

# 対訳文から抽出した対訳句を利用した統計翻訳

池淵堅斗 村上仁一 坂田純 徳久雅人

鳥取大学大学院 工学研究科 情報エレクトロニクス専攻

{s102002, murakami, d112004, tokuhisa} @ ike.tottori-u.ac.jp

## 1 はじめに

近年、機械翻訳の分野で原言語から目的言語に翻訳する統計翻訳が注目されている。統計翻訳においては、対訳データの量が多ければ多いほど翻訳精度が高くなる。しかし、利用できる対訳データには限りがある。

この課題に対して、Maja Popovićらは、セルビア語英語間の翻訳において、対訳文と対訳句を組み合わせた対訳データを用いて統計翻訳を行った。その結果、自動評価において翻訳精度が向上した [1]。

また日野らは、日本語英語間において、同じように対訳文と対訳句を組み合わせた対訳データを用いた統計翻訳を行った。その結果、翻訳精度が向上した [2]。

しかし、Maja Popovićらや日野らの研究の実験環境では、対訳文に対して新たな対訳句を追加したものを学習に用いているため、単に対訳データのデータ量が増加したために翻訳精度が向上した可能性が考えられる。

そこで、本研究では、学習に用いる対訳文から対訳句を抽出することによって、対訳データのデータ量の増加が生じない実験環境における、対訳文と対訳句を組み合わせた対訳データを用いた統計翻訳の有効性の確認を行う。また、対訳句の抽出方法として、手動抽出と自動抽出の2通りの抽出方法を試し、対訳句の抽出精度やその対訳句を用いた場合の統計翻訳の翻訳精度の調査を行う。

## 2 本研究における統計翻訳の流れ

本研究では、翻訳モデル作成の際の学習データとして日英対訳句と日英対訳文を組み合わせた日英対訳データを用いる。図1に日英方向の統計翻訳の流れを示す。統計翻訳の流れの手順を以下に示す。

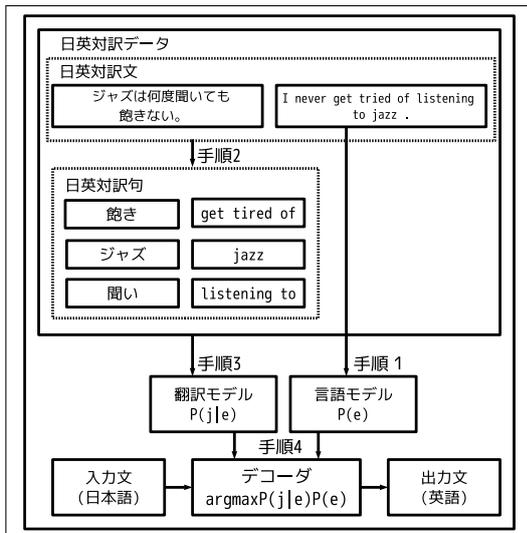


図1 日英方向の統計翻訳の流れ

手順1 英語文を学習データとして言語モデルを作成

する

手順2 日英対訳文と日英対訳文から抽出した日英対訳句を組み合わせて日英対訳データを作成する

手順3 手順2で作成した日英対訳データを学習データとして翻訳モデルを作成する

手順4 手順1と手順3で作成したモデルを用いて統計翻訳を行う

## 3 実験環境

### 3.1 前処理

統計翻訳の前処理として、日本語句と日本語文に対して、MeCab[3]を用いて形態素解析を行う。また、英語句と英語文に対して“tokenizer.sed [4]”を用いて分かち書きを行う。

### 3.2 日英対訳文

本研究では、日英対訳文として重文複文 [5]を用いる。本研究で用いる日英対訳文の内訳を表1に示す。日英対訳文の例を表2に示す。

表1 日英対訳文の内訳

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

表2 日英対訳文の例

あの男はあいさつもせずに帰ってしまった。  
That man left without so much as saying goodbye.

### 3.3 日英対訳句

本研究では、日英対訳句として前節で示した日英対訳文として用いた重文複文から抽出したものをを用いる [6]。日英対訳句は、鳥バンク [6]の手動で作成したパターンと原文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), *ADVP*(形容詞句) の日英対訳句を抽出し作成する。これを, 手動抽出した日英対訳句とする。また, 日英対訳句は基本的に句単位で抽出しているが, 単語単位や節単位の形で抽出されたものも含まれる。

本研究で用いる, 日英対訳句の内訳を表 3 に示す。日英対訳句の例を表 4 に示す。

表 3 手動抽出した日英対訳句の内訳

| 学習データ     | 日本語       | 英語        |
|-----------|-----------|-----------|
| 句数(句)     | 549,694   |           |
| 単語数(語)    | 1,058,949 | 1,031,265 |
| 異なり単語数(語) | 35,613    | 34,823    |
| 平均句長(語)   | 1.93      | 1.88      |

表 4 手動抽出した日英対訳句の例

| 日本語句 | 英語句            |
|------|----------------|
| 帰っ   | left           |
| あの   | That           |
| あの男  | That man       |
| 男    | man            |
| あいさつ | saying goodbye |

### 3.4 手動抽出した日英対訳句の精度評価

3.3 節の手動抽出した日英対訳句からランダムに 100 句抽出して人手評価を行った。評価基準を以下に示す。評価結果を表 5 に示す。評価例を表 6 に示す。

- 適切な対応をとる日英対訳句
- 部分的に適切な対応をとる日英対訳句
- × 不適切な対応をとる日英対訳句

表 5 手動抽出した日英対訳句の人手評価結果

|      |      |     | × |
|------|------|-----|---|
| 83 句 | 12 句 | 5 句 |   |
|      |      |     |   |

表 6 手動抽出した日英対訳句の評価例

|     | 日本語句 | 英語句            |
|-----|------|----------------|
| の例  | この分野 | this field     |
|     | 老人   | senior citizen |
|     | 紋章   | a coat of arms |
| の例  | 私    | My             |
|     | 両社   | companies      |
|     | 記憶し  | keeps          |
| ×の例 | びたっと | all            |
|     | 僕    | I              |
|     | 日本   | it             |

### 3.5 実験ツール

本研究ではデコーダとして“Moses[7]”を用いる。また, パラメータの最適化のために“mert-moses.pl”を用いる。

### 3.6 評価方法

本研究では, 翻訳結果の評価として自動評価と人手評価を行う。自動評価は自動評価法“BLEU[8]”,

“METEOR[9]”, “RIBES[10]”を用いる。人手評価は翻訳結果からランダムに 100 文抽出し, 対比較評価を行う。

## 4 手動抽出した対訳句を用いた翻訳実験

本章の実験では, 翻訳モデル作成の際の学習データとして, 日英対訳文(以降, 対訳文)と手動抽出した日英対訳句(以降, 対訳句)を組み合わせた日英対訳データ(以降, 対訳データ)を用いる。この手法を手動手法とする。翻訳モデル作成の際の学習データとして, 対訳文のみを用いる場合を baseline 手法として手動手法との翻訳精度の比較を行う。

### 4.1 自動評価結果

実験の翻訳結果に対する自動評価結果を表 7 に示す。

表 7 自動評価結果

| 自動評価手法    | baseline 手法 | 手動手法         |
|-----------|-------------|--------------|
| BLEU(%)   | 11.98       | <b>14.18</b> |
| METEOR(%) | 39.62       | <b>44.71</b> |
| RIBES(%)  | 66.54       | <b>69.30</b> |

### 4.2 人手評価結果

人手評価手法として出力結果からランダムに 100 文抽出して対比較評価を用いる。対比較評価結果の表記の説明を以下に示す。

- baseline 手法  
対比較評価において, 比較対象の翻訳結果の翻訳精度よりも baseline 手法がより優れていると判断した場合
- 手動手法  
対比較評価において, 比較対象の翻訳結果の翻訳精度よりも手動手法がより優れていると判断した場合
- 差なし  
対比較評価において, 比較する両実験の翻訳結果の翻訳精度に明確な差がないと判断した場合
- 一致  
対比較評価において, 比較する両実験の翻訳結果が完全に同一出力の場合

baseline 手法と手動手法の翻訳結果の対比較評価結果を表 8 に示す。

表 8 baseline 手法 VS 手動手法

| baseline 手法 | 手動手法 | 差なし  | 一致  |
|-------------|------|------|-----|
| 3 文         | 16 文 | 77 文 | 4 文 |

### 4.3 対比較評価例

baseline 手法と手動手法の翻訳の対比較評価例を表 9 に示す。

### 4.4 実験結果のまとめ

実験結果のまとめを以下に示す。

- 自動評価, 人手評価ともに baseline 手法よりも手動手法の精度がより優れている。

表 9 baseline 手法 VS 手動手法の対比較評価例

| baseline 手法 の例 |  |
|----------------|--|
| 入力文            | あんなやつを負かすのは何でもない。                          |
| 正解文            | I can beat him hands down .                |
| baseline       | on my mettle , I can beat him hands down . |
| 手動手法           | I can beat him .                           |
| 手動手法 の例        |  |
| 入力文            | 警官は私に止まれと合図した。                             |
| 正解文            | The policeman gave me a sign to stop .     |
| baseline       | The policeman to me to stop .              |
| 手動手法           | The policeman signaled to me to stop .     |

## 5 自動抽出した対訳句を用いた翻訳実験

4章の実験結果より、対訳データのデータ量が増加しない実験環境においても、手動手法が翻訳精度の向上に対して有効であることがわかった。しかし、4章の実験で用いた対訳句は対訳文から手動で抽出しているため、コストが大きい。そこで、本章では、統計翻訳の学習において対訳文から自動抽出した対訳句を用いる。これを自動手法と呼ぶ。そして、この自動手法を用いて、コストの削減を図った上で翻訳精度の向上を目指す。

### 5.1 対訳句の自動抽出 [11]

対訳句の自動抽出の手順を以下に示す。詳細は興相らの論文 [11] を参照のこと。

- 手順 1 対訳文と GIZA++ を利用して単語辞書を作成。  
 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辞書を作成。  
 手順 3 対訳文と単語に基づく文パターンを用いて、対訳句を作成。

自動抽出によって得られた対訳句の内訳を表 10 に示す。

| 学習データ      | 日本語    | 英語     |
|------------|--------|--------|
| 句数 (句)     | 18,192 |        |
| 単語数 (語)    | 56,885 | 43,638 |
| 異なり単語数 (語) | 9,003  | 7,770  |
| 平均句長 (語)   | 3.13   | 2.4    |

### 5.2 自動抽出した対訳句の精度評価

自動抽出した対訳句からランダムに 100 句抽出して人手評価を行った。評価結果を表 11 に示す。評価例を表 12 に示す。

| ×    |      |      |
|------|------|------|
| 29 句 | 25 句 | 46 句 |

表 10 より、自動抽出した対訳句数は、表 3 の手動抽出の場合と比べて約 3 % 程であり、少ないことがわかる。また表 11 より、自動抽出した対訳句の精度は、表 5 の

表 12 自動抽出した対訳句の評価例

|     | 日本語句     | 英語句                      |
|-----|----------|--------------------------|
| の例  | 5年間      | five years               |
|     | 難しかった    | hard                     |
|     | 守ってくれる印  | symbol against evil      |
| の例  | 遠く       | long                     |
|     | いかにも精力的  | aggressive               |
|     | クラス全員の成績 | the entire class average |
| ×の例 | は        | his                      |
|     | ある       | people                   |
|     | わたしの主義   | My                       |

手動抽出の場合と比べると明らかに低いことがわかる。

### 5.3 自動評価結果

実験の翻訳結果に対する自動評価結果を表 13 に示す。

表 13 自動評価結果

| 自動評価手法    | baseline 手法 | 手動手法         | 自動手法  |
|-----------|-------------|--------------|-------|
| BLEU(%)   | 11.98       | <b>14.18</b> | 12.54 |
| METEOR(%) | 39.62       | <b>44.71</b> | 42.55 |
| RIBES(%)  | 66.54       | <b>69.30</b> | 68.02 |

### 5.4 人手評価結果

人手評価手法として出力結果からランダムに 100 文抽出して対比較評価を用いる。baseline 手法と自動手法の翻訳の対比較評価結果を表 14、手動手法と自動手法の翻訳の対比較評価結果を表 15 に示す。

表 14 baseline 手法 VS 自動手法

| baseline 手法 | 自動手法        | 差なし  | 一致   |
|-------------|-------------|------|------|
| 7 文         | <b>17 文</b> | 64 文 | 12 文 |

表 15 手動手法 VS 自動手法

| 手動手法 | 自動手法 | 差なし  | 一致  |
|------|------|------|-----|
| 10 文 | 7 文  | 76 文 | 7 文 |

### 5.5 対比較評価例

baseline 手法と自動手法の翻訳の対比較評価例を表 16、手動手法と自動手法の翻訳の対比較評価例を表 17 に示す。

表 16 baseline 手法 VS 自動手法の対比較評価例

| baseline 手法 の例 |                                       |
|----------------|---------------------------------------|
| 入力文            | パンを買うお金をいくら下さい。                       |
| 正解文            | Give me something to buy bread with . |
| baseline       | Please put some money to buy bread .  |
| 自動手法           | Please some money to buy bread .      |
| 自動手法 の例        |                                       |
| 入力文            | どうもそう思えてしょうがない。                       |
| 正解文            | I am strongly inclined to think so .  |
| baseline       | I can not think so .                  |
| 自動手法           | I am strongly inclined to think so .  |

表 17 手動手法 VS 自動手法の対比較評価例

| 手動手法 の例 |   |
|---------|---|
| 入力文     | 方々探したが見つからなかった。                                 |
| 正解文     | I looked everywhere , but I could not find it . |
| 手動手法    | I looked everywhere , but I could not find it . |
| 自動手法    | I looked for people , but could not find it .   |
| 自動手法 の例 |   |
| 入力文     | そんな事をしては弁解の余地がない。                               |
| 正解文     | Such action admits of no excuse .               |
| 手動手法    | If you were to do such a thing .                |
| 自動手法    | There is no excuse to do such thing .           |

## 5.6 実験結果のまとめ

実験結果のまとめを以下に示す。

- 自動抽出した対訳句を用いた自動手法は baseline 手法よりも精度が向上している。
- 自動抽出の場合と手動抽出の場合を比較すると手動抽出した対訳句を用いる場合の方が精度が高い。

## 6 考察

### 6.1 精度向上の原因分析

本研究では、対訳文と、その対訳文から手動抽出した対訳句を組み合わせた対訳データを学習データとして用いることにより、翻訳精度が向上することがわかった。さらに、自動抽出した句を用いた場合でも同様に翻訳精度が向上することがわかった。

この原因は、対訳句を追加することにより、フレーズテーブルの翻訳確率値がより信頼性の高い値になっているためと考えられる。例えば、表 9 の手動手法 の例を見ると、手動手法で得られた出力文では、“合図し”が“signaled”に翻訳されているが、baseline 手法では、“合図し”がうまく翻訳されていない。ここで、手動手法と baseline 手法のフレーズテーブルの確認をしてみると、翻訳確率が baseline 手法で 0.125、手動手法で 0.3111 となっており、翻訳確率がより信頼性の高い値となっていることがわかる。

この結果より、対訳データのデータ量の増加が生じない実験環境において、対訳文と対訳句を組み合わせた対訳データを用いた統計翻訳には有効性があると考えられる。

### 6.2 自動抽出と手動抽出の比較

自動抽出と手動抽出の対訳句の比較を行うと、抽出精度の面では手動抽出の方が優れているが、コストの面では自動抽出の方が優れている。そのため、場合に応じては自動抽出を行う手法は有効であると考えられる。さらに、自動抽出はプログラムのパラメータを調節することなどの工夫をしていくことにより、より多くの対訳句の抽出や、より良い精度の対訳句の抽出が可能であると考えられる。

## 7 おわりに

先行研究では、Maja Popović らや日野らの研究の実験環境において、対訳文に対して新たな対訳句を追加したものを学習に用いているため、単に対訳データのデータ量が増加したために翻訳精度が向上した可能性があると考えられるという問題点があった。

そこで、本研究では、学習に用いる対訳文から対訳句を抽出することによって、対訳データのデータ量の増加が生じない実験環境における、対訳文と対訳句を組み合わせた対訳データを用いた統計翻訳の有効性の確認を行った。

そして、本研究結果より、対訳文と、その対訳文から手動抽出した対訳句を組み合わせた対訳データを学習データとして用いることにより、翻訳精度が向上することがわかった。さらに、自動抽出した対訳句を用いた場合でも同様に翻訳精度が向上することがわかった。

今後は、統計翻訳を行う際に用いる対訳句として、手動抽出の対訳句よりも、より有効的な抽出方法における対訳句を用いて翻訳精度の向上を目指したい。

## 参考文献

- [1] Popović Maja, and Ney Hermann “Statistical Machine Translation with a small amount of bilingual training data”, 5th LREC SALT MIL Workshop on Minority Languages, pp.25-29. 2006.
- [2] 日野聡子：“統計翻訳における対訳句コーパスの効果”，2013 年修論
- [3] MeCab <http://mecab.sourceforge.net/>
- [4] tokenizer.sed <http://www.cis.upenn.edu/treebank/tokenizer.sed>
- [5] 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.
- [6] 鳥バンク <http://unicorn.ike.tottori-u.ac.jp/toribank/>
- [7] Moses: “Open Source Toolkit for Statistical Machine Translation”, Proceedings of the ACL 2007 Demo and Poster Sessions, pp.177-180. 2007.
- [8] BLEU: “a Method for Automatic Evaluation of Machine Translation”, Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), pp.311-318. 2002.
- [9] Meteor: Lavie Alon, and Denkowski Michael “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.
- [10] Hideki Isozaki, “Automatic Evaluation of Translation Quality for Distant Language Pairs”, Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, pp.944-952. 2010.
- [11] 興梠玲架, 村上仁一, 徳久雅人, “パターンに基づく統計翻訳における、変数部の確率の総和を利用した対訳句の抽出”, 言語処理学会第 22 回年次大会, 2016.