

概要

現在、日英機械翻訳の分野において、文型パターンの検索を用いて翻訳に利用する手法が注目されている。パターンを用いた翻訳は、日本語の入力文に合う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。その最たる例がパターン翻訳と言える。そして、パターン翻訳を始めとする従来のパターン翻訳技術の要は、翻訳に利用する文型パターンの検索を行う際に、入力に類似した「構文表現」を検索条件とすることである。これに対して、本研究では「意味」を検索条件に使用したパターンの絞り込みを行う。真理項とは、文章の意味を表わす分類情報であり、統語的分類と意味的分類の2種類の分類に分かれている。そのため、意味的分類を主な検索条件とすることで、入力文の意味が表現できる文型パターンの検索が期待できる。そこで本研究では、類似文翻訳を目標とした、真理項を用いた類似文検索を行う。

具体的な方法として、まず入力文の真理項を利用して文型パターン辞書から文型パターンを抽出する。次に、抽出された文型パターンの英文パターンに入力文の意味に沿う任意の単語を挿入する。そして、英単語を挿入した英文パターンに対して人手で評価を行う。最後にカバー率及び正解率を算出し、翻訳に利用できるパターンの検索条件としてどのような真理項の検索条件が適切であるかを示す。

重文、複文合わせて20文の入力文に対して、1000件の英文パターンの検索の結果、重文ではどの真理項の影響が強いかが判明しなかったが、複文では従属節の影響が強いことが判った。又、最適な検索条件は全ての真理項を検索条件とした場合であり、結果はカバー率 $R1$ は80%、正解率 $P2$ は80から100%であることが示された。

この結果より、本研究では真理項を用いた文型パターンの絞り込みの最適な検索条件は、全真理項を検索条件とすることで、高いカバー率 $R1$ と正解率 $P2$ を得ることを確認した。

目次

1	まえがき	2
2	本研究の目的	3
3	研究手順の概略	4
4	研究対象	5
4.1	パターン入手の手段	5
4.2	日英文型パターン辞書	6
4.3	文型パターン辞書の収録内容の例	7
4.4	入力文の決定基準	10
4.5	入力文入手の具体的方法	11
4.6	本研究で利用した入力文	12
4.7	真理項の検索条件	13
4.8	”意味”検索プログラム	14
5	評価	17
5.1	英文パターンからの英文生成	17
5.1.1	生成案1	17
5.1.2	生成案2	18
5.1.3	生成案3, 本研究の生成方法	19
5.2	評価方法	20
5.2.1	評価Aの例	21
5.2.2	評価Bの例	23
5.2.3	評価Cの例	25
5.2.4	評価Dの例	27
5.3	具体例	29
6	実験結果	32
7	複数の真理項を組み合わせたときの実験	33
7.1	カバー率, 正解率の計算方法	35
7.2	推定結果	36

7.3	実際の検索条件による実験	37
8	考察	38
9	あとがき	41
10	感想	42
10.1	従属節の影響力	42
10.2	主, 従属節の分類強化	43
10.3	意外なメリット	44

表 目 次

1	評価基準	20
2	文種別 1 の評価結果	32
3	文種別 3 の評価結果	32
4	組み合わせ一覧	33
5	文種別 1 の推定値	36
6	文種別 3 の推定値	36
7	文種別 1 の実測値	37
8	文種別 3 の実測値	37

1 まえがき

日英機械翻訳の分野において、用例翻訳、統計翻訳など様々な翻訳方式が提案されている。提案されている多くの翻訳方式の1つにパターン翻訳がある。パターン翻訳は、大量のパターンから入力表現に似たパターンを翻訳に利用する翻訳方式である。

従来、パターンを用いた翻訳を行う際、使用する文型パターンの統語的構造に着目した手法が多く、文構造に問題は無いものの、利用者の意図しない出力が現れる問題がある。又、パターンを用いた翻訳の分野では意味を対象とした検索方法が未解明である。

ところで、日英機械翻訳の分野では大規模な日英文型パターン辞書が作成されている。この辞書は、「類推原理」による「意味的等価変換方式」に基づいて、非線形な言語表現構造に対する文型パターンを意味的に類型化したものである。日英の原文と単語レベル、句レベル、節レベルの文型パターンと真理項が収録されている。真理項は、統語的分類と意味的分類に大別できる。統語的分類には、文種別、統語構造第1分類、統語構造第2分類、第1節間キーワード、第2節間キーワードがあり、意味的分類には、第1節間の意味、第2節間の意味、第1従属節の意味、第2従属節の意味、主節の意味がある。

しかし、研究の時点では、これらの真理項に対し情報付与が完了しているものは、統語構造第2分類、第1節間キーワード、第1節間の意味、第1従属節の意味、主節の意味である。そこで本稿では、統語構造第2分類を統語構造、第1節間キーワードおよび第1節間の意味を、節間のキーワードおよび節間の意味分類、第1従属節の意味を従属節の意味と略称する。

そこで本研究では、現時点で情報付与が完了している真理項を用いた英文パターン検索実験を行い、意味的に適切な英文パターンを出力するための最適な真理項の検索条件について調査する。

重文、複文合わせて20文の入力文に対して、1000件の英文パターンの検索の結果、重文では統語構造の真理項の影響が弱く、複文では従属節の影響が強いことが判った。又、最適な検索条件は全ての真理項を検索条件とした場合であり、結果はカバー率 R1 は 80 %、正解率 P2 は 80 から 100 %であることが示された。

本稿は以下の章で構成されている。2章では本研究の目的、3章では、研究手順の概略、4章では研究に利用した対象、5章では評価方法、6章では実験結果、7章では複数の真理項の組み合わせ実験、8章では考察と今後の課題、9章で結論、10章では本研究に対しての感想を述べる。又、付録には本研究で検索した英文パターン1000件と、真理項の説明を記載する。

2 本研究の目的

従来のパターンを用いた翻訳方式では，入力文の表現構造に一致する文型パターンをパターンパーサから検索して出力しているが，入力文の翻訳に使用出来ない文型パターンが正解パターンと共に多く抽出されてしまう問題がある．この問題が発生する理由として，従来の翻訳方式は，検索条件に入力文の意味を考慮していないからであると思われる．このことより，翻訳に利用するパターンの検索は，文章構造のみを検索条件に用いる場合，困難であると考えられる．そのため，従来の翻訳方式に用いられる検索方法では，入力文の翻訳に使用出来ないパターンも抽出されてしまう．この問題を解決するためには，文の意味が統語構造に与える影響を考慮した検索方法を用いるべきである．つまり，統語構造のみならず，統語構造と意味の両面で検索を行うことで，従来方式の問題の解決を試みる．

ところで，日英機械翻訳の分野において，日英機械翻訳の分野では大規模な日英対訳文型パターン辞書が作成されている．この辞書には，各パターンに対して真理項が付与されている．この真理項は統語的分類と意味的分類の2種類の情報をもっている．そこで，真理項を検索条件に利用することで，入力文と統語的分類と意味的分類の両面から文型パターンの絞りこみを行えるため，意味が類似した文型パターンが検索できる可能性がある．又，本研究にて意味ではなく，“意味”としているのは，文章の統語情報と文章の意味を統合して“意味”と呼んでいるからである．

本研究では，真理項を用いた英文パターン検索実験を行い，意味的に適切な英文パターンを出力するための最適な真理項の検索条件について調査する．

3 研究手順の概略

本研究の検索実験は、日英類似文翻訳を目標としていることから、入力文を日本語文、出力を英語パターンとする。又、従来の研究とは異なり、「意味」による検索を行うため意味についての情報を含む、真理項を検索条件に用いる。この構想を実現するためには、入力、検索条件、出力の入手方法等の研究のストーリーを計画しなくては目標通りの結果を得ることは困難である。

そこで、本節は以降に続くさまざまな説明に先だって、研究手順の概略を示す。ただし、研究手順で誤って解釈してはならないことに、本研究では文型パターン辞書を用いて入力、検索条件、出力を得る考えが前提であるため、他の研究の様に、入力等の設定を計画してから計画に合ったツールを探す、もしくは製作するという手順と、手順が一部逆転している。まず始めにツールとなる文型パターン辞書を入手するのは本研究のような特殊な条件下の場合のみである。

計画段階

1. 入力、検索条件、出力、模範訳が記載された文型パターン辞書を入手する。
2. どのような条件で文型パターン辞書より、入力文と検索条件を得るか計画する。
3. 出力をどのようにして入力文の意味と比較するかを計画する。

収集段階

1. パターン辞書から入力を得る。
2. 計画した検索条件の基、文型パターン辞書より出力パターンを得る。

実験段階

1. 出力パターンを入力文の意味と比較できる参考用の文を生成する。
2. 生成した文を参考に入力文の意味を比較する。

考察段階

1. 比較結果を集計する
2. 集計結果から法則性を読み取る

4 研究対象

4.1 パターン入手の手段

まず，本研究を説明するにあたり，入力と検索条件と出力を明確にする．簡潔にいえば，入力は日本語文であり，検索条件は日本語入力文の真理項（真理項の細かい説明は付録に記載），出力は英語の文型パターンである．これは，入力文の意味を，出力である英文パターンが表現可能であるかによって，検索条件となる真理項の影響力が判別可能になるからである．又，判別の際，英文パターンの変数部に，入力文の意味を表現できる様な英単語を挿入した文を参考にする．

そこで問題となるのは，いかにして入力となる日本語文，検索条件の真理項，出力の英文パターンを得るか，である．しかし，これらの問題はあるパターン辞書により解決する．それが，最近になり，科学技術振興事業団「JST」の戦略的基礎研究推進事業「CREST」における 研究領域「高度メディア社会の生活情報技術」の研究課題「セマンティックタイポロジーによる言語の等価変換と生成技術」の支援により，大規模な日英対訳文型パターン辞書が開発された．このパターン辞書ひとつで，入力文，真理項，出力パターンが収集可能である．それでは，以降に文型パターンについての説明に移る．

4.2 日英文型パターン辞書

前述の通り，本研究で使用する日英対訳文型パターン辞書（以後，文型パターン辞書）は「CREST」支援により開発されたパターンが記載された辞書である．まず，この辞書を手に入れる．本辞書名は「3レベルの文型パターンファイル」ともいわれる．その所以は，辞書には，単語レベル（12万件），句レベル（9万件），節レベル（1.2万件）の3つの文型パターン（合計22万件）が記載されているからである．又，これらのパターンには，必ず日本語文，英語文そして真理項が収録されている．

その結果，本研究の入力である日本語文，検索条件である真理項，出力である英文パターン，これに加え，出力に対する模範英文を得ることができる．

又，本研究では，単語，句，節レベルのパターンのうち，単語レベルのパターンを実験の対象とする．この理由として，以下のことが挙げられる．

1. 句，節レベルのパターンは現在でも未確立分野であるため，句，節レベルでの真理項を用いた研究は見送るべき．
2. 句，節は単語より大きい分類であり，単語レベルで良い結果が得られない場合，句や節で良い結果が得られるとは考えづらい．
3. 英文を構成する最小構成要素はアルファベットであるが，意味を表わす最小構成要素は，単語である．これは，本研究の入力文の意味を出力に反映させる目論見に適った事実である．

以上の理由により，文型パターン辞書から，日本語文（入力），真理項（検索条件），英文の単語レベルパターン（出力），入力文の英文（模範訳）を得る．次に，文型パターン辞書の収録例を示す．

4.3 文型パターン辞書の収録内容の例

例 1 :

1. 日本語パターン

/y \$1^{cf #1[GEN2] /k N3 は } /tcfk NUM4 で \$1 /ytk NUM5 だ。

(日本語パターンの例文)

このタバコは 20 本入りで 1 ドルだ。

2. 英語パターン

#1[AJ2] N3 be NUM5 for NUM4.

(英語パターンの例文)

These cigarettes are a dollar for twenty.

3. 真理項

- (a) 一文種別 : 1
- (b) 統語構造 : 並列
- (c) 節間 KW : 連用中止
- (d) 節間の意味 : 順接的並列
- (e) 従属節の意味 : 判断措定の表現
- (f) 主節の意味 : 判断措定の表現

例 2 :

1. 日本語パターン

/ytcfk N1 は /tcfk N2 から /tcfk N3 も /cf AJV4、 /ytck おまけに /tcfk N5(まで |迄) /cf 多い。

(日本語パターンの例文)

アメリカ人は衣食住から給料も豊かで、おまけに税金まで多い。

2. 英語パターン

N1 be ADV(AJV4) housed, V(N2).past, V(N3)^ed and V(N5).past.

(英語パターンの例文)

American people are abundantly housed, fed, clothed, salaried and taxed.

3. 真理項

- (a) —文種別 : 1
- (b) —統語構造 : 並列
- (c) —節間 K W : 連用中止
- (d) —節間の意味 : 順接的並列
- (e) —従属節の意味 : 判断措定の表現, 幸運不幸, 様相
- (f) —主節の意味 : 判断措定の表現

例 3 :

1. 日本語パターン

$/y \ \$1^{\{ /tcfk \ N1 \ 是 \}} /cf (V2^{\wedge} \ rentai | \ ND2 \ をする) ! \ ことを \ \$1 /cf \ V3.kako.$

(日本語パターンの例文)

彼は出発することを通告した

2. 英語パターン

$N1 \ V3.past \ of \ N1^{\wedge} \ poss \ N(V2 | \ ND2).$

(英語パターンの例文)

He gave notice of his departure.

3. 真理項

- (a) 一文種別 : 3
- (b) 統語構造 : 名詞節, 補足節
- (c) 節間 KW : 連体形, 事
- (d) 節間の意味 : 固有名詞, 一般名詞, 形式名詞
- (e) 従属節の意味 : 開始, 行く出発
- (f) 主節の意味 : 通知報道

4.4 入力文の決定基準

まず、研究する上で入力文をどのような基準で得るかを設定しなければ、入力文を得ることができない。ここで、前述の文型パターン辞書の説明を踏襲した上で、入力文の設定基準を説明する。まず、入力文等は全て文型パターン辞書から得る訳だが、辞書に収録されたパターン全て(22万件)を入力文として実験することだできない。以下にその理由を示す。

1. 人手では労力と時間が掛かり過ぎる。
2. プログラムにより全パターンを入力文としても、コンピュータへの負荷が強過ぎる上、やはり時間が掛かり過ぎる。
3. そもそも全パターンを入力文としたら、検索用の英文パターンが残っていない。

以上のことが問題となる。そこで、まずは何件のパターンを検索の実験に使えば良いかを決めて、次に検索条件がいくつかを決めた上で、入力文の数を逆算する。

ここで、知っているべきことは、時間的な問題と計算時の有効桁数の関係上、翻訳や検索の分野では実験対象の数が1000件というのがひとつの限界であることである。このことより、本研究では、検索条件が5つであることを考慮する。又、各検索条件に対し、評価や集計の行い易さから、検索するパターン数は10文とする(現時点、50件)。さらに、入力が重文であるか複文であるかにより実験結果が大きく異なることを予想し、入力文は重文、複文の2種類を用意したい($50 \times 2 = 100$)。つまり、入力文の数は1000を100で割り、重文、複文についてそれぞれ10文を用意すれば良いことになる。

次に、どのような条件で入力文を選ぶかである。これは、全てのパターンを選ぶことができないことから、「表現が一般的な文」を入力文の対象に留めておくことが考えられる。次に、入力文の真理項は検索条件に用いることから、実験に用いる真理項が全て付与された文を選ぶ必要がある。又、一部のパターンの真理項には「0」や「-」が付与されてはいるが、この真理項では検索条件が無指定になってしまうため、この2つの真理項が付与されたものは除外する。以上のことから、次に入力文の具体的な入手方法を挙げる。

4.5 入力文入手の具体的方法

以下に入力文の入手手順を示す．なお，括弧内は前述の入力文の決定の際の問題を表わす．

1. 文型パターン辞書から，重文及び複文を全て取り出す．(重文と複文がパターンに大きな影響を及ぼしていると考えられるから)
2. 1で得たパターンを収録された真理項でソートする．(一般的な表現の文を入力としたいから)
3. 2.でソートした結果を重文，複文について降順に10文ずつ得る．(実験件数の1000件からの逆算)

以上の方法により，入力文を得る．又，文型パターン辞書には各文に対して真理項が必ず付与されているので，この作業の過程で検索条件に用いる真理項も同時に収集できる．

4.6 本研究で利用した入力文

付録には、検索で得た英文パターン，英単語を挿入して得た評価用英文，模範訳を載せるので、入力文についての理解を深める意味を込めて以下に入力文の一覧を示す。

重文

1. このタバコは20本入りで1ドルだ。
2. たびたびお手数をかけてすみません。
3. 貴社よりご照会をいただき、ありがとうございます。
4. 彼は損をしてプリプリしている。
5. 四方八方から人が集まって来た。
6. 親父が借金を残して死んだ。
7. 彼は当惑面をして帰って来た。
8. 別便で心ばかりの粗品を送らせていただきましたので、ご笑納頂ければ幸甚に存じます。
9. 彼はぼくをばかにしきってふふんという顔をした。
10. ご注文をいただきありがとうございました。

複文

11. 自分のすることには完全主義者です。
12. 私は彼が無罪だと思う。
13. 彼は西南役で有名な大将だ。
14. 今は議論をしている時ではない。
15. この刀は正宗の作だという。
16. 歴史上大きな動乱の時機であった。
17. 人の性は善であると言われている。
18. 小樽は人気の荒い処だ。
19. これは公にすべき事柄でない。
20. 彼の生涯は平凡なものだった。

4.7 真理項の検索条件

本研究の目的は，入力文の意味を表現できる文型パターンの検索条件には，どの真理項が相応しいかを示すことにある．そこで本研究では，重文，複文の違いを表わす文種別を除いた（つまり，入力文選定時点で条件として固定されている）5つの真理項を単独の検索条件とすることで，検索結果からどの真理項が意味に対する影響力があるかを示す．以下に，検索条件を簡潔に示す．

1. 「統語構造」のみを検索条件とする
2. 「節間KW」のみを検索条件とする
3. 「節間の意味」のみを検索条件とする
4. 「従属節の意味」のみを検索条件とする
5. 「主節の意味」のみを検索条件とする

4.8 ”意味”検索プログラム

本研究では，入力文と翻訳に利用する文型パターンの検索を自動的に行うためのプログラムを用いる．このプログラムは，真理項を検索項目として，文型パターン辞書から入力文と真理項が一致する文型パターンを出力するプログラムを”意味”検索プログラムと呼んでいる．

この”意味”検索プログラムは，2種類の入力が可能である．ひとつは，日本文を形態素解析によって各要素に分解し，真理項を抽出する．そして、抽出した真理項のうち，指定した真理項が一致する文型パターンを文型パターン辞書から出力する．もうひとつは，真理項を直接指定し，文型パターン辞書から文型パターンを出力する．

本研究では，入力文は，文型パターン辞書に使用される日本文を利用するため，後者の方式を文型パターン収集の方法に利用する．

又，本プログラムは条件の指定方法が非常に柔軟である．なぜなら，真理項の指定時に，全ての真理項を検索条件とする必要が無いからである．状況に合わせて必要な真理項の項目のみを指定するだけでもパターンを検索することができる．その際，指定しなかった項目は，どのような真理項でも良いとプログラムは解釈する（その真理項項目は「無し」の分類も含む）．この柔軟性により，利用者は自分が望む真理項のパターンを検索することができる．

入力文の真理項を用いて検索した例を次に示す．

1. 入力文 1

この刀は正宗の作だという

2. 入力真理項

文種別 : 3

統語構造 : 補足節

節間 K W : と

節間の意味 : 間接引用

従属節の意味 : 判断措定の表現

主節の意味 : 音声発話，言語表現一般，伝搬拡散

3. 出力された英語文型パターンの一例

- $N1 V4 \#5(\hat{present}|\hat{past})$ to be
 $N3$ of $N2$.
- $N1$ 'is' $\#5(\hat{present}|\hat{past}) V4.past$ to be
 $NUM3$ of $N2$.
- $N1 V5$ that $N3$ be $AJ(N4)$ from $N2$.
- $N1 V4.passive$ to be the birthplace
 $\#2[$ of $N3]$.
- $N2^{\hat{poss}} N1 V4$ of $N2^{\hat{obj}}$ as $N3$.
- $N1$ have $V(V5|ND5).past$ that $AJ2 N3 V(N4).past.passive$.
- \langle You| $N1\rangle$ must $V2 \langle$ yourself| $N1^{\hat{pron}}\hat{reflex}\rangle$ to the decree of Karma.
- \langle I| $N1\rangle V4.past$ as \langle my| $N1^{\hat{pron}}\hat{poss}\rangle N3$.

1. 入力文 2

別便で心ばかりの粗品を送らせていただきましたので、ご笑納頂ければ幸甚に存じます。

2. 入力真理項

文種別 : 1

統語構造 : 副詞節

節間KW : ば

節間の意味 : 因果関係

従属節の意味 : 賞讃崇拜

主節の意味 : 思考活動, 識別理解

3. 出力された英語文型パターンの一例

- $AJ1 N2$ be still $AJ3$ despite appearance.

- #1[AJ2] N3 will V4.*not* in the wash.
- N1 have made #5[AJ(N6)] operational changes to V4 N3 to N2.
- As <one| N1> grow older, <one| N2> V4 N3.
- #1[N2[^] poss] N3 V7.*past* #4[the] policy's N6.
- V4 N3 to V2 N1.
- If N1 would V3 before N2, <we| N5> should have AJ6 N4.
- N4 will V6.*not* unless <you| N1> V3 <it| N5> with N2 and water.

1. 入力文 3

彼の生涯は平凡なものだった。

2. 入力真理項

文種別 : 3
 統語構造 : 名詞節
 節間KW : 連体形
 節間の意味 : 固有名詞, 一般名詞
 従属節の意味 : 判断措定の表現
 主節の意味 : 特徴

3. 出力された英語文型パターンの一例

- N6 turned N7 while <I| N1> be.*past* #2[ADV3] V5[^]ing N4.
- N1 and N2 be #4[AJ5] N6 of N3.
- N1 be AJ(V4) about N3 on ' N2' fashions.
- <I| N1> will tell <you| N3[^] obj> N2 that will make <you| N3[^] pron[^] obj> laugh.
- Let's V6 N5 to V4 #1[N(GEN2)] N3.
- Few of #1[AJ2] N3 will V5 N4.
- A knock sounded on N1.
- ADV(TIME1) N2 be taking N4 on how long N3 lasts.

5 評価

5.1 英文パターンからの英文生成

5.1.1 生成案1

ここで記載する方式は、本研究では「利用されなかった方法」である。そのため、研究には直接関係ないものの、英文パターンの生成方法の参考になるために、記載する。ここで説明する生成方式は、パターンの字面、変数を問わずに品詞の種類に分解し、品詞に合った英単語を挿入する方法である。しかし、この方法では、「字面を無視」していること、全てを品詞に分解しているので、 $S + V + O$ 構文を基に英単語を挿入するという「パターンの形骸化」の2点から本研究の英文パターンの生成方法に「相応しくない」と判断した。下記にその手順と例を示す。

生成手順

1. パターン中の字面と変数部を品詞に分解する。
2. 入力文の意味と品詞の対応を利用して英単語を挿入する。

具体例

入力文：私は彼が無罪だと思う。

・検索例1

検索パ : $N1 V2.past$ the right thing.

品詞順 : N V N P N N

生成後 : I assume the innocence him.

・検索例2

検索パ : $N1$ blamed $N3^{\wedge}obj$ for $N1^{\wedge}poss$ $N2$.

品詞順 : N V N P N N

生成後 : I assume him of innocent man.

・検索例3

検索パ : $\langle I | N1 \rangle$ gave $N2$ credit for more sense.

品詞順 : N V N N P AJ N

生成後 : I assume his innocence for more time.

5.1.2 生成案 2

ここで記載する方式も，本研究では「利用されなかった方法」である．しかし，研究英文パターンの生成方法の参考になるために記載する．ここで提案した方法は先程とは真逆の厳格さと言える．それは，入力文模範訳のパターンと検索パターンの「変数種だけでなく変数番号も対応させる」方法である．しかし，本研究では翻訳を利用する存在は人間であり，結局は検索できた英文パターンを利用者が判断するというプロセス，つまり「人手による英文パターンの評価」が前提である．又，変数番号は，日本文側，英文側の双方に付与されてはいるが，両者は対応関係がとれている訳ではなく，それぞれの特徴に合わせて個別に付与だれている．以上の2点より，この生成案は本研究の英文パターンの生成方法から「見合わせる」ことにした．下記に具体的な手順と例を示す．例を見ると，挿入する英単語に明らかな違和感を感じる．

生成手順

1. 入力文の模範訳パターンと検索した英文パターンの変数種と変数番号が一致するものを挿入し，英文パターン中に挿入しきれないものは任意の英単語を挿入する．

例

入力：彼は当惑面をして帰って来た

模範訳：He returned with a puzzled countenance.

模範パ：N1 V3.past with N2.

・検索例 1

検索パ：V4 AJ1 N2 with N3.

生成後：Return the puzzled countenance with him.

・検索例 2

検索パ：N1 V5.past #2[AJ3] N4 between N1^{pron} obj.

生成後：He returned puzzled countenance between him.

・検索例 3

検索パ：N3 never V(V4| ND4) in N1^{poss} N2.

生成後：You never return in him puzzled countenance.

5.1.3 生成案 3, 本研究の生成方法

ここで説明する方法が本研究で「採用された方法」である。真理項を検索条件とした出力はあくまで英文パターンである。従って、このままでは入力文の意味が表現できるパターンかは判別できない。そこで、英文パターンの変数部 (N や V) に入力文の意味に近くなる様に、英単語を挿入する。このときに英文パターンに英単語を挿入したのでは「パターン」とは言えないので、「生成文」と本稿では言う。この生成文と入力文の模範訳の意味を比較することで、検索条件である真理項の影響力を推定する。以下に英文パターンが生成文になる過程を示す。又、前置詞は後に続く名詞に依存していると考え、前置詞は任意の前置詞に置き換えても良い。ただし、前置詞の中には、動詞と呼応関係のものがある。この呼応関係がある前置詞についても懸念すべきである。以下に生成手順と例を示す。

生成手順

1. 入力文と意味が類似するような英単語を、パターン中の変数部に挿入する。
2. 前置詞を入力文の意味に類似する様に別の前置詞に置き換える。

例

入力文：このタバコは20本入りで1ドルだ。

模範訳：These cigarettes are a dollar for twenty.

・具体例 1

出力パ：<He| $N1$ > $V3.past$ $N2$ and $V5.past$ $N4$.

挿入： $N1$ Cigarettes, $V3$ are, $N2$ twenty, $V5$ cost, $N4$ one dollar

生成文：These cigarettes were twenty and cost one dollar.

・具体例 2

出力パ： $N1$ be $N2$, and with to $V3.passive$ as such.

挿入： $N1$ Cigarettes, $V3$ are, $V3$ be

生成文：Cigarettes are twenty, and with to be cost as such.

・具体例 3

出力パ：In #1[$AJ2$] $N3$, $N4$ be $NUM6$ $N5$.

挿入： $AJ2$ twenty, $N3$ cigarettes, $N4$ it, $NUM6$ one, $N5$ dollars,

生成文：In twenty cigarettes, it is one dollar.

5.2 評価方法

前節の生成方法 3 により，英文パターンに英単語を挿入することで得た生成文を入力文の模範訳と評価する．生成文の評価が良ければ，生成文の基盤となる英文パターンが翻訳に利用できると考えられるからである．又，本研究では，意味による検索を重点に置くため，評価 A の次の評価 B には「パターン」より「意味」が類似で判断すべきだと考え，評価 B には「パターンが類似」ではなく「意味が類似」という基準にした．以下に，評価基準と具体例を示す．なお、() 内は模範訳で [] 内は人手で作成した英文である．

表 1: 評価基準

評価	説明
A	入力パターンと意味，文構造が共に類似
B	入力パターンと意味が通じる
C	辛うじて入力パターンの意味が汲み取れる
D	入力パターンと意味がかけ離れている

5.2.1 評価 A の例

・例 1

入力文：たびたびお手をかけてすみません

模範訳：(I'm sorry to trouble you so often.)

英文パターン：<I|N1> be sorry to trouble
<you|N3[^] obj> so AJ(ADV2).

生成文：[I am sorry to trouble you so often.]

例 1 の入力文と生成文を比較すると、「お手数」が「trouble」、「すみません」が「sorry」に対応しているため、意味が類似している。これに加えて、生成文の構造が模範訳の構造と類似(例 1 では一致)しているため、評価は A が適当である。

・例 2

入力文：彼は西南役で有名な大将だ

模範訳：(He is a general of Seinan Affair fame.)

英文パターン：N1 be.past AJ4 N5 in N3 of N2.

生成文：[He was famous genneral in affair of Seinan.]

例 2 では、過去形になっているものの、入力文の意味を全て表現している上、模範訳とパターンが若干異なるものの、非常に判り易い構造であるため、評価 A が適切と判断した。

・例3

入力文：たびたびお手をかけてすみません

模範訳：(I'm sorry to trouble you so often.)

英文パターン：<I| N1> V3.past N2^poss ADV(V5) N4.

生成文：[I apologized you over the trouble.]

例2では，入力文の「たびたび」を表現できていないものの，入力文の意味が損なわれているわけではなく，パターンも綺麗であるため，評価Aが適切と判断した．

5.2.2 評価 B の例

・例 1

入力文：このタバコは 20 本入りで 1 ドルだ。

模範訳：(These cigarettes are a dollar for twenty.)

英文パターン：In #1[AJ2] N3, N4 be NUM6 N5.

生成文：[In twenty cigarettes, it is one dollar.]

例 1 では、生成文では「20 本のタバコ」、「1 ドルである」という意味を示しているが、模範訳のパターンとは大きく違う上、意味は通じるが、価格を表わすために「is」よりも「cost」を用いるべきである。そのため、この検索できた英文パターンは評価 B の表現が限界であると判断した。

・例 2

入力文：親父が借金を残して死んだ

模範訳：(His father left debts behind him.)

英文パターン：N1 V4.past, V3^{ing} behind

N2.

生成文：[His father died, leaving behind debts.]

例 2 では、パターンが模範訳と大きく違うため、意味にも若干の悪影響が出ている。しかし、「借金を残した上に、死んでしまったこと」がカンマの区切りにより明確になっている。この結果、「パターンは良くない」、「意味は明確」と判断した。厳密な評価 B というよりも、「評価 C 寄りの評価 B」と言った方が正しい。

例 3

入力文：貴社よりご照会をいただき、ありがとうございます。

模範訳：(We thank you for your inquiry.)

英文パターン： *N1V4.pastN3 for N2.*

生成文：[We thanked inquiry from you.]

例 3 では、パターンが模範訳と違うものの、全体的に模範訳よりすっきりして、意味が明確になっている。しかし、感謝しているにも関わらず、文から丁寧さが欠けてしまっているため、評価 B に該当する。

5.2.3 評価 C の例

・例 1

入力文 : 彼は当惑面をして帰って来た

模範訳 : (He returned with a puzzled countenance.)

英文パターン : In $N1^{\wedge}poss$ opinion, $N3$ should
be $N2$.

生成文 : [In his opinion, he should be a puzzled countenance.]

例 1 ではパターンの違いが意味に対して大きな悪影響を与えている。「彼が当惑面である」というニュアンスはあっているものの、「should」によって「彼が現実には当惑面ではないかも知れない」ことを示唆している。さらに生成文には「帰って来る」ことには触れていない。又、「In his opinion」となる部分は明らかに邪魔である。以上のことから、彼が当惑面であるというニュアンスのみが汲めるので評価 C と判断した。

・例 2

入力文 : 自分のすることには完全主義者です

模範訳 : (I am a perfectionist at what I do.)

英文パターン : <We| $N1$ > V($V7$ | $ND7$).past not to $V6$ AJ($N4$) $N5$ #2[to $N3$].

生成文 : [I thought not to be a perfectionist.]

例 2 でも、自分と完全主義者についての関係に触れていることが判るだけで、このパターンでは、入力文の意味が不明瞭になってしまう。その理由として、生成文には「自分のすることには」という限定表現が存在しない上、「not」により否定形になっている。このことから、「自分と完全主義者についての関係」のニュアンスが汲み取れるだけとなり、評価 C が適切と判断した。

・例3

入力文：彼は損をしてプリプリしている

模範訳：(He is sore on the subject of his loss.)

英文パターン：AJ1 N2 N(V4) without so much as V(N3)^grn.

生成文：[He annoys without so much as saying goodbye.]

例3では、「損をする」と「プリプリ」することがわかるが文に余計な部分が含まれているので入力の意味が判りづらくなっているため、評価Cと判断した。

5.2.4 評価 D の例

・例 1

入力文：人の性は善であると言われている

模範訳：(It is said that men are born good.)

英文パターン：N1 V5 #2[AJ(N3)] pretty

N4.

生成文：[People say pretty born.]

例 1 では、英文パターンから「皆は～と言う」という表現しかできない変数の量と配置である。これに加え、「～と言う」の「～」の部分には真理項の関係上、英文パターン中で「pretty」に固定されてしまっている。その結果、「意味が似ていない上に、パターンによる表現方法にも問題がある」としか言えない生成文になってしまった。後述の実験結果で、評価 C や D が圧倒的に多いのは、この例の様に従様部分に字面が残っている、パターン中の変数量では表現しきれない、ということが多いからである。

・例 2

入力文：貴社よりご照会をいただき、ありがとうございます

模範訳：(We thank you for your inquiry.)

英文パターン：<I| N2> be looking forward to our continued friendship N1.

生成文：[You are looking forward to our continued friendship us.]

例 2 は例 1 と同様に、本研究の障害を良く説明してくれている。例 2 ではパターン中の字面があまりにも多過ぎるので、仮にパターンが良いため「検索自体が成功」でも、「単語挿入ができず、翻訳に利用可能か判らない」という問題がある。

そもそも字面が残る理由は、真理項が表わす情報を反映させるには、パターンのこの部分を字面にするしかない、ということである。もし、本方式の大幅改良するとしたら、生成文の作成方法の厳しさを、却下された「生成方法 1」のように緩くするか、真理項とパターンの関係を見直す必要がある。

・例3

入力文：今は議論をしている時ではない。

模範訳：(This is no time for argument.)

英文パターン：N1's N2 be #3[ADV4] thoughtful.

生成文：[This is thoughtful.]

例3の場合、このパターンの規模ではそもそも入力文を表現し切れない上に、文の重要な部分の動詞が字面で固定されているため、完全に入力の意味が判らない状態である。

5.3 具体例

以下に具体的な検索時のデータ例を示す。

1. 入力文

私は彼が無罪だと思う。

[I assume him innocent.]

2. 検索条件となる真理項

文種別 : 3

統語構造 : -

節間KW : -

節間の意味 : -

従属節の意味 : 判断措定の表現

主節の意味 : -

3. 検索英文パターンの一部

英文パターン : $N1 V5\#6(\hat{present}|\hat{past}) to$

$V4 AJ2 N3.$

[I supposed to assume his innocence.]

4. 評価 ... B

主語と目的語の関係に誤りが無く、動作が入力文と同じである。しかし、

「supposed」により、「assume」の動作が消極的になるため、評価を B とする。

1. 入力文

この刀は正宗の作だという

[This sword is ascribed to the forge of Masamune.]

2. 検索条件となる真理項

文種別 : 3

統語構造 : 補足節

節間 K W : -

節間の意味 : -

従属節の意味 : -

主節の意味 : -

3. 検索英文パターンの一部

英文パターン : <I| N1> would like to V5 AJ3 N4 of N2.

[Masamune would like to forge this sword.]

4. 評価 ... C

正宗がこの刀を作りたがっていたことから、結局「実は正宗が作った」と無理に推測することはできるが、実際に正宗が刀を作ったかは不明である上、入力文は伝聞形であるが、生成文は平叙文であるため、入力文の意味をうまく反映していないと考え、評価はCとした。

1. 入力文

小樽は人気の荒い処だ

[Otaru is a lively town.]

2. 検索条件となる真理項

文種別 : 3

統語構造 : -

節間 K W : -

節間の意味 : -

従属節の意味 : 物の性状既定

主節の意味 : -

3. 検索英文パターンの一部

英文パターン : *N1 strike N2 as #3[AJ4] N5.*

[Otaru strike place as lively town.]

4. 評価 ... C

小樽が賑やかであることを説明しているが、このパターンでは、小樽が「賑やかに感じる」のであって、本当に賑やかであるかは判らない。そこで、入力文の意味を一部しか表現できないため、評価を C とした。

6 実験結果

翻訳に利用する文型パターンを表中の5種類の真理項の条件で検索する．出力される英文パターンが非常に多いため，1つの条件で出力される英文パターンからランダムに10パターン選択し，1000パターン(文種別2種類 × 入力文10文 × 条件5種類 × 英文パターン10件)評価する．結果を表2, 3に示す．なお，表内の数字は英文パターン数である．

表 2: 文種別 1 の評価結果

検索条件	A	B	C	D
統語構造	1	6	26	67
節間KW	4	13	22	61
節間意味	2	13	26	59
従属節	2	15	19	64
主節	3	12	23	62

表 3: 文種別 3 の評価結果

検索条件	A	B	C	D
統語構造	1	14	31	54
節間KW	1	12	23	64
節間意味	1	12	27	60
従属節	6	25	25	44
主節	3	17	29	51

実験結果より，文種別1では，いずれの真理項の検索条件も，大きな差が見られないが，文種別3では”従属節の意味”が重要であることが判る．

7 複数の真理項を組み合わせたときの実験

本研究の目的は、意味的に適切な文型パターンを抽出する為の最適な真理項の条件の調査である。しかし、真理項のすべての組み合わせを検索条件として実験することは困難である。

そこで、前章で評価に使用した 1000 パターンを対象に、各パターンに付与されている真理項が、どれだけ入力文の真理項と一致するかを調べることにより、複数の真理項を指定した場合に抽出される英文パターンの品質を推定する。

以下に、入力文の真理項と検索した 1000 パターンの組み合わせと評価結果を表で示す。

表 4: 組み合わせ一覧

統語構造	従属節	主節	節間	kw	A,B	C,D	良評価%	推定値
					19	29	39.58	0.33
				×	1	2	33.33	0.27
			×	×	2	9	18.18	0.24
		×			22	85	20.56	0.26
		×		×	2	8	20.00	0.21
		×	×	×	10	33	23.25	0.18
	×				34	138	19.76	0.24
	×			×	1	9	10.00	0.19
	×		×	×	4	29	12.12	0.16
	×	×			59	390	13.14	0.18
	×	×		×	4	26	13.33	0.12
	×	×	×		0	1	0.00	0.15
	×	×	×	×	3	38	7.31	0.10
×			×	×	1	1	50.00	0.24
×		×	×	×	1	21	4.54	0.19
×	×				0	1	0.00	0.25
×	×		×	×	1	13	7.14	0.17
×	×	×	×	×	0	3	0.00	0.10

表 4 の結果を参考に , 品質の良いパターンが検索できそうな真理項の組を決定した .

1. 条件

文種別+統語構造+節間キーワード+節間の意味分類+従属節の意味+主節の意味

2. 条件

統語構造+従属節の意味+節間の意味+節間キーワード

3. 条件

統語構造+節間の意味

4. 条件

従属節の意味+主節の意味

7.1 カバー率，正解率の計算方法

評価には，カバー率 $R1$ ，正解率 $P1$ ， $P2$ を用いる．カバー率 $R1$ ，正解率 $P1$ ， $P2$ の計算方法を以下に示す．

$$\text{カバー率 (R1)} = \frac{\text{出力がある入力文の数}}{\text{入力文の数}}$$

$$\text{正解率 (P1)} = \frac{\text{全パターン中に含まれる A,B 評価の数}}{\text{全出力英語パターン数}}$$

$$\text{正解率 (P2)} = \frac{\text{A,B 評価が 1 つでもある入力文の数}}{\text{出力がある入力文の数}}$$

7.2 推定結果

実験結果を表5および表6に示す。

表 5: 文種別 1 の推定値

各条件	カバー率 $R1$	正解率 $P1$	正解率 $P2$
条件	80%(8/10)	52.9%(9/17)	50.0%(4/8)
条件	100%(10/10)	20.4%(10/49)	60.0%(6/10)
条件	30%(3/10)	14.8%(4/27)	66.7%(2/3)
条件	20%(2/10)	50.0%(1/2)	50.0%(1/2)

表 6: 文種別 3 の推定値

各条件	カバー率 $R1$	正解率 $P1$	正解率 $P2$
条件	80%(8/10)	36.3%(12/33)	62.5%(5/8)
条件	90%(9/10)	22.5%(14/62)	100%(9/9)
条件	30%(3/10)	0.0%(0/4)	0%(0/3)
条件	0%(0/10)	(0/0)	(0/0)

この結果をみると、6つの真理項をすべて一致させて検索することで、正解率 $P1$ が高く、条件 および ではカバー率 $R1$ と正解率 $P2$ が共に高いことが判る。つまり、大量の翻訳候補の中に1つでも翻訳に利用できる適切なパターンが含まれていることを意味している。

7.3 実際の検索条件による実験

前節の分析結果を実証するために，文種別 1 および文種別 3 において其々 2 文ずつ，出力された英文パターンから，条件 ，条件 ，条件 の条件毎に 10 文を選択する．そして選択された英文パターンの評価を行った．評価結果を表 7 および表 8 に示す．

表 7: 文種別 1 の実測値

各条件	カバー率 $R1$	正解率 $P1$	正解率 $P2$
条件	100%(2/2)	46.6%(14/30)	50%(1/2)
条件	100%(2/2)	3.3%(1/30)	50%(1/2)
条件	100%(2/2)	13.3%(4/30)	50%(1/2)

表 8: 文種別 3 の実測値

各条件	カバー率 $R1$	正解率 $P1$	正解率 $P2$
条件	100%(2/2)	33.3%(10/30)	100%(2/2)
条件	100%(2/2)	16.6%(5/30)	100%(2/2)
条件	100%(2/2)	10.0%(3/30)	100%(2/2)

この表 7，表 8 をみると，表 5，表 6 に近い値が得られている．よって表 5，表 6 は，ある程度信頼性のおける値であると考えている．

8 考察

表 2, 3 の結果からわかる通り, 単一の真理項を検索条件とした場合, 全検索文から翻訳に使えるパターンは少ないことが判る. しかし, 表 7, 8 の結果より, 6 つの全真理項 (文種別+統語構造+節間キーワード+節間の意味分類+従属節の意味+主節の意味) を指定した場合, カバー率 $R1$ が 80% と, ある程度高い値が得られることが示された. また統語構造+従属節の意味+節間の意味+節間キーワードを指定した場合, 正解率 $P2$ が 60.0% から 100% と高い値が得られることがわかった.

カバー率 $R1$ と正解率 $P2$ の高さが本手法のメリットであり, 何らかの入力に対して必ず翻訳に利用できるパターンの出力が可能であることが判る. しかし残念なことに, 現時点では正解率 $P1$ はまだまだ低いので, 本手法の実装には正解率 $P1$ の向上に勤めなければならない. この為には, 評価方法でも記述した様に, 検索した英文パターンの字面の問題解決, もしくは生成文の作成方法の改善のみならず, 真理項の組み合わせによる検索条件の見直しも考えられる.

ここで, 本研究を引き継ぐ際には, 字面の問題はパターン辞書の問題にも繋がるため, この点についてはパターン辞書作成の研究で行うことであり, 本研究で改善に尽力する必要性は少ない. ただし残りの 2 点, 「生成文の作成方法」と「検索条件である真理項の組み合わせ」の再検討が, 正解率向上に対する今後の課題と考えられる. 特に真理項の組み合わせ検索では, 実測した入力文の数が重文, 複文共に 2 文しか無いため, 結果にやや不安が残る. 又, 今回の実験では, 真理項の組み合わせの設定が, 各真理項の影響力を判断するには難しい設定である. 追実験を行うには, 以下の様に単純な設定にする方法が良い.

本研究の結果より, 条件の緩和により, ひとつ前の条件と比較するだけで各真理項の影響力が判る様にする. 特に複文では従属節の真理項の影響が強いため, 早期に組み合わせ条件から除外する.

例 1：複文の真理項調査の場合

1. 全真理項

文種別+統語構造+節間キーワード+節間の意味分類+従属節の意味+主節の意味

2. 前の条件より従属節の意味を除外

文種別+統語構造+節間キーワード+節間の意味分類+主節の意味

3. 前の条件より主節の意味を除外

文種別+統語構造+節間キーワード+節間の意味分類

4. 前の条件より節間の意味を除外

文種別+統語構造+節間キーワード

5. 前の条件より節間キーワードを除外

文種別+統語構造

ただし、先に提案した方法にはある問題が存在する。それは、文章への影響力は各真理項の単独の力では無く、他の真理項と相生した影響力の効果がある場合、先の提案では判別が困難である。しかし、全ての可能性について実験するのは困難であるため、いくつかの条件に限定する他ない。そこで、敢えて提案するとしたら、真理項が統語的分類と意味的分類にわかれていることに着目し、以下の10通りの限定実験も今後の課題に考慮できる。

例 2：真理項の相生効果への対策調査例

1. 統語的構造全て

- ・ 文種別+統語構造+節間キーワード

2. 統語的構造の真理項単独

- ・ 文種別+統語構造
- ・ 文種別+節間キーワード

3. 意味的分類全て

文種別+節間の意味分類+従属節の意味+主節の意味

4. 意味的分類のうち，2つの網羅

文種別+節間の意味分類+従属節の意味

文種別+節間の意味分類+主節の意味

文種別+従属節の意味+主節の意味

5. 意味的分類の真理項単独

文種別+節間の意味分類

文種別+従属節の意味

文種別+主節の意味

9 あとがき

本研究では，真理項を用いた英文パターン検索実験を行い，意味的に適切な英文パターンを出力するための最適な真理項の検索条件について調査した．

その結果，6つの全真理項(文種別+統語構造+節間キーワード+節間の意味分類+従属節の意味+主節の意味)を指定した場合，カバー率 $R1$ が80%と，ある程度高い値が得られることが示された．また統語構造+従属節の意味+節間の意味+節間キーワードを指定した場合，正解率 $P2$ が60.0%から100%と高い値が得られることがわかった．真理項のパラメータには木構造になっているものがあるため，今後の課題としては，検索条件である真理項の組み合わせと，生成文の作成方法の再検討による正解率の向上を考えていきたい．

10 感想

この感想部分では，研究に直接関わることは無いが，本研究を引き継ぐ際の参考の一部となる「感想」のようなものを記載する．

10.1 従属節の影響力

単一の真理項を検索条件とした場合の結果が悪いことは前述した通りである．(この結果に対する改善方法は考察でも述べている)しかし，この結果の中で，複文の従属節だけは影響力が大きいことが判っている．この理由として(私見ではあるが)，重文とは違い，複文では従属節の後に「主節」となる部分が明確には存在しないからであると思われる．ここで，重文と複文の構成に再び触れながら説明しようと思う．

・重文の構成例

1. 連用節 + 主節
2. 副詞節 + 主節
3. 並列節 + 主節

・複文の構成例

1. 連体節 + 名詞 + 述部
2. 補足節 + 述部
3. 名詞修飾節 + 名詞 + 述部

重文では，従属節となる連用，副詞，並列節が主節全体を修飾するが，複文では，従属節がすぐ後にある品詞部分しか修飾していない．その結果，複文では従属節の表現を明確にすることで，他者に文の意味を伝えやすくなる作用があり，表現者が無意識に従属節の表現を明確しているからだと思われる．

10.2 主，従属節の分類強化

次に，従属節と主節の分類方法の強化についての提案を記載する．本研究で用いた真理項はまだ改良の余地があり，(残念ながら，本研究ではその必要性について提案する程の成果を挙げることができなかったが) その結果としてさらなる検索精度向上が考えられるからである．まず，従属節や主節の分類方法について付録に記載してあることとは別の角度で説明する．

そもそも，従属節や主節の「意味」という表現自体が具体的ではない．私見ではあるが，分類方法から見れば，従属節(もしくは主節)の「用言」(動詞，形容詞，形容動詞)という表現の方が利用者にも判り易い．これは「意味」というものを「変化や状態についての情報」という面で(この考え方が正しいか判らないが)考えているからであると思われる．又，体言には無数のバリエーションがあり，体言での意味の特定は困難であると共に，本研究の始動時(2005年)において，体言の掛かり先等の研究が完了していないという側面も考えられる．いずれにしても，本稿では従属節(もしくは主節)の意味とは，節の用言であると思って差し支えないと思われる．

以上を踏まえた上での提案として，分類方法の強化とは分類する種類の増加と厳密化である．例えば「変化の現象」の分類には，抽象表現の「なる」の他にも「見直す」や「入れ替える」等が同一分類されている．この3つの表現を別々の分類にすることができるはずである．具体的には，以下の様にする．

・従来

「変化の現象」: なる，見直す，入れ替える

・分類強化後

1. 「変化の現象」: なる
2. 「修正」: 見直す
3. 「交替」: 入れ替える

10.3 意外なメリット

本手法では、一旦、文章のパターンや字面を、役割や意味の分類に変換するため、文章表現が乱雑であったり、方言があっても支障がほとんど無い。例えば、本手法を利用する際、“意味”検索プログラムに直接、文章を入力するであろうが、仮にスラム的な文章表現でも、形態素解析による解析可能な範囲であれば真理項の抽出ができる。又、地方には「寒い」の動詞形として、「寒かりよる」という表現(形容動詞といった方が良いかも知れない)がある。この様に、あまり利用されない(筆者も使ったことがない)表現でも、「温度変化」の分類に追加してしまえば即座に利用可能である。つまり、本手法の運用は表現自体に左右されない強みがあり、形態素解析が不可能な表現を除いて、方言混じり、くだけた表現等も現時点のロジックで充分に対応できるのである。

謝辞

本研究は、科学技術振興事業団「JST」の戦略的基礎研究推進事業「CREST」における研究領域「高度メディア社会の生活情報技術」の研究課題「セマンティックタイポロジーによる言語の等価変換と生成技術」の支援により行いました。

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，文型パターンの分類方法の検討，研究資料，2005.
- [2] 岡田ほか，多変量解析による最適文型パターンの選択方式，言語処理学会第11回年次大会，pp.25-28，2005.
- [3] 藤本，埋め込み文に対する日英翻訳規則の作成について，言語処理学会第10回年次大会，pp.245-248，2004.
- [4] 前田ほか，パターンを使用した重文複文の日英翻訳の精度，言語処理学会第10回年次大会，pp.237-240，2004.