意味類型を用いた日英文型パターンの対応づけ

斉藤健太郎　池原 悟　村上 仁一

鳥取大学大学院 知能情報工学研究科
〒 680-0947 鳥取県鳥取市湖山町南 4-101
{ksaito,ikehara,murakami}@ike.tottori-u.ac.jp

あらまし:
自然言語処理の技術的発展を図る上で意味解析の重要性が指摘され, 意味を構成する原言語の表現をパターンとし, 目的言語の表現と対応づける方法が提案されている. 柔軟で意味的に適切な翻訳を実現する方法の 1 つとして, 意味類型を用いてパターンを対応づける方法が考えられる. そこで本研究は比較文を例としてとりあげ, 意味的, 構造的に対応する目的言語のパターンを求める仕組みを示し, その中から文脈上適切と考えられる訳語パターンを選択する手法を提案する. さらにオープンテストにより, 性能評価, 検討を行う.

キーワード : 機械翻訳 , パターン, 意味類型

Coresponding a pattern to semantic typology

Kentaro SAITO　Satoru IKEHARA　Jin'ichi MURAKAMI

Graduate School of Information and Knowledge Engineering, Tottori University
4-101, Minami Koyamachou, Tottori city, 680-0947, Japan

Abstract:
In developing natural language processing technology, it is proposed how to make the original language as patterns, which consist of meaning, correspond to the objective language. As one of flexible and appropriate translation method, it is thought how to correspond a pattern to semantic typology. Then this paper treats comparison sentences as examples. And this paper proposes construction searching for a pattern of objective language, a pattern which is made to correspond to meaningly and structurally. Also this paper proposes a method which chooses word expression that is adequate in context. Father more, we evaluate capability computation and results by open test.

Keywords : machine translation, semantic typology, pattern

1. はじめに

自然言語処理の技術的発展を図る上で意味解析の重要性が指摘され, 意味を構成する原言語[1]の表現をパターンとし, 目的言語[2]の表現と対応づける方法 [1] が提案されている. しかしこの手法では, 日本語表現と英語表現の対応が単一的である. そのため, 文脈等を考慮したとき, 対応している表現の中に適切な訳語表現がない場合がある.

柔軟で意味的に適切な翻訳を実現する方法の 1 つとして, 意味類型[3]を介して 1 つの原言語表現に複数の目的言語表現を対応づける方法 [2] が考えられる. しかし意味類型を介した手法を用いると訳文に用いる表現の候補が複数となるため, 訳語表現を一意に決定できないという問題点がある [3].

そこで本研究はまず比較文型を例としてとりあげ, 意味類型を用いて原言語のパターンに意味的・構造的に対応する目的言語のパターンを求める仕組みを示す. また訳語の表現を一意に決定できない問題点を解決するために，パターン外の単語やパターン内の変数に着目し, 複数のパターンの候補の中から文脈より最適と考えられる訳語のパターンを決定する手法を示す.

2. 従来の研究と意味類型の必要性

2.1 研究の背景

従来の機械翻訳では, 文を単語に分解して単語単位に訳語を決定したのち, 単語間の文法的関係に基づいてそれらを組み合わせ訳語を合成する方法や, 連語や慣用表現などを単語に分解せずに翻訳する方法が行われてきた. しかし現実の言語表現では文を単語に分解する過程で, 文全体の意味が失われ元の意味が再現されなくなるという問題点がある．

この問題を解決する方法の 1 つとして, 意味のまとまりとなる原言語の表現をパターンとし, 意味的に目的言語の表現と対応づけることが挙げられる. この方法によって意味的に適切な対応関係が得られると期待される.

2.2 意味類型を用いたパターンの対応づけ

意味類型を用いた手法の例を従来の方法の例と対比して示す.

[1]言語 A から言語 B へ翻訳する際，基となる言語 A
[2]言語 A から言語 B へ翻訳する際，対象となる言語 B
[3]意味ごとにまとめたもの

見出し語動詞	日本語文型パターン	品詞の意味属性	英語文型パターン
入れる	*N1*が*N2*に*N3*を入れる	*N1* < 主体 > *N2* < 施設 > *N3* < 設備、機械 >	*N1* install *N3* in *N2*.
	*N1*が*N2*に*N3*を入れる	*N1* < 人 > *N2* < 飲物 > *N3* < 人 >	*N1* make *N2* for *N3*.

図 1: 日本語語彙体系に定義されたパターン対の例

図 1 は日本語語彙体系 [1] に示されている表現対応づけの例である. この図では「$N1$ が $N2$ に $N3$ をいれる」の場合を例にとると，$N1$ が < 主体 > ，$N2$ が < 施設 > ，$N3$ が < 設備，機械 > の意味属性を持てば「$N1$ install $N3$ in $N2$.」に対応し，$N1$ < 人 > ，$N2$ < 飲物 > ，$N3$ < 人 > の場合は「$N1$ make $N2$ for $N3$.」に対応することを示す . このように日本語語彙体系では, 動詞と名詞の組合わせでパターンが 1 対 1 で対応しており, 単文レベルで大幅の翻訳精度の向上が見られた. しかしパターンの対応が単一的で, 文脈に応じて適切なパターンを選ぶことができないという問題点がある.

これに対して, 図 2 は意味類型を用いて日本語と英語のパターンを対応づける方法の例である. この手法は意味類系を介して, 原言語のパターンと目的言語のパターンが複数対複数で対応されるため，訳語の表現を選択する自由度が増し，文脈に最も適した目的言語の表現を選択できる可能性がある.

図 2: 意味類型を用いた日英パターンの対応例

しかしこの手法は，1 つの原言語のパターンに複数の目的言語のパターンが対応するため, 目的言語のパターンを一意に選択する手法が必要となる.

このような意味類型を介した翻訳を実現するためには，まず 1 つの原言語のパターンに意味的・構造的に対応している複数の目的言語のパターンを対応づける必要がある．そこで本研究では, 比較の文を例としてとりあげ，比較の代表的な意味，細かい付加的な意味，品詞の種類・数に着目して対応づける仕組みを作成する. また，複数の目的言語のパターンの中からパターンを一意に選択するため，パターンの外の単語やパターンの中の変数に着目し文脈上適切と考えられるパターンを選択する手法を示す.

3. 意味類型を用いたパターン対応方法

3.1 用語の定義

以下に本論文で用いる用語の定義を行う.

- 比較文... 2 つ以上の事柄を比べあわせ, そこに認められる異同について考えている文のこと

  以下に比較文の例を示す．例文 3 は対訳の比較文の例である．

  (例文 1) 彼は私に比べ, 背が高い．

  (例文 2) 彼は私くらい疲れている．

  (例文 3) あなたは馬にはかなわない.

  You are no match for a horse.

  上の定義に該当してる文でも, 以下のような文は比較文とせず, 本研究の対象外とする．

  - 例文 4 のように比較の対象が明示されておらず, 暗黙に世間一般の平均値と認められる文

    (例文 4) 彼はどちらかというと背が高い
  - 例文 5 のように比較する 2 物の属性が互いに異なっている文

    (例文 5) 彼は倒れそうなくらい疲れてる
  - 対訳文において, 英文は比較を表す文 (as-as や more-than などを用いた文) であっても, 日本文が比較文の定義にあてはまらない文

    (例文 6) 私はたった 500 円しかもってない.

    I have no more than 500 yen.

- 対応パターン... 原言語のパターンに対応している目的言語のパターンのこと (例：「$N1$ is more K than $N2$.」に対して「$N1$ は $N2$ より K だ」.)
- 同値パターン群... 意味と品詞の種類・数が同じパターンの集合 (例：$N1$ is more K than $N2$.，$N2$ is not as K as $N1$．などのパターン群)
- 意味ラベル... 比較の種類，比較する内容などに着目し，比較の代表的な意味を表したもの. (例：「$N1$ は $N2$ より K だ」の意味ラベルは [同級] となる)
- 概念キーワード... 比較における度合・性質に着目し，意味的な概念を表したもの (例：観点・強調・様態・否定・好印象・多数比較, 一者比較・変化・程度... など [4])
- 付加意味... 概念キーワードをパターンやパターンの外の品詞に付加的につけたもの. (例：「$N1$ is Ker than any other $N2$.」の付加意味は [強調] となる．)
- 品詞列... N[名詞],V[動詞],K[形容詞]，F[副詞] のように，パターンの中の品詞の種類と数を示したもの．

  (例：「$N1$ は $N2$ と同じくらい K だ」の品詞列は [$N1$,$N2$,K] となる)
- パターン外意味... 原文中のパターンに含まれない単語 (副詞等) に概念キーワードを付加したもの．

  (例文 7)：的場君は驚いたことに私より背が高い

  例文 7 は「$N1$ は $N2$ より K だ」というパターンにあてはまる「驚いたことに」のパターン外意味として [強調，感嘆] となる．
- 変数意味情報... パターンの中の変数に概念キーワードを付加したもの．

  (例文 8)：彼は私よりもテニスがうまい

  例文 8 は「$N1$ は $N2$ より K だ」というパターンにあてはまる「うまい」の変数意味情報として，[K：好印象] となる．

3.2 全体の流れ

本研究では，意味類型を用いて原言語のパターンを目的言語のパターンに対応づけることを目的としている．対応している目的言語のパターンを探す手順を以下の図 3 に示す.

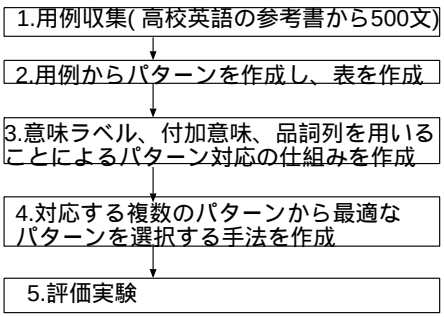

図 3: 本研究の流れ

3.3 パターンの作成

本研究では，パターンを作成するため，比較の定義に従い高校英語の参考書から用例を 500 文程度収集する．また図 3 のパターン対応の仕組みにおいて，構造と意味の面から対応づけるため，パターン表を作成においてパターンを意味ラベル・付加意味・品詞列ごとに分類する. パターン表を作成する手順を以下に示す.

図 3 の手順によって得られるパターン表の一部を例として表 1・2 に示す.

パターンは，同値パターン群ごとにまとめられており, 表中の「 同値番 」の欄の数字は同値パターン群ごとの番号である. また「 同値番 」欄のカッコ内は意味ラベルを表し, 中カッコ内は付加的意味を表す.「 番 」欄の数字はパターンごとの番号を表す．

3.4 意味的・構造的なパターンの対応

意味類型を介してパターンを対応づけるために，原言語のパターンを目的言語のパターンに意味的・構造的に対応づける仕組みが必要となる.

まず意味的に対応しているパターンを網羅的に探すため，原言語のパターンの意味ラベルを目的言語のパターンの意味ラベルと対応づける．意味ラベルを用いてパターンを検索する例を以下に示す．

(例 1) 表 1 から意味ラベル [比較級] に対応する英語パターンを検索すると，同値パターンの番号 5，6，7，8 が出力される．

次に意味的に対応しているパターンを詳細に探すため，原言語のパターンの付加意味を目的言語

表 1: 日本語のパターン表の一部

同値番	番	パターン	品詞列
1(同級)	1A	$N1$ は $N2$ と同じように K だ	$N1$,$N2$,K
	1B	$N1$ は $N2$ くらい K だ	$N1$,$N2$,K
2(同級)	2A	$N1$ は $N2$ と同じ程度 FV する	$N1$,$N2$,F,V
3(同然)	3A	$N1$ と $N2$ は同じだ	$N1$,$N2$
	3B	$N1$ は $N2$ も同然だ	$N1$,$N2$
	3C	$N1$ は $N2$ のようなものだ	$N1$,$N2$
	3D	$N1$ と $N2$ は同じものだ	$N1$,$N2$
4(同級)	4A	$N1$ は $K1$ と同じくらい $K2$ だ	$N1$,$K1$,$K2$
[一者]	4B	$N1$ は $K1$ であるように $K2$ だ	$N1$,$K1$,$K2$
5(比較)	5A	$N1$ は $N2$ ほど K でない	$N1$,$N2$,K
	5B	$N2$ は $N1$ に比べて K だ	$N1$,$N2$,K
	5C	$N2$ は $N1$ より K だ	$N1$,$N2$,K
	5D	$N2$ は $N1$ に対して $K1$ だ	$N1$,$N2$,K
6(比較)	6A	$N1$ は $N2$ ほど FV しない	$N1$,$N2$,F,V
7(比較)	7A	$N1$ は $N2$ より $N3$ が劣る	$N1$,$N2$,$N3$
[否定]	7B	$N1$ は $N2$ より $N3$ がよくない	$N1$,$N2$,$N3$
8(比較)	8A	$N1$ は $N2$ より上で V する	$N1$,$N2$,V
[位置]	8B	$N1$ は $N2$ の上で V する	$N1$,$N2$,V

表 2: 英語のパターン表の一部

同値番	番	パターン	品詞列
1(同級)	1a	$N1$ is as K as $N2$.	$N1$,$N2$,K
	1b	$N1$ is no less K than $N2$.	$N1$,$N2$,K
	1c	$N1$ is K equivalent to $N2$.	$N1$,$N2$,K
2(同級)	2a	$N1$ V as F as $N2$.	$N1$,$N2$,F,V
3(関係)	3a	$N1$ is no more $N2$ than $N3$.	$N1$,$N2$,$N3$
4(同然)	4a	$N1$ is the same as $N2$.	$N1$,$N2$
	4b	$N1$ is as good as $N2$.	$N1$,$N2$
	4c	$N1$ is almost like $N2$.	$N1$,$N2$
	4d	$N1$ is $N2$ in effect.	$N1$,$N2$
5(比較)	5a	$N2$ not V as F as $N1$.	$N1$,$N2$,V,F
	5b	$N1$ V more F than $N2$.[4]	$N1$,$N2$,V,F
6(比較)	6a	$N1$ is inferior to $N2$ in $N3$.	$N1$,$N2$,$N3$
[否定]	6b	$N1$ is worse than $N2$ in $N3$.	$N1$,$N2$,$N3$
7(比較)	7a	$N1$ excel over $N2$.	$N1$,$N2$
[肯定]	7b	$N1$ is superior to $N2$ in $N3$.	$N1$,$N2$,$N3$
8(比較)	8a	$N2$ is not K enough for $N1$.	$N1$,$N2$,K
	8b	$N2$ is not as K as $N1$.	$N1$,$N2$,K
	8c	$N2$ is less K than $N1$.	$N1$,$N2$,K
	8d	$N1$ is more K than $N2$.	$N1$,$N2$,K

のパターンの付加意味と対応づける．付加意味を用いてパターンをさがしだす例を以下に示す．

(例 2) 表 1 から意味ラベル [比較級] の中からさらに付加意味が [否定] のパターンを検索すると，同値パターンの番号 6 が出力される．

また構造的に適切なパターンを細かく探すため，原言語のパターンの品詞列を目的言語のパターンの品詞列と対応づける．品詞列を用いてパターンをさがしだす例を以下に示す．

(例 3) 表 1 から意味ラベル [比較級] の中から，さらに品詞列 [$N1$,$N2$,F,V] のパターンを検索すると，同値パターンの番号 5 が出力される．

このように本研究では，原言語のパターンの意味ラベル，付加意味，品詞列を目的言語のパターンに対応づけることで，意味的・構造的に対応している同値パターン群を一意に求めることができる．

3.5 パターンを対応づける仕組み

原言語のパターンを目的言語のパターンと対応づける手順を図 4 に示す.

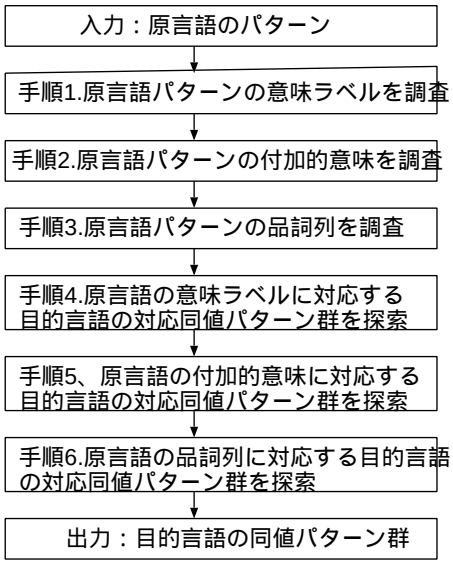

図 4: パターンを対応づける仕組み

表 1・2 を用いて，目的言語のパターンをしぼりこむ例を以下に示す. 原言語のパターンに対応している目的言語のパターンの総数を残存したパターン数とする．ここでは残存したパターン数は表 1 の英語パターンの総数，20 からはじまる．

(例 4) 日本語パターン 2A「$N1$ は $N2$ と同じ程度 FV する」に対応する英語パターンをもとめる手順を以下に示す．
手順 1：2A の意味ラベル = 同級
手順 2：2A の付加的意味情報 = ナシ
手順 3：2A の構造 (品詞列) = $N1$,$N2$,F,V
手順 4 : 意味ラベル「同級」に対応する英語同値パターン番号 = 1,2
(残存したパターン数 20 → 5)
手順 5：省略
手順 6：手順 4 で出力された結果の中で, 品詞列が $N1$,$N2$,F,V の英語同値パターン番号 = 2
(残存したパターン数 5 → 2)

日本語パターン 2A に対応する同値パターン番号は 2 で, 対応パターンは 2a「$N1$ V as F as $N2$.」と 2b「$N1$ V no less F than $N2$.」の 2 つとなる.

例 4 より手順 4 ー 6 の各手順で残存したパターン数が減少し, 同値パターン一つまでしぼりこまれていることがわかる. しかし図 4 の全ての手順が終了した後も複数のパターンが対応している場合が多いため, その中から文脈に応じた最適なパターンを選択する方法が必要となる.

4. パターンを文脈に応じて選択する方法

文脈に応じて最適なパターンを選択する手法の流れを以下の図 5 に示す.

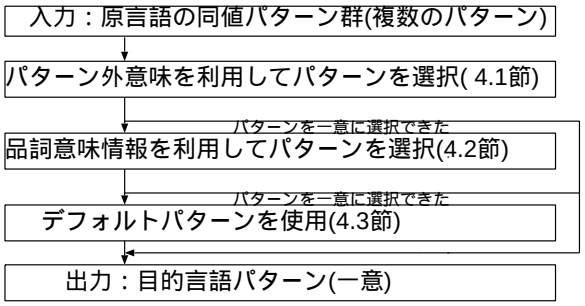

図 5: パターンを一意に選択する手法の流れ

図 5 中のパターン外意味を利用してパターンを選択する方法を 4.1 で示す．品詞意味情報を利用してパターンを選択する方法を 4.2 で示す．またデフォルトパターンの設定方法を 4.3 で示す．

4.1 パターン外意味によるパターンを選択

文脈に応じた最適なパターンを選択する時に，原文中にパターンに含まれない単語があれば，この単語の意味を利用して目的言語のパターンを選択できる．

そこで本研究は，パターン外意味とパターンの付加意味を比較することで，文脈上最も適切なパターンを選択する．以下にパターン外意味を利用してパターンを選択する方式の例を示す.

(例文 9)「私の考えも結局のところ君と同じだ」

例文 9 は表 1 より日本語パターン 5B「$N1$ は $N2$ と同じだ」にあてはまる. パターン 5B の対応パターンとして, パターン対応の仕組みによって表 2 から 4 パターンが出力される. 出力されたパターンを付加意味とともに以下に示す.

- 4a.$N1$ is the same as $N2$.[-]
- 4b.$N1$ is as good as $N2$.[重々しい]
- 4c.$N1$ is almost like $N2$.[やわらげ]
- 4d.$N1$ is $N2$ in effect.[強調，重々しい]

原文のパターンに含まれない単語として「結局のところ」には，[強調，重々しい] の意味があるため, 上記の 4 パターンのうち最も意味的に一致した 4d を選択する.

4.2 品詞意味情報によるパターンを選択

単文には，パターンに含まれない単語がない場合がある．この理由からパターン外意味だけでの最適なパターンの選択は，不十分だと考えられる. このため，本研究ではパターンに含まれない単語がない場合，原文中の変数意味情報を目的言語のパターンの変数意味情報と比較し, 最も一致したパターンを意味的に最適なパターンとして選択する.

以下に品詞意味情報を用いた手法の例を示す.

(例文 10)「父の命令は法律同然であった」

例文 10 は，表 1 より日本語パターン 3B 「$N1$ は $N2$ と同然である」にあてはまる. パターン 3B に対応するパターンとして, パターン対応の仕組みによって表 2 から 4 パターンが出力される. 出力されたパターンを変数意味情報とともに以下に示す.

- 4a.$N1$ is the same as $N2$.[-]
- 4b.$N1$ is as good as $N2$.[$N1$ 重々しい,$N2$ 重々しい]
- 4c.$N1$ is almost like $N2$.[$N2$ 強調, 重々しい]
- 4d.$N1$ is $N2$ in effect.[$N2$ 重々しい]

例文 10 には，パターン外に含まれない単語がないため, パターン外の意味を利用してパターンを選択することができない. しかし原文中の「$N1$: 父の命令」「$N2$: 法律」の変数意味情報が [$N1$ 重々しい,$N2$ 重々しい] となるため, 上記の 4 パターンのうち, 変数意味情報が最も一致した 4b.「$N1$ is as good as $N2$.」を選択することができる.

4.3 デフォルトパターンの設定

4.1 と 4.2 で示した手法を用いて一意にパターンを選択できなかった場合, あらかじめパターンごとに登録しておいたデフォルトパターンを用いてパターンを一意に選択する. 以下にデフォルトパターンの設定方法を示す.

- 用例中で原言語のパターンに対応している頻度の最も高い目的言語のパターンをデフォルトパターンとして設定
- 上記の設定方法において，用例が同数の場合や用例の存在しないパターンの場合, 人手で設定

以下にデフォルトパターンの設定の例を示す．

例 5 :「$N1$ は $N2$ と同じように K だ」に対応するパターンと「$N1$ は $N2$ と同じように K だ」を用いた用例の数を以下に示す．

中カッコ内の数字は「$N1$ は $N2$ と同じように K だ」の用例中に対訳の表現として出て来た回数である．対応している頻度の最も高い 1a をデフォルトパターンとする．

- 1a．$\underline{N1 \text{ is as } K \text{ as } N2.}$[6]
- 1b．$N1$ is no less K than $N2$.[1]
- 1c．$N1$ is K equivalent to $N2$.[0]

5 評価実験

本手法の性能評価のため，評価実験を行う．

5.1 パターンの総数と対応の仕組みのパターンしぼりこみ割合の検討

パターンの数を調べるため，比較文 500 文を高校参考書から収集する．作成したパターンの総数を調査する．パターンの総数を表 3 に示す.

表 3 : パターンの総数

	日本語	英語
パターン総数	361	252

また，図 4 の手順ごとのしぼりこみの割合を調べるため, 対応パターンの数が減少している割合を調査する．評価として手順後の対応パターンの数を残存したパターン数として求め，パターンの減少している割合を以下の式で表す．その結果を表 4 に示す

$$\text{減少率} = \frac{\text{手順を行った後の残存したパターン数}}{\text{手順を行う前の残存したパターン数}}$$

表 4 : 対応の仕組みの残存したパターン数と減少率

手順名	英→日	日→英
手順前 (デフォルト数)	361(-)	252(-)
手順 4(意味ラベル使用)	23.1(93 %)	16.2(93 %)
手順 5(付加意味使用)	6.8(71 %)	4.8(70 %)
手順 6(品詞列使用)	4.2(64 %)	2.8(69 %)

表 4 中の数値は，残存したパターンの数でカッコ内の数値が減少率を表す．表 4 より手順 4 ー 6 の各手順で残存したパターンの数が減少し, パターンの候補がしぼりこまれていることがわかる. しかし手順の 4 から 6 でしぼりこんだ後も残存したパターン数が 1 になっていないことから，最適なパターンを一意に選択する手法が必要となることがわかる．

5.2 オープンテスト

5.2.1 オープンテストの流れ

本手法の性能を調査するため，新聞記事, アンカー英和辞典・和英辞典・学研英和辞典・基本語用例辞典・英語表現辞典の対訳コーパスから比較単文を中心に 192 文抽出し，評価した．以下にオープンテストの手順を図 6 に示す.

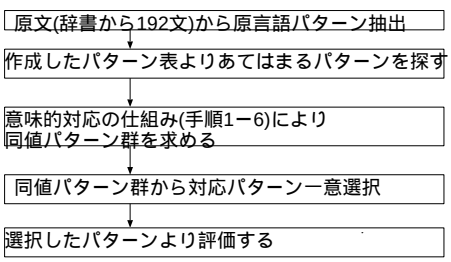

図 6: 意味類型を用いたパターン対応・選択の流れ

オープンテストの評価方法を以下に示す．

$$\text{カバー率} = \frac{\text{パターン表に適合できたパターン数}}{\text{全パターン数}}$$

「平均同値パターン数」は，対応の仕組みによって同値パターン群の対応パターンの数と定義する．

「一意対応割合」は，原言語のパターンと目的言語のパターンが 1 対 1 で対応していたものであり，パターンを一意に選択する手法を用いる必要がなかった割合と定義する．

また「パターン外使用割合・変数情報使用割合」は，パターン外意味・変数意味情報を利用して文脈からパターンをした割合と定義する．

5.2.2 オープンテストの結果

オープンテストの結果を表 5 に示す.

表 5 : オープンテストの結果

	英→日	日→英
カバー率	68 % (129/192)	74 % (141/192)
平均同値パターン数	5.1	4.5
一意対応割合	0 % (0/129)	8 % (14/141)
パターン外使用割合	42 % (54/129)	43 % (62/141)
変数情報使用割合	9 % (12/129)	16 % (22/141)

表 5 の「平均同値パターン数」と「一意対応割合」より 90 %以上の文でパターンが一意に決定せず，平均 5 パターン程度の中から一意に選択する必要があることがわかる．また本手法のパターンを文脈に応じて選択する手法を英語から日本語は 51 %, 日本語から英語は 59 %の割合で適用できたことがわかる．

文脈に応じてパターンを選択する手法とデフォルトパターンを用いる手法のどちらが最適なパターンを選択するのに適しているかを調べるため，文脈に応じて選択したパターンとデフォルトパターンのどちらが文脈上適しているかを比較した．その結果を表 6 に示す．

表 6 : 文脈から選択とデフォルトの比較

文脈より選択したパターン	59 %
デフォルトパターン	41 %

表 6 より文脈に応じてパターンを選択した方が最適なパターンを選択できた割合が高い．これにより意味類型を用いて最適なパターンを文脈より選択できる可能性がある．しかし最適なパターン

の適否は著者の判断によって決められており，今後はネイティブによるさらに正確な判断が必要となる．

6 考察

6.1 文脈によって選択できなかった例

文脈よりパターンを選択できなかった例を示す.

1. 意味的に参考になる単語が存在しない

(例文 11)「中国に比べると日本は小さい」

例文 11 の原文は，表 1 より 5B「$N1$ は $N2$ に比べると K だ」という日本語パターンにあてはまる. しかし例文 11 は，パターン外にも変数にも付加的な意味をもつ単語が存在しないため, 文脈からパターンを選択できない．

2. パターンを一意に選択できない

(例文 12)「彼の答えは拒絶も同然だ」

例文 12 の原文は，3B「$N1$ は $N2$ も同然だ」というパターンにあてはまる. 対応の仕組みを用いると, 表 2 より同値パターン群として 4 パターンを出力される. 以下にそのパターンを示す．

- 4a.$N1$ is the same as $N2$.[-]
- 4b.$N1$ is as good as $N2$.[$N1$ 重々しい,$N2$ 重々しい]
- 4c.$N1$ is almost like $N2$.[$N2$ 強調, 重々しい]
- 4d.$N1$ is $N2$ in effect.[$N2$ 重々]

原文中のパターンに含まれない単語がなく「$N2$: 拒絶」の変数意味情報が [$N2$ 重々] であるため, 4b と 4c と 4d が対応している. しかし本手法では, これ以上のしぼりこみができない.

1，2 における問題点を解決するため, 変数の意味属性を用いて対応づけるなどの手法が必要となる.

6.2 文脈による選択とデフォルトとの比較

表 6 より文脈に応じて選択したパターンとデフォルトパターンを比較した結果，文脈に応じて選択したパターンが適しているが高いことがわかった．しかしデフォルトパターンの方が適しているパターンとなる場合もあった．その例を以下に示す．

(例文 13)「昨日より今日の方がやや暖かい」

例文 13 の原文は，5C「$N1$ は $N2$ より K だ」というパターンにあてはまる. パターン外意味として「やや」が [やわらげ] であるため「$N2$ is not K enough for $N1$.[やわらげ，否定]」を文脈より選択することが可能である．それに対しデフォルトパターンとして求められるのは「$N1$ is more K than $N2$. となる．

例文 13 の場合，デフォルトパターンの方が意味的に適切と考えられる．この理由として，パターン外意味とパターンの付加意味に一致していない部分があるため，意味的に適切でないパターンが選択されたことが挙げられる．このような場合，付加意味を重みづけし，異なる意味を選んだ場合の得点を差し引くことで，異なる意味をもつパターンを選択しない方法が考えられる.

7. おわりに

本稿では, 比較文を例として取り上げ, 原言語のパターンに意味的・構造的に対応する目的言語のパターンを探しだす仕組みを作成した. またパターン外の単語やパターンの変数の意味に着目し, 複数のパターンの候補の中から文脈に応じて最適なパターンを選択する手法を提案した.

本手法の性能調査のため, 対訳コーパスを用いてオープンテストを行った. その結果, 英語から日本語は 52 %, 日本語から英語は 59 %の割合で, パターンを文脈から選択できることがわかった. またデフォルトパターンとの比較を行った結果，文脈から選択したパターンの方が適しているパターンである割合が高いことがわかった．これにより意味類型を用いて最適なパターンを選択できる可能性が見えたと言える．

今後は, ネイティブの判断によって，最適なパターンの選択する手法の正確な精度を調べる必要がある，またその精度を高めるため付加意味に重みづけするなどの手法を用いる必要がある.

参考文献

[1] 池原, 宮崎, 白井, 横尾, 中岩, 小倉, 大山, 林：日本語語彙体系, 岩波書店 (1997)
[2] 池原 悟：意味類型に着目した日英意味辞書の研究企画, 辞書プロジェクト会議 (2000/4/1-2)
[3] 斉藤健太郎, 池原 悟, 村上仁一:日英比較表現のパターン化と意味的対応関係, 第 7 回言語処理学会,P241-244(2001.3)
[4] グループ・ジャマシイ (1998)：日本語文型辞典, くろしお出版