句を追加する手法を2通り設定する。2通りの対訳句追加手法を以下に示す。

- + **人手対訳句** 日英 NMT の流れの手順 1 において、対訳文から対訳句を作成する際に  
人手による方法を用いた場合の手法
- + **自動対訳句** 日英 NMT の流れの手順 1 において、対訳文から対訳句を作成する際に  
自動による方法を用いた場合の手法

## 5 実験環境

### 5.1 実験データ

本研究では、実験データとして日英重文複文コーパス [17] を用いる。実験で用いるデータの内訳を表 5.1 に、実験データにおける対訳文の例を表 5.2 に示す。

表 5.1: 実験データの内訳

| 学習データ (重文複文)       | 日本語         | 英語          |
|--------------------|-------------|-------------|
| 文数                 | 100,000 文   |             |
| 単語数                | 1,404,366 語 | 1,137,163 語 |
| 異なり単語数             | 38,530 語    | 37,035 語    |
| 平均文長               | 14.0 語      | 11.4 語      |
| テストデータ (重文複文)      | 日本語         | 英語          |
| 文数                 | 10,000 文    |             |
| 単語数                | 140,671 語   | 113,586 語   |
| 異なり単語数             | 13,765 語    | 12,176 語    |
| 平均文長               | 14.1 語      | 11.4 語      |
| デベロップメントデータ (重文複文) | 日本語         | 英語          |
| 文数                 | 1,000 文     |             |
| 単語数                | 14,237 語    | 11,591 語    |
| 異なり単語数             | 3,438 語     | 3,047 語     |
| 平均文長               | 14.3 語      | 11.6 語      |

表 5.2: 対訳文の例

|  |
|--|
| 彼女はにっこり笑ってぼくにあいさつした。<br>She greeted me with a big smile .  |
| 僕は君と同じ時計を持っている。<br>I have the same watch as yours .  |
| この本は値段が付けられないほど貴重な物です。<br>This book is so valuable it would be impossible to put a price on it . |

## 5.2 ニューラル機械翻訳の実験設定

ニューラル機械翻訳のツールキットには OpenNMT [18] を用い、モデルは Luong ら [14] により提案された Global Attention を用いる。Encoder, Decoder の LSTM は 2 層とし、ユニット数は 500、単語の分散表現のベクトルサイズは 500 を設定する。ミニバッチサイズは 40 とし、モデルの訓練は最大 32 エポック行う。

## 5.3 評価実験

実験では NMT の学習データに対訳文のみを用いる手法をベースラインとし、対訳文に人手対訳句を追加した学習データを用いる手法 (+人手対訳句)、対訳文に自動対訳句を追加した学習データを用いる手法 (+自動対訳句) の 2 つの対訳句追加手法をそれぞれベースラインと比較する。

翻訳精度の評価方法として、自動評価では実験において得られた 10,000 文の出力文に対して、4 つの自動評価指標 (BLEU, METEOR, RIBES, WER<sup>1</sup>) に基づいて評価を行う。人手対比較評価では出力文 10,000 文より無作為に抽出した 100 文に対して、正確性 (adequacy: 入力文の意味をどれだけ正確に翻訳文より読み取れるか) に基づいて人手対比較評価を行う。

---

<sup>1</sup>WER は誤り率に基づく指標であり、値が低いほど翻訳精度が高いことを示す

## 6 対訳句の抽出

本研究では，5章で示した対訳文より作成した対訳句を利用する．本章では，翻訳実験において用いる対訳句の作成を行い，作成した対訳句の数や精度を調査する．具体的な対訳句の説明としては，対訳文から人手で作成する手法により作成した人手対訳句と，対訳文から自動で作成する手法により作成した自動対訳句がある．

### 6.1 人手対訳句

人手対訳句として，鳥バンク [19] において人手により作成したパターンを対訳文と比較し，変数部の品詞が *AJ*(形容詞)，*V*(動詞)，*N*(名詞)，*ADV*(副詞)，*AJV*(形容動詞)，*AJP*(形容詞句)，*VP*(動詞句)，*AJVP*(形容動詞句) となるものを利用する．本研究で用いる人手対訳句の内訳を表 6.1 に，人手対訳句の例を表 6.2 に示す．

表 6.1: 人手対訳句の内訳

|        | 日本語         | 英語          |
|--------|-------------|-------------|
| 句数     | 548,694 句   |             |
| 単語数    | 1,058,949 語 | 1,031,265 語 |
| 異なり単語数 | 35,613 語    | 34,823 語    |
| 平均句長   | 1.93 語      | 1.88 語      |

表 6.2: 人手対訳句の例

| 日本語句       | 英語句         |
|------------|-------------|
| 彼女         | She         |
| にっこり 笑っ    | a big smile |
| ぼく         | me          |
| ぼくに あいさつ し | greeted me  |
| あいさつ し     | greeted     |

人手対訳句の精度を測るため、全ての人手対訳句から無作為に 100 対を抽出し、それぞれの対応の適切さを人手により評価した。評価基準を以下に示す。評価結果を表 6.3 に、評価例を表 6.4 に示す。

- 適切な対応をとる対訳句
- △ 部分的に適切な対応をとる対訳句
- × 不適切な対応をとる対訳句

表 6.3: 人手対訳句の評価結果

| ○    | △    | ×   |
|------|------|-----|
| 86 句 | 10 句 | 4 句 |

表 6.4: 人手対訳句の評価例

| 評価 | 対訳句例       |                     |
|----|------------|---------------------|
| ○  | 条件         | the conditions      |
|    | 演説し        | made a speech       |
|    | あんな 賢い 子ども | such a clever child |
| △  | 暇          | will be free        |
|    | 歌詞 を 書け    | Fit the words       |
|    | 勝負         | victory             |
| ×  | 山田 さん      | he                  |
|    | 自社         | their               |
|    | ブラウス       | it                  |

## 6.2 自動対訳句

自動対訳句の作成手順 [20] を以下に示す。

手順 1 対訳文と GIZA++を利用して単語辞書を作成する。

手順 2 対訳文と作成した単語辞書を用いて単語レベル文パターンを作成する。

手順 3 対訳文と単語レベル文パターンを用いて、対訳句を作成する。

実験において得られた自動対訳句の内訳を表 6.5 に示す.

表 6.5: 自動対訳句の内訳

|        | 日本語       | 英語        |
|--------|-----------|-----------|
| 句数     | 26,757 句  |           |
| 単語数    | 127,931 語 | 106,224 語 |
| 異なり単語数 | 15,459 語  | 12,976 語  |
| 平均句長   | 4.78 語    | 3.97 語    |

また, 6.1 節と同様に, 全ての自動対訳句から無作為に 100 対を抽出し, それぞれの対応の適切さを人手により評価した. 評価結果を表 6.6 に, 評価例を表 6.7 に示す.

表 6.6: 自動対訳句の評価結果

| ○    | △    | ×    |
|------|------|------|
| 36 句 | 44 句 | 20 句 |

表 6.7: 自動対訳句の評価例

| 評価 | 対訳句例  |
|----|---|
| ○  | 容赦のない残忍さ<br>食事の用意をし<br>同じ道<br>remorseless brutality<br>preparing the meals<br>same path |
| △  | の痛みが和らい<br>はしごを掛け<br>フラスコ<br>pain from the<br>a ladder<br>flask and heat                |
| ×  | 入りて<br>の卑しい<br>で給水の<br>Ogata<br>him in birth<br>on its                                  |

人手対訳句と自動対訳句の精度の比較では, 自動対訳句の精度は適切な対応をとる句が 100 句中 36 句となり, 適切な対応をとる句が 100 句中 86 句である人手対訳句と比較して, 低くなっていることが確認できる.

## 7 翻訳実験

本章では，日英 NMT の学習データとして，対訳文と対訳文から作成した対訳句を追加する手法の翻訳実験を行う．また，追加する対訳句には，人手対訳句および自動対訳句を利用する．

### 7.1 自動評価結果

各手法の出力文 10,000 文に関して自動評価を行った．評価結果を表 7.1 に示す．

表 7.1: ベースライン，+人手対訳句，+自動対訳句の自動評価結果 (10,000 文)

| 手法     | BLEU   | METEOR | RIBES  | WER    |
|--------|--------|--------|--------|--------|
| ベースライン | 0.1534 | 0.4147 | 0.7312 | 0.7107 |
| +人手対訳句 | 0.1701 | 0.4430 | 0.7411 | 0.6873 |
| +自動対訳句 | 0.1585 | 0.4194 | 0.7332 | 0.7017 |

表 7.1 より，自動評価の結果では，+人手対訳句，+自動対訳句，ベースラインの順に翻訳精度が高くなっていることが確認できる．

### 7.2 人手対比較評価結果

各手法の出力文 10,000 文より無作為に抽出した 100 文に対して人手対比較評価を行った．評価結果を表 7.2 および表 7.3 に示す．また，人手対比較評価に関する表記の説明を以下に示す．

- ベースライン○

比較する出力文に対してベースラインの出力文の方がより入力文の意味に近い

- +人手対訳句○

比較する出力文に対して+人手対訳句の出力文の方がより入力文の意味に近い

- +自動対訳句○

比較する出力文に対して+自動対訳句の出力文の方がより入力文の意味に近い

- 差なし

比較する両手法の出力文には正確性において差がない

- 一致

比較する両手法の出力文が完全に同一

表 7.2: ベースラインと+人手対訳句の人手対比較評価結果 (100 文)

| ベースライン○ | +人手対訳句○ | 差なし  | 一致  |
|---------|---------|------|-----|
| 14 文    | 44 文    | 38 文 | 4 文 |

表 7.3: ベースラインと+自動対訳句の人手対比較評価結果 (100 文)

| ベースライン○ | +自動対訳句○ | 差なし  | 一致  |
|---------|---------|------|-----|
| 10 文    | 29 文    | 56 文 | 5 文 |

表 7.2 および表 7.3 より，人手対比較評価では，2つの対訳句追加手法はいずれもベースラインと比較して高い翻訳精度であることがわかる．さらに，翻訳精度の向上は，+人手対訳句の方が+自動対訳句と比較して大きいことが確認できる．また，+自動対訳句において，自動評価の向上はあまり大きくない一方で，人手対比較評価では大きく向上したことが確認できる．

## 7.3 人手対比較評価例

この節では，7.2 節の人手対比較評価における具体的な評価例を示す．

### 7.3.1 ベースラインと+人手対訳句の人手対比較評価例

ベースラインと+人手対訳句との対比較評価において，+人手対訳句○とされた例を表 7.4 に，ベースライン○とされた例を表 7.5 に，差なしとされた例を表 7.6 に示す．

表 7.4: +人手対訳句○

|        |   |
|--------|---|
| 入力文    | 彼は自尊心が強すぎてとても引き返せなかった。  |
| 参照文    | He was far too proud to go back .   |
| ベースライン | He was too proud for his pride .  |
| +人手対訳句 | He was too proud to turn back .   |
| 入力文    | いくら洗ってもそのようなよごれは取れないでしょう。   |
| 参照文    | No amount of washing will remove such stains .  |
| ベースライン | No matter how much you wash it , you will not have such stain .                           |
| +人手対訳句 | No matter how much you wash it , that stain will not come off .                           |
| 入力文    | 技術系専門職の職員は高度な自主・独立性を求めたがる。  |
| 参照文    | Technical professionals want a high degree of autonomy .                                  |
| ベースライン | A employee 's staff want to want highly active independence .                             |
| +人手対訳句 | The workers in technical expertise are apt to appeal for an advanced spirit independent . |

表 7.5: +ベースライン○

|        |   |
|--------|---|
| 入力文    | それを見て私はどきどきした。  |
| 参照文    | I throbbed at the sight .   |
| ベースライン | The sight made me feel nervous .                                      |
| +人手対訳句 | The sight thumped me .  |
| 入力文    | そのうちによい機会を捉えて留学したいと思っている。   |
| 参照文    | I hope I can get the chance to study abroad someday .                 |
| ベースライン | I want to let the best opportunity to study abroad .                  |
| +人手対訳句 | I would like to take my chance to study in time for the opportunity . |
| 入力文    | 弾丸が心臓を貫き、彼はこと切れた。   |
| 参照文    | A bullet through the heart ended him .                                |
| ベースライン | The bullet bombarded his heart .                                      |
| +人手対訳句 | His heartbeat gave him a jolt .                                       |

表 7.6: 差なし

|        |  |
|--------|--|
| 入力文    | 彼は助けを求めてさげんでいた。                                  |
| 参照文    | He was crying for help .                         |
| ベースライン | He was sulking for help .                        |
| +人手対訳句 | He was rumped for help .                         |
| 入力文    | この問題に答えられた者は一人もいなかった。                            |
| 参照文    | None of them were able to answer this question . |
| ベースライン | There was no one to answer this question .       |
| +人手対訳句 | No one was able to answer this problem .         |
| 入力文    | そんなことすると口が干上がってしまう。                              |
| 参照文    | That will take the bread out of my mouth .       |
| ベースライン | That would interfere in the mouth .              |
| +人手対訳句 | It will lapse my mouth .                         |

### 7.3.2 ベースラインと+自動対訳句の人手対比較評価例

ベースラインと+自動対訳句との対比較評価において、+自動対訳句○とされた例を表 7.7 に、ベースライン○とされた例を表 7.8 に、差なしとされた例を表 7.9 に示す。

表 7.7: +自動対訳句○

|        |  |
|--------|--|
| 入力文    | 警官に護送されて来た。  |
| 参照文    | He came under police escort .                              |
| ベースライン | He has been sent to the police .                           |
| +自動対訳句 | He came under police escort .                              |
| 入力文    | 今が彼に追いつく絶好のチャンスだ。  |
| 参照文    | This is a good occasion to overtake him .                  |
| ベースライン | This is a good chance to get him .                         |
| +自動対訳句 | This is the best chance to surpass him .                   |
| 入力文    | これらの威しにひるむことなく仕事を続けた。                                      |
| 参照文    | Undaunted by these threats , he went on with his task .    |
| ベースライン | They continued to work without threats to these threats .  |
| +自動対訳句 | He continued his work without flinching of these threats . |

表 7.8: +ベースライン○

|        |  |
|--------|--|
| 入力文    | 両国は兵力を1万に抑えることで合意した。                                   |
| 参照文    | Both countries agreed to limit troops to 10,000.       |
| ベースライン | The two countries agreed to restrict troops to 10,000. |
| +自動対訳句 | The two countries agreed to chase to 10,000.           |
| 入力文    | 食卓がいっぱいでパンを置く所がない。                                     |
| 参照文    | There is no room on the table for the bread.           |
| ベースライン | The table is full and there is no place for bread.     |
| +自動対訳句 | There is no place for bread to keep the bread.         |
| 入力文    | 彼は甘んじて恥辱を受ける決心をした。                                     |
| 参照文    | He decided to submit to disgrace.                      |
| ベースライン | He made up his mind to take a humiliation.             |
| +自動対訳句 | He was resolved to get his bite.                       |

表 7.9: +差なし

|        |   |
|--------|---|
| 入力文    | 彼は誇りが高すぎてそんなに易しい曲は演奏しない。                                      |
| 参照文    | He is too proud to play such an easy piece.                   |
| ベースライン | He is too proud to play so hard.                              |
| +自動対訳句 | He is too proud to play so much.                              |
| 入力文    | なるべく人と衝突を避けるが好い。  |
| 参照文    | You ought to avoid collision with people as much as possible. |
| ベースライン | You had better avoid a collision with a person.               |
| +自動対訳句 | You ought to avoid a collision with a man.                    |
| 入力文    | 彼は英語を知っているから用いられる。  |
| 参照文    | His knowledge of English makes him useful.                    |
| ベースライン | He is used in teaching English.                               |
| +自動対訳句 | He is used to learn English.                                  |

## 8 考察

### 8.1 対訳句利用の効果

結果として、2つの対訳句追加手法において翻訳精度の向上が確認された。利用した2種類の対訳句はいずれも対訳文より作成し、対訳文外の語彙などの情報を有しないものである。これより、文レベルの対応情報に加えて句レベルの対応情報を学習データに付与することがNMTにおいて語句の対応情報をモデル化するAttentionの強化につながり、翻訳精度が向上したと考えられる。

また、2つの対訳句追加手法のうち、句の対応が比較的適切かつ多量である人手対訳句を利用する手法がより高い翻訳精度を示した。したがって、翻訳精度の向上にはより高精度かつ多量の対訳句を利用する必要があると推察される。

### 8.2 翻訳結果と対訳句の分析

翻訳結果と対訳句の分析から、正しい対訳句を追加したことにより翻訳精度が向上したと見られる例が確認された。ここでは、まず翻訳精度が向上した結果と追加した対訳句との関係性について、+人手対訳句の実験、+自動対訳句の実験の順に述べ、次に翻訳精度が低下した結果と追加した対訳句との関係性について、+人手対訳句の実験、+自動対訳句の実験の順に述べる。

#### 8.2.1 翻訳精度が向上した結果

+人手対訳句の結果について、表7.4の例において、入力文中の単語「引き返せ」に対して、+人手対訳句では比較的正しい翻訳となっている。これに関して、追加した人手対訳句を調査すると、「引き返せ」に対して正しい訳語「turn back」を含む対訳句が存在し、正しい句レベルの対応を学習したために正確性が向上したと考えられる。

また、+自動対訳句の結果についても同様のことが確認できる。表7.7の1つ目の例において、入力文中の語句「警官に護送されて」に対して、自動対訳句中に比較的正しい訳語「under police escort」が含まれている。これにより、適切な語句を選択するように学習が強化され、より高い精度の翻訳が得られたと考えられる。

### 8.2.2 翻訳精度が低下した結果

一方で、対訳句を追加することで翻訳精度が低下したと考えられる例も存在する。+人手対訳句の結果について、表 7.5 の例において、+人手対訳句では出力文中に「thumped me」という誤った翻訳が生じている。これに関して、入力文中の日本語単語「どきどき」を含む人手対訳句を調査すると、英語側に「thumped」を含むものが複数確認された。これは、「胸がどきどきする」という日本語に対して「my heart thumped」という翻訳を含む対訳文から作られた対訳句である。このような対訳句の対応は必ずしも誤りとは言えないものの、表 7.5 の例のような文中の用法としては不適切なものとなっている。このように、対訳句を追加する手法では、句レベルの対応を重視した結果、文全体としては誤った表現を導く可能性があることが課題としてあげられる。

また、+自動対訳句の結果について、表 7.8 の例では、入力文中の語句「兵力を 1 万に抑える」に対して、ベースラインでは「restrict troops to 10,000」と、比較的正しい翻訳が得られている一方で、+自動対訳句では誤った翻訳となっている。これについて、入力文中の日本語句と自動対訳句との相関を調査すると、自動対訳句中に日本語句「兵力を 1 万」に対して「to 10,000」という英語句の対が確認された。この対では、日本語側の単語「兵力」に対して正しい訳語が含まれておらず、これを用いた NMT の学習では不十分な対応が強調された可能性がある。このために、+自動対訳句の翻訳文においても正しい訳語を含まない出力文が生成されたと推察される。

## 8.3 人手対訳句と自動対訳句の比較

人手対訳句と自動対訳句との比較として、自動対訳句と比較して人手対訳句の方が高精度かつ多量に得られており、翻訳精度についても人手対訳句を追加する手法は自動対訳句を追加する手法と比較してより高い結果を示した。しかしながら、人手対訳句の作成には大きなコストがかかる問題がある一方で、自動対訳句の作成におけるコストは小さい。また、自動対訳句はプログラムやパラメータの改良を行うことで、より適切な対訳句の作成や、より多くの対訳句の作成が見込めるという利点がある。

## 9 追加実験：外部辞書データより抽出した対訳句

7章の実験結果より，NMTの学習データに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．本研究では，対訳句を利用してNMTの翻訳精度を向上させるための他の手法による実験を行い，対訳句を利用してより有効にNMTの翻訳精度を向上させる手法について検証する．

### 9.1 外部辞書データより抽出した対訳句の利用

7章の実験では，追加する対訳句には対訳学習文から人手もしくは自動で作成したものを利用した．これより，NMTにおいて，対訳学習文外の情報を有しない対訳句利用した場合であっても，翻訳精度は向上することが確認できた．

本節では，対訳文外の情報を含むデータを有する外部辞書より抽出した対訳句を対訳文に追加する手法について実験を行う．この手法の利点および欠点について，対訳文から作成された対訳句を利用する手法と比較して，以下に示す．

- 利点 1. 既に作成された外部辞書などの対訳句コーパスを有していれば，対訳文から対訳句を作成する必要がなく，特に人手対訳句を利用する手法と比較してコストが小さい．
- 利点 2. 対訳文外の情報を有する対訳句コーパスを利用することで，対訳文に存在しない語彙などの情報を拡張した翻訳の実現が見込める．これに対して，対訳文より対訳句を作成する手法では翻訳文中に未知語（対訳文に存在しない単語）の翻訳は理論上不可能である．
- 欠点 1. 対訳文とは別の対訳句コーパスを備えるコストが必要となる．
- 欠点 2. 対訳文コーパスと対訳句コーパスのドメインが乖離している場合には翻訳精度が低下する可能性がある．また，この問題が利用するコーパスのペアにおいて生じるか否かは実験的に検証する必要がある．
- 欠点 3. 対訳文外の対訳句コーパスを利用することで，学習におけるNMTの語彙数が大きくなり，出力における分類問題が複雑になるため，精度低下や計算コストの増大が生じる可能性がある．

## 9.2 実験設定

今回実験において対訳句として利用する外部辞書データには英辞郎 [21] を用いる。表 9.1 に英辞郎の内訳を，表 9.2 に英辞郎の例を示す。

なお，実験に利用する対訳文，テストデータ，ディベロップメントデータには 5 章と同じ重文複文対訳コーパスを用いる。

表 9.1: 英辞郎の内訳

|        | 日本語         | 英語          |
|--------|-------------|-------------|
| 句数     | 550,000 句   |             |
| 単語数    | 1,641,461 語 | 1,405,828 語 |
| 異なり単語数 | 110,292 語   | 140,143 語   |
| 平均句長   | 2.98 語      | 2.56 語      |

表 9.2: 英辞郎の例

| 日本語句   | 英語句                   |
|--------|-----------------------|
| シャンデリア | chandelier            |
| ドミニク   | Dominique             |
| 継承する   | inherit               |
| 激烈な論争  | bitter controversy    |
| 私的諮問機関 | private advisory body |

## 9.3 英辞郎の精度

6.1 節と同様に，全ての英辞郎の対訳句から無作為に 100 対を抽出し，それぞれの対応の適切さを人手により評価した。評価結果を表 9.3 に，評価例を表 9.4 に示す。

表 9.3: 英辞郎の評価結果

| ○    | △   | ×   |
|------|-----|-----|
| 98 句 | 1 句 | 1 句 |

表 9.4: 英辞郎の評価例

| 評価 | 対訳句例      |                              |
|----|-----------|------------------------------|
| ○  | 陣笠 議員     | backbencher                  |
|    | 礼拝によく行く人  | chapelgoer                   |
|    | モギレフスキー   | Mogilevskii                  |
| △  | 分解 写真 撮影  | suspended motion photography |
| ×  | サッと 機首を振る | jank                         |

表 9.4 において、英辞郎の対訳句中に対応が部分的に不適切と評価された句が 1 対、対応が不適切とされた句が 1 対存在した。対応が部分的に不適切と評価された句について、日本語の「分解写真」は連続写真を意味する語とされ、これに対して「suspended motion」は、評価者が調査した限りにおいて「動作が停止された様子」以上の意味は認められず、△の評価とされた。また、対応が不適切とされた句について、日本語側および英語側のいずれの単語の意味も不鮮明であると評価されたため、×の評価となった。これら 2 例については評価者の異なる場合に、必ずしも同様の評価結果が得られるとは限らないと考えられる。

## 9.4 NMT 学習時の語彙数の制限

表 9.1 より、英辞郎を追加する手法では学習データにおける異なり単語数が非常に多くなるのがわかる。一般に、NMT の学習では、学習データの異なり単語数を語彙数と呼び、学習時の語彙数が多くなる場合には利用する語彙を頻度などを基準に制限する手法が用いられる [1, 22]。したがって、今回は日英両言語の学習時の語彙数を頻度上位 60,000 語と設定し、それ以外の単語は全て特殊記号 (<unk>) に置き換えて実験を行った。語彙数を制限した学習データの例を表 9.5 に示す。

表 9.5: 語彙数を制限した学習データの例

|        | 日本語側                  | 英語側                               |
|--------|-----------------------|-----------------------------------|
| 元データ   | <u>ろば</u> は優しい目をしている。 | <u>Donkeys</u> have gentle eyes . |
| 語彙数制限後 | <unk> は優しい目をしている。     | <unk> have gentle eyes .          |
| 元データ   | <u>サンディア</u> 山脈       | <u>Sandia</u> Mountains           |
| 語彙数制限後 | <unk> 山脈              | <unk> Mountains                   |
| 元データ   | 汽車に乗ると酔う。             | I get <u>trainsick</u> .          |
| 語彙数制限後 | 汽車に乗ると酔う。             | I get <unk> .                     |
| 元データ   | <u>造石</u> 税           | tax on brewage                    |
| 語彙数制限後 | <unk> 税               | tax on brewage                    |

## 9.5 自動評価結果

入力文 10,000 文に対して得られた各手法の出力文 10,000 文について、自動評価を行った。結果を表 9.6 に示す。

表 9.6: ベースライン, +英辞郎の自動評価結果 (10,000 文)

| 手法     | BLEU   | METEOR | RIBES  | WER    |
|--------|--------|--------|--------|--------|
| ベースライン | 0.1534 | 0.4147 | 0.7312 | 0.7107 |
| +英辞郎   | 0.1624 | 0.4376 | 0.7412 | 0.6864 |

表 9.6 より、+英辞郎の手法における自動評価では、全ての指標においてベースラインを上回る結果となった。

また、7章の対訳句追加手法と比較して、全ての指標において+自動対訳句よりも高く、RIBES, WER の指標において+人手対訳句より高く、BLEU, METEOR の指標において+人手対訳句より低い結果となった。

## 9.6 人手対比較評価結果

出力文10,000文より100文を抽出し、人手対比較評価を行った。人手対比較評価は、+英辞郎の出力文とベースラインの出力文とを比較し、より入力文の意味が読み取れる(正確である)と考えられる文を○とした。結果を表9.7に示す。

表 9.7: ベースラインと+英辞郎の人手対比較評価結果 (100 文)

| ベースライン○ | +英辞郎○ | 差なし | 同一出力 |
|---------|-------|-----|------|
| 14      | 34    | 49  | 3    |

人手対比較評価の結果から、ベースラインと比較して+自動対訳句の翻訳精度が高いことが確認できる。

## 9.7 人手対比較評価例

人手対比較評価の結果、に+英辞郎の出力文が比較的正確であると評価された例を表9.8、ベースラインの出力文が比較的正確であると評価された例を表9.9に、2つの出力文は正確さにおいて差がないと評価された例を表9.10に示す。

表 9.8: +英辞郎○

|        |  |
|--------|--|
| 入力文    | 信用度の高い車がほしい。                               |
| 参照文    | I want a car with cachet .                 |
| ベースライン | I want a generous car .                    |
| +英辞郎   | I want a high credit car .                 |
| 入力文    | 彼は その知らせを聞いて 顔面 そう白になった。                   |
| 参照文    | He lost color at the news .                |
| ベースライン | He looked down at the news .               |
| +英辞郎   | When he heard the news , he turned white . |
| 入力文    | 故人は 誠実な方でした。                               |
| 参照文    | The deceased was an honest man .           |
| ベースライン | The man was a faithful man .               |
| +英辞郎   | The deceased was a sincere one .           |

表 9.9: ベースライン○

|        |  |
|--------|--|
| 入力文    | その 宝石 は どう 低く 見積もって も 5 万ドル の 価値 が ある だろ う 。                               |
| 参照文    | Even the lowest estimate would put the value of the jewel at \$ 50 , 000 . |
| ベースライン | The jewels will have a \$ worth \$ 50 , 000 .                              |
| +英辞郎   | The jewel will cost \$ 5 , 000 .   |
| 入力文    | その うち に よい 機会 を 捉え て 留学 し たい と 思っ て いる 。                                   |
| 参照文    | I hope I can get the chance to study abroad someday .                      |
| ベースライン | I want to let the best opportunity to study abroad .                       |
| +英辞郎   | I am thinking of reading the good opportunity for <u>studying</u> .        |
| 入力文    | 新 時代 が 近づい て いる の を はっきり 見て 取っ た 。   |
| 参照文    | He saw clearly that new times were coming .                                |
| ベースライン | I clearly discerned the approach of a new era .                            |
| +英辞郎   | I felt a vivid glimpse of the new age .                                    |

表 9.10: 差なし

|        |  |
|--------|--|
| 入力文    | 私 は 大学生 の 中 に 紛れ込ん で 授業 を 聞いて 来 た 。  |
| 参照文    | I blended in with the university students and attended the class .           |
| ベースライン | I mixed up in the university and listened to class .                         |
| +英辞郎   | I mixed into the university university and attended the class .              |
| 入力文    | 馬 は あちこち に 散らばっ て 草 を はん で いた 。  |
| 参照文    | The horses spread out to graze .   |
| ベースライン | The horse scattered around and occupied the grass .                          |
| +英辞郎   | The horse scattered scattered and scattered grass .                          |
| 入力文    | もし 来 る こ と が で き な け れ ば 、 前 の 晩 に 電 話 で 知 ら せ て くだ さ い 。                    |
| 参照文    | If you are unable to come , please let me know by phone the evening before . |
| ベースライン | If you can not come , please let me know before the evening .                |
| +英辞郎   | If you can not come , please let me know by tomorrow .                       |

## 9.8 追加した外部辞書と翻訳結果の分析

対訳学習文に英辞郎を追加する手法について、出力文と英辞郎の関係性について調査したところ、8.2節の考察と同様に追加した対訳句と+英辞郎の出力文に相関性があることが認められた。

表 9.8 の 1 つ目の例について、入力文中の語句「信用度の高い」に対してベースラインでは誤った訳語「generous」が出力されている一方、+英辞郎では比較的正しい訳語「high credit」が得られている。これについて、学習データ中の日本語単語「信用度」を含む対を調査した。「信用度」を含む対訳学習文中の対を表 9.11 に、「信用度」を含む英辞郎中の対訳句を表 9.12 に示す。

表 9.11: 「信用度」を含む対訳学習文中の対

| 日本語側  | 英語側  |
|---|--|
| 同社は、そのサービスのターゲットを信用度の高い人、つまりクレジットカードによる借金をすぐに払ってくれる人に置いた。 | The company targeted its offering to people with good <u>credit</u> , people who usually pay off credit card debt quickly. |
| かならず手形を期日に支払うので彼の信用度は高い。                                  | He never fails to honor his bills and his <u>credit</u> is good.   |

表 9.12: 「信用度」を含む英辞郎中の対訳句

| 日本語側      | 英語側                              |
|-----------|----------------------------------|
| 信用度       | credit raking                    |
| 本人の信用度    | credit worthiness of the subject |
| の信用度を評価する | evaluate one's credit            |
| 信用度       | credit rating                    |
| 信用度       | credit rating                    |

対訳学習文において「信用度」は表 9.11 に示す 2 文のみに出現する低頻度語となっている。NMT では統計的信頼性の低い低頻度語を含む文や長い文を含む処理において Attention の精度が低下するため、例 1 ではベースラインの出力文において誤りが生じたと考えられる。

一方で表 9.12 より、英辞郎において「信用度」を含む対訳句は 5 対 (全て 5 単語以下) ある上、全ての英語対訳句は比較的正しい訳語「credit」を含んでいることがわかる。したがって、入力文中の単語「信用度」に対して適切な対応をより多く含む英辞郎を用いる手法では推定の精度が向上し、比較的正しい単語「credit」が訳出されたと推察される。

## 10 追加実験：1変数対訳文パターンを用いた対訳句

7章の実験結果より，NMTの学習データに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．本研究では，対訳句を利用してNMTの翻訳精度を向上させるための他の手法による実験を行い，対訳句を利用してより有効にNMTの翻訳精度を向上させる手法について検証する．

### 10.1 1変数対訳文パターンを用いた対訳句の抽出手法

1変数対訳文パターンによる対訳句抽出手法 [23] とは，全対訳文パターンより，変数が1つである対訳文パターンのみを利用して対訳句の作成を行う手法である．1変数対訳文パターンによる対訳句作成手法では，7章において利用している，複数の変数が含まれる対訳文パターンを用いる手法により作成された対訳句 (以下，多変数対訳句と呼ぶ) と比較して，非常に高い精度で対訳句を抽出することが可能であるという報告がある．よって，本節では1変数対訳文パターンによる手法を用いて対訳文より作成した対訳句 (以下，1変数対訳句と呼ぶ) をNMTの学習データに追加する実験を行う．

ただし，森本らの提案した1変数対訳文パターンによる対訳句作成手法では，現状，変数部と対訳文パターンの照合が比較的難しい重文複文においては対訳句の作成が困難であるという課題点がある．したがって，本節では7章の実験で利用していた重文複文コーパスは利用せず，単文コーパスを用いて実験を行い，1変数対訳句を追加する手法の有効性を検証する．

## 10.2 実験設定

本節における実験では、NMT の学習データに多変数対訳句を追加する手法 (以下, +多変数対訳句と呼ぶ) および NMT の学習データに 1 変数対訳句を追加する手法 (以下, +1 変数対訳句と呼ぶ) を, 対訳文のみを学習データに追加するベースラインと比較する.

また, +1 変数対訳句の手法において, 重文複文コーパスが利用できない課題があるため, 実験データには電子辞書などから抽出した単文コーパス [17] を用いる. 表 10.1 に 1 変数対訳文パターンおよび自動対訳句作成に使用した実験データの内訳を示す.

表 10.1: 実験データの内訳

| 学習データ (単文)       | 日本語         | 英語          |
|------------------|-------------|-------------|
| 文数               | 159,998 文   |             |
| 単語数              | 1,620,994 語 | 1,395,630 語 |
| 異なり単語数           | 42,756 語    | 45,634 語    |
| 平均文長             | 10.1 語      | 8.7 語       |
| テストデータ (単文)      | 日本語         | 英語          |
| 文数               | 10,000 文    |             |
| 単語数              | 101,349 語   | 86,978 語    |
| 異なり単語数           | 12,624 語    | 11,977 語    |
| 平均文長             | 10.1 語      | 8.7 語       |
| デベロップメントデータ (単文) | 日本語         | 英語          |
| 文数               | 1,000 文     |             |
| 単語数              | 10,059 語    | 8,537 語     |
| 異なり単語数           | 2,791 語     | 2,771 語     |
| 平均文長             | 10.1 語      | 8.5 語       |

### 10.3 多変数対訳句および1変数対訳句の作成結果

表 10.1 の対訳学習文を用いて，多変数対訳句1変数対訳句の抽出を行った．表 10.2 に作成した多変数対訳句の内訳を，表 10.3 に抽出した1変数対訳句の内訳を示す．

表 10.2: 多変数対訳句の内訳

|        | 日本語       | 英語        |
|--------|-----------|-----------|
| 句数     | 35,127 句  |           |
| 単語数    | 176,515 語 | 157,263 語 |
| 異なり単語数 | 19,374 語  | 17,618 語  |
| 平均句長   | 4.5 語     | 4.5 語     |

表 10.3: 1変数対訳句の内訳

|        | 日本語      | 英語       |
|--------|----------|----------|
| 句数     | 17,522 句 |          |
| 単語数    | 75,258 語 | 65,546 語 |
| 異なり単語数 | 10,124 語 | 8,442 語  |
| 平均句長   | 4.3 語    | 3.7 語    |

### 10.4 多変数対訳句および1変数対訳句の精度

6.1 節と同様に，全ての1変数対訳句から無作為に100対を抽出し，それぞれの対応の適切さを人手により評価した．多変数対訳句における評価結果を表 10.4 に，1変数対訳句における評価結果を表 10.5 に示す．また，多変数対訳句における評価例を表 10.6 に，1変数対訳句における評価例を表 10.7 に示す．

表 10.4: 多変数対訳句の評価結果

| ○    | △    | ×    |
|------|------|------|
| 51 句 | 38 句 | 11 句 |

表 10.5: 1変数対訳句の評価結果

| ○    | △   | ×    |
|------|-----|------|
| 81 句 | 8 句 | 11 句 |

表 10.6: 多変数対訳句の評価例

| 評価 | 対訳句例  |
|----|---|
| ○  | 救急車で病院へ<br>それより品が落ちる<br>空中ぶらんこ  |
|    | to the hospital by ambulance<br>inferior to that in quality<br>the trapeze                        |
| △  | 昨夜の雨で川<br><br>谷間には全く日<br>英国では事情   |
|    | river has turned muddy from the<br>rain last<br>valley was completely<br>Matters are different in |
| ×  | 番号で注文し<br>が水たまり<br>の花道を   |
|    | filters by<br>sun dried up<br>the last days   |

表 10.7: 1変数対訳句の評価例

| 評価 | 対訳句例  |
|----|---|
| ○  | 彼を夕食に招待した<br>京都で一晩泊まった<br>家に由緒ある刀   |
|    | invited him to dinner<br>stayed overnight in Kyoto<br>a sword with a history in my<br>house |
| △  | 仕事がいやでいつも逃げ回っ<br>ている<br>彼にはおかしな面<br>5 フィート 2 インチ ある   |
|    | always shirking work<br>an eccentric side<br>five foot two                                  |
| ×  | を少しも<br>独立<br>わき  |
|    | I<br>boss<br>away   |

表 10.4 および表 10.5 より, 1 変数対訳句を用いた手法では, 多変数対訳句を用いた手法と比較して, 高い精度で対訳句を作成することが可能であるとわかる.

## 10.5 自動評価結果

入力文 10,000 文に対して得られた各手法の出力文 10,000 文について、自動評価を行った結果を表 10.8 に示す。

表 10.8: ベースライン, +多変数対訳句, +1 変数対訳句の自動評価結果 (10,000 文)

| モデル      | BLEU   | METEOR | RIBES  | WER    |
|----------|--------|--------|--------|--------|
| ベースライン   | 0.1960 | 0.4723 | 0.7737 | 0.6153 |
| +多変数対訳句  | 0.1945 | 0.4731 | 0.7739 | 0.6158 |
| +1 変数対訳句 | 0.1967 | 0.4764 | 0.7770 | 0.6152 |

## 10.6 人手対比較評価結果

出力文 10,000 文より 100 文を抽出し、人手対比較評価を行った。人手対比較評価は、+多変数対訳句および+1 変数対訳句の出力文とベースラインの出力文とを比較し、より入力文の意味が読み取れる (正確である) と考えられる文を○とした。+多変数対訳句の人手対比較評価結果を表 10.9 に、+1 変数対訳句の人手対比較評価結果を表 10.10 に示す。

表 10.9: ベースラインと+多変数対訳句の人手対比較評価結果 (100 文)

| ベースライン○ | +多変数対訳句○ | 差なし | 同一出力 |
|---------|----------|-----|------|
| 15      | 24       | 51  | 8    |

表 10.10: ベースラインと+1 変数対訳句の人手対比較評価結果 (100 文)

| ベースライン○ | +1 変数対訳句○ | 差なし | 同一出力 |
|---------|-----------|-----|------|
| 17      | 30        | 37  | 16   |

結果として、+1 変数対訳句の手法の翻訳精度について、自動評価ではほとんど差が見られなかったが、人手対比較評価ではベースライン○の 17 に対して+自動対訳句 (森本ら)○は 30 となり、精度が向上していることが確認できる。

また、+多変数対訳句の手法と+1 変数対訳句の手法との比較では、+1 変数対訳句の方が+多変数対訳句よりも翻訳精度の向上が大きいことが確認できる。

## 10.7 人手対比較評価例

人手対比較評価の結果，+1 変数対訳句の出力文が比較的正確であると評価された例を表 10.11 に，ベースラインの出力文が比較的正確であると評価された例を表 10.12 に，2 つの出力文は正確さにおいて差がないと評価された例を表 10.13 に示す。

表 10.11: +1 変数対訳句○

|          |  |
|----------|--|
| 入力文      | 幸い彼は軽いけがで済んだ。  |
| 参照文      | Luckily his injury was not serious .                       |
| ベースライン   | Fortunately he was badly injured .                         |
| +1 変数対訳句 | Fortunately he got off with a slight injury .              |
| 入力文      | 停電が東京都から埼玉県に及んだ。   |
| 参照文      | The blackout extended from Tokyo to Saitama Prefecture .   |
| ベースライン   | The power failure led to Saitama Prefecture .              |
| +1 変数対訳句 | The power failure extended from Tokyo to Saitama , Tokyo . |
| 入力文      | 容疑者は家宅捜索を受けた。  |
| 参照文      | The police searched the suspect's house .                  |
| ベースライン   | The suspect received a premises .                          |
| +1 変数対訳句 | The suspect had a house searched .                         |

表 10.12: ベースライン○

|          |  |
|----------|--|
| 入力文      | もちろんパーティーには伺います。                           |
| 参照文      | Of course I'll come to the party .         |
| ベースライン   | Of course I will come to the party .       |
| +1 変数対訳句 | I'll come to see you at the party .        |
| 入力文      | このにおいは目を刺激する。                              |
| 参照文      | This smell is irritating to the eyes .     |
| ベースライン   | This scent stimulates my eyes .            |
| +1 変数対訳句 | This scent excites eyes .                  |
| 入力文      | 彼の名前は戦死者名簿に載っている。                          |
| 参照文      | His name is on the roll of honor .         |
| ベースライン   | His name is on the list of war victims .   |
| +1 変数対訳句 | His name is listed in the visitors' book . |

表 10.13: 差なし

|          |  |
|----------|--|
| 入力文      | 彼の求婚を退けた。  |
| 参照文      | She rejected his suit .                          |
| ベースライン   | I rejected his suit .                            |
| +1 変数対訳句 | She rejected his suit .                          |
| 入力文      | 彼は昼間は働き夜学校へ行く。                                   |
| 参照文      | He works days and goes to school nights .        |
| ベースライン   | He goes to school at night at night .            |
| +1 変数対訳句 | He goes to school night at night .               |
| 入力文      | これで彼のいたずらをする能力もそがれるはずだ。                          |
| 参照文      | This should decrease his capacity for mischief . |
| ベースライン   | This will start the abilities of his mischief .  |
| +1 変数対訳句 | This will catch the abilities of his mischief .  |

## 10.8 1変数対訳句を用いた繰り返し処理

1変数対訳文パターンによる対訳句抽出手法では、1度作成した対訳句を用いて新たに対訳文パターンを作成することが可能であり、この処理を繰り返し行うことで多くの対訳句を作成することができる。したがって、本節ではこの繰り返し処理を用いて作成した対訳句を利用する手法についても実験を行った。以下、繰り返しを1回行う処理により作成した対訳句をループ1対訳句、繰り返しを2回行う処理により作成した対訳句をループ2対訳句と呼ぶ。また、本節における実験設定以下にを示す。

+ループ1対訳句 1変数対訳句を用いて1変数対訳文パターンを作成し、作成した1変数対訳文パターンより新たな対訳句を作成する「ループ処理」により作成した対訳句をNMTの学習データに追加する

+ループ2対訳句 ループ1対訳句を用いてループ処理を再び行い、作成した対訳句をNMTの学習データに追加する

### 10.8.1 1変数対訳文パターンと繰り返し処理による対訳句の作成結果

表10.1の対訳学習文を用いて、1変数対訳文パターンと繰り返し処理による対訳句の作成を行った。表10.14に作成した対訳句の内訳を示す。

表 10.14: 1変数対訳文パターンと繰り返し処理により作成した対訳句の内訳

| ループ1対訳句 | 日本語         | 英語        |
|---------|-------------|-----------|
| 句数      | 91,158 句    |           |
| 単語数     | 573,177 語   | 446,184 語 |
| 異なり単語数  | 21,342 語    | 17,577 語  |
| 平均句長    | 6.3 語       | 4.9 語     |
| ループ2対訳句 | 日本語         | 英語        |
| 句数      | 163,454 句   |           |
| 単語数     | 1,143,920 語 | 880,902 語 |
| 異なり単語数  | 25,015      | 21,318    |
| 平均句長    | 7.0 語       | 5.4 語     |

表10.14より、繰り返し処理を行うごとに多くの対訳句が作成され、また作成された対訳句の平均句長は長くなることが確認できる。

## 10.8.2 1変数対訳文パターンと繰り返し処理による対訳句の精度

作成したループ1対訳句とループ2対訳句について、6.1節と同様に全ての1変数対訳句から無作為に100対を抽出し、それぞれの対応の適切さを人手により評価した。ループ1対訳句の評価結果を表10.15に、ループ2対訳句の評価結果を表10.16に示す。また、ループ1対訳句の評価例を表10.17に、ループ2対訳句の評価例を表10.18に示す。

表 10.15: ループ1対訳句の評価結果

| ○   | △   | ×  |
|-----|-----|----|
| 84句 | 16句 | 0句 |

表 10.16: ループ2対訳句の評価結果

| ○   | △   | ×  |
|-----|-----|----|
| 82句 | 18句 | 0句 |

表 10.17: ループ1対訳句の評価例

| 評価 | 対訳句例  |
|----|---|
| ○  | 1月に2週間暖かい時期があった<br>had a two-week thaw in January<br>うれしい気持ちの半分も言い表せません<br>cannot tell half of my pleasure<br>クラスの首席である<br>top of her class |
| △  | 彼は飲み食いによくお金を使う<br>spends a lot of money on food and drink<br>我々はアユを川へ放流した<br>stocked the river with ayu<br>よい<br>fine fruit                 |

表 10.18: ループ 2 対訳句の評価例

| 評価 | 対訳句例         |   |
|----|--------------|---|
| ○  | テレビの音を消した    | turned off the sound of the television            |
|    | 商用でマレーシア     | Malaysia on business                              |
|    | 自分の責任を僕に負わせた | shuffled off his responsibility upon my shoulders |
| △  | 彼の木訥さが私の心をつ  | simplicity touched me                             |
|    | た            |   |
|    | 母の代わりに私が来た   | come in place of my mother                        |
|    | 果敢無い夢        | dream   |

表 10.15 および表 10.16 より、繰り返し処理により作成した対訳句の精度は、1 変数対訳句とあまり変わらないことが確認できる。

### 10.8.3 自動評価結果

入力文 10,000 文に対して得られた各手法の出力文 10,000 文について、自動評価を行った結果を表 10.19 に示す。

表 10.19: ベースライン, +ループ 1 対訳句, +ループ 2 対訳句の自動評価結果 (10,000 文)

| モデル        | NIST   | BLEU   | METEOR | TER    | WER    | RIBES  |
|------------|--------|--------|--------|--------|--------|--------|
| ベースライン     | 5.7271 | 0.1960 | 0.4723 | 0.5997 | 0.6153 | 0.7737 |
| +ループ 1 対訳句 | 5.6663 | 0.1928 | 0.4686 | 0.6056 | 0.6218 | 0.7722 |
| +ループ 2 対訳句 | 5.5840 | 0.1874 | 0.4652 | 0.6118 | 0.6278 | 0.7686 |

### 10.8.4 人手対比較評価結果

+ループ 1 対訳句および+ループ 2 対訳句の実験について、出力文 10,000 文より 100 文を抽出し、人手対比較評価を行った。人手対比較評価は、+ループ 1 対訳句および+ループ 2 対訳句の出力文とベースラインの出力文とを比較し、より入力文の意味が読み取れる (正確である) と考えられる文を○とした。+ループ 1 対訳句の人手対比較評価結果を表 10.20 に、+ループ 2 対訳句の人手対比較評価結果を表 10.21 に示す。

表 10.20: ベースラインと+ループ1 対訳句の人手対比較評価結果 (100 文)

| ベースライン○ | +ループ1 対訳句○ | 差なし | 同一出力 |
|---------|------------|-----|------|
| 26      | 19         | 40  | 15   |

表 10.21: ベースラインと+ループ2 対訳句の人手対比較評価結果 (100 文)

| ベースライン○ | +ループ2 対訳句○ | 差なし | 同一出力 |
|---------|------------|-----|------|
| 27      | 20         | 38  | 15   |

自動評価および人手対比較評価の結果として、+ループ1 対訳句および+ループ2 対訳句の手法の翻訳精度について、翻訳精度の向上は認められなかった。

#### 10.8.5 人手対比較評価例

ループ1 対訳句とベースラインとの対比較評価において、表 10.22 に+ループ1 対訳句の出力文が比較的正確であると評価された例を、表 10.23 にベースラインの出力文が比較的正確であると評価された例を、表 10.24 に2つの出力文は正確さにおいて差がないと評価された例を示す。

表 10.22: +ループ1 対訳句○

|           |  |
|-----------|--|
| 入力文       | サルが石でクルミの殻を砕いた。  |
| 参照文       | The monkey crushed the walnut shell with a stone .           |
| ベースライン    | The monkey shattered the shell of the walnuts with a stone . |
| +ループ1 対訳句 | The monkey broke off the shell of the walnuts with stone .   |

表 10.23: ベースライン○

|           |                               |
|-----------|-------------------------------|
| 入力文       | 雨季に入った。                       |
| 参照文       | The rainy season has set in . |
| ベースライン    | The rainy season has set in . |
| +ループ1 対訳句 | We went on the rainy season . |

表 10.24: 差なし

|           |  |
|-----------|--|
| 入力文       | 川は雪解け水で増水していた。                                   |
| 参照文       | The rivers were swollen with the melting snows . |
| ベースライン    | The river was swollen with melting water .       |
| +ループ1 対訳句 | The river was swollen with snow water .          |

ループ2 対訳句とベースラインとの対比較評価において、表 10.25 に+ループ2 対訳句の出力文が比較的正確であると評価された例を、表 10.26 にベースラインの出力文が比較的正確であると評価された例を、表 10.27 に2つの出力文は正確さにおいて差がないと評価された例を示す。

表 10.25: +ループ2 対訳句○

|           |   |
|-----------|---|
| 入力文       | 彼はこの仕事から手を引いた。                                |
| 参照文       | He withdrew from involvement with this work . |
| ベースライン    | He received his hands from this work .        |
| +ループ2 対訳句 | He quit this business .                       |

表 10.26: ベースライン○

|           |  |
|-----------|--|
| 入力文       | 全身状態が好転する。                                     |
| 参照文       | The general condition takes a favorable turn . |
| ベースライン    | The whole body takes a favorable turn .        |
| +ループ2 対訳句 | The whole body is changing .                   |

表 10.27: 差なし

|           |  |
|-----------|--|
| 入力文       | 彼女は壁に身をもたせた。                             |
| 参照文       | She leaned against the wall .            |
| ベースライン    | She scattered herself against the wall . |
| +ループ2 対訳句 | She hoisted herself up to the wall .     |

## 10.9 1変数対訳文パターンを用いた対訳句のまとめ

結果として、1変数対訳文パターンを用いた手法において、ベースラインと比較して翻訳精度の向上が認められた。また、従来の多変数対訳文パターンによる手法と比較して、向上の幅は大きいことも確認できた。これについて、追加した1変数対訳句は多変数対訳句と比較して正しい対応を多く含むために、翻訳精度の比較的大きな向上につながったと考えられる。ただし、より高い精度の対訳句を利用する1変数対訳文パターンを用いた手法における翻訳精度の向上が小幅となった理由として、対訳句の数が考えられる。1変数対訳句では作成可能な対訳句の数が限られているため、対訳句の数は今回の比較においても多変数対訳句より少なくなっている。このため、大幅な翻訳精度の向上には至らなかった可能性がある。

一方、高い精度かつ比較的多量の対訳句である、繰り返し処理により作成した対訳句において、翻訳精度の向上は認められなかった。この理由として、1変数対訳句や、多変数対訳句と比較して、繰り返し処理により作成した対訳句では対訳句が長いことが一因となっている可能性が考えられる。繰り返し処理を行う手法では、対訳句抽出に用いる対訳文パターン中の変数部が長くなり、これを用いて対訳句を抽出するために作成される対訳句が長くなる課題がある。このため、強調すべき語句の対応の範囲が広がり(文レベルに近づき)、したがってモデルが学習する語句の対応を強化するという対訳句追加の本来の効果が減少した可能性がある。したがって、翻訳精度の向上には、高い精度で、多量かつ、比較的短い対訳句を利用する必要があると推察される。

## 11 おわりに

ニューラル機械翻訳において、入力文と出力文における語句の対応情報である Attention の学習に誤りがある場合に、翻訳精度が低下するという問題が指摘されていた。そこで、本研究では、日英 NMT の翻訳精度を向上する目的で、NMT において対訳文に対訳文より作成した対訳句を追加する 2通りの手法を提案し、手法の翻訳精度向上への有効性を検証した。具体的な提案手法としては、まず、人手で作成した対訳句を追加する手法として、対訳文から人手で作成した対訳文パターンを用いて対訳句を作成し、NMT の学習データに追加する手法を提案した。また、自動で作成した対訳句を追加する手法として、自動的に作成した対訳文パターンを用いて対訳句を作成し、NMT の学習データを追加する手法を提案した。

結果として、人手で作成した対訳句を用いた手法の評価では、ベースラインの方が正確であると評価された文が出力文 100 文中 14 文であるのに対して、人手で作成した対訳句を追加する手法の方が正確であると評価された文が 44 文となった。また、自動で作成した対訳句を用いた手法では、ベースラインの方が正確であると評価された文が出力文 100 文中 10 文であるのに対して、自動で作成した対訳句を追加する手法の方が正確であると評価された文が 29 文となった。これより、ベースラインと比較して対訳句を追加する 2つの手法では翻訳精度が向上への有効性が確認された。本研究で用いた 2種類の対訳句はいずれも対訳文から作成され、対訳文外の語彙などの情報を有しないものである。したがって、対訳句を追加することにより、NMT において語句の対応をモデル化する Attention が強化されたため、翻訳精度が向上したと考えられる。

今後はより高精度かつ多量の対訳句を得る手法や、提案手法を他のニューラル機械翻訳システムと組み合わせた手法を検討し、更なる翻訳精度の向上を試みたい。

## 謝辞

最後に、三年間に渡り、本研究のご指導をいただきました鳥取大学持続性社会創生科学研究科自然言語処理研究室の村上仁一准教授、村田真樹教授に深く感謝すると共に、厚く御礼申し上げます。そして、日常の議論を通じて多くの知識や示唆を頂いた同研究室の皆様へ深謝いたします。また、参考にさせていただいた論文の著者の方々に対して、深く感謝申し上げます。

## 参考文献

- [1] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *In Proceedings of International Conference on Learning Representations*, 2015.
- [2] Maja Popović and Hermann Ney. Statistical machine translation with a small amount of bilingual training data. *5th LREC SALT MIL Workshop on Minority Languages*, pp. 25–29, 2006.
- [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. 言語処理学会第19回年次大会, pp. 556–559, 2013.
- [4] 池淵堅斗. 対訳文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.
- [5] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *In Proceedings of Conference on Empirical Methods in Natural Language Processing*, 2014.
- [6] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. *In Advances in Neural Information Processing Systems*, 2014.
- [7] Yaohua Tang, Fandong Meng, Zhengdong Lu, Hang Li, and Philip L. H. Yu. Neural machine translation with external phrase memory. *Computing Research Repository*, 2016.
- [8] Chong Wang, Yining Wang, Po-Sen Huang, Abdel rahman Mohamed, Dengyong Zhou, and Li Deng. Sequence modeling via segmentations. *ArXiv*, 2017.
- [9] Po-Sen Huang, Chong Wang, Dengyong Zhou, and Li Deng. Neural phrase-based machine translation. *Computing Research Repository*, 2017.
- [10] Jiangtao Feng, Lingpeng Kong, Po-Sen Huang, Chong Wang, Da Huang, Jiayuan Mao, Kan Qiao, and Dengyong Zhou. Neural phrase-to-phrase machine translation. *Computing Research Repository*, 2018.

- [11] Philipp Koehn, Marcello Federico, Brooke Cowan, Richard Zens, Chris Dyer, Richard Zens, Chris Dyer, Ondřej Bojar, Alexandra Constantin, and Evan Herbst. Moses: Open source toolkit for statistical machine translation. *Proceedings of the ACL 2007 Demo and Poster Sessions*, pp. 177–180, 2007.
- [12] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Computing Research Repository*, 2017.
- [13] Phi Xuan Nguyen and Shafiq R. Joty. Phrase-based attentions. *Computing Research Repository*, 2018.
- [14] Minh-Thang Luong, Hieu Pham, and Christopher D. Manning. Effective approaches to attention-based neural machine translation. *In Proceedings of Conference on Empirical Methods in Natural Language Processing*, pp. 1412–1421, 2015.
- [15] 池原悟, 阿部さつき, 徳久雅人, 村上仁一. 非線形な表現構造に着目した重文と複文の日英文型パターン化. 言語処理学会論文誌, Vol. 11, No. 3, pp. 69–95, 2004.
- [16] Franz Josef Och and Hermann Ney. A systematic comparison of various statistical alignment models. *Computational linguistics*, pp. 19–51, 2003.
- [17] 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.
- [18] Guillaume Klein, Yoon Kim, Yuntian Deng, Jean Senellart, and Alexander M. Rush. Opennmt: Open-source toolkit for neural machine translation, 2017.
- [19] 鳥バンク. <http://unicorn.ike.tottori-u.ac.jp/toribank/>.
- [20] 興梶玲架. パターンに基づく統計翻訳において変数部の確率の総和を使った対訳句の抽出. 2015年度鳥取大学卒業論文, 2015.
- [21] 英辞郎. <http://www.alc.co.jp/>.
- [22] 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.

- [23] 森本智喜, 村上仁一. 1 変数対訳文パターンを用いた対訳句の抽出調査. 言語処理学会第 25 回年次大会, pp. 328–331, 2019.