

# 実用文を対象とした格フレーム選択における 大規模格パターン辞書の効果と問題分析

石原 啓史 宮崎 正弘

新潟大学自然科学研究科

{hiroshii, miyazaki}@nlp.ie.niigata-u.ac.jp

## 要旨

文の意味を解析する際に、格フレームの同定問題は重要である。格パターンマッチ処理の結果を左右する大きな要素として格パターン辞書とパターンマッチ評価関数が挙げられる。格パターンマッチに関する既存の手法には、経験則から得られた複雑な評価関数を用いた手法が目立ち、格パターン辞書自体の性能が評価困難であった。本稿では、単純な評価関数を用いることによって、辞書本来の性能を実用文を用いて評価し問題を分析する。同時に本研究で用いた日本語語彙大系構文体系と IPAL 基本動詞辞書の性能差についても論ずる。

## An Evaluation of Japanese case frame selection using large-scale case patterns dictionary for ordinary sentence

Ishihara Hiroshi, Miyazaki Masahiro

Graduate School of Science and Technology, Niigata University

{hiroshii, miyazaki}@nlp.ie.niigata-u.ac.jp

## Abstract

It is necessary to identify a case frame on semantic analysis of Japanese sentence. On case pattern matching, dictionary and evaluation function have a great influence on the precision. Ordinary heuristic method were too complex to evaluate the performance of case patterns dictionaries. We take Goi-Taikei-Japanese Lexicon and IPAL basic verbs dictionary for evaluation. We analyze case frame of ordinary sentences using a simple evaluation function, and compare the ability of dictionaries.

## 1 はじめに

格パターンの同定問題は、自然言語処理の重要なテーマの1つである。この問題に対していくつかの手法が提案されているが、実用文を試験文とした評価についてはほとんど報告されていない。

また格パターンマッチングアルゴリズムについても辞書に特化したものが多く、マッチングシステムの比較などが行いにくい。そのため辞書やアルゴリズムの比較評価が難しい。

本稿は辞書に最適化した複雑な評価関数を用いる格パターンマッチングではなく、十分に単純な評価関数を用いたマッチングを行ない、その違いから格パターン辞書自体の性能を評価分析する。

そして大規模格パターン辞書として日本語語彙大

系構文体系[1]、IPAL 基本動詞辞書[2]を取り上げ、実用文にどれだけ対応できるかについて評価し、問題点を分析する。

## 2 一般的な格パターンマッチ

一般的なマッチング手法としては、単文に対して各格要素と動詞を抽出する。そして、その動詞をキーとして候補となる格パターンを各パターン辞書から取得する。次に、格パターンに付与されている格毎の名詞制約に例文における同じ格の名詞を当てはめる。これを必要な格要素について行い、得点化することによって格パターンを選択する。

この名詞制約については大きく2通りがあり、単語を記述する手法と、単語グループを定義し、その

グループの包含関係を記述する手法である。主に後者の方が制約を照合する際に扱いやすい。

また、そのシソーラスに配置された名詞のグループを利用する場合、所属グループ間の「距離」を得点化することもある。最後に算出した得点を基に候補となった格パターンから最適格パターンを選択する。なお、この時に得点を扱うアルゴリズムについて、今後「評価関数」と呼称する。

### 3 評価関数が最適化された格パターンマッチング手法

本稿は辞書が公開されており再現実験可能なマッチング手法 [3] (以降「黒橋方式」と呼称) を比較対象とした。

この手法では、まず名詞の制約として分類語彙表 [4] を用いて、格パターンには IPAL 基本動詞辞書 (以降 IPAL と呼称) を利用している。なお IPAL の格パターン収録形態は以下のとおりとなる。

表 1: 「走る」に関する IPAL の格パターン (一部)

文型	N1 が N2 ヲ
格形式 1	ガ
意味素性 1	HUM/ANI
名詞句 2	彼/猫
格形式 2	ヲ
意味素性 2	CON/LOC
名詞句 2	彼女, 彼女の裸, 彼女の後ろ姿, 山, 金魚, 海, 壁, 景色, 空/上, 下, 横, 彼女の方

このように格パターンに対して各格要素に出現する名詞の例が「名詞句」として示されており、マッチングでは、この名詞句の格要素に対して比較が行われ、それを点数化する。

点数化には 2 つの要素、4 つのパラメータが存在する。1 つ目の要素は格要素の名詞における得点化で、これから得られるパラメータを類似度と呼んでいる。類似度は分類語彙表の 5 桁の分類番号を利用しており、関係は表 2 に示す。

レベル	0	1	2	3	4	5	6	一致
類似度	0	0	5	7	8	9	10	11

2 つ目の要素は格の種類と数に関して、パターン中の例文と入力文の共通格の数について得点化し 3 つのパラメータを設定する。なお、格の種類としては「が」「を」「に」の 3 格を必須格、それ以外を任意の格として区別し扱う。

- l: 入力文の全格数から、必須格以外でかつ例文の格要素に存在しない格数を引いた格数
- m: 例文の全格数から、必須格以外でかつ入力文の格要素に存在しない格数を引いた格数
- n: 入力文と例文の共通格要素数

以上の 4 つの変数を以下の評価関数に代入して得られる値がそのパターンの得点となる。

$$\text{得点} = \frac{(\text{類似度の和})}{\sqrt{l}} \cdot \sqrt{\frac{n}{l}} \cdot \sqrt{\frac{n}{m}}$$

### 4 単純な評価関数を用いる格パターンマッチング手法

ここでは格パターン辞書本来の性能評価を前提としたマッチング機構の評価を目的として、複雑なパラメータや評価関数の最適化をするのではなく、極力簡素な手法でのマッチングを試みる (以降「本方式」と呼称)。

本方式は名詞制約は日本語語彙大系意味体系 (以後「意味体系辞書」と呼称)、格パターンには同辞書の構文体系 (以後「構文体系辞書」と呼称) を用いている。なお、構文体系辞書の格パターン収録形体は表 3 を参照されたい。

表 3: 構文体系辞書「走る」の一部

種類	文型	制約条件
一般	結果 (動作 受身不可)	N1(4 人) が N2(1252 夢) を見る
字面	身体動作 (動作)	N1(4 人) が N2(“辞書/字引/字引き/辞典/字典”) を見る
慣用	感情動作 (状態 受身不可)	N1(4 人) が 馬鹿を見る

この辞書は名詞の意味制約として、カテゴリで表示する「一般表現」(カテゴリ表現) と字面で示す「慣用表現」、括弧内の字面で示す「字面表現」の 3 種類ある。今回は「字面表現」は「一般表現」と同様

に扱う。また、慣用表現についての扱いには若干注意が必要となる。その点については本方式のマッチングアルゴリズムで示す。意味体系辞書は上位・下位関係のシソーラス構造(付図1を参照)をしており、その全てのノードにはカテゴリ番号が付与されている。この番号を基に名詞の制約条件を得る。マッチングアルゴリズムは以下に基づく。

- 格要素の評価は得点制とし、1 マッチにつき1点。減点は考慮しない。
- マッチを調べる格要素は、パターンに含まれる格要素のみ。
- 取得得点の一番高いもの(複数可)を最適パターンとする。
- 慣用表現マッチがあった例文に対しては、それを含むパターンを優先し、複数の慣用表現パターンにマッチした場合は、一般の格と同様にマッチした格について1点づつの加算で決定する。
- 最適パターンは1点以上とする。

また、この評価関数を式に示すと以下となる。

$$\text{得点} = (\text{マッチした格数})$$

なお、カテゴリのマッチ条件はそのカテゴリの概念とその下位概念に含まれる単語を対象とする。以上のことを考慮した場合のマッチングの例を表4に示す。

入力文: 彼(23,48)が農業(1959)で食べる

表4: 「彼が農業で食べる」のマッチング状況

得点	文型	制約条件
2	状態(状態受身不可)	N1(4-361人)が N2(1936-2028仕事)で食べる
1	身体動作(動作)	N1(4-361人/535-670動物)が N2(2838-862食料/534-705生物)を食べる

また、慣用表現のパターンに含まれる格要素が発見されると自動的に慣用表現と判断されるため、慣用表現である「店で油を売る」を「店で油を販売する」の意味で捉えることはできない。これは文脈解析を必要とするため、今回の研究対象からは除外する。

## 5 第三の手法

上記の2つの手法は、「手法」「辞書」がともに異なっているため、性能差がどちらから引き起こされているか判断が難しい。従って「辞書」「手法」の一方が異なった手法を同時に比較する必要がある。ただし黒橋方式の評価関数を意味体系辞書の対応させるのは辞書の構造の違いから難しく、またパラメータの設定などにおいて妥当性の証明も困難である。

それぞれの手法について表5に示す。今回は「本研究のマッチング方式」+「黒橋方式の名詞制約・格パターン辞書」の組み合わせを試みた(以降「第三方式」と呼称)。ただ、IPALの格パターン制約には分類語彙表の分類番号を名詞制約に用いておらず、独自の「意味素性」を付与している。この問題を解決するため、分類語彙表で発見された単語を意味体系辞書で再度検索し、意味体系辞書におけるカテゴリ番号とIPAL意味素性のマッピングテーブル[5]を利用することにより意味素性を得ることとした。

なお、分類語彙表と意味体系辞書のシソーラスの精度の違いについては、カテゴリ番号と意味素性の対応を行なっている点でシソーラスの段数による細分効果が減少する。

さらに、名詞制約の評価を目的に、分類語彙表で発見された単語ではなく、意味体系辞書そのものを参照した方式も用意した(以降「第四方式」と呼称)。

マッチングの評価に関しては、本方式のマッチング方式に近づけるため、格パターンの意味素性と入力文の意味素性が一致することに1点とし、その累計によって格パターンを選択する。

表5: 格手法の違い

方式	評価関数(手法)	格パターン
黒橋方式	$\frac{(\text{類似度の和})}{\sqrt{n}} \cdot \sqrt{\frac{n}{l}} \cdot \sqrt{\frac{n}{m}}$	IPAL
本方式	マッチした格数	構文体系辞書
第三方式	マッチした格数	IPAL
第四方式	マッチした格数	IPAL
第五方式	$\frac{(\text{類似度の和})}{\sqrt{n}} \cdot \sqrt{\frac{n}{l}} \cdot \sqrt{\frac{n}{m}}$	構文体系辞書

例文「彼(HUM/23,48)がグランド(ORG/367,LOC/462)を走る。」についての評価は表6となる。例文、入力文について複数の意味素性があった場合その組み合わせのどれかでマッチすれば1点とみなす。最高点は2点となり、意味記述「(足を使って)早くすすむ」「何かが早く動く」の2通りが同点候補となる。

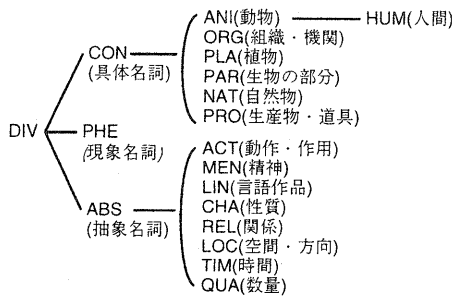


図 1: IPAL 意味素性の大系

表 6: 第三の手法におけるマッチング状況

得点	意味記述	格の制約
2	(足を使って)早く進む	(HUM/ANI/PRO)が(LOC/ABS)を(LOC)から(LOC)に/へ
2	何かが早く動く	(PHE/CON)が(LOC)を/に
1	ある目的のために動き回る	(HUM/ANI)が(ACT)に
1	好ましくない傾向に進む	(HUM/ORG)が(ABS)に

なお、各辞書の違いについては、付表2を参照されたい。

## 6 実験

### 6.1 実験用例文について

本稿の実験で使用した例文は、日経新聞社説B('97 9/7)全文(以降日経新聞と省略)及び朝日新聞社説('97 12/11)全文(以降朝日新聞と省略)である。この例文は、構文解析システム評価を目的として人手で形態素解析、構文解析レベルの解析ミスが修正が行なわれている。

なお実際のマッチング処理は単文として行なわれるため、単文数は例文中の動詞に依存する。また動詞に係る格要素が存在しない単文は対象外とする。それぞれの数は以下の通り。

表 7: 例文について

例文名	例文数	動詞数	単文数
日経新聞	57	127	111
朝日新聞	69	138	120

### 6.2 評価の判断

この例文を上記の3つの手法について格パターンマッチ処理を行ない、その結果が正しいかどうかは人手で判断する。評価の判断分けは以下の6通りを定めた。

- 一意選択正解——一意の解が得られ、そのパターンが正解
- 複数選択正解——複数の解が得られ、その中に正解パターンが存在
- 一意選択不正解——一意の解が得られ、そのパターンが不正解
- 複数選択不正解——複数の解が得られ、その中に正解パターンが存在しない
- パターンなし——動詞のエントリ存在するが、格要素がどのパターンとも一致しない
- エントリなし——パターン辞書にその動詞が存在していない

### 6.3 受け身、使役等の扱い

受け身や使役の助動詞「れる」「られる」「せる」「させる」が付いた動詞に対しての格パターンは用意されていない。そこで、「に」「を」格を「が」格に変換するなどの格変化が伴う処理が必要となる。しかし、これを機械処理するための方策はヒューリスティックに頼るところが大きい。単純に出現頻度からどれかに決定してしまう方法(「れる」「られる」の多くは受け身として使用される)も可能であるが、これは辞書の性能評価をする上では妨げとなる。

ここでは、方式の比較実験であることを考慮し、実験の公平性は保てるため、この問題は無視することとする。

### 6.4 パターンを決定づける格

パターンを決定づける格が不足している場合、例えば「彼が走る」といった場合、表6におけるパターン全てに可能性がある。この場合の正解は一意に決定するのは人間にも難しい。これは文脈や状況の情報が必要となり、本実験の対象からは除外とし、どれも正解と判断する。

### 6.5 実験結果

実験結果を表8に示す。なお、正解/不正解の項目にある括弧の中の値はマッチング結果を出力した単文数を全体とした場合の比率である。

表 8: 実験結果

方式		黒橋	本方式	第三	第四
日経	一意選択正解	18.9% (42.9%)	32.4% (50.0%)	6.3% (14.3%)	6.3% (13.7%)
	複数選択正解	10.8% (24.5%)	18.9% (29.2%)	27.0% (61.2%)	27.9% (13.7%)
	一意選択不正解	10.8% (24.5%)	2.7% (4.2%)	3.6% (8.2%)	4.5% (9.8%)
	複数選択不正解	3.6% (8.2%)	10.8% (16.7%)	7.2% (16.3%)	7.2% (15.7%)
	パターンなし	11.7%	12.6%	11.7%	9.9%
	エントリなし	44.1%	22.5%	44.1%	44.1%
朝日	一意選択正解	23.3% (48.3%)	30.0% (42.4%)	7.5% (17.0%)	7.5% (16.4%)
	複数選択正解	5.0% (10.3%)	20.8% (29.4%)	21.7% (49.1%)	21.7% (47.3%)
	一意選択不正解	13.3% (27.6%)	5.8% (8.2%)	6.7% (15.1%)	7.5% (16.4%)
	複数選択不正解	6.7% (13.8%)	14.2% (20.0%)	8.3% (18.9%)	9.2% (20.0%)
	パターンなし	18.3%	11.7%	22.5%	20.8%
	エントリなし	33.3%	17.5%	33.3%	33.3%

## 7 評価

### 7.1 名詞制約

全体を評価する前に、第三、第四方式の評価を行なう。この2方式の違いは名詞制約の辞書の収録語数だけである。結果にはほとんど違いはみられず、収録語の違いが解析結果に現れていない。よって単純に辞書の収録語数を増やしても解析率の向上にはつながらない。よって今後は第三方式のみを比較する。

### 7.2 一意正解

本来、格パターンのマッチングを処理に利用する上で一意に選択されない場合を考慮する必要はない。パターン選択結果における一意正解率では黒橋方式と本方式では若干本方式が上まわり、それに対して第三方式は大きく下回る。第三方式と黒橋方式は辞書が同じであることを考えると、最適化された評価関数によってかなり解析率の向上が得られる。しかし一意決定には不向きであるマッチング法を用いた本方式と黒橋方式の差を考れば格パターンマッチの一意選択に最も有効なことはパターン辞書と名詞ソーラスの高精度化であると言える。

### 7.3 複数選択正解を含む正解率

「一意正解率」+「複数選択正解」の点から評価すると、本方式は、黒橋方式、第三方式に比べて大きく向上している。これも格パターン辞書と名詞ソーラスの高精度化が大きく寄与している。なお、黒橋方式と第三方式の違いについては、ほぼ同様の実験が[3]の論文中に掲載されていたが、同様に差がないことから、この実験が「評価関数の最適化が正解率の向上」には影響しないことを、実用文について実証したことになる。

### 7.4 不正解について

本実験で目立って不正解の原因となったのが、動詞「する」「ある」「なる」である(表9)。特に構文体系辞書では相当の多義が発生してしまう。この大量の多義の場合に単に「複数解答」と言える状況ではない。今回は7を基準として、それより多い多義は不正解としている。この7という値についての根拠は、以下の「ある」「する」「なる」について、7パターン以上出力した場合には、人間が見てパターンを決定していると言えない結果であることから決定している。ちなみにこの3動詞が抽出単文の10%程度を占めている。

表 9: 多義の多い動詞

名詞	する	ある	なる
IPAL 辞書	19	15	23
構文体系辞書	319	219	203

また動詞が受身表現だったものに関しては、本実験は格変化処理をせずに受け身、使役の助動詞を除去しただけでマッチング処理しているため、そのほとんどが不正解に含まれる。なお、受け身、使役表現は抽出単文全体の5%~10%である。

### 7.5 マッチせず/エントリ無しについて

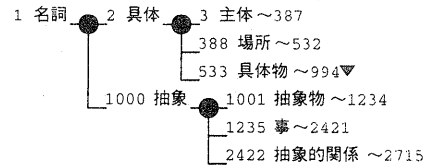
格パターンマッチの結果を出力できない状態が「マッチせず」と「エントリ無し」であり、これは本方式に対して黒橋方式、第三方式が大きく上回っている。これは格パターン辞書の規模の違いである。

特に「エントリ無し」については、動詞のエントリ自体が存在しない訳であり、これは格パターン辞書の問題である。

「マッチせず」については、動詞についての適切

な格パターンが無い場合と、名詞制約におけるミスの2つの可能性がある。前者の問題は、やはり格パターン辞書を充実することで解決するが、後者については名詞シソーラスとの兼ね合いがある。例えば、「プログラムで食べる」の「プログラム」は意味体系辞書では「目録」としての意味しかない。同様に分類語彙表についても「辞書・目録・暦」といったカテゴリに分類されている。よって名詞シソーラスそのものを改良する他ない。

## A 付録



付図 1: 意味体系辞書の構造

## 8 おわりに

- IPAL 動詞辞書に比べて構文体系辞書はパターンの網羅性が高い
- 名詞シソーラスの収録語数について、マッチ処理を行なう上では意味体系辞書と分類語彙表との間で大きな差はない
- 経験則による複雑な評価関数は一意選択率を向上させるが、辞書性能によっても一意選択性が同程度に向上する
- 経験則の評価関数が解析率を向上させることはない
- 実用文においては、どの辞書を用いても3割程度の未収録の動詞、パターンが発生する
- 名詞シソーラスの分解能を十分考慮した格パターン辞書における名詞制約の設定が重要である

## 9 謝辞

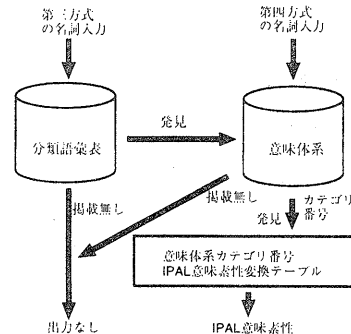
日本語語彙大系の利用許可を頂いた NTT コミュニケーション科学基礎研究所の関係各位に感謝します。

## 参考文献

- [1] 池原, 宮崎, 白井, 横尾, 中岩, 小倉, 大山, 林: 日本語語彙大系, 岩波書店 (1997)
- [2] 計算機用日本語基本動詞辞書 IPAL(Basic Verbs) 説明書, 情報処理振興事業協会技術センター (1987)
- [3] 黒橋禎夫, 長尾真: 格フレーム選択における意味マーカと例文の有効性について, 情報処理学会自然言語処理研究会報告, 91-11(1992)
- [4] 国立国語研究所: 分類語彙表, 秀英出版 (1964)
- [5] 小林佳徳, 宮崎正弘: 格パターンと用例を用いたハイブリッド型の多義動詞意味解析法, 第 54 回情報処理学会全国大会, 2-85(1997)

付表 1: カテゴリ番号と意味素性の対応表

カテゴリ番号	意味素性	カテゴリ番号	意味素性
535 ~ 551	ANI	4 ~ 361	HUM
362 ~ 387	ORG	671 ~ 705	PLA
552 ~ 670	PAR	707 ~ 759	NAT
760 ~ 994	PRO	2304 ~ 2421	PHE
1237 ~ 1559	MEN	1002 ~ 1234	LIN
2483 ~ 2582	CHA	2423 ~ 2482	REL
2670 ~ 2715	TIM	2585 ~ 2609	QUA
388 ~ 532	LOC	2610 ~ 2669	LOC



付図 2: 名詞辞書の検索方式

付表 2: 辞書の規模と精度

辞書	用言数	総エントリ
格パターン辞書		
IPAL 動詞辞書	861	3,379
構文体系辞書	6,000	14,800
名詞シソーラス	単語数	シソーラス段数
分類語彙表	32,600	5
意味体系辞書	400,000	12