

概要

近年，日英機械翻訳方式として，文型パターンを用いて翻訳を行うパターン翻訳が注目されている．パターン翻訳は，日本語の入力文に合った日本語パターンを文型パターンデータベースから検索し，検索した日本語パターンに対応する英語パターンを用いて出力英文を作成する翻訳方式である．パターン翻訳を行う際，大量の文型パターンを用意する必要がある．過去の研究では，パターン作成における変数化の問題などがあり，意味的に独立したパターンを大量に作成する事が困難であった．これに対して参照文献 [1] で提案した方法により，大量のパターンを作る事が可能になった．現在，日本語の重複文 12 万文に対してパターンが作成されている．

そこで本研究では，入力文対象を重複文とし，パターン翻訳を用いた翻訳精度の検証を行う．

具体的な検証方法として，まず文型パターンパーサを利用して候補の日本語パターンを抽出する．次に人手で最適な日本語パターンを候補から選択し，対応する英語パターンの変数へ適切な英単語を人手で適合させ，出力英文を作成する．最後に得られた英文に対し，人手と機械で評価する．

検証の結果，入力文 100 文中 48 文が 1 つ以上の日本語パターンの候補を抽出出来た．次に候補を抽出出来た 48 文に対して翻訳精度を調査した所，人手による評価では A と判定された文が 32%，B と判定された文が 10%，C と判定された文が 15%，D と判定された文が 43%であった．

よって，本研究では 48 文中，42%の品質の高い英文が作成出来た事を確認した．

目次

1	はじめに	1
2	パターンを利用した翻訳実験	2
2.1	利用するデータベースとパターンとプログラム	2
2.1.1	日英対訳データベース	2
2.1.2	評価用試験文	3
2.1.3	文型パターンデータベース	4
2.1.4	文型パターンパーサ [3]	5
3	実験方法	7
3.1	重複文の翻訳精度実験	7
3.2	評価方法	9
3.2.1	人手による評価	9
3.2.2	機械による評価	10
4	実験結果	11
4.1	評価 A の例	12
4.2	評価 B の例	13
4.3	評価 C の例	14
4.4	評価 D の例	15
4.5	翻訳実験の結果	16
5	考察	17
5.1	実験結果の考察	17
5.2	パターンの問題	18
5.2.1	被覆率の問題	18
5.2.2	前置詞の問題	19
5.3	パターン適合率	20
5.4	句レベル，節レベルの利用	21
6	おわりに	22

表 目 次

1	実験結果 (人手による評価)	16
2	実験結果 (機械による評価)	16

1 はじめに

日英機械翻訳において、トランスファー方式、用例翻訳、確率翻訳などが提案されている。これらの翻訳方式は単文に対しては有効であるが、重複文に対しては、翻訳精度が低下する傾向がある。

ところで、提案されている翻訳方式の1つに、パターン翻訳がある。パターン翻訳は、日英対訳パターンを用いて翻訳を行う翻訳方式である。しかし、パターン作成における変数化の問題などがあり、意味的に独立したパターンを大量に作成する事が困難であった。これに対して参照文献 [1] で提案した方法により、大量のパターンを作る事が可能になった。現在、日本語の重複文 12 万文に対してパターンが作成されている。

そこで本研究では、12 万パターンを使用して重複文の日英翻訳を試みる。まず、文型パターンパーサを利用して候補の日本語パターンを抽出する。次に人手で最適な日本語パターンを候補から選択し、対応する英語パターンの変数へ適切な英単語を人手で適合させ、出力英文を作成する。この様に得られた英文に対し、人手と機械で評価する。

本論文の構成は以下の通りである。2 章では翻訳精度実験に使用するデータベース、パターン、プログラムについて説明する。3 章では実験方法の手順を例を使用して説明し、さらに評価方法について述べる。4 章では評価 A から評価 D の例と実験結果を示す。5 章では実験の考察とパターンの問題点を述べる。6 章では結論と今後の課題を述べる。

2 パターンを利用した翻訳実験

2.1 利用するデータベースとパターンとプログラム

実験に利用する日英対訳データベース，評価用試験文，文型パターンデータベース，プログラムを示す．

2.1.1 日英対訳データベース

日英機械翻訳の研究は以前より行われている．しかし，この研究において必須と考えられている大量の対訳データベースを一般で入手するのは困難であった．しかし，最近では電子辞書が出版され，これらから対訳を抽出出来るようになった．本研究では，電子辞書の対訳から重複文 12 万文を収集した日英対訳データベース [2] を使用する．

2.1.2 評価用試験文

評価用試験文は、日英対訳データベースよりランダムに選択した100文(平均単語数:12)を使用する。例文を以下に示す。

- 試験文例 = 彼にはその任務を果たせるだけの能力がなかった。
- 模範訳 = He was not equal to the task.

2.1.3 文型パターンデータベース

文型パターンは、日英言語の表現対において、線形要素を変数記号に書き換えたパターンである [1]。文型パターンは日英対訳データベースから作られている。本研究は文型パターン辞書に収録された単語レベルの文型パターン (12 万対) を使用する。この辞書には、単語レベルの他、句レベル、節レベルが収録されており、その中の単語レベルは表現に含まれる名詞、動詞、形容詞、副詞などの自立語の線形な要素を変数化している。文型パターンの例を以下に示す。

- 原文 (日)=母は赤ん坊をあやして笑わせた
- 原文 (英)=Mother played with the baby and got him to smile.
- 日本語パターン = $N1$ は/ $N2$ を/ $V3$ て/笑わせた。
- 英語パターン = $N1$ $V3$.past $N2$ and got $N2$.pron.obj to smile.

N は名詞、 V は動詞を表している。英語パターンは日本語パターンを元に、原文 (英)を変数化して作成されている。

2.1.4 文型パターンパーサ [3]

文型パターンパーサは，入力文 (日本語の形態素解析結果) と 12 万の日本語パターンを ATN を用いて照合し，入力文に適合する日本語パターンとパターン適合率を出力するプログラムである [3] .

パターン適合率

パターン適合率とは，入力文と適合したパターンの文字単位での一致率である．一般的にパターン適合率が高い日本語パターンを選択した方が，品質の高い出力英文を作れる可能性が高い．例を以下に示す．

- 入力文 = 私は子どもの将来を思うと切ない。

(a) パターン適合率 100%の例

- パターン適合率 100%の日本語パターン = $N1$ は/ $N2$ の/ $TIME3$ を/ $V4$ と/切ない。

(b) パターン適合率 56%以外の例

- パターン適合率 56%の日本語パターン = $N1$ は/ $V2$ と/ $AJ3$ 。

3 実験方法

入力文 100 文に対して，日英対訳パターン (12 万対) を使用した翻訳精度実験を行う．
翻訳方法と評価方法を示す．

3.1 重複文の翻訳精度実験

本研究で用いた翻訳の手順を例を使用して以下に示す．

1. 入力文と模範訳

入力文と模範訳を用意する．以下に例を示す

- 入力文 = 道路を横断するときは車に注意しなさい。
- 模範訳 = Watch out for the traffic when you cross the street.

2. 文型照合プログラムによる日本語パターン候補の抽出

文型照合プログラムに入力文を代入し，日本語パターンの候補を抽出する．この
時複数の候補が出力される．以下に例を示す．

(a) 候補の日本語パターン 1

<N1 は>/N2 を/V3/ときは/N4 に/V5.meirei。

N2:‘道路’，V3:‘横断する’，N4:‘車’，

V5:‘注意する’

(b) 候補の日本語パターン 2

N1 を/V2/ときは/注意しなさい。

N1:‘道路’，V2:‘横断する’

(c) 候補の日本語パターン 3

N1 を/V2 て/V3.meirei。

N1:‘道路’，V2:‘横断する’，V3:‘注意する’

3. 人手による最適パターンの選択

抽出した日本語パターンの候補から人手で最適な日本語パターンを選択する．以下に最適な日本語パターンを示す．

(a) 最適な日本語パターン

<N1 は>/N2 を/V3/ときは/N4 に/V5.meirei。

N2:‘道路’，V3:‘横断する’，N4:‘車’，

V5:‘注意する’

4. 変数における辞書引き

(a) のパターンの変数部分に対応する英単語を人手で検索する．英単語を検索する際，研究社新英和・和英中辞典を英単語辞書として使用する．該当する英単語が複数あった場合，人手で最適な英単語を選択する．以下に代入する英単語を示す．

- 代入する英単語: V5 = pay attention , N4 = the traffic , V3 = cross , N2 = the street

5. 英語パターンを使った出力英文の作成

英語パターンの変数部分に，英単語を代入し，出力英文を作成する．以下に英語パターンと出力英文を示す．

- 英語パターン = V5 to N4 when (N1 or you) V3 N2.
- 出力文 = Pay attention to the traffic when you cross the street.

6. 出力英文の評価

出力英文の評価を行い(3.2節参照)，入力文に対する翻訳精度を調査する．

3.2 評価方法

評価方法は，人手による評価と機械による評価で行う．

3.2.1 人手による評価

出力英文を人出によって評価する．日英機械翻訳において，翻訳精度を検証する評価方法は様々あるが，本研究では文献 [4] を参考にし，評価基準を以下の4段階とする．

- A...意味的に正しくそのまま英文として使用出来る
- B...重要でない情報が欠落しているか，文法的に正しくない所があるが簡単に理解出来る
- C...失敗だが，努力すれば理解出来る
- D...重要な情報が間違って訳されている

3.2.2 機械による評価

機械による評価は最近盛んになってきている．本研究では以下の2つの評価方法で行う．

a) BLEU による評価

BLEU[5] は米国 IBM が提案した方法である．基本的には出力英文と模範訳を 4-gram で評価する．BLEU によるスコアの範囲は 0 から 1 である．本研究では，BLEU スコアを出すために NIST MT evaluation kit version 9 を使用する [6]．BLEU スコアを出す際，一般的に模範訳を複数用意するが，本研究では模範訳を 1 文とする．

b) NIST による評価

NIST[6] は米国 NIST が提案した方法である．基本的には出力英文とそれに対する模範訳を 5-gram で評価するプログラムである．NIST スコアを出すためのプログラムは BLEU スコアを出すプログラムと同一である．また，BLEU と同様に模範訳を 1 文とする．

4 実験結果

入力文 100 文 (平均単語数:12) と文型パターンデータベース (12 万パターン) を照合した所, 48 文が日本語パターンの候補を 1 つ以上抽出出来た。また, 48 文の平均候補抽出数は 16 パターンであった。入力文は文型パターンデータベースには含まれていない。したがって, 本実験はオープンテストとなる。

人手による評価は評価者 2 人によって行い, 平均値を実験結果とした。4.1 節から 4.4 節で A から D の各評価の例を, 4.5 節で翻訳実験の結果をそれぞれ示す。

4.1 評価 A の例

入力文と模範訳を以下に示す。

- 入力文 = 警官が来て騒ぎを鎮めた。
- 模範訳 = The policemen came and got the things under control.

入力文を文型パターンパーサに代入し、日本語パターン候補を抽出する。次に、得られた候補から最適な日本語パターンを手で選択する。得られた最適な日本語パターンを以下に示す。

- 最適な日本語パターン = #1[REN2]/N3 が/V4 て/N5 を/V6.kako。

最適な日本語パターンを選択後、日本語パターンの変数部分に対応する英単語を辞書を使って検索する。代入する英単語を以下に示す。

- 代入英単語：N3 = the policemen , V4 = came , V6 = suppressed , N5 = the disturbance

最後に、英語パターンに英単語を代入し、出力英文を作成する。英語パターンと出力英文を以下に示す。

- 英語パターン = #1[REN2] N3 V4.past and V6.past N5.
- 出力英文 = The policemen came and suppressed the disturbance.

4.2 評価 B の例

4.1 節と同様の手順で出力英文を作成する．入力文，模範訳，最適な日本語パターン，代入英単語，英語パターン，出力英文をそれぞれ以下に示す．

- 入力文 = 他人の成功をうらやむのは一般の通弊である。
- 模範訳 = People are too apt to envy the success of others.
- 最適な日本語パターン = #1[N2]/V3/のは/#4[REN5]/N6.da。
- 代入英単語 : N6 = a common evil , V3 = envy , N2 = others
- 英語パターン = It is #4[AJ5] N6 to V3 #1[N2].
- 出力英文 = It is a common evil to envy others.

4.3 評価 C の例

4.1 節と同様の手順で出力英文を作成する．入力文，模範訳，最適な日本語パターン，代入英単語，英語パターン，出力英文をそれぞれ以下に示す．

- 入力文 = 枕に顔を押し当てて泣いた。
- 模範訳 = He sobbed into the pillow.
- 最適な日本語パターン = <N1 は>/N2 に/N3 を/
#4[*NUM5*]/V6 て/<N1 は>/#7[*NUM8*]/V9.kako.
- 代入英単語 : V9 = cried , V6 = pressed , N3 = the face , N2 = pillow
- 英語パターン = (N1orI) V9.past #7[*of NUM8*] when (N1orI).pron V6.past #4[*NUM5*]
N3 to (N1orI).pron.poss N2.
- 出力英文 = I cried when I pressed the face to my pillow.

4.4 評価 D の例

4.1 節と同様の手順で出力英文を作成する．入力文，模範訳，最適な日本語パターン，英語パターン，代入英単語，出力英文をそれぞれ以下に示す．

- 入力文 = 今晩は家にいると思います。
- 模範訳 = We will very likely stay home this evening.
- 最適な日本語パターン = <N1 は>/N2 は/#3[TIME4]/V5 と/V6.teinei.
- 代入英単語 : V6 = expect , N2 = this evening ,
V5 = stay
- 英語パターン = (N1orI) V6 that N2 will V5 #3[TIME4].
- 出力英文 = I expect that this evening will stay.

4.5 翻訳実験の結果

入力文 100 文に対して，48 文がパターン候補を選択出来た．翻訳実験の結果を表 1，表 2 にまとめる．表中の () 内は文数を表している．

表 1: 実験結果 (人手による評価)

調査数 (文)	A	B	C	D
48	32% (16)	10% (5)	15% (7)	43% (21)

表 2: 実験結果 (機械による評価)

調査数 (文)	BLEU	NIST
48	0.19	3.3

表 1 より，48 文中 A と判定された文は 32%，B と判定された文は 10%，C と判定された文は 15%，D と判定された文は 43%であった．2 人の評価は，C 判定の数に若干の差が出た．本研究では評価 A と評価 B を品質の高い英文としている．よって，本研究では品質の高い英文が 42%(A+B) 作成出来る事が示された．また，表 2 より，BLEU スコアは 0.19，NIST スコアは 3.3 であった．

5 考察

5.1 実験結果の考察

品質の高い英文は48文中の42%作成出来たが、全体(100文)では21%となり、BLEU、NISTスコアの結果も併せて考えると、まだ翻訳精度は低い。翻訳精度を向上するためには、以下の問題を解決する必要があると考えられる。

パターンの問題

1. 被覆率の問題
2. 前置詞の問題

5.2 パターンの問題

品質の高い出力英文を出せない原因として、日英対訳パターンの作り方に問題があることが分かった。問題点を以下に示す。

5.2.1 被覆率の問題

翻訳精度を下げる原因の1つに被覆率の問題がある。被覆率とは、入力文に対してパターン候補を抽出出来た文数を表す。本研究では、入力文100文に対して、文型パターンパーサを使ってパターン候補を抽出出来なかった文が52文あった。被覆率を下げる原因の1つとして0型代名詞の問題がある。以下に0型代名詞を持つ文の例を示す。

0型代名詞を持つ文

(a) 口車に乗せられて承諾した。

(a)の文に関して文型パターンデータベースを調べた所、意味的に一致している日本語パターンがあるにもかかわらず、文型パターンパーサで抽出されていない事が分かった。意味的に一致している日本語パターンと日本語パターンの原文(日)以下に示す。

意味的に一致している日本語パターン

(1) $M1$ は/口車に/乗せられて/承諾した。

意味的に一致している日本語パターンの原文(日)

(b) 私は口車に乗せられて承諾した。

上記の(a)と(b)の文を比べると、意味的には同じである。しかし、変数 $M1$ の部分が必須格であるため(1)の日本語パターンが文型パターンパーサで抽出出来ない。解決策として、 $M1$ を任意格に変更を行えば(1)の日本語パターンに対応出来る。変更後の日本語パターンを以下に示す。

変更後の日本語パターン

(2) < $M1$ は>/口車に/乗せられて/承諾した。

5.2.2 前置詞の問題

英語パターンに前置詞が明示してある場合の問題点を以下に示す。

(a) 前置詞が名詞に依存する場合

入力文，模範訳，英語パターン及び出力文を以下に示す。

- 入力文 = 村外れに古びた社がある。
- 模範訳 = There is an old Shinto shrine at the outskirts of the village.
- 英語パターン = There is #2[AJ3] N4 in N1.
- 出力文 = There is a Shinto shrine in the outskirts of the village.

上記の例では，出力英文に必ず‘in’が出る。しかし，N1が場所を表す名詞の場合，前置詞は名詞によって決まる。したがって，この例ではatが適切となる。よって，英語パターンに前置詞があると出力英文の品質が低下する可能性がある。この問題を解決するために，前置詞を決定する名詞の意味属性を利用して，名詞によって前置詞を変化させるパターンを作る事が考えられる。

(b) 前置詞が動詞に依存する場合

英語パターンを以下に示す。

- (N1orI) V5.past N4 to V3 up N2.

上記の例では，出力英文に必ず前置詞‘up’が出る。前置詞が動詞によって決定される場合，前置詞は動詞に対応してさまざまな前置詞に変化する問題がある。解決策として，前置詞を決定する動詞の意味属性を利用して，前置詞を変化させるパターンを作る必要がある。

5.3 パターン適合率

本研究では，100 文中 48 文に対して日本語パターンの抽出が出来た．このパターン適合率の平均は 77%であった．また，A 判定のパターン適合率の平均は 98%，B 判定では 83%，C 判定では 75%，D 判定では 60%であった．

よって，パターン適合率が高いほど品質の高い英文を作成出来る傾向にある．

5.4 句レベル，節レベルの利用

文型パターンには単語レベルの他に句レベル，節レベルがある．句レベルおよび節レベルの文型パターンは，原文(日)を元に句，節の線形要素を変数に置き換えて作成されている．例を以下に示す．

- 原文(日) = 私の祖母は私に女は人前に入るものではないと教えた。
- 単語レベルの日本語パターン = #1[N2 の]/N3 は/N4 に/N5 は/N6 に/出るものではないと/V7.kako。
- 句レベルの日本語パターン = NP1 は/N2 に/N3 は/N4 に/出るものではないと/V5.kako。
- 節レベルの日本語パターン = NP1 は/N2 に/CL3 と/V4.kako。

句レベルの日本語パターン中の NP1 は名詞句を，節レベルの日本語パターンの CL3 は節を表している．単語レベルの文型パターンに比べ，句レベル，節レベルの文型パターンは句や節を変数化している．

本研究では，単語レベルの文型パターンを用いた実験において，100 文中 48 しか日本語パターン候補を抽出出来なかった．今後，句レベル，節レベルの文型パターンを用いる事で被覆率は上昇すると考えている．

6 おわりに

本研究では、重複文に対して、12万対の文型パターンを使用した翻訳精度について検証した。この結果、人手による評価は、48文中でAと判定された文が32%、Bと判定された文が10%、Cと判定された文15%、Dと判定された文が43%であった。また、機械による評価は、BLEUスコアが0.19、NISTスコアが3.3であった。しかし、単語レベルでのパターン翻訳は、重複文の翻訳において質の高い翻訳は21%(21/100)に留まっている。今後は、さらに翻訳精度を向上させる方法を考えていく。また、単語レベルだけではなく、句レベル、節レベルの文型パターンを使った翻訳精度の検証を行う予定である。

謝辞

本論文作成に際して，多大なる検討と助言をしてくださった池原悟教授ならびに村上仁一助教授，徳久雅人助手そして計算機工学C研究室の方々に深く感謝します．

また，参考にさせて頂いた文献の著者の方々に対して感謝します．

参考文献

- [1] 池原悟:非線形な言語表現と文型パターンによる意味の記述，情報処理学会，自然言語処理研究会，2004-NL-159，pp.139-146，2004-1
- [2] 村上ほか:日本語英語の文対応の対訳データベース，「言語・認識・表現」，第7回年次研究会，2002-12
- [3] 徳久ほか:文型パターンパーサの試作，言語処理学会第10回年次大会，2004(発表予定)
- [4] E.Sumita:Example-based machine translation using DP-matching between word sequences，DDMT workshop of 39th ACL，2001
- [5] K.Papineni，et al:BLEU:a Method for Automatic Evaluation of Machine Translation，IBM Research Report，September17，2001
- [6] NIST:Automatic Evaluation of Machine Translation Quality Using N-gram Co-Occurrence Statistics，<http://www.nist.gov/speech/tests/mt/mt2001/resource/>，2002

付録

入力文 100 文と入力文の模範訳

48 文の精度実験結果

入力文 100 文と入力文の模範訳

表は以下の通りに構成されている。

入力文 100 文と入力文に対応する模範訳 100 文を記載している。

48文の精度実験結果

表は以下の通りに構成されている。

ページ1からページ2は入力文と模範訳，ページ3からページ4は出力英文，評価1，評価2，原文(日)(英)．ページ5からページ6は最適な日本語パターンと日本語パターンに対応する英語パターンをそれぞれ記載している．また，評価1，評価2とは，2人の評価者それぞれの評価である．