

述語と項の位置関係ごとの候補比較による日本語述語項構造解析

林部 祐太[†]・小町 守^{††}・松本 裕治[†]

一般に、項は述語に近いところにあるという特性がある。そのため、従来の述語項構造解析の研究では、候補を述語との位置関係でグループ分けし、あらかじめ求めておいたグループ間の優先順序に従って正解項を探索してきた。しかしながら、その方法には異なるグループに属する候補同士の比較ができないという問題がある。そこで我々は、異なるグループごとに最尤候補を選出し、それらの中から最終的な出力を決めるモデルを提案する。このモデルは優先度の高いグループに属する候補以外も参照することによって最終的な決定を行うことができ、全体的な最適化が可能である。実験では、提案手法は優先順序に従う解析よりも精度が向上することを確認した。

キーワード：述語項構造解析, 項と述語の位置関係, 探索先行分類型モデル

Japanese Predicate Argument Structure Analysis by Comparing Candidates in Different Positional Relations between Predicate and Arguments

YUTA HAYASHIBE[†], MAMORU KOMACHI^{††} and YUJI MATSUMOTO[†]

In general, arguments are located near the predicate. A previous study has exploited this characteristic to group candidates by positional relations between a predicate and its candidate arguments and then searched for the final candidate using a pre-determined priority list of the groups. However, in such an analysis, candidates in different groups cannot be compared. Therefore, we propose a Japanese predicate argument structure analysis model that gathers the most likely candidates from all the groups and then selects the final candidate amongst them. We can account for candidates with less priority before making a final decision to perform global optimization. Experimental results show that our model outperforms deterministic models.

Key Words: *predicate argument structure analysis, positional relations between predicates and arguments, selection-then-classification model*

[†] 奈良先端科学技術大学院大学, Nara Institute of Science and Technology

^{††} 首都大学東京, Tokyo Metropolitan University

1 はじめに

述語項構造解析の目的は、述語とそれらの項を文の意味的な構成単位として、文章から「誰が何をどうした」という意味的な関係を抽出することである。これは、機械翻訳や自動要約などの自然言語処理の応用において重要なタスクの1つである (Surdeanu, Harabagiu, Williams, and Aarseth 2003; Wu and Fung 2009)。

述語は文の主要部で、他の要素とともに文を構成する (日本語記述文法研究会 2010)。日本語では、述語は品詞によって、形容詞述語・動詞述語・名詞述語の3種類に分けられる。述語が意味をなすためには、補語 (主語を含む) が必要であり、それらは項と呼ばれる。また、述語と項の意味的な関係を表すラベルを格と呼ぶ。

項は前後文脈から推測できるとき省略¹されることがあり、省略された項をゼロ代名詞、ゼロ代名詞が指示する要素を先行詞と呼ぶ。この言語現象はゼロ照応と呼ばれ、日本語では項の省略がたびたび起きることから、述語項構造解析はゼロ照応解析としても扱われてきた (河原, 黒橋 2004; 笹野, 黒橋 2011)。

本稿では、項と述語の位置関係の種類を次の4種類に分類する。述語と同一文内にあり係り受け関係にある項²、(ゼロ代名詞の先行詞として同一文中に存在する) 文内ゼロ、(ゼロ代名詞の先行詞として述語とは異なる文中に存在する) 文間ゼロ、および (文章中には存在しない) 外界項である。本稿では、それぞれ *INTRA_D*, *INTRA_Z*, *INTER*, *EXO* と呼ぶ。ある述語がある格にて項を持たないときは、その述語の項は ARG_{NULL} だとし、その述語と ARG_{NULL} は *NULL* という位置関係にあるとして考える。本稿では、*EXO* と *NULL* を総称して *NO-ARG* と呼ぶ。例えば、例1において、「受け取った」と「食べた」のヲ格項「コロッケ」はそれぞれ *INTRA_D*・*INTRA_Z*、「飲んだ」のガ格項「彼女」は *INTER* で、ニ格項は ARG_{NULL} である。

コロッケを受け取った彼女は、急いで食べた。
(ϕ が) ジュースも飲んだ。 (1)

一般に、項は述語に近いところにあるという特性 (近距離特性) を持つ。そのため、これまでの述語項構造解析の研究では、この特性の利用を様々な形で試みてきた。

河原, 黒橋 (2004) や Taira, Fujita, and Nagata (2008) は項候補と述語の係り受け関係の種類ごとに項へのなりやすさの順序を定義し、その順序に従って項の探索を行った。また、Iida, Inui, and Matsumoto (2007) は述語と同一文内の候補を優先的に探索した。これらの先行研究ではあらかじめ定めておいた項の位置関係に基づく順序に従った探索を行い、項らしいものが見つかれば以降の探索はしない。そのため、異なる位置関係にある候補との「どちらがより項

¹ 本稿では、省略を項が述語と直接係り受け関係にないことと定義する。

² ここでの関係は向きを持たない。複数の項が同一の述語と関係を持つこともありうる。

らしいか」という相対的な比較は行えず、述語と項候補の情報から「どのくらい項としてふさわしいか」という絶対的な判断を行わなければならないという問題点がある。

そこで、本稿では、項の位置関係ごとに独立に最尤候補を選出した後、それらの中から最尤候補を1つ選出するというモデルを提案する。位置関係ごとに解析モデルを分けることで、柔軟に素性やモデルを設計できるようになる。また、位置関係の優先順序だけでなく、その他の情報（素性）も用いて総合的にどちらがより“項らしい”かが判断できるようになる。

本稿の実験では、まず、全ての候補を参照してから解析するモデルと、特定の候補を優先して探索するモデルを比較して、決定的な解析の良し悪しを分析する。また、陽に項の位置関係ごとの比較を行わないモデルや、優先順序に則った決定的な解析モデルと提案モデルを比較して、ガ格・ヲ格ではより高い性能を達成できたことも示す。

本稿の構成は以下のようになっている。まず2章で述語項構造解析の先行研究での位置関係と項へのなりやすさの優先順序の扱いについて紹介する。3章では提案手法について詳述し、4章では評価実験の設定について述べる。5章・6章では実験結果の分析を行い、7章でまとめを行う。

2 関連研究

ここでは、述語項構造解析の先行研究における、位置関係と項へのなりやすさの優先順序の扱いについて紹介する。先行研究と提案手法の概要を表1にまとめた。

2.1 決定的な解析を行う方法

2.1.1 優先順序を統計的に求める方法

河原、黒橋 (2004) は、解析をゼロ代名詞検出と先行詞同定の2段階に分け、統計的に求めた優先順序を先行詞同定の際に用いた。彼らの手法では、まず、格フレーム辞書に基づく格解析

表 1 先行研究と提案手法の概要

	項の位置関係の数	位置関係ごとに項同定モデルを分ける	決定的な解析
河原, 黒橋 (2004)	20		✓
Iida et al. (2007)	2	✓	✓
Taira et al. (2008)	7	✓	✓
笹野, 黒橋 (2011)	—		✓
Imamura, Saito, and Izumi (2009)	—		
吉川, 浅原, 松本 (2013) ³	—		
提案手法	3	✓	

³ 文間項は探索しない。

によって、ゼロ代名詞の検出を行う。そして、項が存在すると判断された場合は、あらかじめ求めておいた優先順序に従って候補を探索し、候補と格フレーム用例の類似度が閾値以上かつ分類器でも正例と分類される候補を先行詞として同定する。分類器は項の位置関係に関わらず、共通のものを作成した。素性には、格フレームとの類似度や品詞などを用いた。

彼らは、従属節、主節、埋め込み文などといった文・文章中の構造をもとに、項の位置関係(彼らは「位置カテゴリ」と呼んだ)を20種類に分類した。

彼らは、位置カテゴリごとに、先行詞の取りやすさを

$$\frac{\text{先行詞がその位置カテゴリにある回数}}{\text{その位置カテゴリにある先行詞候補の数}} \quad (1)$$

でスコア化した。そして、位置カテゴリごとに、京都大学テキストコーパス (Kawahara, Kurohashi, and Hasida 2002) からスコアを算出し、得られたスコアを降順にソートしてそれぞれの格について優先順序を得た。

2.1.2 文内候補を優先的に探索する方法

Iida et al. (2007) は、先行詞候補とゼロ代名詞の統語的關係をパターン化するために、木を分類するブースティングアルゴリズム BACT (工藤, 松本 2004) を用いた。BACT は木構造データを入力とし、全ての部分木の中から分類に寄与する部分木に対して大きな重みをつける。彼らは、先行詞候補とゼロ代名詞間の係り受け木や、関係を表す素性を、根ノードに子としてつなげて BACT の入力とした。

文間先行詞の同定には係り受け関係を利用できないため、彼らは先行詞の同定モデルを文内と文間に分け、文内候補を優先的に探索する以下の方法をとった。

- (1) 最尤先行詞同定モデル M_{10} で、文内最尤先行詞 C_1^* を求める
- (2) 照応性判定モデル M_{11} で、 C_1^* の先行詞らしさのスコア p_1 を求める。あらかじめ定めておいた閾値 θ_{intra} に対して、 $p_1 \geq \theta_{\text{intra}}$ であれば、 C_1^* を先行詞として決定する。そうでなければ (3) に進む。
- (3) 最尤先行詞同定モデル M_{20} で、文間最尤先行詞 C_2^* を求める
- (4) 照応性判定モデル M_{21} で、 C_2^* の先行詞らしさのスコア p_2 を求める。あらかじめ定めておいた閾値 θ_{inter} に対して、 $p_2 \geq \theta_{\text{inter}}$ であれば、 C_2^* を先行詞として決定する。そうでなければ、先行詞なしとする。

$M_{10} \dots M_{21}$ はそれぞれ BACT を使って学習・分類し、パラメータ θ_{intra} と θ_{inter} は、開発データを用いて最適なものを求める。この手法では、文内の最尤先行詞同定や照応性判定には文間の候補の情報は参照せずに、決定的に解析している。

2.1.3 優先順位を経験的に決める方法

Taira et al. (2008) は、決定リストを用いて全ての格の解析を同時に行う方法を提案した。決定リストは規則の集合に適用順位を付けたものであり、機械学習の結果を人が分析しやすいという特長がある。彼らは項の位置関係やヴォイス・機能語に加えて、単語の出現形・日本語語彙大系(池原, 宮崎, 白井, 横尾, 中岩, 小倉, 大山, 林 1997) から得られる意味カテゴリ・品詞のいずれか1つを加えたものを組として扱い、それぞれの組を1つの素性とした。そして、述語ごとに Support Vector Machine の学習で素性の重みを得て、素性を重みでソートしたものを決定リストとした。すなわち、1つの素性を1つの決定リストのルールとして扱った。

彼らは項の単位を単語とし、項の位置関係を係り受け関係に基づいて次の7種類に定義している。なお、fw と bw は追加的な種類で、その他の種類と兼ねることができる。

- Incoming Connection Type (ic): 項を含む文節が述語を含む文節に係っている
日米交渉_{ガ:進展} が 進展した
- Outgoing Connection Type (oc): 述語を含む文節が項を含む文節に係っている
衝動買_{ガ:買い}いた 新刊本
- Within the Same Phrase Type (sc): 項が述語と同じ文節内にある
日米交渉_{ガ:日} が
- Connection into Other Case role Types (ga_c, wo_c, ni_c): 項を含む文節が述語を含む文節に、他の格の項を介して係っている
トム_{ヲ:説得,ga_c} への 友人_{ガ:説得} による説得
- Non-connection Type (nc): 項が述語とは異なる文にある
- Forward Type (fw): 文章内にて、項が述語の前方にある
- Backward Type (bw): 文章内にて、項が述語の後方にある

実際の解析は、各述語について次の手順で行った。

- (1) ic, oc, ga_c, wo_c, ni_c について、決定リストを用いて項を決定する
- (2) (1) で決まらなかった格について、sc の決定リストを用いて項を決定する
- (3) 対象の述語が項を持つ確率が50%以上であれば(4)に進む
- (4) nc, fw, bw に関する決定リストを用いて項を決定する

この手法は経験的に、優先順位を

ic, oc, ga_c, wo_c, ni_c > sc >> nc, fw, bw

のように定めたといえる。ic, oc, ga_c, wo_c, ni_c 間での、探索の優先関係はない。

この方法は、格と項の位置関係を考慮しつつ、項になりやすいものから決めていくのが特徴である。ただし、着目している候補と述語の情報のみを用いて項らしいかどうかを判断していくため、必ずしも全ての候補を参照してから最終的な出力を決定するわけではなく、候補間でどれが項らしいかの相対的な判断は行われない。

2.1.4 述語と係り受け関係にある候補を優先的に項であるとみなす方法

笹野, 黒橋 (2011) は, 解析対象述語の格フレーム候補それぞれに対して, 次の手順で格フレームと談話要素の対応付け候補を生成した.

- 解析対象述語と直接係り受け関係にある談話要素を, 選ばれた格フレームの格スロットと対応付ける. 談話要素が係助詞をともなって出現した場合や, 被連体修飾節に出現した場合など, 複数の格スロットとの対応付けが考えられる場合は, 考えうるすべての対応付けを生成する.
- 上記の処理で生成された対応付け候補に対し, 対応付けられなかったガ格・ヲ格・ニ格と, 解析対象述語と係り受け関係にない談話要素の対応付けを行う.

そして, 対数線形モデルにて最も確率的評価が高い対応付けを解析結果として出力した. 素性には, 意味クラスや固有表現情報の他に, 出現格と出現位置に関する 85 個の 2 値素性も用いた.

この手法では, 格ごとに独立に解析を行なっているのではなく, 同時に解析を行なう. しかし, 述語と係り受け関係にある候補を優先的に項であるとみなすため, 係り受け関係にある候補と, 係り受け関係にない候補または他の文にある候補との比較は行えない.

2.2 優先順序を素性として表現する方法

位置関係と項へのなりやすさの関係を優先順序として利用し決定的な解析を行うのではなく, 素性として利用した研究もある.

2.2.1 最大エントロピー法を用いる方法

Imamura et al. (2009) は, 最大エントロピー法に基づく識別的モデルを用いた. 彼らは, 位置関係ごとにモデルを分けるのではなく, 素性として, 述語と候補の位置関係, 係り受け関係を用いた. そして候補集合に, 項を持たないことを示す特別な名詞句 NULL を加え, その中から最尤候補を同定するというモデル化を行った. なお, 候補数削減のため, 文間項候補は述語を含む文の直前の文に出現したものと, これまでの解析ですでに項として同定されたものに限定している.

この方法では格ごとにモデルは 1 つだけ学習すればよい. ただし, この手法では, 候補間の関係を素性として用いることはできない.

2.2.2 Markov Logic を用いる方法

吉川 他 (2013) は, Markov Logic を利用して, 文内の複数の述語の項構造解析を同時に行う手法を提案した. Markov Logic は一階述語論理と Markov Networks を組み合わせたもので, 一階述語論理式の矛盾をある程度のペナルティの上で認めることができる統計的關係学習の枠組

みである。彼らは項同定・項候補削減・格ラベル付与を同時に行うモデルを提案した。

彼らは、文間の項候補を加えるのは計算量の問題から困難だとしている。素性（観測述語）は述語と候補の係り受け関係などを用いた。

3 述語と項の位置関係ごとの候補比較による日本語述語項構造解析

先行研究では、優先順位の低い位置関係にある候補は参照されずに、解析が行われていた。この方法は、優先順位の高い位置関係にある項の同定の性能は上げることができるが、優先順位の低い位置関係にある候補の再現率は下げってしまうという問題点がある。また、優先順位の低い位置関係にある候補も参照してから最終的な決定を行った方が、全体的な解析性能が向上すると考える。

そこで我々は、探索とトーナメントの2つのフェーズからなる、位置関係ごとに最尤候補を求めてから最終的な出力を決めるモデルを提案する。これは、「探索」・「分類」という2つのフェーズを持つ探索先行分類型モデル (Iida, Inui, and Matsumoto 2005) に着想を得て、後半の分類フェーズをトーナメント式に置き換えたものである。なお、このモデルは格ごとに解析器を学習・使用する。

3.1 項構造解析における探索先行トーナメントモデル

3.1.1 探索

はじめのフェーズでは任意の項同定モデルを用いて INTRA_D, INTRA_Z, INTER の最尤候補を選出する。それぞれ異なる素性やモデルを用いてもよい。モデルには、述語と探索対象の候補を入力として与え、探索対象の候補の中の1つを出力させる。

3.1.2 トーナメント

次のフェーズでは探索フェーズで得られた3つの最尤候補を入力とし、そのうちの1つか“NO-ARG”を出力する。これにより、最尤候補のうちどれが正解項であるか、もしくは項を持たないかを判断する。

このフェーズは図1に示したように(a)から(c)の3つの2値分類モデルで構成される。なお、予備実験にて異なる順序を試したが、文内最尤候補同士を(a)にて直接比較できるこの順序の性能が最も高かった。

- (a) INTRA_D と INTRA_Z を比較して、よりその述語の項らしい方を選ぶ
- (b) INTER と (a) で選出された候補を比較して、よりその述語の項らしい方を選ぶ
- (c) (b) で選出された候補と “NO-ARG” を比較して、よりその述語の項らしい方を選ぶ

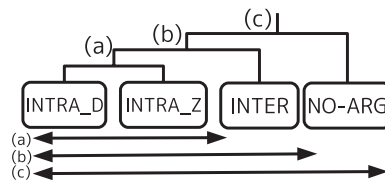


図 1 トーナメントフェーズでの、位置関係が異なる候補からの項の同定

(a) から (c) の分類器の学習事例には、Algorithm 1 で示すように探索フェーズで得られた最尤候補を用いる。

3.2 提案手法の関連研究

提案手法は2つのモデルを参考にしている。

1つ目は名詞句の照応解析における探索先行分類型モデル (*selection-then-classification model*) (Iida et al. 2005) である。このモデルは最初に、最尤先行詞を求める(彼らはこれを“探索”と呼んだ)。次に、その最尤先行詞を用いて、名詞句が実際に照応詞であるかどうかを判定する(彼らはこれを“分類”と呼んだ)。このモデルの利点は、照応性を持たない名詞句も学習事例の生成に使えることである。彼らは実験で、最尤先行詞を用いて照応性判定を行ったほうが、最尤先行詞を用いない場合よりも高い性能が出ることを確かめた。提案手法も、位置関係ごとに最尤候補を求めた後、どの候補が実際に項であるのかを判定する。最尤候補の探索を先に行なうことで、位置関係ごとの最尤候補を学習事例の生成に用いることができる。

2つ目はゼロ照応解析におけるトーナメントモデル(飯田, 乾, 松本 2004) である。そのモデルは、全ての先行詞候補(実際には先行する全ての名詞句)のペアに対して、どちらがより先行詞らしいかの2値分類を繰り返す。トーナメントモデルの利点は候補間の関係性の素性を使うことができる点である。提案手法のトーナメントフェーズでも同様に、トーナメントモデルを用いて、位置関係ごとに選出された最尤候補のペアからどちらが正解項らしいかの2値分類を繰り返し、候補間の比較を行うことができる。

4 評価実験

4.1 実験データセット

評価実験には NAIST テキストコーパス 1.4β (飯田, 小町, 井之上, 乾, 松本 2010) を用いた。これは京都大学テキストコーパス 3.0⁴ を基にしており、述語項構造, 事態性名詞の項構造,

⁴ <http://nlp.ist.i.kyoto-u.ac.jp/nl-resource/corpus/KyotoCorpus3.0.tar.gz>

Algorithm 1 分類器 (a) classifier_a, (b) classifier_b, (c) classifier_c の学習事例の作成アルゴリズム

procedure TRAIN(predicate, gold_argument, candidates)

gold_argument_type ← getArgumentsType(predicate, gold_argument)

▷ 正解項の位置関係を取得する

▷ 位置関係ごとに最尤候補を取得する

most_likely_candidate_INTRA_D ← getMostLikelyCandidate(predicate, candidates, INTRA_D)

most_likely_candidate_INTRA_Z ← getMostLikelyCandidate(predicate, candidates, INTRA_Z)

most_likely_candidate_INTER ← getMostLikelyCandidate(predicate, candidates, INTER)

if gold_argument_type = NO_ARG **then**

MakeExample(classifier_c, NO_ARG, predicate, most_likely_candidate_INTRA_D)

MakeExample(classifier_c, NO_ARG, predicate, most_likely_candidate_INTRA_Z)

MakeExample(classifier_c, NO_ARG, predicate, most_likely_candidate_INTER)

return

end if

MakeExample(classifier_c, HAVE_ARG, predicate, gold_argument)

if gold_argument_type = INTRA_D **then**

MakeExample(classifier_a, INTRA_D, predicate, gold_argument,
most_likely_candidate_INTRA_Z)

MakeExample(classifier_b, INTRA, predicate, gold_argument, most_likely_candidate_INTER)

else if gold_argument_type = INTRA_Z **then**

MakeExample(classifier_a, INTRA_Z, predicate, gold_argument,
most_likely_candidate_INTRA_D)

MakeExample(classifier_b, INTRA, predicate, gold_argument, most_likely_candidate_INTER)

else if gold_argument_type = INTER **then**

MakeExample(classifier_b, INTER, predicate, gold_argument, most_likely_candidate_INTRA_D)

MakeExample(classifier_b, INTER, predicate, gold_argument, most_likely_candidate_INTRA_Z)

end if

return

end procedure

procedure MAKEEXAMPLE(classifier, label, predicate, candidate1, candidate2)

▷ candidate2 は省略できる

項候補 candidate1, candidate2 が照応関係にあれば事例は作成しない。

述語 predicate と項候補 candidate1, candidate2 に対して, 素性集合 F を取得する。

学習器 classifier に対して, F を用いて, label をラベルとする学習事例を 1 つ作成する。

end procedure

共参照に関する情報が約 40,000 文の新聞記事にわたって付与されている。なお、アノテーションの誤りのため 6 記事⁵を除外した。このコーパスの記事を表 2 で示すように学習・開発（パラメータチューニング）・評価のために 3 分割した。これは、Taira et al. (2008) や吉川 他 (2013) と同じ分割方法である。表 3 に項の分布の統計情報を示す⁶。

4.2 実験設定

実験では、MeCab0.996・IPADIC-2.7.0-20070801 で解析して得られた形態素情報、京都大学テキストコーパス 3.0 で付与されている文節情報、CaboCha0.66 で解析して得られた係り受け関係を用いた。項の候補は文節単位で抽出した。解析は文頭から文末の順で行い、述語を含む文以降からは項候補を抽出しない。なお、ある述語の格についての解析結果は同じ述語の他の格についての解析に影響を及ぼさない。

本稿では項同定に焦点を絞るため、述語同定タスクには取り組まない。言い換えると、どれが述語であるかはあらかじめシステムに与えておく。述語には軽動詞「する」や複合動詞も含む。

最尤候補同定には、トーナメントモデル（飯田 他 2004）を用いた。その際、最尤候補の探索範囲ごとに異なるモデルを作成し、モデルの学習方法も（飯田 他 2004）に従った。例えば、提案手法は探索フェーズでは INTRA_D, INTRA_Z, INTER の最尤候補を同定するが、それぞれ

表 2 NAIST テキストコーパスの統計情報

	ニュース記事の範囲	社説記事の範囲	記事数	文数	述語数
学習	1月1日～11日	1月～8月	1,673	23,150	64,173
開発	1月12日～13日	9月	458	4,634	13,105
評価	1月14日～17日	10月～12月	662	8,795	24,296

表 3 NAIST テキストコーパスにおける項の分布

	格	SAME_BS	INTRA_D	INTRA_Z	INTER	EXO	NULL
学習	ガ格	128	33,775	12,057	7,436	10,529	248
	ヲ格	58	22,869	2,051	806	60	38,329
	ニ格	527	10,427	555	278	25	52,361
開発	ガ格	39	6,954	2,700	1,819	1,501	92
	ヲ格	10	4,724	445	147	18	7,761
	ニ格	51	2,673	212	80	5	10,084
評価	ガ格	40	12,805	4,829	3,241	3,140	241
	ヲ格	18	8,665	829	291	14	14,479
	ニ格	132	5,023	358	155	12	18,616

⁵ 除外した文書 ID : 951230038, 951225057, 950106156, 950106034, 951221047, 950106211

⁶ SAME_BS は項と述語が同一文節であることを示す。

異なる合計 3 つの解析モデルを最尤候補同定に用いる。

4.3 分類器と素性

探索フェーズ・トーナメントフェーズで用いる分類器には、Support Vector Machine (Cortes and Vapnik 1995) を線形カーネルで用いた。具体的には LIBLINEAR1.93⁷の実装を用い、開発データを用いたパラメータチューニングを行った。

素性には Imamura et al. (2009) で用いられたものとはほぼ同一の素性を用いた。

- 述語・項候補の主辞・機能語・その他の語の出現形・形態素情報
- 述語が受け身の助動詞を含むときはその原形
- 係り受け木上の述語と項候補の関係⁸

係り受け木上の項候補ノード N_a と述語ノード N_p からそれぞれ ROOT 方向に辿っていくときに初めて交叉するノードを N_c とし、 N_a から N_c までの道のりに含むノード列を $A_{a\dots c}$ 、 N_p から N_c までの道のりに含むノード列を $A_{p\dots c}$ とする。また、 N_c から木の ROOT までの道のりに含むノード列を A_{c_1, c_2, \dots, c_r} とする。本実験では、ノード列の文字列表現として、

- 主辞の原形
- 主辞の品詞
- 機能語の原形
- 機能語の品詞
- 機能語の原形 + 機能語の品詞

の 5 通りを用いた。 $A_{a\dots c}$ の文字列表現を $S_{a\dots c}$ 、 $A_{p\dots c}$ の文字列表現を $S_{p\dots c}$ とし、それらの連結を $S_{a\dots c} + S_{p\dots c}$ とする。

素性には、 $S_{a\dots c} + S_{p\dots c}$ 、 $S_{a\dots c} + S_{p\dots c} + S_{c_1}$ 、 $S_{a\dots c} + S_{p\dots c} + S_{c_1, c_2}$ 、 \dots 、 $S_{a\dots c} + S_{p\dots c} + S_{c_1, c_2, \dots, c_r}$ の $r+1$ 個の文字列を用いた。つまり、述語と項候補の関係を $5(r+1)$ 個の文字列で表現した。

- 係り受け木上の 2 つの項候補の関係
上と同様の素性表現を行った。
- 述語と項候補・2 つの項候補間の距離（文節単位・文単位ともに）
- 「述語・項候補の主辞・助詞」のコーパス中の共起スコア⁹

動詞と項の共起のモデル化は (藤田, 乾, 松本 2004) に従った。名詞 n が格助詞 c を介し

⁷ <http://www.csie.ntu.edu.tw/~cjlin/liblinear/>

⁸ Imamura et al. (2009) では、どのような素性表現に落とし込んだかは詳述されていない。

⁹ Imamura et al. (2009) では、これに相当するものとして、Good Turing スムージングを施した共起確率を用いている。計算は NAIST テキストコーパス相当部分を除いた 1991~2002 年の毎日新聞を用いた。

て動詞 v に係っているときの共起確率 $P(\langle v, c, n \rangle)$ を推定するため、 $\langle v, c, n \rangle$ を $\langle v, c \rangle$ と n の共起とみなす。

共起尺度には自己相互情報量 (Hindle 1990) を用いた。

$$PMI(\langle v, c \rangle, n) = \log \frac{P(\langle v, c, n \rangle)}{P(\langle v, c \rangle)P(n)}$$

なお、スムージングは行わなかった。自己相互情報量の算出には次の 2 つのコーパスを用い、2 つの値をそれぞれ二値素性として¹⁰用いた。

NEWS: 1995 年を除く 1991 年から 2003 年までの毎日新聞約 1,800 万文。MeCab0.98¹¹ で形態素解析を行い CaboCha0.60pre4¹² で係り受け解析を行った。辞書は NAIST Japanese Dictionary 0.6.3¹³ を用いた。約 2,700 万対の〈動詞, 格助詞, 名詞〉の組を抽出した¹⁴。

WEB: Kawahara and Kurohashi (2006) がウェブから収集した日本語約 5 億文。JUMAN¹⁵ で形態素解析を行い、KNP¹⁶ で係り受け解析を行なっている。KNP の項構造解析結果から約 53 億対の〈述語, 格助詞, 項〉の組を抽出した¹⁷。

- 項候補が以前の項構造解析で項となったか否かを示す 2 値情報
- 項候補の主辞の Salient Reference List (Nariyama 2002) における順位

4.4 比較対象

先行研究では、我々のものと異なる素性や機械学習の手法を使っており実験設定が異なる。そのため、ベースラインモデルとして IIDA2005、比較対象モデルとして IIDA2007・IIDA2007+・PPR- を実装し、位置関係ごとに最尤候補を求めてから最終的な出力を決める提案モデル PPR (Preferences based on Positional Relations) と比較する。

4.4.1 IIDA2005

位置関係に関わらずに、全ての候補の中から最尤の候補を探索フェーズで 1 つ選出した後、トーナメントフェーズでそれが項としてふさわしいか否かを判断するモデル。(Iida et al. 2005) の探索先行分類型モデルである。

全ての候補の中から 1 つを選ぶという点で (Imamura et al. 2009) とほぼ同等のモデルである。

¹⁰ 値が x 以下のときのみ発火する素性。実際には、 x を -4 から 4 まで 0.1 刻みで変化させた素性を用いた。

¹¹ <https://code.google.com/p/mecab/>

¹² <https://code.google.com/p/cabocha/>

¹³ <http://sourceforge.jp/projects/naist-jdic/>

¹⁴ 動詞が約 3 万種、名詞が約 32 万種で、ユニーク数は約 700 万組。

¹⁵ <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

¹⁶ <http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

¹⁷ 動詞が約 8 億種、項が約 2.8 億種で、ユニーク数は約 1.6 億組。

彼らのモデルと異なる主な点は、最尤候補同定と照応性判定を異なるモデルで行う点と、最尤候補同定時に2候補間の関係性も素性として用いる点である。

このベースラインモデルとその他のモデルと比較することで、項の位置関係によって探索の優先順序をつけることの効果や、位置関係ごとに最尤候補同定モデルを作り最尤候補同士の比較を陽に行う効果を調べる。

4.4.2 IIDA2007

文内最尤候補を選出した後、分類器が項としてふさわしいと判断すればそれを項として出力し、そうでなければ同様に文間候補の探索を行うモデル。2.1.2節で述べた (Iida et al. 2007) の文内候補を優先的に探索するモデルである。

彼らのモデルと異なる主な点は、最尤候補同定や候補の適格性判定を行う分類器に BACT ではなく SVM を用いる点である。

IIDA2005 と比較することで、文内候補を優先的に探索することの効果調べる。

4.4.3 IIDA2007+

INTRA_D の探索後、最尤候補が項としてふさわしいかどうかの判断（適格性判定）を行う。適格であればそれを出力し終了する。非適格であれば INTRA_Z の探索を行い、同様に適格性判定を行う。それも非適格であれば INTER の探索を行い、適格であればそれを出力し、非適格であれば項は無いと判断する。IIDA2005 と IIDA2007 の自然な拡張で、述語から統語的な距離の近いものを優先的に探索する。

IIDA2007 と比較することで、文内候補を細かく INTRA_D と INTRA_Z に分けて優先順序をつけることの効果を調べる。

4.4.4 PPR-

このモデルは、提案モデルとほぼ同じモデルであるが、INTRA_D と INTRA_Z を区別せずに、位置関係が INTRA と INTER の2グループであると仮定する。図1の(b)と(c)で示すようにトーナメントフェーズは2つの2値分類モデルからなる。分類器(c)は INTRA と INTER の候補のどちらが最尤候補であるかを判断する。

PPR と比較することで、文内の項の位置関係を細かく INTRA_D と INTRA_Z に分けて最尤候補同定モデルを作り、最尤候補同士の比較を行うことの効果を調べる。

4.4.5 比較対象とする先行研究

NAIST テキストコーパスを使い、全ての項の位置関係で実験を行なっている (Taira et al. 2008) と (Imamura et al. 2009) との比較も行う。ただし、本実験とは微妙に実験設定が異なるため、

厳密な比較はできないことに注意してほしい。

Taira et al. (2008) の実験では 19,501 個の述語をテストに、49,527 個を学習に、11,023 個を開発に使っている。また学習では京都大学テキストコーパス 4.0 で付与されている係り受け情報と形態素情報を用いているが、テストでは独自の係り受け解析器を用いている。

Imamura et al. (2009) の実験では、25,500 個の述語をテストに、67,145 個を学習に、13,594 個を開発に使っている。我々は京都大学テキストコーパス 3.0 を用いたが、Imamura et al. (2009) は京都大学テキストコーパス 4.0 で付与されている係り受け情報と形態素情報を学習とテストに用いている。

4.4.6 その他の先行研究

笹野, 黒橋 (2011) は、提案システムは表層格の解析を行うことから、受け身・使役形である述語は評価から除外しており、本稿では比較対象としない。

吉川 他 (2013) は、文間項は解析対象としていないため、本稿では比較対象としない。

(渡邊, 浅原, 松本 2010) は述語語義と項の意味役割の依存関係を考慮しながら、双方を同時に学習、解析を行う構造予測モデルを提案している。しかし、本稿とは異なるデータセットを用いていることから、比較対象とはしない。

4.5 評価尺度

Precision, Recall, F 値で位置関係ごとに評価を行う。

システムが出力した位置関係が T であるもののうち、正しく同定できているものの数を $tp_{(T)}$ 、できていないものの数を $fp_{(T)}$ 、システムに同定されなかった項のうち位置関係が T であるものの数を $fn_{(T)}$ とすると、

$$Precision = \frac{tp_{(T)}}{tp_{(T)} + fp_{(T)}}, \quad Recall = \frac{tp_{(T)}}{tp_{(T)} + fn_{(T)}}, \quad F = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

と定義できる。

また、システム全体 (ALL) の tp, fp, fn と Precision, Recall, F 値も、同様に定義できる。

5 議論

表 4, 5, 6 にガ格・ヲ格・ニ格の実験結果を示す。 P, R, F, A_M はそれぞれ Precision, Recall, F 値, F 値のマクロ平均 (INTRA_D, INTRA_Z, INTER の F 値の算術平均) を示す。

ALL の F 値に関して、PPR- と PPR が IIDA2007 と比較して有意差があるかどうかの検定を

表 4 ガ格の述語項構造解析の比較

	INTRA_D			INTRA_Z			INTER			ALL			
	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	A_M
IIDA2005	80.32	85.49	82.82	45.86	48.33	47.07	27.05	17.37	21.16	66.56	88.54	75.99	50.35
IIDA2007	82.93	87.40	85.11	51.12	50.11	50.61	40.49	2.56	4.82	74.29	79.59	76.85	46.85
IIDA2007+	85.22	85.85	85.53	59.63	33.53	42.92	31.74	2.25	4.21	80.06	68.72	73.95	44.22
PPR-	82.11	87.88	84.90	50.22	51.31	50.76	26.65	14.22	18.55	69.68	87.73	77.67*	51.40
PPR	84.25	86.97	85.59	51.81	50.67	51.24	25.78	15.86	19.64	70.72	85.95	77.59*	52.15
Taira et al. (2008)	—	—	75.53	—	—	30.15	—	—	23.45	—	—	57.4	43.04
Imamura et al. (2009)	85.2	88.8	87.0	58.8	43.4	50.0	47.5	7.6	13.1	79.4	68.0	73.2	50.03

表 5 ヲ格の述語項構造解析の比較

	INTRA_D			INTRA_Z			INTER			ALL			
	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	A_M
IIDA2005	92.30	92.42	92.36	42.41	32.33	36.69	13.13	8.93	10.63	87.33	89.29	88.30	46.56
IIDA2007	92.48	92.57	92.53	42.51	31.48	36.17	22.22	0.69	1.33	89.11	88.18	88.65	43.34
IIDA2007+	92.87	92.21	92.54	50.29	10.62	17.53	22.22	0.69	1.33	91.95	84.46	88.05	37.13
PPR-	92.25	92.72	92.48	41.60	31.97	36.15	14.94	4.47	6.88	88.12	89.21	88.66	45.17
PPR	92.94	92.72	92.83	46.29	29.31	35.89	15.96	5.15	7.79	89.52	88.40	88.96*	45.51
Taira et al. (2008)	—	—	88.20	—	—	11.41	—	—	9.32	—	—	79.5	36.31
Imamura et al. (2009)	95.60	92.20	93.90	53.70	21.60	30.80	25.00	0.40	0.70	94.3	84.7	89.2	41.80

表 6 ニ格の述語項構造解析の比較

	INTRA_D			INTRA_Z			INTER			ALL			
	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	A_M
IIDA2005	90.18	71.49	79.76	40.85	8.10	13.52	8.57	1.94	3.16	88.63	64.75	74.83	32.14
IIDA2007	90.27	71.27	79.65	41.33	8.66	14.32	0.00	0.00	nan	89.25	64.47	74.86	nan
IIDA2007+	89.95	71.45	79.64	72.00	5.03	9.40	0.00	0.00	nan	89.73	64.22	74.86	nan
PPR-	90.18	71.49	79.76	39.39	7.26	12.26	10.00	0.65	1.21	89.16	64.71	74.99	31.08
PPR	90.15	71.57	79.79	51.11	6.42	11.41	7.14	0.65	1.18	89.42	64.59	75.01	30.80
Taira et al. (2008)	—	—	89.51	—	—	3.66	—	—	11.76	—	—	83.15	34.98
Imamura et al. (2009)	91.10	72.60	80.80	0.00	0.00	nan	0.00	0.00	nan	91.1	66.1	76.6	nan

Takamura によるスクリプト¹⁸を用いて Approximate Randomization Test (Chinchor, Hirschman, and Lewis 1993) を行った¹⁹. 0.05 水準で有意であったものに、記号 * を付記した.

5.1 決定的に項を同定していくモデルの比較

IIDA2005, IIDA2007, IIDA2007+の ALL の F 値を比較することで、システム全体の性能について論じる.

¹⁸ http://www.lr.pi.titech.ac.jp/~takamura/pubs/randtest_fm.pl

¹⁹ この Test を行うためにはシステムの出力によらずに事例の正解ラベルを定める必要があるため、「項あり」のときにシステムが誤った出力した場合は、*fp* ではなく、*fn* として扱った.

5.1.1 ガ格の性能

ALL の性能を比較すると、ガ格の性能は IIDA2007>IIDA2005>IIDA2007+ である。

IIDA2007 と IIDA2005 の性能を比較すると、Precision は IIDA2007 の方が高く、Recall は IIDA2005 の方が高い。探索範囲を文内に限定することで、Precision が上がることが分かる。IIDA2007 の INTER の Recall は減少しているが、文間項よりも文内項の方が3倍以上多いため、システム全体の性能としては向上することが分かる。

IIDA2005 と IIDA2007+ の性能を比較すると、INTRA_D を優先的に探索することで、INTRA_D の Precision が上昇し、F 値も上昇することが分かる。INTRA_Z の Precision も上昇するが、Recall は悪化し、INTRA_Z の分量が相当数あるため、全体としては性能が悪化することが分かる。

5.1.2 ヲ格の性能

ガ格と同様であるが、INTRA_Z の数は比較的少ないため INTRA_D を優先的に探索しても、精度はガ格ほど悪化しない。

5.1.3 ニ格の性能

ニ格の性能はガ格・ヲ格とは異なり、IIDA2007+ \simeq IIDA2007>IIDA2005 である。

この傾向は項の分布が影響している。ニ格は表3によると全ての項のうち、全体の90%以上が INTRA_D である。このため、INTRA_D の探索を優先し、INTRA_D の Recall を上昇させることで、全体としての性能を上昇させることができる。

5.2 提案手法の効果

決定的な解析では優先度の低い位置関係にある候補の再現率と F 値が低下するため、優先順序をつけるほどマクロ平均は下がっていく。しかし、提案手法は全ての位置関係について最尤候補を比較するので、マクロ平均を大きく下げずにマイクロ平均 (ALL の F 値) も向上させることができている。

PPR と PPR- のいずれも、IIDA2005・IIDA2007・IIDA2007+ より性能が向上している。そのため、トーナメントフェーズで最尤候補を陽に比較する提案モデルは、決定的に項を同定していくモデルよりも効果があるといえる。

また、PPR は PPR- と比較して、ガ格・ニ格では性能はほとんど変わらないが、ヲ格では INTRA_D の Precision が向上したため、全体の性能も向上していることが分かる。そのため、文内項も INTRA_D と INTRA_Z で、最尤候補の同定モデルを分けて陽に比較することで、さらに性能を向上することがあると分かる。

5.3 先行研究との比較

ガ格において、提案手法は Taira et al. (2008) と Imamura et al. (2009) の性能を上回っている。Imamura et al. (2009) は候補同士の比較をせず、Taira et al. (2008) は優先順序を用いた決定的な解析を行っており、それらが、提案手法と比べて性能が低い原因であると考えられる。ヲ格では、提案手法は Taira et al. (2008) の性能を上回っており、Imamura et al. (2009) と同程度の性能を達成している。

しかしながら、ニ格では、Taira et al. (2008) が最も性能が高い。Imamura et al. (2009) も、ガ格・ヲ格では Taira et al. (2008) を上回る性能を発揮しているのにも関わらず、ニ格では Taira et al. (2008) よりも性能が低い。この理由として、ニ格は INTRA_D が最も多く、他の格の解析結果に依存することが挙げられる。一般に、1つの述語に対して異なる格で項を共有することはない。しかし、提案手法も Imamura et al. (2009) も各格で独立に解析を行っており、他の格の解析結果の利用ができない。一方、Taira et al. (2008) は「項を含む文節が述語を含む文節に、他の格の項を介して係っている」という関係をモデル化 (ga_c, wo_c, ni_c) し、他の格の解析結果を利用して同時に解析を行っている。そのため、INTRA_D の解析性能が高いと考えられる。

6 事例分析

6.1 成功事例

6.1.1 特定の位置関係を優先する決定的な解析モデル (IIDA) では解析できず、提案モデル (PPR) で解析できた事例

位置関係の優先順序を用いる決定的な解析モデルの中で、全体的な性能が最も高い IIDA2007 と、優先順位を持たない提案モデル (PPR-・PPR) を比較すると、INTER の Precision が少し低下しているが、Recall は上昇し、F 値も上昇している。ガ格の解析にて、IIDA2007・PPR-・PPR が解析に誤った事例の内訳を表 7 に Confusion Matrix で示した。PPR- や PPR では、誤って INTER を出力した事例が増えており (3 列目を参照)、一方で、誤って「項なし」と判断した事例が減っていることが分かる (4 列目を参照)。IIDA2007 は文間の候補を参照せずに、文内

表 7 IIDA2007 (各セル左側)・PPR- (同中央)・PPR (同右側) のガ格の誤り事例の Confusion Matrix

出力 実際	INTRA_D	INTRA_Z	INTER	NO-ARG
INTRA_D	468/478/434	464/457/485	16/194/231	665/423/518
INTRA_Z	686/717/642	733/748/699	23/266/319	967/620/722
INTER	632/679/527	615/677/579	43/489/564	1868/935/1057
NO-ARG	506/563/465	486/558/499	40/317/361	0/0/0

最尤候補が項らしいか否かを判定しなければならないが、PPR- や PPR は文内最尤候補と文間最尤候補を比較した上で、項として何が適切かを判断できるため、INTER の Recall を上昇させることができたと考える。そして、これが全体の性能に影響している。

6.1.2 2種類の最尤候補を用いるモデル (PPR-) では解析できず、3種類の最尤候補を用いる提案モデル (PPR) で解析できた事例

PPR- と PPR を比較すると、ガ格は INTRA_D と INTRA_Z の Precision と F 値が上昇しており、ヲ格は INTRA_D の Precision と F 値が、上昇している。

PPR は INTRA_D の最尤候補同定モデルと INTRA_Z の最尤候補同定モデルの2つの異なるモデルで INTRA_D と INTRA_Z の最尤候補を選んでから、陽に INTRA_D の INTRA_Z のどちらが項らしいかを比較することで、正解項を同定しやすくなっていると考えられる。これは、特に（候補数が増加する）長い文の中にある文内項の同定に効果があった。

一九五二年以来の不平等が続いている「日米航空協定」の平等化を実現するため、
「政府_ガが米側に、米航空会社の新規路線開設を今後認めない強硬方針を通告し (2)
ていたことが、十三日明らかになった。

「認める」のガ格に対して、PPR- では誤って「方針」を項として出力したが、PPR は正しく「政府」を出力した。

6.2 誤り分析

項構造解析に失敗した事例を分析したところ、誤り理由の上位3つは次のものであった。

1つ目は、談話の理解が必要な場合である。

以下の文で、「絡みつく」の二格は「ユリカモメ」である。しかし、システムは二格は項なしと判断してしまった。

東京・上野の不忍池で、無残な姿の鳥が目立つ。片足が切れたユリカモメ。釣り糸を引っ掛けて取れなくなって、そのうちに足を切断してしまうケースが多い。竹ぐしが右の首に突き刺さったユリカモメ_ニも。くしが十センチほど体の外にの (3)
ぞく。水面に浮かんだゴム_ガが絡み付き、もがくうちに首まで入ってしまったらしい。

「ユリカモメ」が話題の中心であることが捉えられなかったことが解析に失敗した理由として考えられる。今回の実験で、談話を捉えるために、Salient Reference List を用いたが、「絡みつく」の解析時に「ユリカモメ」は List には無いため、うまくいかない。これを解析するためには、「ユリカモメは負傷している」「絡みつくは負傷に関する述語である」という知識のもとで、「ユリカモメが絡みつくの二格である」という推論が必要となる。その知識を本文中から取

得するには、「鳥」や2回出てくる「ユリカモメ」が照応関係にあるという知識も必要となることから、固有表現解析や共参照解析などと推論を用いた述語項構造解析を同時に行うことで互いに精度を高めあうことができると考える。

2つ目は、格フレームなどの情報を使った格の同時解析が必要な場合である。次の文の「書く」の二格は「日記」・ヲ格は「矛盾」とアノテートされているが、システムは二格は「項なし」・ヲ格は「日記」と判断した。

日記_二には、小説の読後感や将来への夢、希望などをつづるようになり、高校生になると、大学受験のこと、沖縄における政治の矛盾_ヲなども書くようになった。(4)

一般に、「書く」の二格に「日記」が来ることは少ない。しかし、京都大学格フレーム(河原, 黒橋 2005)²⁰のような格フレーム辞書を用いれば、「書く」は「日記」を二格にとりうるのがわかる。表8に京都大学格フレームにおける「書く」の第1格フレームと第3格フレームを示した。この表は、それぞれの格フレームを構成する格がどのような項をどのくらい取るのかを、WEBコーパス内の頻度付きで表している。表8より、ヲ格に“補文”(ここでは「沖縄における政治の矛盾」)をとれば、「問い」を二格にとりうる、とわかる。

3つ目は、一般の述語とは異なる扱いをすべき述語の場合である。NAISTテキストコーパスでは名詞述語『名詞句+コピュラ「だ」』も述語としてアノテーションされている。

欧州連合_ガが十五カ国に拡大して初の交渉となる。昨年は欧州市場での乗用車の売れ行き回復を受け、規制枠を若干上方修正したが、今年については「昨年の新車登録台数集計を踏まえて対応したい」と慎重姿勢だ。(5)

しかしながら、名詞述語の振る舞いは他の述語とは明らかに異なり、同一の素性・モデルで項を同定するのは難しい。そのため、他の述語の解析モデルと分けるべきであると考え。

実際に、PPRを、名詞述語とそれ以外の述語で単純に解析モデルを分けて学習・テストした

表8 京都大学格フレームにおける「書く」の第1・第3格フレーム

	第1フレーム	第3格フレーム
ガ格	人：208, 私：203, 僕：59, 誰：45, 自分：38, 子供：35,	私：13, 誰：9, 彼：4, 僕：4, 新聞：4
ヲ格	日記：44461, 記事：34560, 文章：32990, 手紙：18234,	“補文”：85701
ニ格	気に入り：1797, “補文”：1439, ブログ：759, メール：402, 紙：393	日記：354, 紙：242, ブログ：204, ノート：202, 先：98
デ格	英語：478, ブログ：258, テーマ：226, 中：220, タイトル：206,	承知：158, 覚悟：82, ブログ：46, 言葉：34, 英語：30
ノ格	自分：1353, メール：638, 内容：525, 人：434, 本：367,	
修飾	“補文”：972, 久し振りだ：626, “数量”：612, 為：458,	そのまま：483, “数量”：312, つらつら：188,

²⁰ <http://www.gsk.or.jp/catalog/GSK2008-B/catalog.html>

表 9 名詞述語とそれ以外の述語とでモデルを分けた場合のガ格の性能の比較

モデル	対象の述語	INTRA.D			INTRA.Z			INTER			ALL		
		P	R	F	P	R	F	P	R	F	P	R	F
PPR	名詞述語	86.79	90.73	88.72	47.88	48.71	48.29	14.16	11.52	12.70	74.70	98.65	85.02
	その他述語	83.76	86.26	84.99	52.25	50.88	51.56	27.21	16.25	20.35	70.06	84.04	76.41
	全ての述語	84.25	86.97	85.59	51.81	50.67	51.24	25.78	15.86	19.64	70.72	85.95	77.59
モデルを分けた場合	名詞述語	87.72	92.68	90.13	50.81	47.20	48.94	11.27	8.92	9.96	76.26	97.76	85.69
	その他述語	83.67	86.34	84.99	52.22	51.09	51.65	27.48	16.49	20.61	70.00	84.31	76.49
	全ての述語	84.33	87.36	85.82	52.10	50.71	51.40	25.75	15.86	19.63	70.88	86.10	77.75

ところ、表 9 に示したように²¹ガ格の ALL の F 値が 77.59 から 77.75 と 0.16 ポイント上昇した。

大きな上昇がみられなかったのは、項と名詞述語の意味的關係を既存の素性ではうまく捉えられないためだと考える。名詞述語文の働きは様々で、「ラッセルは哲学者だ」のようにある事物がどのような範疇に属するのかを述べたり、「この部屋の温度は 19 度だ」のように記述を満たす値がどれなのかを述べたりする(今田 2010)。このような関係は 4.3 節での素性では捉えられない。そのため、京都大学名詞格フレーム(笹野, 河原, 黒橋 2005)や日本語語彙大系(池原他 1997)などの名詞間の関係を捉える知識を用いる必要があると考える。

また、動詞にも一般動詞とは異なる振る舞いをする動詞「なる」の解析誤りも多かった。

山花氏らにとっては、社会党が離脱を認めるかどうか_ガが、最初の関門_ニとなる。
(6)

長さ_ガ40メートル_ニにもなる3両編成の大型トラック、ロードトレインに便乗して大乾燥地帯に行く蛭子。
(7)

福井市の中心から足羽川を上流へ十キロたどると、そこ_ガはもうひなびた農村のたたずまい_ニとなる。
(8)

これらの事例の「なる」自体には意味はあまり持たず、ニ格が名詞述語相当の意味を持っているとも言える。そのため、名詞述語同様、解析モデルを分けるべきであると考えられる。

7 おわりに

本稿では、位置関係ごとに最尤候補同定モデルを作成し、実際の解析時には、各位置関係の最尤候補の中から最終的な出力を選ぶモデルを提案した。従来の研究では位置関係ごとに優先

²¹ 「全ての述語」は「名詞述語」と「その他の述語」からなる。

順位をつけ、決定的な解析を行ってきたが、それよりも提案手法が精度良く解析できることを確かめた。

今後の課題は、複数の格の解析を同時に行う手法と、本手法を統合させることを考えている。これまでに、同時解析を行うモデルは Taira et al. (2008) や 笹野, 黒橋 (2011) によって提案されてきたが²², いずれも、特定の位置関係を優先的に決定する手法である。それらの手法を、異なる位置関係の候補を参照するように発展させることを考えている。

また、名詞述語などの特殊な述語については、一般の述語とは解析モデルを分けることで、精度向上を目指すことも考えている。これらは名詞間の意味的知識がなければ解析が難しいことが分かったので、日本語語彙大系などのシソーラスを活用することを考えている。

謝 辞

ウェブから収集した日本語文データを使用させてくださった河原大輔氏に感謝いたします。また、Taira et al. (2008) の詳細なアルゴリズムを教えてくださった平博順氏にお礼申し上げます。そして、多数の有益なコメントをくださった匿名の3名の査読者に深謝いたします。

参考文献

- Chinchor, N., Hirschman, L., and Lewis, D. D. (1993). “Evaluating Message Understanding Systems: An Analysis of the Third Message Understanding Conference (MUG-3).” *Computational Linguistics*, **19** (3), pp. 409–449.
- Cortes, C. and Vapnik, V. (1995). “Support-Vector Networks.” *Machine learning*, **20** (3), pp. 273–297.
- 藤田篤, 乾健太郎, 松本裕治 (2004). 自動生成された言い換え文における不適格な動詞格構造の検出. 情報処理学会論文誌, **45** (4), pp. 1176–1187.
- Hindle, D. (1990). “Noun Classification from Predicate-Argument Structures.” In *Proceedings of the 28th Annual Meeting on Association for Computational Linguistics*, pp. 268–275. Association for Computational Linguistics.
- 飯田龍, 乾健太郎, 松本裕治 (2004). 文脈の手がかりを考慮した機械学習による日本語ゼロ代名詞の先行詞同定. 情報処理学会論文誌, **45** (3), pp. 906–918.
- 飯田龍, 小町守, 井之上直也, 乾健太郎, 松本裕治 (2010). 述語項構造と照応関係のアノテーション: NAIST テキストコーパス構築の経験から. 自然言語処理, **17** (2), pp. 25–50.

²² 吉川 他 (2013) を文間候補を考慮するように発展させるのは計算量の問題から困難だと考える。

- Iida, R., Inui, K., and Matsumoto, Y. (2005). “Anaphora Resolution by Antecedent Identification Followed by Anaphoricity Determination.” *ACM Transactions on Asian Language Information Processing*, **4** (4), pp. 417–434.
- Iida, R., Inui, K., and Matsumoto, Y. (2007). “Zero-anaphora Resolution by Learning Rich Syntactic Pattern Features.” *ACM Transactions on Asian Language Information Processing*, **6** (4), pp. 1:1–1:22.
- 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 (1997). 日本語語彙大系. 岩波書店.
- 今田水穂 (2010). 日本語名詞述語文の意味論的・機能論的分析. 博士 (言語学) 学位論文, 筑波大学.
- Imamura, K., Saito, K., and Izumi, T. (2009). “Discriminative Approach to Predicate-Argument Structure Analysis with Zero-Anaphora Resolution.” In *Proceedings of the Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing*, pp. 85–88.
- 河原大輔, 黒橋禎夫 (2004). 自動構築した格フレーム辞書と先行詞の位置選好順序を用いた省略解析. 自然言語処理, **11** (3), pp. 3–19.
- 河原大輔, 黒橋禎夫 (2005). 格フレーム辞書の漸次的自動構築. 自然言語処理, **12** (2), pp. 109–131.
- Kawahara, D. and Kurohashi, S. (2006). “Case Frame Compilation from the Web using High-Performance Computing.” In *Proceedings of the 5th International Conference on Language Resources and Evaluation*, pp. 1344–1347.
- Kawahara, D., Kurohashi, S., and Hasida, K. (2002). “Construction of a Japanese Relevance-tagged Corpus.” In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, pp. 2008–2013.
- 工藤拓, 松本裕治 (2004). 半構造化テキストの分類のためのブースティングアルゴリズム. 情報処理学会論文誌, **45** (9), pp. 2146–2156.
- Nariyama, S. (2002). “Grammar for Ellipsis Resolution in Japanese.” In *Proceedings of the 9th International Conference on Theoretical and Methodological Issues in Machine Translation*, pp. 135–145.
- 日本語記述文法研究会 (2010). 現代日本語文法 1. くろしお出版.
- 笹野遼平, 河原大輔, 黒橋禎夫 (2005). 名詞格フレーム辞書の自動構築とそれを用いた名詞句の関係解析. 自然言語処理, **12** (3), pp. 129–144.
- 笹野遼平, 黒橋禎夫 (2011). 大規模格フレームを用いた識別モデルに基づく日本語ゼロ照応解析. 情報処理学会論文誌, **52** (12), pp. 3328–3337.

- Surdeanu, M., Harabagiu, S., Williams, J., and Aarseth, P. (2003). “Using Predicate-Argument Structures for Information Extraction.” In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics*, Vol. 1, pp. 8–15.
- Taira, H., Fujita, S., and Nagata, M. (2008). “A Japanese Predicate Argument Structure Analysis Using Decision Lists.” In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 523–532.
- 渡邊陽太郎, 浅原正幸, 松本裕治 (2010). 述語語義と意味役割の結合学習のための構造予測モデル. 人工知能学会論文誌, **25** (2), pp. 252–261.
- Wu, D. and Fung, P. (2009). “Can Semantic Role Labeling Improve SMT?” In *Proceedings of the 13th Annual Conference of the European Association for Machine Translation*, pp. 218–225.
- 吉川克正, 浅原正幸, 松本裕治 (2013). Markov Logic による日本語述語項構造解析. 自然言語処理, **20** (2), pp. 251–271.

略歴

- 林部 祐太**：2009年大阪大学基礎工学部情報科学科中途退学。2011年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。現在、同研究科博士後期課程在籍。修士（工学）。意味解析とその応用に興味をもつ。
- 小町 守**：2005年東京大学教養学部基礎科学科科学史・科学哲学分科卒。2010年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。博士（工学）。同研究科助教を経て、2013年より首都大学東京システムデザイン学部准教授。現在に至る。大規模なコーパスを用いた意味解析および統計的自然言語処理に関心がある。人工知能学会、情報処理学会、ACL 各会員。
- 松本 裕治**：1977年京都大学工学部情報工学科卒。1979年同大学大学院工学研究科修士課程情報工学専攻修了。同年電子技術総合研究所入所。1984～1985年英国インペリアルカレッジ客員研究員。1985～1987年財団法人新世代コンピュータ技術開発機構に出向。京都大学助教授を経て、1993年より奈良先端科学技術大学院大学教授。現在に至る。工学博士。専門は自然言語処理。言語処理学会、情報処理学会、人工知能学会、認知科学会、AAAI, ACL, ACM 各会員。情報処理学会フェロー、ACL Fellow。

(2013年7月5日 受付)

(2013年9月16日 再受付)

(2013年10月13日 再々受付)

(2013年10月27日 採録)