

# 意味的等価変換方式に基づく単語レベルパターン翻訳方式の評価

坂田純 徳久雅人 村上仁一 村田真樹

鳥取大学大学院工学研究科情報エレクトロニクス専攻

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

## 1 はじめに

要素合成法を基本とする従来の機械翻訳方式には、文全体の構成要素への分解過程において意味が消失し、目的言語の生成過程で意味が復元されなくなる問題がある。この問題の解決を目的として、意味的等価変換方式が考案されている [1]。この変換方式においては、文の構造が抽象的な意味を包含するとみなし、それらの意味が保存されるように文パターンを記述する。具体的には、他の要素に置き換えても抽象的な意味の変わらない線型部分と、置き換えにより抽象的な意味が変容してしまう非線型部分との区別を付けて、文パターン化を行う。機械翻訳への利用にあたっては、線型部分に対する局所翻訳を行い、非線型部分と組み合わせて文全体の訳出を行う。

意味的等価変換方式の日英機械翻訳への利用を目的として、現在、日英重文複文文型パターン辞書が作成されている [2]。文パターンは、線型要素のレベルに応じて、単語、句、節レベルの3つのレベルで記述されている。すでに、単語レベル文パターンを用いる単語レベルパターン翻訳方式 (ITM-w)、句レベル文パターンを用いる句レベルパターン翻訳方式 (ITM-p) が実装されている [3],[4]。このうち、単語レベルパターン翻訳方式は、クローズドテストによる翻訳精度の評価しか行われていない。そこで本論文において、オープンテストによる翻訳精度の調査および問題解析を行う。

## 2 日英重文複文文型パターン辞書

重文、複文の日英対訳文約 12 万対を対象に、3 レベルの文型パターン対が作成されている。本研究では、単語レベル文パターン約 12 万対を用いる。文パターンは、字面、変数、関数、記号の4種類の要素を用いて記述されている。これらの要素の内、主に字面と関数が非線型部分にあたり、変数が線型部分にあたる。また、記号は適合範囲の拡大等の役割を持つ。表 1 に単語レベル文パターンの例を示す。

日本語の助詞や接続詞は、文の構造と密接な関係を有しており、文末表現“なかつた”が否定形かつ過去形といった抽象の意味を持つように、人がその文を理解するのに決定的な働きをもつ場合が多い。英語文との対応においては、語型の変化や語順に対応することも多く、その場合、単純な語と語の対応をとることはできない。

このような部分を非線型部分と呼び、字面、関数、さらに語順として、文パターン対の中に記述することによって、抽象的な意味を保存することができる。

具体的には、日本語パターンの“.hitei”は動詞変数 V4 が否定形であることを意味し、英語パターンの字面“never”と対応している。また、“.kako”は過去形であることを意味し、英語関数“^past”と対応している。この例では、V5 照合部分の局所翻訳結果を、過去形に変換することを指示している。日本語パターンの“<N1 は>”と英語パターンの“<I|N1>”は、主語省略文に対応するための記号である。主語 N1 に対応する語が入力文にあるときは、英語パターンにおいて N1 が選択され、そうでないときは、“I”が選択されることを意味している。

表 1 重文複文文型パターン辞書の記述例

日原文	彼のお母さんがああ若いとは思わなかった。
英原文	I never expected his mother to be so young.
日パターン	<N1 は>N2 の N3 が ああ AJ4 とは V5.hitei.kako.
英パターン	<I N1> never V5^past N2^poss N3 to be so AJ4.

## 3 単語レベルパターン翻訳

### 3.1 文パターン検索

使用文パターンの検索には、SPM[5]を用いる。SPM は、Augmented Transition Network(ATN) method[6]を用いる、日本語パターン検索システムである。本論文では SPM への入力に、名詞と用言の意味属性コード [7] を付与した形態素解析結果を用いる。SPM は、照合パターンと共に、照合されたパターンの線形要素に関する情報を出力する。入力文「彼のお母さんがああ若いとは思わなかった。」の出力例を、形態素解析結果、パターン照合結果の順に示す。

1. /彼 (1710,{NI:23,NI:48})
2. + の (7410)
3. /お母さん (1100,{NI:80,NI:49})
4. + が (7410)
5. /ああ (1110)
6. /若い (3106,{NY:5})
7. + と (7420)
8. + は (7530)
9. /思わ (2392, 思う, 思わ,{NY:32,NY:31})
10. + なかつ (7184, ない, なかつ)
11. + た (7216)
12. +。 (0110)
13. /nil

図 1 形態素解析結果

図 1 の 1 行目を例にとると，“彼”が形態素であり，{NI:23,NI:48} が名詞意味属性コードを示している。

表 2 文パターン照合結果

パターン ID	WJAC000004-00
日パターン	N2, の, N3, が, ああ, AJ4, とは, V5, .hitei, .kako, .
英パターン	<I N1> never V5^past N2^poss N3 to be so AJ4.
照合形態素	[1,2,3,4,5,6,7,8,9,10,11,12]=12
変数情報	N2=彼 [1], N3=お母さん [3], AJ4=若い [6], V5=思わ(思う)[9]

表 2 において，変数情報の N2=彼 [1] を例にとると，形態素番号 1 の形態素 “彼”，が名詞変数 N2 に照合されていることを示している。

### 3.2 単語レベルパターン翻訳システム ITM-w

ITM-w により，入力文にパターン照合された単語レベル文パターンを用いて翻訳英文を生成する。変数照合部分の日本語に対して局所翻訳を行い，英語パターンの変数部分にその局所翻訳結果を挿入し，訳出文を得る。その際，局所翻訳候補から訳語を一つ選択する必要がある。これまで tri-gram 値を用いて訳語選択を行っていた。本論文ではさらに，各単語の翻訳確率も用いて訳語選択を行う。なお，英語パターン中の字面には翻訳確率 1.0 を与える。ITM-w による翻訳の具体的な手順を以下に示す。

1. SPM による日本語パターン照合
2. 変数に照合された日本語の局所翻訳
3. 局所翻訳結果の英語パターンへの非決定的代入
4. 翻訳スコアによる局所翻訳結果の決定
5. 訳出文の出力

### 3.3 翻訳候補文からの出力文の選択

入力文に複数パターンが照合された場合，複数の翻訳候補文が得られる。本論文では，まず意味属性コードを制約として用い，使用パターンを制限する (3.4 節で説明)。複数パターンが残る場合は，絞り込まれたパターン全てを翻訳に使用し，得られた複数候補文から出力文を一つ選択する。ここで，一文全体の tri-gram 乗算値と各単語翻訳確率を乗算した値を，翻訳候補文の翻訳スコアと呼ぶことにする。そして翻訳スコア最大の候補文を出力文として選択する。

### 3.4 意味属性コードを用いたパターンの制限

パターンにおける変数は，他の要素に置き換え可能な部位である。しかし，どのような要素にも置き換え可能なわけではなく，置き換え可能な要素の範囲には，一定の制限がある。その制限を無視すると，不適切なパターンを翻訳に使用する可能性が高い。不適切なパターンの使用を除去するため，形態素の意味属性コードを用いて，使用文パターンの絞り込みを行う。意味属性コードは，日本語語彙大系 [1] のコードを用いる。

入力文の変数照合部分における形態素の意味属性コードと，照合パターンの変数部分における，日本語原文の形態素の意味属性コードを比較し，両コードが大きく異なる照合パターンは使用パターンから削除する。[1] の意味属性は，対象の見方や捉え方が，一般名詞意味属性，固有名詞意味属性，用言意味属性として約 3,000 属性に分類，体系化されている。このうち一般名詞意味属性と用言意味属性を，使用文パターンの制約として用いる。一般名詞意味属性は 12 階層，2,715 分類，用言意味属性は 4 階層，36 分類で記述されている。以下の手順で使用パターンの制約をかける。

1. 入力文と照合パターンに対し，変数部分の形態素の意味属性コードを抽出
2. 入力側の意味属性コードに対し，そのコードの親コードに包含されるコード全て (親コードの子孫全て) を，適合範囲として決定
3. パターン側の意味属性コードが，入力側の適合範囲に含まれるか調べる
4. 全ての変数において，意味属性コードが適合範囲に含まれるパターンを，使用パターンとして選択

### 3.5 ITM による翻訳例

システム全体のモデル図を図 2 に示す。

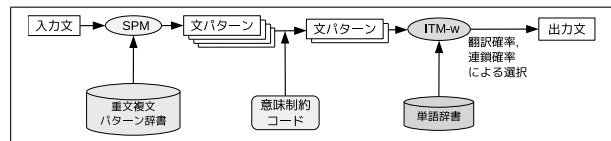


図 2 ITM による句レベルパターン翻訳

次に翻訳例を示す。入力文を「彼のお母さんがああ若いとは思わなかった。」とする。照合パターンとその変数情報は表 2 のものを用いる。各変数照合部分に対して，辞書引きを行い複数の局所翻訳結果を得る。例えば，V5^past の照合部分は「思う」であり，辞書引き結果に関数を適用した結果，局所翻訳の候補「thought, expected, ...」が得られる。同様に N2 に対し「he, he's, N3 に対し「mother, mama, mummy, ...」,そして AJ4 に対して「young, younger, ...」が得られる。最後に，tri-gram 値と翻訳確率を用いて訳語を選択し，出力文「I never expected his mother to be so young .」を得る。訳語選択の例を図 3 に示す。

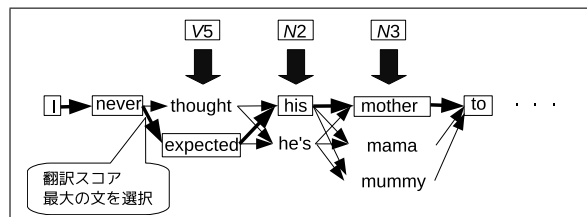


図 3 翻訳スコアによる訳語選択の例

## 4 翻訳実験

翻訳精度の調査と問題解析を行う。

### 4.1 実験方法

辞書 [2] の原文対約 12 万から 10 万文を抽出し、言語モデルの学習および単語辞書の抽出に用いる。残り 2 万文から 5,000 文を抽出しテスト文に用いる。ベースラインにはフレーズベース Moses を用い、ITM-w の結果と比較する。パラメータは、 $\text{distortion-limit}=-1$ 、 $\text{ttable-limit}=60$  とし、パラメータチューニングを行う。提案手法、ベースライン共に言語モデルの学習には英単文約 18 万文を加える。

翻訳精度の評価は、自動評価と人手評価により行う。

### 4.2 評価結果

入力 5,000 文のうち、文パターン照合に成功した文は 1,028 文であった。意味制約による絞込みにより、照合パターンを持つ文は 502 文に減少した。この 502 文の翻訳結果の翻訳精度を調べた。自動評価結果を表 3 に示す。

表 3 自動評価結果 (制約あり 502 文)

	BLEU	TER	METEOR	RIBES
ベースライン	0.292	0.562	0.569	0.778
ITM-w	0.329	0.520	0.611	0.796

表 3 より、全ての評価法において ITM-w がベースラインを上回っていることがわかる。

次に、人手評価結果について述べる。502 文からランダムに 100 文抽出し人手評価を行った。人手評価基準を表 4、人手評価結果を表 5 に示す。

表 4 人手評価基準

評価 5	原文の意味を正しく理解できる
評価 4	一部不適切な部分があるが、概ね理解できる
評価 3	原文の意味が何となく読み取れる
評価 2	部分的な理解にとどまる、または意味が原文と異なる
評価 1	理解不能

表 5 人手評価結果 (制約あり)

	評価 1	評価 2	評価 3	評価 4	評価 5	平均
ベース	21	36	12	9	22	2.77
ITM	4	26	12	14	44	3.68

表 5 の平均値より、ITM-w がベースラインを大きく上回っていることがわかる。各評価の文数をみると、ITM-w はベースラインに比べ、評価 1 が非常に少なく、代わりに評価 5 が多い。また、ベースラインと ITM-w 共に評価値は高い。意味制約をかけたことにより、入力文と類似したパターンのみを使用していることになる。よってベースラインも高い翻訳精度を持つが、ITM-w の方がよりよい翻訳精度が得られている。

### 4.3 評価 5 の例

評価 5 の例を表 6 に示す。

表 6 評価 5 の例

入力文	彼女のスカートはくぎにひっかかって裂けた。
参照文	Her skirt tore on a nail.
使用日バ	\$1^{*}\{ \#1[N2]N3 \text{は} \#2[くぎに] \text{(ひっかかって   引っ掛かって)} \text{て} \\$1V5.kako. \}\$
使用英バ	$\#1[N2^{*}poss] N3 V5^{*}past \text{ on } (a an) \#2[nail] .$
出力文	Her skirt tore on a nail .

## 5 問題分析

ITM-w の人手評価結果の低い 42 文 (評価 1 から 3) に対し、その原因を調査した。翻訳精度の低い原因として、主に次の 3 つが予想される。

1. 使用文パターンが不適切
2. 局所翻訳精度が低い
3. 候補文から翻訳精度の高い文を選択できない

出力文の翻訳精度が低い原因と、その内訳を表 7 に示す。

表 7 翻訳精度の低い原因

原因	文数
パターン不適切	10
パターン記述ミス	2
局所翻訳精度が低い	24
形態素解析、コード付与失敗	6

表 7 より、局所翻訳精度が低いことが 24 例で、半数以上を占めることがわかる。不適切な文パターンの使用が 10 例で、次に多い。

### 5.1 使用文パターンの問題

パターンが不適切な例を表 8 に示す。

表 8 パターンが不適切な例

入力文	彼はそれをする約束した。
参照文	He engaged himself to do it.
使用日バ	$\#1\{N1 \text{は}, N2 \text{を}\} V3 \text{と} < N4 \text{は} > V5.kako.$
使用英バ	$< I N4 > V5^{*}past N1 \text{ would } V3^{*}base N2 .$
変数情報	$N1 = \text{彼}, N2 = \text{それ}, V3 = \text{する}$ $V5 = \text{約束し (約束する)}$
出力文	I promised he would do it .
スコア等	翻訳スコア=-57.5717, 評価 2
日原文	彼はそれを買うと思った。
日英原文	I thought he would buy it.

入力文、使用パターン原文共に「彼は～すると～した」という文構造を持つ。しかし入力文では「約束をした」のは「彼」であり、パターン原文で「思った」のは省略主語の「私」である。よって補完記号  $< I|N4 >$  の適用により、「私」が約束をした」という意味の出力文となっている。文中の主語が述語の主語であるかを見分けることは非常に難しく、個々の事例を元にして意味上の主語を見分けるような処理ルーチンがなければ、このようなパターンの選択は困難と思われる。

### 5.2 局所翻訳精度の低い問題

局所翻訳精度の低い主な原因を示す。

1. 複合的な日本語表現の翻訳精度が低い
2. 辞書引きの失敗
3. 翻訳スコアによる不適切な表現の選択

原因 1 の例を示す。

表 9 局所翻訳精度の低い例

入力文	彼女は文句を言わずに金を払った。
参照文	She paid without protest.
使用日バ	\$1^{N1}\$ は \$N2\$ を \$1V3\$ hitei に \$N4\$ を \$1V5\$ kako.
使用英バ	\$N1\$ \$V5\$ ^past \$N4\$ without \$V3\$ ^ing \$N2\$ .
変数情報	\$N1\$=彼女, \$N2\$=文句, \$V3\$=言わ(言う) \$N4\$=金, \$V5\$=払っ(払う)
出力文	She paid the money without saying a word .
スコア等	翻訳スコア=-67.5873, 評価値 = 3

出力文をみると、「文句を言う」の局所翻訳結果が「saying a word」になっている。「文句を言う」は、英語では一般的に「complain」であり、使用英パターンの「V3(言う)^ing N2(文句)」のように、「言う」と「文句」を別々に辞書引きしたのでは、精度のよい翻訳は難しい。このような入力文の翻訳は、単語レベル文パターンでは難しいが、句レベル文パターンを用いることにより、よい翻訳精度が得られる可能性がある。

2, 3 の問題は省略する。

### 5.3 候補文からの出力文の選択

句レベルパターン翻訳 (ITM-p) において、多数の照合パターンから適切なパターンを選択できない問題が報告されている [4]。しかし ITM-w においては、照合パターン数が 10 を越えるような入力文は少数であった。また評価 1 から 3 の文において、候補文に評価 4, 5 の文が含まれていたのは、わずか 2/42 文であった。よって ITM-w では、翻訳候補文からの出力文の選択は、翻訳精度にほとんど影響を与えていないとみなせる。

### 5.4 問題分析のまとめ

翻訳精度の低い最大の原因は、局所翻訳精度が低いことであった。

意味属性を用いた制約により、不適切なパターンを取り除いたため、翻訳候補文からの選択に大きな問題はみられなかった。ただし制約をかけてなお、不適切なパターンの使用もみられた。

## 6 意味属性による制約の効果 (追加実験)

使用パターンの絞込み効果を調べるため、制約をけなかった場合の翻訳精度を調べた。自動評価結果を表 10 に、人手評価結果を表 11 に示す。

表 10 自動評価結果 (制約なし 1028 文)

	BLEU	TER	METEOR	RIBES
ベースライン	0.219	0.635	0.484	0.736
ITM-w	0.219	0.645	0.504	0.736

表 10 より、ベースラインと ITM-w の評価値にほとんど差はみられない。また表 3 の値と比べると、全ての評価法で「制約あり」の評価値の方が高い。

表 11 人手評価結果 (制約なし)

	評価 1	評価 2	評価 3	評価 4	評価 5	平均
ベース	25	33	13	4	25	2.71
ITM	12	29	13	14	32	3.21

表 11 より、ITM-w の結果がベースラインを上回っていることがわかる。しかし表 5 の結果と比べると、ベースラインではほとんど平均値に変化がない (2.77 から 2.71) が、ITM-w では平均値が 3.68 から 3.21 に大きく減少している。

よって意味属性コードを用いたパターン絞込みの効果を確認された。ただし制約により、出力文が約半数 (照合率約 20% から約 10%) に減少する副作用もみられた。

## 7 おわりに

本論文において、単語レベルパターン翻訳方式の評価と問題解析を行った。実験により、自動評価と人手評価共に、本方式が高い翻訳精度を持つことがあきらかになった。特に意味属性コードによるパターン絞込みを行った場合、人手評価において、ベースライン 2.77 に対し本方式 3.68 の高い評価が得られた。翻訳精度の低い最大の原因は、局所翻訳精度が低いことであった。局所翻訳精度の低い主要因の一つに、複合的表現の翻訳が難しいことが挙げられる。これはパターン照合率 (約 10%) の問題と共に、句レベルパターン翻訳方式により解決可能と考えられる。今後は句レベルパターン翻訳の調査を行う予定である。

## 参考文献

- [1] 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復文の日英文型パターン化. 自然言語処理, Vol.11, No.3, pp.69-95, 2004.
- [2] 池原悟. 鳥バンク, 日本語表現意味辞書 -重文複文編-. 2007. (<http://unicorn.ike.tottori-u.ac.jp/toribank>)
- [3] 石上真理子, 徳久雅人, 村上仁一, 池原悟. 結合価パターンを用いた動詞句の翻訳可能性の調査. 言語処理学会第 11 回年次大会, pp.364-367, 2005.
- [4] Jun Sakata, Masato Tokuhisa, Jin'ichi Murakami. Phrase-Level Pattern-Based Machine Translation Based on Analogical Mapping Method. 8th International Conference on NLP(japTAL2012), Oct. 2012.
- [5] 徳久雅人, 村上仁一, 池原悟. 重文・複文文型パターン辞書からの構造照合型パターン検索. 情報処理学会研究報告, Vol.2006, No.124, pp.9-16, 2006
- [6] Stuart C.Shapiro. Generalized Augmented Transition Network Grammars For Generation From Semantic Networks. Computational Linguistics archive. Volume 8 Issue 1, January-March, Pages 12-25, 1982.
- [7] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦. 日本語語彙大系. 岩波書店, 1997.