

統計翻訳における文型パターン辞書の効果

吉田大蔵 村上仁一
鳥取大学大学院 工学研究科 情報エレクトロニクス専攻
{s062062, murakami} @ ike.tottori-u.ac.jp

1 はじめに

現在, 日英機械翻訳において統計翻訳の研究が盛んに行われている [1]. 統計翻訳の一般的な問題として, 対訳文から人手によって作成できるフレーズ対と, 機械的な学習によって作成されるフレーズ対に違いがある. これは, 機械的な学習によって対訳文の単語アライメントに間違った対応が取られ, その対応を元にしてフレーズ対が作成されることが原因として考えられる. 先行研究では, 人手によって作成されたフレーズ対を用いることによる翻訳精度の向上が確認されている [2]. また, 単語アライメントのアルゴリズムを改良することによる翻訳精度の向上が確認されている [3]. そこで, 本稿では, あらかじめ文の要素が人手によって変数化されている文型パターン辞書をコーパスに加える. この手法によって単語アライメントを改善し, 翻訳精度の向上を図る.

2 従来手法における問題点

日英対訳文と学習によって抽出されたフレーズ対の例を表1に示す.

表1 日英対訳文からフレーズ対を抽出した例

日英対訳文
ああいう人と付き合っ ては だめだ。 you shouldn 't associate with that kind of person .
抽出したフレーズ対
ああいう人と 付き合っ て associate with
ああいう人と 付き合っ て associate with that
ああいう人と 付き合っ て associate with that kind
ああいう人と 付き合っ て associate with that kind of
ああいう人と 付き合っ て associate with that kind of person

「ああいう人と付き合っ て」と「associate with that kind of person」が対応するのは適切であるが, それ以外と対応するのは不適切である. このように, 統計翻訳における翻訳モデルの学習では, 不適切なフレーズ対を抽出することが多い.

3 提案手法

本稿では, 日英統計翻訳で使用するコーパスに, 人手によって作成された文型パターン辞書を加える手法を提案する. 図1に, 文型パターン辞書の例を示す. また, 文型パターン辞書を加えることによる効果を 3.1 節に示す.

日英対訳文:

ああいう人と付き合っ ては だめだ。
you shouldn 't associate with that kind of person .

文型パターン対:

X1 は ああいう X2 と 付き合っ ては だめだ。
X1 shouldn 't associate with that kind of X2 .

図1 文型パターン辞書の例

3.1 単語アライメント取得時の効果

本稿では, 日英対訳文の単語アライメントの自動取得に GIZA++[4] を用いる. 提案手法では, あらかじめ人手によって対応が取られた変数が文型パターン辞書に大量に存在するので, 変数同士の単語アライメントが取られる可能性が高くなる. これにより, 変数以外の単語と変数の単語アライメントは取られ難くなり, 変数以外の単語同士の単語アライメントが取られる可能性が高くなる. したがって, 単語アライメントが改善され, 適切なフレーズ対が作成されることが考えられる. この結果, 翻訳精度の向上が見込める.

3.2 フレーズ対抽出時の処理

翻訳時には, 変数を含むフレーズ対は使用しないので, 表2に示す, 変数が含まれたフレーズ対は削除する.

表2 削除するフレーズ対の例

削除対象
ああいう X2 と 付き合っ て associate with that kind of X2

4 実験データ

本稿に用いるデータは, 統計翻訳の前処理として, 日本語文には chasen[5] を用いて形態素解析を行う. 英語文に対しては tokenizer.perl を用いて分かち書きを行う. 本稿では, 学習データ, テストデータ, Development データに文型パターン辞書として「鳥バンク」を用いる.

4.1 鳥バンク

「鳥バンク」の「日本語表現意味辞書(重文・複文編)」[6]には, 日本語の重文・複文とその対訳英文対を約12万文対, および, その文対から作成された「意味類型パターン(22.7万件)」が収録されている. 簡略化した例を図2に示す. 図2では, 日英対訳文, それに対応する文型パターン対が記述されている. また, この文型パター

ン対には、単語レベル、句レベル、節レベルの3つのレベルがある。文型パターン対は日英対訳文から生成されており、それぞれのレベルに応じて対応可能な要素が変数化されている。また文型パターン対の変数の対応は、変数末尾の値によって取られている。提案手法では、日英対訳文とすべてのレベルの文型パターン対を学習データに用いる。

日英対訳文:
あの建物はああ見えても新しい。
That building is still new despite appearance.

単語レベルパターン:
GEN1 N2 はああ見えても AJ3。
AJ1 N2 be still AJ3 despite appearance.

句レベルパターン:
NP1 はああ見えても AJ2。
NP1 be still AJ2 despite appearance.

節レベルパターン:
NP1 はああ見えても CL2。
CL2 despite appearance.

図2 日本語表現意味辞書における文型パターン対の例

4.2 変数名の変更

本稿では、文型パターン対の変数名の情報は利用しない。したがって、「鳥バンク」における文型パターン対の変数を表3に示す形式で統一する。

日英対訳文:
あの建物はああ見えても新しい。
That building is still new despite appearance.

単語レベルパターン:
X1 X2 はああ見えても X3
X1 X2 be still X3 despite appearance.

句レベルパターン:
X1 はああ見えても X2。
X1 be still X2 despite appearance.

節レベルパターン:
X1 はああ見えても X2。
X2 despite appearance.

図3 文型パターン対の変数名の変換

4.3 学習データ

学習データには、文型パターン辞書から抽出した日英対訳文と文型パターン対の合計 296,815 文対を用いる。内訳として日英対訳文は 110,787 文対、文型パターン対は 186,028 文対である。なお、言語モデルの生成には、日英対訳文中の英文のみを用いる。

4.4 テストデータ

テストデータには、文型パターン辞書から抽出した日英対訳文 10,000 文対を用いる。このテストデータは、学習データに対してオープンデータである。また、テストデータに対応する文型パターン対は学習データに使用しない。

4.5 Development データ

Development データには、文型パターン辞書から抽出した日英対訳文 1,000 文対を用いる。この Development データは、学習データとテストデータに対してオープンデータである。また、Development データに対応する文型パターン対は学習データに使用しない。

5 実験環境

5.1 翻訳モデルの学習

本稿では、フレーズテーブルの作成に、Moses 付属の train-model.perl を用いる。ただし、変数を含むフレーズ対は削除する。

5.2 言語モデルの学習

言語モデルには、 N -gram モデルを用いる。 N -gram モデルの学習には、“SRILM” [7] を用いる。本稿では、5-gram を用い、スムージングには kndiscount を用いる。

5.3 デコーダのパラメータ

本稿では、デコーダとして、“Moses” [8] を用いる。また翻訳実験ではパラメータの最適化 [9] を行う。

6 翻訳実験

本稿では、ベースラインと提案手法の2つの実験を行う。

6.1 ベースライン

ベースラインでは学習データとして、日英対訳文 110,787 文対のみを用いる。

6.2 提案手法

提案手法では学習データとして、日英対訳文と文型パターン対の合計 296,815 文対を用いる。

7 評価方法

7.1 自動評価

本稿では、出力文の自動評価として、“BLEU” [10] , “NIST” [11] , “METEOR” [12] を用いる。

7.2 人手評価

本稿では、出力文の人手評価として、ベースラインと提案手法の翻訳出力の対比較評価を行う。ベースラインと提案手法の出力文からそれぞれ 100 文づつランダムに抽出し比較を行う。判断基準を以下に示す。

提案手法	提案手法の出力がベースラインの出力より優れている場合
提案手法 ×	提案手法の出力がベースラインの出力より劣っている場合
差なし	提案手法の出力とベースラインの出力の表現に差がない場合
同一出力	提案手法の出力とベースラインの出力が同一の場合

8 実験結果

8.1 自動評価

自動評価の結果を表 3 に示す。

表 3 自動評価結果

	BLEU	NIST	METEOR
ベースライン	0.1282	4.7429	0.4313
提案手法	0.1343	4.7217	0.4266

実験を行った結果, BLEU スコアは向上したが, NIST スコアおよび METEOR スコアは若干低下した。

8.2 人手評価 (対比較評価)

対比較評価の結果を表 4 に示す。

表 4 対比較評価結果

提案手法	提案手法 ×	差なし	同一出力
23	22	44	11

実験を行った結果, 提案手法 と提案手法 × の件数がほぼ同じであった。

8.2.1 提案手法 の例

提案手法 の例を表 5 に示す。

表 5 提案手法 の例

入力文	日本 から 北 へ 行け ば 行く ほど 寒く なる 。
正解文	The farther north you go in japan , the colder it gets .
ベースライン	As you go if you go north from japan , the colder it becomes .
提案手法	As you go farther north from japan , the colder it becomes .

ベースラインの出力では, 「As you go if you go」と同じ意味を持った節が連続して並んでおり, 不適切である。したがって, 提案手法 と判断した。

8.2.2 提案手法 × の例

提案手法 × の例を表 6 に示す。

表 6 提案手法 × の例

入力文	井上 さん は わたし の 理想 に かなっ た 人 です 。
正解文	Inoue is my ideal type .
ベースライン	Mr. inoue is ideal of a man .
提案手法	Mr. inoue beau ideal of me .

提案手法の出力では述語動詞が存在していない。また, ベースラインの出力の方が入力文の意味に近いと考え, 提案手法 × と判断した。

9 考察

提案手法 の例 (表 5) において, ベースラインと提案手法で翻訳に使われたフレーズ対を表 7 に示す。

表 7 提案手法 の例において翻訳に使用されたフレーズ対

ベースライン
日本 から from japan
北 へ 行け you go north
ば if
行く ほど as you go
寒く なる 。 , the colder it becomes .
提案手法
日本 から from japan
北へ 行け ば farther north
行く ほど as you go
寒く なる 。 , the colder it becomes .

表 7 より, 提案手法の方がベースラインより, 出力に適したフレーズ対で翻訳されている。この例では, 提案手法で, 出力に適した翻訳モデルが作成されたと考えられる。しかし, 今回の実験において, BLEU 以外の評価ではあまり効果が得られなかった。つまり, 全体の出力文の翻訳精度は, ベースラインと提案手法で差がないと考える。考えられる要因を以下に示す。

文型パターン辞書の表記

提案手法とベースラインを比べて, 単語アライメントが良くならなかった場合があると考えられる。その原因としてパターンの表記の問題がある。本稿で用いた文型パターン対では, 表 8 のように be 動詞が原形で表記されてある。したがって, 日英対訳文では取られなかった日文の要素と「be」の対応が取られ, 正しい単語アライメントの作成につながらなかったと考えられる。

表 8 be 動詞を含む文型パターン対の例

X1 X2 は ああ 見えて も X3
X1 X2 be still X3 despite appearance.

10 階層的句に基づく統計翻訳

本稿では, 追加実験として, 階層的句に基づく統計翻訳 [13] を用いて提案手法の調査を行う。翻訳モデルの学習には, Moses 付属の train-model-new.perl を用いる。その他の実験データと実験環境は, 第 4 章および第 5 章と同じである。

10.1 追加実験の結果

追加実験において, 自動評価の結果を表 9, 対比較評価の結果を表 10 に示す。

表 9 自動評価結果

	BLEU	NIST	METEOR
ベースライン	0.1306	4.6976	0.4262
提案手法	0.1367	4.7500	0.4343

表 10 対比較評価結果

提案手法	提案手法 ×	差なし	同一出力
17	14	66	3

表 9 より、提案手法は、ベースラインに対して、すべての自動評価のスコアが向上している。表 10 より、対比較評価においては、若干、提案手法 が提案手法 × より多い。

10.2 追加実験の考察

追加実験において、若干ではあるが、提案手法の評価が向上した。したがって、提案手法は、階層的句に基づく統計翻訳に効果があると考えている。また、第 6 章で行った実験の評価結果に比べ、追加実験の評価では、良い結果が得られた。したがって、提案手法は、階層的句に基づく統計翻訳において、句に基づく統計翻訳より有効である可能性がある。考えられる要因を以下に示す。

ルール作成に対する単語アライメント改善の効果
 階層的句に基づく翻訳では、翻訳モデルの学習において、非終端記号を含むルールを作成する。特に、文型パターン対からルールを作成する場合は、変数の対応が非終端記号に置き換えられて抽出される。したがって、より良いルールが作成された可能性がある。

10.3 句に基づく統計翻訳と階層的句に基づく統計翻訳の比較

句に基づく統計翻訳と、階層的句に基づく統計翻訳において、提案手法を用いた場合の対比較評価を行った。評価基準は以下に示し、結果を表 11 に示す。

フレーズ	句に基づく統計翻訳の出力が階層的句に基づく統計翻訳の出力より優れている場合
階層的句	階層的句に基づく統計翻訳が句に基づく統計翻訳の出力より優れている場合
差なし	句に基づく統計翻訳と階層的句に基づく統計翻訳の出力の表現に差がない場合
同一出力	句に基づく統計翻訳と階層的句に基づく統計翻訳の出力が同一の場合

表 11 対比較評価結果

フレーズ	階層的句	差なし	同一出力
11	15	61	13

階層的句に基づく統計翻訳の方が、句に基づく統計翻訳より良い結果となった。

11 おわりに

本稿では、統計翻訳のコーパスに文型パターン辞書を追加する手法を提案した。実験を行った結果、BLEU スコアは上昇したが、それ以外の評価では、差がなかった。しかし、階層的句に基づく翻訳に提案手法を用いた結果、全体的のスコアが僅かに向上した。ゆえに、提案手法は、階層的句に基づく統計翻訳において、句に基づく統計翻訳より有効であると考えている。

12 謝辞

本研究を行うにあたって、様々な助言、御指導をいただきました須藤克仁さんをはじめ、NTT コミュニケーション科学基礎研究所創言 G の皆様に深く御礼申し上げます。

参考文献

- [1] Richard Zens, Franz Josef Och, Hermann Ney “Phrase-based Statistical Machine Translation”, KI 2002, pp35-56, 2002.
- [2] 東江恵介, 村上仁一, 徳久雅人, 池原悟, “日英統計翻訳における英辞郎の効果”, 言語処理学会 第 16 回年次大会, pp641-644, 2010.
- [3] Kuzman Ganchev, João V. Graca, and Ben Taskar, “Better alignments = Better translations?”, In Proceedings of the 46th Annual Meeting of ACL, pp986-993, 2008.
- [4] GIZA++, <http://www.fjoch.com/GIZA++>
- [5] chasen, 日本語形態素解析器, <http://chasen-legacy.sourceforge.jp/>
- [6] 鳥バンク, 日本語表現意味辞書 - 重文複文編 -, <http://unicorn.ike.tottori-u.ac.jp/toribank>
- [7] SRILM, The SRI Language Modeling Toolkit, <http://www-speech.sri.com/projects/srilm/>
- [8] Philipp Koehn, Marcello Federico, Brooke Cowan, Richard Zens, Chris Dyer, Ondej Bojar, Alexandra Constantin, Evan Herbst, “Moses: Open Source Toolkit for Statistical Machine Translation”, Proceedings of the ACL 2007 Demo and Poster Sessions, pp177-180, 2007.
- [9] Franz Josef Och, “Minimum Error Rate Training in Statistical Machine Translation”, Association for Computational Linguistics, pp160-167, 2003.
- [10] Kishore Papineni, Salim Roukos, Todd Ward, Wei-Jing Zhu, “BLEU: a Method for Automatic Evaluation of Machine Translation”, Association for Computational Linguistics, pp311-318, 2002.
- [11] NIST Open MT scoring, <http://www.nist.gov/speech/tests/mt/>
- [12] METEOR Automatic Machine Translation Evaluation System, <http://www2.cs.smu.edu/~alavie/METEOR/>
- [13] David Chiang, “Hierarchical phrase-based translation”, Computational Linguistics, pp201-228, 2007.