大規模格フレームによる解候補削減を用いた ニューラルネットゼロ照応解析

山城 颯太[†]・西川 仁[†]・徳永 健伸[†]

本論文では日本語文内・文間ゼロ照応解析モデルを提案する。文間ゼロ照応解析において複数格の同時推定を行う際、複数の文をまたぐ大量の格要素の組合せ候補を取り扱う必要があり、これはゼロ照応解析モデルの訓練、解析に際して重大な障害となる。この問題に対して、我々は格フレームの情報を用いた効果的な解候補削減手法を提案する。提案解候補削減を用いて複数格を同時推定したモデルと解候補削減を用いずにそれぞれの格を独立に推定したモデルを日本語均衡コーパス上で比較し、0.056の精度向上を確認した。また、ローカルアテンション付きRNNを導入することで、文間ゼロ照応解析の精度が上昇することも確認した。

キーワード：ゼロ照応解析, 述語項構造解析, 解候補削減, 格フレーム

Neural Japanese Zero Anaphora Resolution with Candidate Reduction Using Large-scale Case Frames

SOUTA YAMASHIRO[†], HITOSHI NISHIKAWA[†] and TAKENOBU TOKUNAGA[†]

This paper presents a model for Japanese zero anaphora resolution that deals with both intra- and inter-sentential zero anaphora. Our model resolves anaphora for multiple cases simultaneously by utilising and comparing information from other cases. This simultaneous resolution requires the consideration of many combinations of antecedent candidates, which could be a crucial obstacle in both the training and resolving phases. To cope with this problem, we have proposed an effective candidate pruning method using case frame information. We compared the model, which estimates multiple cases simultaneously, by using our proposed candidate pruning method and model, which estimates each case independently without a candidate reduction method in a Japanese balanced corpus. The results confirmed a 0.056-point increase in accuracy. Furthermore, we also confirmed that the introduction of local attention Recurrent Neural Network increases the accuracy of inter-sentential anaphora resolution.

Key Words: *Zero Anaphora Resolution, Predicate-argument Structure Analysis, Candidate Reduction, Case Frames*

[†] 東京工業大学情報理工学院, School of Computing Tokyo Institute of Technology

1 はじめに

ゼロ照応解析とは、テキスト中の述語の省略された項（ゼロ代名詞）を検出し、項として埋めべき格要素を同定するタスクである。格要素は先行詞としてテキスト中で言及されている場合もあれば、言及されていない場合もある。前者の場合、先行詞は述語と同じ文中にある（文内ゼロ照応）か、先行する文中にある（文間ゼロ照応）¹。後者（外界ゼロ照応）の例として、テキストの著者である主語が明示的に言及されない場合などがある。

(1) 大岡山商店街でも (ϕ ガ) お洒落な建物を見かけるようになった。カフェテリアが特に多くて、今月も新しく (ϕ ガ)(ϕ ニ) オープンしてる。

例 (1) では「見かける」のガ格と「オープンしてる」のガ格、ニ格が省略されている。「オープンしてる」のガ格の格要素は同文中に言及されている「カフェテリア」であり（文内ゼロ照応）、ニ格の格要素は前文で言及されている「大岡山商店街」である（文間ゼロ照応）。一方、「見かける」のガ格の格要素はテキスト中では明示的に言及されていない「著者」である（外界ゼロ照応）。

本論文では特に日本語のゼロ照応解析を取り扱うが、項の省略が起こる pro-drop 言語は日本語だけではなく、他に中国語、イタリア語、スペイン語などがあり、各言語で日本語ゼロ照応解析と類似したタスクに取り組む研究が数多くある (Chen and Ng 2016; Yin, Zhang, Zhang, and Liu 2017; Yin, Zhang, Zhang, Liu, and Wang 2018a, 2018b; Iida and Poesio 2011; Rello, Baeza-Yates, and Mitkov 2012)。また英語では意味役割付与タスクがゼロ照応解析に似た研究として挙げられる (Zhou and Xu 2015; He, Lee, Lewis, and Zettlemoyer 2017)。日本語ゼロ照応解析は、日本語述語項構造解析の部分問題であり、自動要約 (Yamada, Teufel, and Tokunaga 2017) や情報抽出 (Sudo, Sekine, and Grishman 2001)、機械翻訳 (Kudo, Ichikawa, and Kazawa 2014) など様々な自然言語処理アプリケーションの精度改善にとって重要であるため、緊急に解決されるべき課題として盛んに研究されている (Sasano and Kurohashi 2011; Hangyo, Kawahara, and Kurohashi 2013; Ouchi, Shindo, and Matsumoto 2017; Matsubayashi and Inui 2017, 2018; Shibata and Kurohashi 2018; Kurita, Kawahara, and Kurohashi 2018)。

本研究の貢献は大きく二つに分けられる。第一に大規模均衡コーパス上で日本語ゼロ照応解析を行い評価したことと、第二にこの大規模均衡コーパス上で文内・文間ゼロ照応解析を可能にするための解候補削減手法を提案したことの二点である。

従来のゼロ照応解析研究は、新聞記事のみからなる『NAIST テキストコーパス』(NTC) (Iida, Komachi, Inui, and Matsumoto 2007) で評価を行うものが多かった。従って、それらの評価ではテキストドメインの違いによる影響が考慮されていない。ゼロ照応解析結果の応用を考えた

¹ この研究では後方照応は扱わない。

時, 複数ドメインの文書に対しても頑健なゼロ照応解析手法の有用性は高い. 我々は『現代日本語書き言葉均衡コーパス』(BCCWJ) (Maekawa, Yamazaki, Ogiso, Maruyama, Ogura, Kashino, Koiso, Yamaguchi, Tanaka, and Den 2014) を評価実験に使用した. BCCWJは13ドメインにまたがって構築された約一億語からなる日本語均衡コーパスである. このうちの約100分の2にあたる約二百万語からなるコアデータに対しては, 人手による述語項構造と照応関係の付与がされている. また, BCCWJは新聞, 雑誌, 書籍, 白書, Yahoo!知恵袋, Yahoo!ブログの6ドメインにまたがったテキストを含んでいる. ドメインによるゼロ照応解析の性能の違いを調べるために, 我々はBCCWJを使用した.

表1はBCCWJコアデータセットの述語と格ごとの格要素の距離の分布を示している. ここの距離は述語と格要素の間の文数である. 距離0は文内照応を示しており, 距離1以上は文間ゼロ照応を示している. この表から, 文中に格要素の出現するゼロ照応のうち, 半数以上のゼロ照応が文間ゼロ照応であることがわかる. 表2はテキストドメインごとに分類した述語とガ格の格要素との距離の分布を示している. この表から, 文内, 文間ゼロ照応のドメインごとの違いが確認できる. OW(白書)のガ格には文内ゼロが多く出現している一方で, PM(雑誌), OC(QA)のガ格には文内ゼロ比率が少なく, 5文前までの文間ゼロが多く出現している. また, PM(雑誌), OY(ブログ)のガ格は7文以上前の文間ゼロが他のドメインと比較して多く出現しており, このようなドメインごとの述語と格要素間の距離の違いが文間ゼロ照応解析の精度に影響を及ぼす原因だと予測できる. これらの観察から, 異なるタイプのテキスト上で評価実験を行うことの重要性が示唆される.

表1に示すとおり現実の文書には文間ゼロ照応が頻出するが, 従来のゼロ照応解析研究の多くは, 文内ゼロ照応のみに焦点を絞っている (Iida, Torisawa, Hashimoto, Oh, and Kloetzer 2015; Shibata, Kawahara, and Kurohashi 2016; Ouchi et al. 2017; Matsubayashi and Inui 2017, 2018;

表1 格要素と述語の距離の分布

距離	ガ格	ヲ格	ニ格	total	%
0	16,621	4,545	2,059	23,225	32.6
1	8,231	1,764	1,113	11,108	15.6
2	3,396	599	430	4,425	6.2
3	1,792	317	227	2,336	3.3
4	1,020	172	126	1,318	1.8
5	690	83	84	857	1.2
6	414	45	51	510	0.7
≥ 7	1,917	217	178	2,312	3.2
外界	23,344	925	909	25,178	35.3
total	57,425	8,667	5,177	71,269	

表2 文書ドメインごとのガ格ゼロ照応の分布 (%)

距離	OW	PB	PN	PM	OC	OY
0	72.3	49.5	51.1	40.3	38.8	49.8
1	15.1	25.0	24.4	23.9	29.1	23.1
2	5.6	9.8	9.6	11.2	13.4	8.7
3	2.4	5.0	4.8	6.7	6.9	4.5
4	0.9	2.7	2.6	4.3	3.9	3.5
5	0.9	1.3	2.3	2.7	2.9	1.6
6	0.4	1.1	1.0	1.7	1.4	1.7
≥ 7	2.5	5.5	4.4	9.3	3.6	7.2

OW: 白書, PB: 書籍, PN: 新聞,
PM: 雑誌, OC: QA, OY: ブログ

Kurita et al. 2018). Ouchi et al. (2017) は、文内ゼロ照応のみを取り扱う理由として、探索範囲の問題を指摘している。文間ゼロ照応では、格要素候補をテキスト全体から探す必要があるため、文内ゼロ照応解析に比べて探索範囲が拡大する。Matsubayashi and Inui (2017, 2018) と Kurita et al. (2018) は解析に際して文脈素性を取り入れるために、リカレントニューラルネットワーク (RNN) を導入し、格要素候補と述語が含まれる文を読み込ませている。しかしこれと同じ手法を文間ゼロ照応解析において適用しようとする、テキスト全体を RNN に入力として与える必要がある。長距離の文脈を記憶する仕組みを持つ LSTM や GRU を使用しても、システムがテキスト全体における長距離の依存関係を十分に学習できるとは限らない。また、テキスト全体を記憶しなくても、選択的に抽出された文脈情報のみで解析できる可能性がある。

先述の研究と異なり、Sasano and Kurohashi (2011) と Hangyo et al. (2013), Shibata and Kurohashi (2018) は、文内・文間ゼロ照応解析手法を提案している。しかし、彼らは新聞コーパスと Web コーパスを用いて評価実験を行っており、複数ドメインコーパスを用いた評価実験は行っていない。また、Shibata and Kurohashi (2018) は各 entity ごとに embedding を割り当て、現在解析対象としている述語より文書の前方における RNN の解析結果を用いて、それぞれの embedding を逐次更新することで文脈情報を使用している。

これら 2 つの問題に対して、本研究では様々なドメインの文書への対応を可能とするために大規模格フレームを利用し、述語が取りうる複数の格要素の組合せから最適なものを選ぶ。大規模コーパスから得られた述語の格フレームに関する統計的情報を使用することで、特定のドメインにのみ出現する事例数の少ない格フレームに対しても対処できるようになると考えられる。解候補となる格要素の組合せが膨大になる問題に対しては、格フレームを使用した解候補削減手法を取り入れることで、より汎用性の高い文内・文間ゼロ照応解析モデルを提案する。

ひとつのモデルで文内・文間のゼロ照応解析を同時に行う際、各格に対してそれぞれ独立に解析を行うより、他の格の情報を利用して複数格を同時に解析する方がより良い精度が得られると考えられる。渡邊、浅原、松本 (2010) は CoNLL-2009 Shared Task データを用いた述語項構造解析において、述語項構造の大域的な情報を取り扱う因子を素性として取り入れることで、日本語を含む 7 種類すべての言語の精度が改善することを確認している。しかし複数格を同時に解析する際には、先行詞の広大な探索範囲の問題に対処する必要がある。特に機械学習を適用する際、正解の候補となる名詞の組合せが大幅に増加する。BCCWJ の場合では、格フレーム候補が平均 20 個、ガヨニ格の格要素候補が平均 10 個出現するため、組合せは $20 \times 10 \times 10 \times 10 =$ 約 20,000 通り考えられ、このうち正しい格フレームと格要素の組合せは一つのみである。結果としてデータ中の正例と負例の比率が約 1 対 20,000 と著しく不均衡となる。このような偏った訓練データは不必要に計算量を増幅させ、かつモデルの汎化を妨げる要因となる。我々は、学習に不要な負例を削減するために、解析対象述語に対応する格フレームを用いた効率的な解候補削減手法を提案する。この提案手法により、正解を候補に残しつつ、約 1,000 分の 1 にまで解候補を削

減することに成功した。また、我々は RNN にローカルアテンション機構 (Luong, Pham, and Manning 2015) を導入することで、前文中のどの部分に注意を向けて解析するかをシステムに学習させた。本研究の提案手法は二つの構成要素からなり、一つは格フレーム内の単語分散表現を使用した解候補削減アルゴリズムで、もう一つは解候補削減に使用した分散表現を利用するニューラルネットゼロ照応解析モデルである²。なお、BCCWJ を訓練・テストに用いた文内・文間のガヲニ格を対象とするゼロ照応解析は、本研究が初の試みである。

2 関連研究

2.1 日本語ゼロ照応解析

表 3 は、タスクの種類、使用しているコーパスのドメイン、コーパスのサイズ、手法の観点から関連研究をまとめたものである。解候補の列はガヲニ格をそれぞれ独立の解析器で解析しているのか、一つの解析機を用いて組合せで解析しているのかを示している。Hangyo et al. (2013) はランキング SVM を用いて、Web コーパスに対して文内、文間、外界のゼロ照応解析を同時に行っている。この Web コーパスは 1,000 文書からなり、それぞれ Web ページの冒頭 3 文を抜き出したものである (Hangyo, Kawahara, and Kurohashi 2012)。Shibata and Kurohashi (2018) はフィードフォワードニューラルネットワーク (FNN) とリカレントニューラルネットワーク (RNN) を組合せて用い、Web コーパス (Hangyo et al. 2012) と新聞コーパスに対して直接の係り受け関係と文内・文間・外界のゼロ照応解析を同時に行っている。Matsubayashi and Inui (2018) は FNN と RNN を組合せて用いることで、NTC に対して直接の係り受け関係と文内のゼロ照応解析を同時に行い、直接の係り受け関係と文内ゼロ照応解析の state-of-the-art を達成し

表 3 関連研究

	タスク				ドメイン			サイズ (文数)	手法			解候補
	係り受け	文内	文間	外界	新聞	Web	etc.		線形	NN	+att	
Imamura, Saito, and Izumi (2009)	○	○	○		○			40,000	○			独立
Hangyo et al. (2013)		○	○	○		○		3,000	○			組合せ
Ouchi, Shindo, Duh, and Matsumoto (2015)	○	○			○			40,000	○			組合せ
Shibata et al. (2016)	○	○		○		○		15,000		○		組合せ
Iida, Torisawa, Oh, Kruengkrai, and Kloetzer (2016)		○			○			40,000		○		独立
Ouchi et al. (2017)	○	○			○			40,000		○		独立
Matsubayashi and Inui (2017)	○	○			○			40,000		○		独立
Shibata and Kurohashi (2018)	○	○	○	○	○	○		15,000		○		組合せ
Kurita et al. (2018)	○	○				○		15,000		○	○	独立
Matsubayashi and Inui (2018)	○	○			○			40,000		○	○	独立
Sasano and Kurohashi (2011)		○	○			○		1,000	○			組合せ
提案手法		○	○		○	○	○	60,000	○	○	○	組合せ

² https://github.com/yamashiros/Japanese_zero_anaphora

た. Sasano and Kurohashi (2011) は対数線形モデルを用いて, 979 文からなる Web コーパスと NTC に対して文内と文間のゼロ照応解析を同時に行い, NTC における文内・文間ゼロ照応解析の state-of-the-art を達成した. これらに対して我々は, ランキング SVM³ (Joachims 2006) モデルと FNN と RNN の組合せモデルを用いて, 『現代日本語書き言葉均衡コーパス』 (BCCWJ) (Maekawa et al. 2014) に対して, 文内・文間のゼロ照応解析を同時に行う.

2.2 大規模格フレーム

格フレームとは述語とその述語が取りうる格要素を述語の格パターンごと, 格ごとに整理した共起情報である. 表 4 のように格パターンに基づいて格フレームを分けることで, 述語と格要素間の語彙的選好の知識を照応解析に利用することができる (Sasano, Kawahara, and Kurohashi 2008; Sasano and Kurohashi 2011; Hangyo et al. 2013).

格フレームの構築に関しては 河原, 黒橋 (2005) が Web テキストから格フレームを自動構築する手法を提案している. 林部, 河原, 黒橋 (2015) は格の出現の仕方を考慮したクラスタリングを行うことで格パターンを抽出し, 述語に対してより頑健に格フレームを割り当てることに成功している. これら手法を用いて約 100 億文の大規模 Web コーパスから取得, 整理された格フレーム知識は京大格フレーム⁴として公開されており, 我々はこれを使用した.

日本語ゼロ照応解析において, 格フレームは解析の手がかりとして広く使用されている (Sasano and Kurohashi 2011; 松林, 乾 2014). しかし, 我々が調べたところ, BCCWJ 中の動詞の正解述語項構造に対する京大格フレームのカバレッジは約 27%と十分ではない. この問題に対して, Sasano and Kurohashi (2011), 松林, 乾 (2014) は意味クラスを用いて格フレーム中の用例の汎化を行うことで対処している. 我々は分散表現と cosine 距離を用いることでカバレッジの問題に対処している. また, 我々はそれぞれの述語と格フレームに対して事例の分散表現の平均ベ

表 4 「オープンしてる」の格フレーム例

格フレーム	ガ格	出現数	ヲ格	出現数	ニ格	出現数
オープンしてる:動 ₁	店	129	—	—	近く	6
	カフェ	38	—	—	跡地	2
	レストラン	14	—	—	ところ	2
	...	...	—	—	...	...
オープンしてる:動 ₂	ブランド	12	ショップ	59	—	—
	専門家	8	サロン	18	—	—
	オーナー	4	ブティック	13	—	—
	...	...	...	...	—	—

³ https://www.cs.cornell.edu/people/tj/svm_light/svm_rank.html

⁴ <http://www.gsk.or.jp/catalog/gsk2018-b/>

クトルを導入することで、格フレーム中の統計情報を汎化し、解候補削減にも使用している。

2.3 解候補削減

文間ゼロ照応に際して、いくつかの先行研究ではそれぞれに解候補削減の基準を設定している。Sasano and Kurohashi (2011), Hangyo et al. (2013) は述語の複数の格を同時に推定しており、述語が含まれる文より 3 文前までに出現する格要素の先行詞候補をすべて含めている。3 文より前の文にも先行詞は出現しうるが、Hangyo et al. (2013) は NTC の述語に対する格要素のうち 82.9% が 3 文中に出現すると報告した。Imamura et al. (2009) は述語の複数の格をそれぞれ独立に推定しており、一文前までに現れる述語の格要素として選ばれた名詞のみを解析の対象としている。この制限によって、NTC 中、何も制限のない状態では平均 102.2 語の名詞を解候補としなければならなかったところを平均 3.2 語まで抑え、ゼロ代名詞の格要素のうち 62.5% をカバー出来たと彼らは報告している。

3 モデルと素性

3.1 モデル

解析対象述語 p が含まれる文を S_0 とし、入力文書 t に含まれる S_0 から h 文前までの文をそれぞれ $S_{-1}, S_{-2}, \dots, S_{-h}$ とする。 S_0 から S_{-h} までに含まれるすべての名詞の集合を $E_p = \{e_1, e_2, \dots, e_n\}$ とする。これらに加えて『照応なし』または『外界照応』を意味する e_{none} を E_p に追加する。述語 p に対応する京大格フレーム中の格フレーム群を $CF_p = \{cf_1^p, cf_2^p, \dots, cf_m^p\}$ とする。1つの格フレーム cf_l^p には、それぞれの格 $c \in \{\text{ガ格}, \text{ヲ格}, \text{ニ格}\}$ に対応する 3 つの格スロットがあり、 E_p 中に含まれるいずれかの名詞がそれぞれの格スロットに対応する格要素である。格スロットと格要素の対応付けを $a = \langle \text{ガ格} : e_i, \text{ヲ格} : e_j, \text{ニ格} : e_k \rangle$ とする。述語項構造候補を (cf_l^p, a) とし、これを表現する素性ベクトルを $f(cf_l^p, a, t)$ とする。このモデルの出力は以下の式 (1) で表せる。訓練時にはパラメータ w が正例のスコアが負例のスコアより大きくなるように学習され、テスト時には最適な格フレームと格要素の組合せを選択するために用いられる。このモデルは、Hangyo et al. (2013) のモデルをベースとしている。

$$cf_l^{p*}, a^* = \underset{cf_l^p, a}{\operatorname{argmax}} w \cdot f(cf_l^p, a, t) \quad (1)$$

3.2 素性

素性ベクトル $f(cf_l^p, a, t)$ は以下 5 タイプの素性の組合せからなる：ベースモデル素性、格要素分散表現、述語分散表現、格フレーム内平均ベクトル (MVC)、文脈ベクトル。我々はランキング SVM と FNN を比較するが、ランキング SVM はベースモデル素性のみを使用し、その他

の素性は FNN でのみ用いる.

ベースモデル素性 ベースモデル素性 ϕ_{BMF} の各要素は実数かバイナリ値である. ベースモデル素性 ϕ_{BMF} は Sasano et al. (2008) の確率的格解析モデルから得られる表層の係り受けの確率と Hangyo et al. (2013) が提案する素性群からなる. Hangyo et al. (2013) の素性は格フレーム素性, 述語素性, 文脈素性の 3 種類からなる. 例えば, ある格要素がその格フレームの格スロットに埋まるかどうかの確率は格フレーム素性の一つである. 付録 A に Hangyo et al. (2013) の素性の一覧を示す.

格要素分散表現 格要素分散表現 ϕ_e は各格 c の格要素 e_c に対応する 3 つの分散表現から構成される. 語の分散表現を生成するモデルとしては word2vec (Mikolov, Chen, Corrado, and Dean 2013) を使用した⁵.

述語分散表現 述語分散表現は word2vec を使って生成された解析対象述語の単語分散表現である.

格フレーム内平均ベクトル (MVC) 表 4 に示すように京大格フレーム内では, 述語 p に対するそれぞれの格フレーム cf_l^p は各格 c に対応する単語リストから構成される. 例えば, 「オープンしてる:動 1」のガ格には『店』, 『カフェ』, 『レストラン』などが格納されている. $W_{cf_l^p(c)}$ を格フレーム cf_l^p と格 c に対応して京大格フレーム中に出現する格要素の全体とする. 例えば $W_{オープンしてる:動_1(ガ)}$ の要素は前述の『店』, 『カフェ』, 『レストラン』などである.

ϕ_w を語 $w \in W_{cf_l^p(c)}$ の分散表現ベクトル, $\text{count}(cf_l^p, c, w)$ を語 w が格フレーム cf_l^p の格 c の格要素として出現する回数とする. この時, 格フレーム内平均ベクトル (MVC) $\bar{\phi}_{cf_l^p(c)}$ は, 格フレーム cf_l^p 中の各格 c の分散表現ベクトル $W_{cf_l^p(c)}$ の重み付き平均として計算される.

$$\bar{\phi}_{cf_l^p(c)} = \frac{\sum_{w \in W_{cf_l^p(c)}} \text{count}(cf_l^p, c, w) \cdot \phi_w}{\sum_{w \in W_{cf_l^p(c)}} \text{count}(cf_l^p, c, w)} \quad (2)$$

例えば表 4 において, 格フレーム「オープンしてる:動 1」はガ格に『店』を 129 回取っている, $\bar{\phi}_{オープンしてる:動_1(ガ)}$ は以下のように計算される.

$$\bar{\phi}_{オープンしてる:動_1(ガ)} = \frac{129 \cdot \phi_{店} + 38 \cdot \phi_{カフェ} + 14 \cdot \phi_{レストラン} + \dots}{129 + 38 + 14 + \dots}. \quad (3)$$

$\bar{\phi}_{cf_l^p(ガ)}$, $\bar{\phi}_{cf_l^p(ヲ)}$, $\bar{\phi}_{cf_l^p(ニ)}$ を結合して $\bar{\phi}_{cf_l^p}$ を生成する. $\bar{\phi}_{cf_l^p}$ を使って, a と cf_l^p の関連 (選択選好) を測り, 尤もらしい組合せを探索する. なお我々は, MVC を照応解析, 解候補削減の両方で使用する.

文脈ベクトル 文脈ベクトル $c_{cf_l^p, a, t}$ はローカルアテンション機構付き RNN の出力である. こ

⁵ 日本語 wikipedia (2016-09-20) の本文全文から取得した約 100 万記事に対して, 次元数を 500, window を 15 として学習させることで得られたモデルを使用した.

の RNN は解析対象述語を含んだ文とその前方 h 文を受け取り, 対象述語に対する文脈をモデリングする. 図 1 の左下部分がこれに該当する. $\text{Enc}(s_i)$ を, $S_{-h:0}$ 中の i 番目の語 s_i を入力として与えられた時の RNN エンコーダの出力する隠れ状態とする. この隠れ状態は RNN が読み込んだ文中の語と同じ数だけ存在し, 各語と一対一に対応する RNN の出力である. $a_t(\cdot)$ は一般的なアテンション重みを表し, $\text{score}(\cdot)$ には訓練データから学習される複数の重みパラメータを使用した. $\text{Local}(\cdot)$ はローカル重みを表し, これを導入することで文全体を散漫と注視するのを防ぎ, 局所的な一部分のみに着目させる. $\text{Len}(S_{-h:0})$ は入力文 $S_{-h:0}$ 中の語数を表し, $\mathbf{v}_p$, $\mathbf{W}_p$ は訓練データから学習されるパラメータである.

我々のアテンション機構モデルは他の素性ベクトルの連結を手がかりとしてアライメント重みベクトルを推論する. この重みベクトルはアテンション機構が文中のどの語へと注意を向けるかという情報に対応する. 図 1 の左上部分がこれに該当する. 文脈ベクトル $c_{cf_l^p, a, t}$ はこのアライメント重みベクトルに基づくエンコーダの出力 $\text{Enc}(S_{-h:0})$ の重み平均として計算する. 図 1 右側の Context embedding がこれに該当する.

$$c_{cf_l^p, a, t} = \sum_{s_i \in S_{-h:0}} a_t(s_i) \text{Local}(s_i) \text{Enc}(s_i) \quad (4)$$

$$a_t(s_i) = \frac{\exp(\text{score}([\phi_{BMF}; \phi_e; \bar{\phi}_{cf_l^p}], \text{Enc}(s_i)))}{\sum_{s_j} \exp(\text{score}([\phi_{BMF}; \phi_e; \bar{\phi}_{cf_l^p}], \text{Enc}(s_j)))} \quad (5)$$

$$\text{Local}(s_i) = \exp\left(-\frac{(i - \text{Len}(S_{-h:0}) \cdot p)^2}{2\sigma^2}\right) \quad (6)$$

$$p = \text{sigmoid}(\mathbf{v}_p^T \tanh(\mathbf{W}_p[\phi_{BMF}; \phi_e; \bar{\phi}_{cf_l^p}])) \quad (7)$$

直感的には, この機構は我々のモデルがアライメントベクトルを介して述語から離れた長文脈中の語を格要素として識別することを可能にしている. 格フレームがその述語から離れた名

図 1 アテンション付き RNN 追加 FNN モデルのネットワーク構造 (モデル F3)

詞を格要素として取るようなケースに対して、我々はこのメカニズムが直接的にその現象をモデル化することを期待している。

4 格フレーム中の分散表現を利用した解候補削減

ゼロ代名詞となる項の先行詞候補を網羅的に探索すれば、列挙される述語項構造候補 (cf_l^p, a) の集合は爆発的な規模となり、探索範囲は非実用的なものとなる。我々は Sasano and Kurohashi (2011) の基準を参考に、ゼロ代名詞となる項の先行詞候補は述語が含まれる文より 3 文前までのみを範囲として解候補削減を行っている。つまり 3.1 節の h を 3 とした。BCCWJ 中の格要素の分布は表 1 のようになっているため、この制限によってゼロ代名詞の 89.16% をカバーできることがわかる。

n と m をそれぞれ E_p 中の名詞句数、対象述語の格フレーム数とすると、この制限を用いてもなお、候補の数は $O(n^3m)$ となり、BCCWJ 中の各動詞に対して約 20,000 個の述語項構造候補が出現する。

4.1 述語内平均ベクトル (MVP)

我々は、格フレーム候補と格要素候補の組合せについて 3.2 で提案した MVC と、述語内平均ベクトル (MVP) $\bar{\phi}_{p(c)}$ の二種類の平均ベクトルを使用した効率的な解候補削減手法を提案する。MVP は各格 c について述語 p に対応するすべての格フレームに渡って MVC $\bar{\phi}_{cf_l^p(c)}$ の重み平均を取ったベクトルである。重みは京大格フレーム中の各格フレームの頻度に基づく。我々の解候補削減手法は Ouchi et al. (2015) の山登り法を参考に、格フレーム候補と格要素候補の組合せ数を削減する。この解候補削減は計算効率のみを目的とするのではなく、訓練データ中の正例・負例のデータ数の非対称性の解消も目的とする。我々のケースでは、1 つの正例に対して 20,000 の負例が生じるため、これに対処している。前述したように、我々は訓練データ中のほとんどの負例は訓練に貢献しないと考え、解候補削減を行う。

4.2 アルゴリズム

我々の提案する解候補削減手法をアルゴリズム 1 に示す。ある述語 p には、文脈に対するその語義の曖昧性を反映した複数の格フレーム CF_p が存在する。それぞれの格フレーム cf_l^p に対応する格フレーム内平均ベクトル $\bar{\phi}_{cf_l^p(c)}$ はその格フレームの選択選好を反映しているため、これと格要素候補ベクトル ϕ_e の距離が近いほど、その格要素候補 e は対象格フレーム cf_l^p の格スロット c に埋まりやすいと言える。このアルゴリズムは与えられた述語に対して、二つのベクトル間の距離が最も近くなる格フレームと格要素候補の組合せを探索する。しかしながら、京大格フレームは自動的手法で構築されているので、本来別々の格フレームが一つの格フレーム

アルゴリズム 1 解候補削減アルゴリズム

Input:

- a predicate p to be analyzed,
- a set of case frames CF_p corresponding to p ,
- a set of cases $C = \{ \text{ガ格, ヲ格, ニ格} \}$,
- a set of nouns E_p appearing within the h preceding sentences.

Output:

optimal cf_l^{p*} , e_c^* for the analyzed p and each case $c \in C$.

```

1: for each case  $c \in C$  do
2:    $e_c^{(0)} \leftarrow \operatorname{argmax}_{e \in E_p} \cos(\bar{\phi}_{p(c)}, \phi_e)$   $\triangleright \phi_e$  is the vector of the argument candidate  $e$ .  $\bar{\phi}_{p(c)}$  is the MVP
3: end for
4:
5:  $cf_l^{(0)} \leftarrow \operatorname{argmax}_{cf_l^p \in CF_p} \sum_{c \in C} \text{PSEUDO-SCORE}(cf_l^p, e_c^{(0)})$ 
6:  $t \leftarrow 0$ 
7: repeat
8:   for each case  $c \in C$  do
9:      $e_c^{(t+1)} \leftarrow \operatorname{argmax}_{e \in E_p} \text{PSEUDO-SCORE}(cf_l^{(t)}, e_c)$ 
10:   end for
11:
12:    $cf_l^{(t+1)} \leftarrow \operatorname{argmax}_{cf_l^p \in CF_p} \sum_{c \in C} \text{PSEUDO-SCORE}(cf_l^p, e_c^{(t+1)})$ 
13:    $t \leftarrow t + 1$ 
14: until  $e_c^{(t)} = e_c^{(t+1)}$  and  $cf_l^{(t)} = cf_l^{(t+1)}$ 
15: return  $cf_l^{p*} \leftarrow cf_l^{(t)}$ ,  $e_c^* \leftarrow e_c^{(t)}$  for each case  $c \in C$ 
16:
17: function PSEUDO-SCORE( $cf_l^p, e$ )
18:    $score \leftarrow 0$ 
19:   for each case  $c \in C$  do
20:      $score \leftarrow score + P(p, cf_l^p, e, c) + \cos(\bar{\phi}_{cf_l^p(c)}, \phi_e) + 0.5 \times d_{p,e}$   $\triangleright \bar{\phi}_{cf_l^p(c)}$  is the MVC
21:   end for  $\triangleright d_{p,e}$  is the distance between  $p$  and  $e$ 
22:
23:   return  $score$ 
24: end function

```

としてまとめられてしまっている、あるいは同じ一つの格フレームが別々に分断されてしまっている可能性がある。この問題に対処するために、我々は提案する解候補削減手法に二種類の平均ベクトルを導入した。MVCはある述語に対する格フレームの違いを区別し、MVPは格フレームの違いを考慮せず述語のみを考慮する。

まず初期値として各格 $c \in C$ に埋まりうる格要素 $e_c^{(0)}$ を仮に定める（行 1-3）。MVP $\bar{\phi}_{p(c)}$ と格要素候補の分散表現 ϕ_e とのコサイン距離を求め、これが最小となる、すなわち対象述語に埋まる格要素群に最も近い格要素を初期格要素とする。この段階では、MVPを使用することで特定の格フレームではなく述語のみを考慮している。格フレーム候補と初期格要素の組合せを入力とした PSEUDO-SCORE（行 17-26）の返すスコアに基づいて、これらの初期格要素に対して最適な格フレーム $cf_l^{(0)}$ を格フレームの初期値とする（行 5）。PSEUDO-SCORE については Sasano et al. (2008) を参考に、我々は以下の 3 つの要素を考慮した。(1) 京大格フレームに基づく（述語、格フレーム、深層格、格要素）の組合せの出現確率、(2) 格フレーム内平均ベクトル (MVC) と格要素候補の間のコサイン類似度、および (3) 述語と格要素候補の間の文数、である。

このスコアの係数は経験的に定めた。以降、格フレーム $cf^{(t)}$ を固定して格要素 $e_c^{(t+1)}$ を探索するフェーズ（行 8–10）と格要素 $e_c^{(t+1)}$ を固定して格フレーム $cf^{(t+1)}$ を探索するフェーズ（行 12）を繰り返し、格フレームと格要素が更新されなくなればループを抜ける（行 6–15）。このアルゴリズムでは返り値として最もスコアの高い格フレームと格要素の組合せを返すが、実際にはループ中の毎回の探索過程で計算した格要素候補のうち 3 ベストまでを候補として保存する。最終的な出力は探索の過程で保存されたすべての格フレームと格要素の組合せである。提案した解候補削減手法により、約 70% の正解を候補に残しつつ、約 1,000 分の 1 まで解候補を削減することができた。

5 評価実験

5.1 ゼロ照応解析手法

学習手法にはランキング SVM と FNN を使用しそれぞれ比較した。先行研究 (Sasano and Kurohashi 2011; Hangyo et al. 2013) と同様に、まず文書全体に対して形態素解析、固有表現抽出、構文解析を行う。これには JUMAN Ver.7.01⁶, KNP Ver.4.16⁷, CaboCha Ver.0.69⁸ を用いた。

S0 提案する解候補削減を行い、ベースモデル素性を使用して SVM モデルを実装した。ランキング学習には SVM^{rank} (Joachims 2006) を使用した。カーネルは線形である。このモデルは正例と負例から識別関数を学習し、この識別関数が最も高い解候補を一つ出力する。

S0_each 提案する解候補削減手法の効果を評価するためには、解候補削減を用いないモデルと比較することが自然である。しかしながら、前方 3 文までに先行詞候補の探索範囲を制限しても、述語一つあたりに対して 20,000 の述語項構造候補が出現するため、訓練時の計算複雑性は現実的ではない。これは複数の格を同時推定するために、格要素候補同士の組合せを考慮していることが原因である。そこで我々は、それぞれの格を独立に解析することで、解候補削減が必要ない単一格解析手法を用意した。この手法では、3 つの格に対してそれぞれ別の SVM モデルを用意し、これらを独立に学習させて、評価の際は各格に対応するモデルのそれぞれの出力を組合せて最終的な出力とした。この時、述語一つあたりに対して、各格約 200 の格フレームと格要素の組合せが出現し、我々の提案手法に比べて計算量は膨大ではあるものの、計算可能な範囲である。この各格に対して独立の SVM を用いて学習を行ったモデルを S0_each とする。

S0' 複数格の同時推定のために我々は単純な解候補削減手法を用意した。この手法では、解析対象述語に近い方から先行する n 個の名詞のみを格要素候補として選ぶ。この時の各格に対す

⁶ <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

⁷ <http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

⁸ <https://taku910.github.io/cabocha/>

表 5 候補削減・素性の組合せ

候補削減・素性	S0_each	S0'	S0	F0	F1	F2	F3
5-best 候補削減		○					
提案候補削減			○	○	○	○	○
ベースモデル素性	○	○	○	○	○	○	○
格要素分散表現					○	○	○
述語分散表現					○		
MVC						○	○
文脈ベクトル							○

る格要素候補数は、提案手法と同程度の格要素の組み合わせ数となるよう調整した値であり、今回は $n = 5$ とした。この単純な解候補削減手法を適用した上で SVM を用いて学習を行ったモデルを $S0'$ とする。

F0 ベースモデル素性を使用して FNN モデルを実装した。FNN の設計に際しては Matsubayashi and Inui (2017) を参考に、誤差関数にはソフトマックスクロスエントロピーを用い、各隠れ層には batch 正則化 と ReLU 活性化関数を使用した。

F1 格要素候補の分散表現と解析対象述語の分散表現を素性に追加することで、F0 を拡張した。

F2 F1 の述語分散表現を格フレーム内平均ベクトル (MVC) に置き換えた。

F3 F2 に文脈ベクトルとして RNN の出力を追加した。RNN には GRU を使用した。図 1 に F3 の全容を示す。

表 5 にそれぞれの素性組合せを示す。なお、MVP は入力素性として使用していない。

5.2 データセット

実験データとして、『現代日本語書き言葉均衡コーパス』(BCCWJ) (Maekawa et al. 2014) のコアデータ⁹を使用した。BCCWJ のコアデータ約 2,000 文書に対しては、人手による述語項構造と照応関係が付与されており、これは新聞、雑誌、書籍、白書、Yahoo!知恵袋、Yahoo!ブログの 6 ドメインにまたがっている。ドメインの偏りに注意し、全体の約 4/5 を訓練用データ、約 1/20 を開発用データとし、残りを評価用データとして使用した。複数の格要素が同じ対象を指示している場合 (共参照)、本研究ではコーパスに付与された共参照情報をもとに出力を評価し、正しい照応先と共参照関係にある先行詞のいずれかを対応付けることが出来ていたならば正解とした。本研究で対象とした述語は動詞のみで、形容詞、事態性名詞は扱っていない。この理由として、動詞のゼロ照応解析において必要とされる手がかりと、形容詞や事態性名詞のゼロ照応解析において必要とされる手がかり、格要素と述語の出現傾向が異なるという点が挙げら

⁹ http://pj.ninjal.ac.jp/corpus_center/bccwj/

れる。松林, 吉野, 林部, 中山 (2015) が示しているように, 例えば, 形容詞は二格やヲ格をほとんど取らない。また, 事態性名詞は他の述語を項に取る場合があり, 格要素候補には名詞以外の品詞も含めて解析を行う必要がある。このため, 述語の種類を考慮しない単純な拡張による同時解析は解析精度の低下を招くと考えた。

6 結果と議論

6.1 結果

複数格同時推定の効果 表 6 は BCCWJ におけるゼロ照応解析の実験結果である。S0_each と S0 を比較すると, 多くの列において, S0 が S0_each より高い精度を示していることがわかる。ただし文間ガ格, 全体ガ格の列においては, S0_each が S0 より高い精度を示している。これは単格の推定では, 比較的精度の高いガ格の推定が他の格における誤りから影響を受けないため, 複数格同時推定の時より値が良くなっているのだと考えられる。一方で, 二格, ヲ格については, S0_each は他の格の情報が使えないため比較的精度が低く, 全体としての精度も, 複数格同時推定を行っている S0 に劣っている。我々はこの結果に対して有意水準 0.1% で McNemar 検定を行い, 統計的有意差を確認した。

また, 候補削減を行わない単一格解析の FNN モデルも用意し同条件で訓練し, テストを行ったものの学習が収束せず, 極めて低い精度を示した。このことから, 候補削減の重要性が示唆される。

解候補削減の効果 表 6 で, S0' と S0 を比較すると, すべての列において, S0 が S0' より精度が高い。我々はこの結果に対して有意水準 0.1% で McNemar 検定を行い, 統計的有意差を確認した。このことから我々の提案した解候補削減手法がうまく機能しているといえる。

分散表現と MVC の効果 ベースラインモデル (F0) に格要素と述語の分散表現を導入すると (F1), 全体の精度が低下した。しかしながら, 述語の分散表現を MVC に置き換えることで (F2)

表 6 BCCWJ における結果 (F 値)

モデル \	格 例数	文内				文間				All			
		ガ格	ヲ格	二格	All	ガ格	ヲ格	二格	All	ガ格	ヲ格	二格	All
S0_each (ベース)		.570	.730	.757	.643	.085	.016	.144	.080	.397	.602	.660	.480
S0' (ベース)		.490	.712	.725	.589	.032	.016	.140	.038	.331	.584	.632	.435
S0 (ベース)		.575	.758	.777	.661	.044	.016	.145	.048	.390	.628	.679	.491
F0 (ベース)		.523	.736	.775	.623	.054	.019	.151	.057	.356	.610	.677	.462
F1 (ベース, 格要素, 述語)		.470	.682	.762	.564	.141	.041	.138	.126	.342	.537	.659	.416
F2 (ベース, 格要素, MVC)		.563	.707	.773	.641	.103	.063	.154	.099	.394	.565	.674	.479
F3 (ベース, 格要素, MVC, 文脈)		.562	.751	.775	.647	.139	.038	.152	.126	.402	.620	.678	.488

精度は上がり, F0 を上回っている. これは述語の情報 (F1) の代わりに, 格フレーム情報を使う (F2) 方がより効果的であることを示している.

文脈ベクトルの効果 ローカルアテンション付き RNN モデルを使用し文脈情報を導入することで (F3), F2 に比べて文内・文間ともに改善が見られた. これは, モデルが前方文脈情報を効率的に学習できていることを示唆している. F3 は全体の精度においては S0 に劣っているが, 文間照応においては様々な素性を入力としたことにより, S0 より高い精度を示している.

優先度学習の有効性 様々な素性を使用した FNN モデルと比較して, ランキング SVM を用いた S0 のほうが全体では最も良い精度を示している. 磯崎, 賀沢, 平尾 (2006), 萩行, 河原, 黒橋 (2014) は日本語ゼロ照応解析に際して, 優先度学習を用いることは精度向上に効果があると指摘している. これは一般にゼロ照応解析がマルチインスタンス学習の枠組みで設定されていることによる. マルチインスタンス学習では, 一つの実体が複数のインスタンス (素性ベクトル) を持ち, 各実体が少なくとも一つの正例インスタンスを持つか否かはわかるがどのインスタンスが正例かはわからないという設定で学習を行う. マルチインスタンス学習のタスクに対して普通の二値分類を行った場合, 比較的解きやすい実体に含まれる正負インスタンスと解きにくい実体に含まれる正負インスタンスを区別なく同時に学習することになってしまい, 精度の悪化を招くと考えられる.

我々は Hangyo et al. (2013), 萩行 他 (2014) 同様, ランキング SVM を用いることで S0 には優先度学習を取り入れているが, 今回使用した FNN モデルにはこれに類する複雑な処理を取り入れておらず, 述語すべてのそれぞれの組合せ候補に対して区別なく二値分類の学習を行っている. その結果, 比較的解きやすい述語の多い文内においては S0 が高い精度を示し, 比較的解きにくい述語の多い文間においてはより多くの素性が与えられた FNN モデルが高い精度を示している. そのため, BCCWJ においては文間ゼロ照応より文内ゼロ照応の事例数のほうが多いことから, 全体の精度としては S0 が上回っている. これは, より多くの素性を与えられた FNN モデルがより難しい述語に適合しようとして簡単な述語に対する精度を下げてしまった可能性がある. 対策として, FNN モデルにも RankNet などの優先度学習を取り入れることで文内の精度が改善される可能性がある.

6.2 議論

候補削減の成功例 候補削減の成功例として BCCWJ の QA 文書である『00020_A_OC08_00002』の 2 行目に出現する「知ってたら」を解析対象述語とした際の候補リストを表 7 に示す.

電動歯ブラシの替えブラシって高い！ですね！？

家計に苦しいので復活&長持ちする方法知ってたら教えて下さい！

表 7 候補削減の成功例

格フレーム	ガ	ヲ	ニ
知る/しる~テ形+る/る:動 13	歯	方法	None
知る/しる~テ形+る/る:動 13	復活	方法	None
知る/しる~テ形+る/る:動 13	家計	方法	None
知る/しる~テ形+る/る:動 13	None	方法	None
知る/しる~テ形+る/る:動 13	歯	家計	None
知る/しる~テ形+る/る:動 13	方法	家計	None
知る/しる~テ形+る/る:動 13	復活	家計	None
知る/しる~テ形+る/る:動 13	None	家計	None
知る/しる~テ形+る/る:動 13	歯	長持ち	None
知る/しる~テ形+る/る:動 13	方法	長持ち	None
知る/しる~テ形+る/る:動 13	復活	長持ち	None
知る/しる~テ形+る/る:動 13	家計	長持ち	None
知る/しる~テ形+る/る:動 13	None	長持ち	None
知る/しる~テ形+る/る:動 13	歯	None	None
知る/しる~テ形+る/る:動 13	方法	None	None
知る/しる~テ形+る/る:動 13	復活	None	None
知る/しる~テ形+る/る:動 13	家計	None	None
知る/しる~テ形+る/る:動 13	None	None	None
知る/しる~テ形+る/る:動 27	歯	None	方法
知る/しる~テ形+る/る:動 27	None	None	方法
知る/しる~テ形+る/る:動 27	歯	None	None
知る/しる~テ形+る/る:動 27	方法	None	None
知る/しる~テ形+る/る:動 27	None	None	None

ガ, ヲ, ニ格に埋まる格要素のうち「知ってたら」と直接係り受け関係にある名詞はない。一見すると「方法」が「知ってたら」に係っているように思えるが, BCCWJ によれば, 「方法」は「教えて下さい!」に係っている。従ってここではガ格, ヲ格にゼロ代名詞が存在し, ガ格の格要素は外界二人称, ヲ格は「方法」, ニ格の格要素は照応なしであると定める。

京大格フレームには動詞「知る/しる~テ形+る/る」の格フレームが 32 種類ある。解析対象述語「知ってたら」を含む文から 3 文前まで（実際には 2 行目の述語なので, 1 文前まで）に含まれる名詞は「電動」, 「歯」, 「ブラシ」, 「替え」, 「ブラシ」, 「家計」, 「復活」, 「長持ち」, 「方法」の 9 個あり, これらが格要素候補となる。これらの名詞 9 個に加えて, 「外界照応 or 照応なし」を示す「None」を含む 10 個が実際の候補削減の対象となる格要素候補となる。これらの組合せは, $32 \times (10 \times 9 \times 8 + 3 \times 9 + 1) = 23,936$ 候補となる（「None」は重複を許容するため）。これに対して我々の提案した候補削減を施し, 23 個まで削減した結果が表 7 である。表 7 は正解の組合せ（「知る/しる~テ形+る/る:動 13」, 「None」, 「方法」, 「None」）を含むので上手く候

補削減ができています。

成功の理由を考察する。京大格フレーム中の「知る/知る~テ形+る/る:動 13」のヲ格の行で最も出現数が多い語は、「方法」が 715 回, 「術」が 207 回, 「使い方」が 144 回, 「コツ」が 111 回, 「仕方」が 88 回である。これらに対して正解のヲ格である「方法」は MVC との cosine 距離、出現確率によって高いスコアを割り当てられている。「知る/知る~テ形+る/る:動 13」のガ格の行で最も出現数が多い語は、「方」が 2 回, 「司祭」が 2 回, 「人」が 2 回, 他 22 種の語が 1 回となっていて、不正解のガ格である「菌」, 「復活」, 「家計」, 「方法」は候補リスト中出现してはいるものの、比較的低いスコアを割り当てられている。

候補削減の失敗例 候補削減の失敗例として BCCWJ の雑誌文書である『00016_A_PM12_00020』の 3 行目中出现する「入ってくる」を解析対象述語とした際の候補リストを表 8 に示す。

17

創は倒れている棚倉を見ても、顔色を変えなかった。

生死を冷静に確認するように観察してから、礼拝堂の中に入ってくる。

ガ, ヲ, ニ格に埋まる格要素のうち「中」は「入ってくる」と直接係り受け関係にある。従って、

表 8 候補削減の失敗例

格フレーム	ガ	ヲ	ニ
入る/はいる~テ形+くる/くる:動 3	中	None	観察
入る/はいる~テ形+くる/くる:動 3	生死	None	観察
入る/はいる~テ形+くる/くる:動 3	None	None	観察
入る/はいる~テ形+くる/くる:動 3	観察	None	中
入る/はいる~テ形+くる/くる:動 3	生死	None	中
入る/はいる~テ形+くる/くる:動 3	None	None	中
入る/はいる~テ形+くる/くる:動 3	観察	None	確認
入る/はいる~テ形+くる/くる:動 3	中	None	確認
入る/はいる~テ形+くる/くる:動 3	生死	None	確認
入る/はいる~テ形+くる/くる:動 3	None	None	確認
入る/はいる~テ形+くる/くる:動 3	観察	None	None
入る/はいる~テ形+くる/くる:動 3	中	None	None
入る/はいる~テ形+くる/くる:動 3	生死	None	None
入る/はいる~テ形+くる/くる:動 3	None	None	None
入る/はいる~テ形+くる/くる:動 11	中	None	None
入る/はいる~テ形+くる/くる:動 11	None	None	None
入る/はいる~テ形+くる/くる:動 17	観察	None	None
入る/はいる~テ形+くる/くる:動 17	None	None	None

本来であれば二格には「中」以外埋まらないものと制限して候補削減を行う。しかし、BCCWJのアノテーションによれば、「入ってくる」は「入って」と「くる」に分割され、「中に」は「くる」に係っているため、「入って」とは直接係り受け関係にないとされ、ゼロ代名詞とされている。ここではBCCWJのアノテーションに従って、ガ格、二格にゼロ代名詞が存在し、ガ格の格要素は「創」、ヲ格は照応なし、二格の格要素は「中」とであると定める。

京大格フレームには動詞「入る/はいる~テ形+くる/くる」の格フレームが24種類ある。解析対象述語「入ってくる」を含む文から3文前まで（実際には3行目の述語なので、2文前まで）に含まれる名詞は「17」、「創」、「棚倉」、「顔色」、「生死」、「確認」、「観察」、「礼拝」、「堂」、「中」の10個あり、これらが格要素候補となる。これらの名詞10個に加えて、「外界照応 or 照応なし」を示す「None」を含む11個が実際の候補削減の対象となる格要素候補となる。これらの組合せは、 $24 \times (11 \times 10 \times 9 + 3 \times 10 + 1) = 24,504$ 候補となる（「None」は重複を許容するため）。これに対して我々の提案した候補削減を施し、18個まで削減された結果が表8である。表8は正解の組合せ（「入る/はいる~テ形+くる/くる:動3」、「創」、「None」、「中」）を含まないため、候補削減としては失敗の例である。

京大格フレーム中の「入る/はいる~テ形+くる/くる:動3」のガ格の行で最も出現数が多い語は、「手」が205回、「男」が61回、「誰」が55回、「蜂」が30回、「彼」が28回である。これらに対して正解のガ格である「創」は我々の解析においては「創造」、「創作」などに近い一般名詞として捉えられてしまい、人名としての分散表現が与えられず、候補削減アルゴリズム中で低いスコアが割り当てられて、候補リストから漏れてしまっている。この候補削減失敗への対策としては、「創」を人名と判定し、人の意味を含む分散ベクトルを割り当てることが考えられる。「入る/はいる~テ形+くる/くる:動3」の二格の行で最も出現数が多い語は、「中」が17,328回、「間」が650回、「内」が361回、「下」が154回、「奥」が154回となっていて、正解の二格である「中」は比較的高いスコアを割り当てられて、候補リスト中にも出現が確認できる。

ドメイン別のゼロ照応解析の精度 我々は複数ドメインコーパスであるBCCWJ上で評価を行っている。テスト事例に対する解析結果をドメインごとに表9に示す。

表9 BCCWJのドメインごとの実験結果のF値

格	文内				文間				All			
	ガ格	ヲ格	二格	All	ガ格	ヲ格	二格	All	ガ格	ヲ格	二格	All
OW	.545	.707	.798	.626	.037	.000	.316	.056	.459	.654	.756	.549
PB	.556	.673	.775	.624	.140	.039	.149	.130	.390	.565	.670	.467
PN	.553	.721	.768	.641	.100	.033	.245	.098	.410	.621	.720	.511
PM	.583	.707	.756	.641	.148	.000	.208	.132	.379	.561	.651	.446
OC	.677	.629	.605	.648	.161	.028	.062	.117	.423	.418	.442	.425
OY	.589	.705	.785	.660	.069	.050	.000	.057	.422	.560	.618	.496

OW (白書), PN (新聞) は一文の語数が多く, 格要素候補の名詞も多数出現するため, 組合せ候補数が増加する傾向にある. 専門的な語や他のドメインでは出現頻度の低い述語が出現しやすい. しかし文法的には比較的正しい書き方がなされており, 文意が取れないほどの省略は少なく, 直接係り受け関係にある格要素が多く存在する. また, 表 2 に示した通り, ガ格ゼロ照応のうち OW は 72%, PN は 51% が文内ゼロであり, 直接係り受け関係にある格要素も比較的多く存在する. そのため, 我々の組合せを用いる解析手法では, 他の格の直接係り受け関係にある格要素を格フレーム選択の手がかりとすることで, 他のドメインにおける精度と比較して全体で高いスコアを示している. また, 多量の格要素候補に対しても提案候補削減が上手く機能し精度の向上に貢献している.

PB (書籍), PM (雑誌) は一文が短く, 文法的に砕けた文が多く見られ, 一読しただけでは文意が捉えきれないような省略が散見される. 倒置表現や, 表 2 に示した通り 3 文より前に格要素が出現することもあり. そのため, 他のドメインに比べて文間の精度は高いものの, 全体の精度は低い. 特に PB, PM に関しては文書ごとに取り扱う題材や書き手に多様性があるため, ドメイン適応の手法を用いても精度の改善につながらない. 我々の手法では格要素候補として述語を含む文から 3 文前までの文中に出現する名詞のみを収集しているが, これをより広範囲に広げることが対策として考えられる. また, 使用する大規模格フレームを現在は Web コーパスのみから構築しているが, より多くのドメインから構築することで, 多様な述語に対して適切な格フレームを選択することができると考えられる. その際, 爆発的に増加する格要素候補と格フレーム候補の組合せに対して, 我々の提案する候補削減手法はより有効に働くものと予想される.

OC (QA) は文内, 文間のガ格において高い精度が出ている. これは, QA におけるガ格の多くが一人称か二人称であり, 他のドメインに比べて簡単なタスクとなっているためだと思われる. 一方, 表 2 に示した通り OC は文間ゼロの比率が最も高く, 3 文より前に格要素が出現することもあるため, 全体の精度は最も低い. OC に出現する文は PB, PM 以上に文法的に砕けた文が多く見られ, 一読しただけでは文意が捉えきれないような省略が散見される. このドメインにおけるゼロ照応解析の精度を上げるためには, 書き手の交代や質問文であるか否かの OC に特有の素性や, 文書全体の構造を手がかりとすることが必要だと考えられる.

OY (ブログ) はタイトルや箇条書き, メールアドレス, URL, 日時, 空白記号が多く出現し, 書かれている内容も日記, 政治ニュースに関する意見, 技術的な説明文など様々である. 我々の手法はより細分化された大規模格フレームを組合せで用いているため, 出現述語の表記ゆれや助動詞の変化にも頑健で, 他のドメインと比較して文内の全体において最も高いスコアが出ている.

係り受け関係とゼロ照応解析の適合率の関係 我々の提案するモデルは複数の異なる格要素を同時に同定する. その効果を見るために, 我々はこの格がすでに係り受け解析によって埋まっ

ているかに基づいてテスト事例を分類した。表 10 はそれぞれの組合せにおける適合率である。列はすでに直接係り受け関係で埋められた格を示し、行はシステムによって埋められるべき格を示す。例えば、『ガ格』行、『ヲ格』列の数字は、係り受け解析によって『ヲ格』の格要素が与えられた上で、『ガ格』の格要素を同定した適合率である。

直感的には、すでに直接係り受け関係で埋められた格が多いほど適合率は上がるものと予想されるが、『ガ格』の行で特に顕著なように、すでに直接係り受け関係で埋められた格が多いほど適合率が下がる傾向が見られる場合もある。これは、今回使用している京大格フレームが Web 上の文における直接係り受け関係の情報のみを収集していて、ゼロ照応の情報はほとんど考慮していないため、我々がゼロ照応解析を行うに際しては、その偏りの影響を受けているのだと考えられる。例えば、実際の日本語文においては、ガヲニ格を取りうる述語であっても、『ガとヲのみ直接係り受け関係で格要素が存在し、ニ格が省略されている文』と『ガヲニすべて直接係り受け関係で格要素が存在する文』では、前者のほうが多く出現し京大格フレーム中の用例が充実している可能性がある。この時、前者の文の出現によって京大格フレームに収録されるのはガ格とヲ格のみが出現する格パターンの用例であり、後者の文の出現によって収録されるのはガ格とヲ格とニ格が出現する格パターンの用例であるため、後者の格パターンが比較的スパスになる。そのような偏りを持つ京大格フレームを解析に用いた結果、『ガ格』行『ヲ格』列は、京大格フレーム中の豊富な用例を用いることができるため比較的適合率が高く、一方で、『ガ格』行『ヲ格、ニ格』列は、用例がスパスであるため比較的適合率が低くなる、と考えられる。従って、解析に際して直接係り受け関係の格要素のみでなくゼロ照応の格要素も考慮した格フレームを用いたならば、直感に反せず、すでに直接係り受け関係で埋められた格が多いほど適合率は上がるものと予想される。

また、特に下線付きのセルが事例数が多いにも関わらず適合率が低いいため、これらの下線付きセルの述語に関するパフォーマンスを改善することが、ゼロ照応解析全体の精度を上昇させることに對して重要である。

表 10 係り受け関係とゼロ照応解析の関係

ゼロ照応として 埋められるべき格	すでに直接係り受け関係で埋められた格						
	項なし	ガ格	ヲ格	ニ格	ガ格, ヲ格	ヲ格, ニ格	ガ格, ニ格
ガ格	.313 (1,645)	—	.282 (1,870)	.287 (683)	—	.243 (292)	—
ヲ格	.257 (416)	.384 (656)	—	.247 (81)	—	—	.750 (1,222)
ニ格	.505 (111)	.430 (337)	.319 (47)	—	.419 (129)	—	—
ガ格, ヲ格	.112 (492)	—	—	.144 (90)	—	—	—
ヲ格, ニ格	.091 (33)	.057 (35)	—	—	—	—	—
ガ格, ニ格	.228 (281)	—	.189 (122)	—	—	—	—
ガ格, ヲ格, ニ格	.000 (21)	—	—	—	—	—	—

NAIST テキストコーパス (NTC) による実験 表 3 に示した通り, 日本語文間ゼロ照応を取り扱う研究としては Imamura et al. (2009), Sasano and Kurohashi (2011), Hangyo et al. (2013), Shibata and Kurohashi (2018) が挙げられる. このうち我々が対象とする文内・文間のゼロ照応を行っているのは Sasano and Kurohashi (2011) のみであり, かつ彼らは NTC 上の評価実験で文内・文間の state-of-the-art を達成している. 我々は NTC を用いて評価実験を行うことで, 本稿の提案手法を Sasano and Kurohashi (2011) と比較した. 表 11 はその実験結果を示す. Sasano and Kurohashi (2011) の設定と揃えるために, ここで対象とした述語は動詞と形容詞とする.

我々の提案手法は比較的精度が出にくい文内のヲ格, ニ格において Sasano and Kurohashi (2011) より高い精度を示している. これは我々が使用する京大格フレームが Sasano and Kurohashi (2011) の使用した版より巨大かつ細分化されているため, 解析対象述語の語彙により適した格フレームを選ぶことで, 適切な格要素を同定できているのだと考えられる.

一方で, 文内, 文間のガ格においては Sasano and Kurohashi (2011) より低い精度を示している. これは Sasano and Kurohashi (2011) が格要素候補に対して, 100 万の名詞を 2,000 個の意味クラスに分類した結果を使用し, 汎化することで格フレームのカバレッジを補っている効果によるものと思われる. 我々の手法は word2vec を使用した分散表現同士の cosine 距離を使用することで京大格フレームの汎化を行っているが, 格要素の詳細な分類が述語に対してより重要な意味を持つガ格のゼロ照応解析においては, Sasano and Kurohashi (2011) の手法のほうが優れていると考えられる. ただしこれらの手法は排他的でなく, 同時に使用することも可能であるため, 双方を候補削減, ゼロ照応解析の素性として格ごとに使い分けることでさらなる精度の向上が期待できる.

『Project NEXT 述語項構造タスク』との比較 松林 他 (2015) は複数の述語項構造解析器の評価を BCCWJ 上で行っている. ただし彼らは NTC 上で解析器の訓練を行い, BCCWJ の一部データにおいてテストをしているため, 我々の提案手法との厳密な比較はできない. しかしできる限り条件を揃えて, 彼らのスコアと S0 を比較した. 表 12 にその結果をまとめる. 松林 他 (2015) は BCCWJ の以下 22 文書 1,625 文をテストデータとして使用している.

- OW: OW6X_00000 OW6X_00010 の 2 記事
- OY: OY01_00082 OY04_00001 OY04_00017 OY04_00027 OY10_00067 OY12_00005 の 6 記事

表 11 NTC を用いた実験結果の F 値

格 例数	文内				文間				All			
	ガ格	ヲ格	ニ格	All	ガ格	ヲ格	ニ格	All	ガ格	ヲ格	ニ格	All
S0 (ベース)	.315	.339	.095	.295	.078	.006	.000	.072	.281	.331	.093	.272
Sasano and Kurohashi (2011)	.395	.175	.089		.244	.066	.026					

表 12 『Project NEXT 述語項構造タスク』の指定する BCCWJ 文書を用いた実験結果の F 値

格 例数	文内				文間				All			
	ガ格	ヲ格	ニ格	All	ガ格	ヲ格	ニ格	All	ガ格	ヲ格	ニ格	All
松林, 乾 (2014)	.482	.143	.125	.398								
Yoshino, Mori, and Kawahara (2013)	.316	.238	.348	.309	.094	.000	.039	.078				
KNP	.411	.199	.236	.359	.164	.087	.022	.140				
S0 (ベース)	.547	.756	.738	.634	.045	.000	.270	.058	.391	.661	.664	.492

- PB: PB12_00001 PB2n_00003 PB40_00003 PB42_00003 PB50_00003 PB59_00001 の 6 記事
- PM: PM12_00006 PM24_00003 PM25_00001 PM26_00004 の 4 記事
- PN: PN1b_00002 PN1c_00001 PN1d_00001 PN3b_00001 の 4 記事

S0 の訓練にはこれらを除くすべての BCCWJ データを使用した。松林 他 (2015) が取り上げている解析器は 松林, 乾 (2014) の解析器, Yoshino et al. (2013) の解析器, KNP (笹野, 河原, 黒橋, 奥村 2013) の 3 種類で, これらを様々な条件下で比較している。ここでは BCCWJ において述語項構造をアノテーションされているすべて動詞から「れる」, 「せる」, 「できる/出来る/可能」, 「である」を伴うものを除き, Yoshino et al. (2013) の解析器には係り受け情報を与え, KNP には-anaphora モードでベストアラインメントを用いて解析を行った結果のスコアである。具体的には 松林 他 (2015) 中の表 50, 表 138, 表 295 の数字をここでは引用している。ただし, 我々の提案手法では京大格フレームとの整合性を保つために, BCCWJ のアノテーションに基づいて, 補助動詞「てあげる, てもらう, など」, 可能「できる, れる, られる」, 受け身「れる, られる」, 使役「せる, させる」, 願望「たい」を伴う動詞を除いている。そのため, 松林 他 (2015) の示す各格の事例数とは少し数字が異なる。表 12 中の数字は我々の基準における例数である。また, 松林, 乾 (2014) の解析器, Yoshino et al. (2013) の解析器, KNP は直接係り受け関係にある格要素の解析も同時に行っている。KNP は述語位置の判定も行っていて, 一部の動詞については述語位置を誤って判定しているため, その分だけ精度が下がっている。

結果を比較すると, 文内において我々の提案手法は他の解析器に大きく差を付けて高い精度を示している。この理由として, 提案手法においては直接係り受け関係にある格要素の情報を使用しているため, より正確に格フレームを選んでいる点が挙げられる。一方で, 文間のガ格で Yoshino et al. (2013) の解析器のほうが精度が高いのは, 彼らの解析器は格ごとに解析を行っているため, S0_each 同様, 他の格の解析エラーの影響を受けないためであると思われる。また, 文間のガ格で KNP の精度が高いのは KNP が Sasano and Kurohashi (2011) と同様に意味クラスを使用した文間照応を行っている効果によるものと思われる。

7 結論

本論文では分散表現で平均化した格フレームによる解候補削減を用いた日本語文内・文間ゼロ照応モデルを提案した。提案した解候補削減アルゴリズムによって大規模な複数ドメインコーパスによる効率的な訓練を可能とした。また、ローカルアテンション機構付き RNN と FNN を組合せて使用し様々な素性を取り入れることで、文間ゼロ照応解析においてより高い精度が出ることを確認した。

我々の今後の課題はタスクの対象を形容詞、事態性名詞にまで拡張し、より実用的なモデルを構築することである。また、対象格要素候補の名詞が先行文脈中の述語にどの格の格要素として取られたかの情報を、我々の提案した解候補削減アルゴリズムに取り入れることで、より良い解候補削減が行えるよう改善する予定である。また、解候補削減を行わない場合、提案 FNN モデルの学習が発散することから、本研究で提案したニューラルネットを用いた解析アーキテクチャそのものに改善の余地があると考えられる。

謝 辞

Hangyo et al. (2013) に関して詳細な情報をご教示くださった萩行正嗣氏、Ouchi et al. (2017) の全体像についてご教示くださった大内啓樹氏に厚く御礼申し上げます。本論文の内容の一部は、The 32nd Pacific Asia Conference on Language, Information and Computation, 第5回自然言語処理シンポジウム（第238回自然言語処理研究発表会）で発表したものである（Yamashiro, Nishikawa, and Tokunaga 2018; 山城, 西川, 徳永 2018）。The authors would like to thank Enago (www.enago.jp) for the English language review.

参考文献

- Chen, C. and Ng, V. (2016). “Chinese Zero Pronoun Resolution with Deep Neural Networks.” In *ACL*, pp. 778–788.
- Hangyo, M., Kawahara, D., and Kurohashi, S. (2012). “Building a Diverse Document Leads Corpus Annotated with Semantic Relations.” In *PACLIC*, pp. 535–544.
- Hangyo, M., Kawahara, D., and Kurohashi, S. (2013). “Japanese Zero Reference Resolution Considering Exophora and Author/Reader Mentions.” In *EMNLP*, pp. 924–934.
- 萩行正嗣, 河原大輔, 黒橋禎夫 (2014). 外界照応および著者・読者表現を考慮した日本語ゼロ照応解析. 自然言語処理, **21** (3), pp. 563–600.
- 林部祐太, 河原大輔, 黒橋禎夫 (2015). 格パターンの多様性に頑健な日本語格フレーム構築. 情

報処理学会第 224 回自然言語処理研究会, NL-224-14, pp. 1–8.

He, L., Lee, K., Lewis, M., and Zettlemoyer, L. (2017). “Deep Semantic Role Labeling: What Works and What’s Next.” In *ACL*, pp. 473–483.

Iida, R., Komachi, M., Inui, K., and Matsumoto, Y. (2007). “Annotating a Japanese Text Corpus with Predicate-Argument and Coreference Relations.” In *Proceedings of the Linguistic Annotation Workshop*, pp. 132–139.

Iida, R. and Poesio, M. (2011). “A Cross-Lingual ILP Solution to Zero Anaphora Resolution.” In *ACL*, pp. 804–813.

Iida, R., Torisawa, K., Hashimoto, C., Oh, J.-H., and Kloetzer, J. (2015). “Intra-sentential Zero Anaphora Resolution using Subject Sharing Recognition.” In *EMNLP*, pp. 2179–2189.

Iida, R., Torisawa, K., Oh, J.-H., Kruengkrai, C., and Kloetzer, J. (2016). “Intra-Sentential Subject Zero Anaphora Resolution using Multi-Column Convolutional Neural Network.” In *EMNLP*, pp. 1244–1254.

Imamura, K., Saito, K., and Izumi, T. (2009). “Discriminative Approach to Predicate-Argument Structure Analysis with Zero-Anaphora Resolution.” In *ACL-IJCNLP*, pp. 85–88.

磯崎秀樹, 賀沢秀人, 平尾努 (2006). 辞書式順序を持つペナルティによるゼロ代名詞解消. 情報処理学会論文誌, **47** (7), pp. 2279–2294.

Joachims, T. (2006). “Training Linear SVMs in Linear Time.” In *Proceedings of the 12th ACM SIGKDD*, pp. 217–226.

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.

Kudo, T., Ichikawa, H., and Kazawa, H. (2014). “A Joint Inference of Deep Case Analysis and Zero Subject Generation for Japanese-to-English Statistical Machine Translation.” In *ACL*, pp. 557–562.

Kurita, S., Kawahara, D., and Kurohashi, S. (2018). “Neural Adversarial Training for Semi-supervised Japanese Predicate-argument Structure Analysis.” In *ACL*, pp. 474–484. Association for Computational Linguistics.

Luong, M.-T., Pham, H., and Manning, C. D. (2015). “Effective Approaches to Attention-based Neural Machine Translation.” In *EMNLP*, pp. 1412–1421.

Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., Koiso, H., Yamaguchi, M., Tanaka, M., and Den, Y. (2014). “Balanced Corpus of Contemporary Written Japanese.” *Language Resources and Evaluation*, **48** (2), pp. 345–371.

松林優一郎, 乾健太郎 (2014). 統計的日本語述語項構造解析のための素性設計再考. 言語処理学会第 20 回年次大会発表論文集, pp. 360–363.

松林優一郎, 吉野幸一郎, 林部祐太, 中山周 (2015). Project Next 述語項構造タスク. 言語処理

- 学会第 21 回年次大会発表論文集, pp. 1–157.
- Matsubayashi, Y. and Inui, K. (2017). “Revisiting the Design Issues of Local Models for Japanese Predicate-Argument Structure Analysis.” In *IJCNLP*, pp. 128–133.
- Matsubayashi, Y. and Inui, K. (2018). “Distance-Free Modeling of Multi-Predicate Interactions in End-to-End Japanese Predicate-Argument Structure Analysis.” In *COLING*, pp. 94–106. Association for Computational Linguistics.
- Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). “Efficient Estimation of Word Representations in Vector Space.” *arXiv preprint arXiv:1301.3781*.
- Ouchi, H., Shindo, H., Duh, K., and Matsumoto, Y. (2015). “Joint Case Argument Identification for Japanese Predicate Argument Structure Analysis.” In *ACL-IJCNLP*, pp. 961–970.
- Ouchi, H., Shindo, H., and Matsumoto, Y. (2017). “Neural Modeling of Multi-Predicate Interactions for Japanese Predicate Argument Structure Analysis.” In *ACL*, pp. 1591–1600.
- Rello, L., Baeza-Yates, R., and Mitkov, R. (2012). “Elliphant: Improved Automatic Detection of Zero Subjects and Impersonal Constructions in Spanish.” In *EACL*, pp. 706–715.
- Sasano, R., Kawahara, D., and Kurohashi, S. (2008). “A Fully-Lexicalized Probabilistic Model for Japanese Zero Anaphora Resolution.” In *COLING*, pp. 769–776.
- 笹野遼平, 河原大輔, 黒橋禎夫, 奥村学 (2013). 構文・述語項構造解析システム KNP の解析の流れと特徴. 言語処理学会第 19 回年次大会発表論文集, pp. 110–113.
- Sasano, R. and Kurohashi, S. (2011). “A Discriminative Approach to Japanese Zero Anaphora Resolution with Large-scale Lexicalized Case Frames.” In *IJCNLP*, pp. 758–766.
- Shibata, T., Kawahara, D., and Kurohashi, S. (2016). “Neural Network-Based Model for Japanese Predicate Argument Structure Analysis.” In *ACL*, pp. 1235–1244.
- Shibata, T. and Kurohashi, S. (2018). “Entity-Centric Joint Modeling of Japanese Coreference Resolution and Predicate Argument Structure Analysis.” In *ACL*, pp. 579–589. Association for Computational Linguistics.
- Sudo, K., Sekine, S., and Grishman, R. (2001). “Automatic pattern acquisition for Japanese information extraction.” In *HLT*, pp. 1–7.
- 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.
- Yamada, H., Teufel, S., and Tokunaga, T. (2017). “Designing an Annotation Scheme for Summarizing Japanese Judgment Documents.” In *KSE*, pp. 275–280.
- 山城颯太, 西川仁, 徳永健伸 (2018). 大規模格フレームによる解析候補削減を用いたニューラルネットゼロ照応解析. 情報処理学会第 238 回自然言語処理研究会, NL-238-5, pp. 1–9.
- Yamashiro, S., Nishikawa, H., and Tokunaga, T. (2018). “Neural Japanese Zero Anaphora Res-

olution using Smoothed Large-scale Case Frames with Word Embedding.” In *PACLIC*.

Yin, Q., Zhang, Y., Zhang, W., and Liu, T. (2017). “Chinese Zero Pronoun Resolution with Deep Memory Network.” In *EMNLP*, pp. 1309–1318.

Yin, Q., Zhang, Y., Zhang, W., Liu, T., and Wang, W. Y. (2018a). “Zero Pronoun Resolution with Attention-based Neural Network.” In *COLING*, pp. 13–23. Association for Computational Linguistics.

Yin, Q., Zhang, Y., Zhang, W.-N., Liu, T., and Wang, W. Y. (2018b). “Deep Reinforcement Learning for Chinese Zero Pronoun Resolution.” In *ACL*, pp. 569–578. Association for Computational Linguistics.

Yoshino, K., Mori, S., and Kawahara, T. (2013). “Predicate Argument Structure Analysis using Partially Annotated Corpora.” In *IJCNLP*, pp. 957–961.

Zhou, J. and Xu, W. (2015). “End-to-end Learning of Semantic Role Labeling Using Recurrent Neural Networks.” In *ACL*, pp. 1127–1137.

付録

A 素性詳細

表 13 ベースモデル素性一覧

タイプ	値	説明
格フレーム素性	log	e が cf の c に対応付く確率
	log	e の持つ JUMAN カテゴリが cf の c に対応付く確率
	log	e の持つ固有表現が cf の c に対応付く確率
	log	e の出現確率
	log	e の持つカテゴリの出現確率
	log	e の固有表現の出現確率
	log	e と cf の c の間の自己相互情報量
	log	e の持つ JUMAN カテゴリと cf の c の間の自己相互情報量
	log	e の持つ固有表現と cf の c の間の自己相互情報量
	log	上記 3 つの自己相互情報量のうち最大のものの値
	log	cf の c が何らかの要素で対応付けられる確率
	log	cf が持つ用例のうち c の用例の割合
用言素性	バイナリ	cf の c が必須格かどうか
	バイナリ	用言 p の持つモダリティ
	バイナリ	用言 p の持つ敬語表現
	バイナリ	用言 p の持つ時制
	バイナリ	用言 p が可能表現かどうか

	バイナリ	用言 p の係り受けタイプ (連用, 連体, 文末)
	バイナリ	用言 p が動態述語か状態述語か
文脈素性	バイナリ	e が持つ固有表現の種類
	整数値	e が入力文 t で言及された回数
	整数値	e が入力文 t で用言 p より前方で言及された回数
	整数値	e が入力文 t で用言 p より後方で言及された回数
	バイナリ	e が対象の文で助詞「は」を伴って言及されたか
	バイナリ	e が何文前で言及されたか
	バイナリ	e が言及された位置の位置カテゴリ (Sasano and Kurohashi 2011)
	バイナリ	e が先頭の文で言及されたか
	バイナリ	e が先頭の文で助詞「は」を伴って言及されたか
	バイナリ	e が先頭の文で助詞「が」を伴って言及されたか
	バイナリ	e が先頭の文の文頭で言及されたか
	バイナリ	e が先頭の文の文頭で助詞「は」を伴って言及されたか
	バイナリ	e が先頭の文の文末で言及されたか
	バイナリ	e が先頭の文の文末で判定詞を伴って言及されたか
	バイナリ	e が先頭の文の文末で体言止めで言及されたか
	バイナリ	e の権限性スコアが 1 以上あるか (Sasano and Kurohashi 2011)
その他	バイナリ	c が対応付けられた

略歴

山城 颯太：2017 年東京工業大学工学部情報工学科卒業。2019 年同大学情報理工学院情報工学系修了。同年、ヤフー株式会社入社。

西川 仁：2006 年慶應義塾大学総合政策学部卒業。2008 年同大学大学院政策・メディア研究科修士課程修了。同年、日本電信電話株式会社入社。2013 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。NTT メディアインテリジェンス研究所研究員を経て、2015 年より東京工業大学情報理工学院助教。2017 年 IE 経営大学院修士課程修了。博士 (工学)。自然言語処理、特に自動要約、自然言語生成の研究に従事。The Association for Computational Linguistics, 言語処理学会, 人工知能学会, 情報処理学会各会員。

徳永 健伸：1983 年東京工業大学工学部情報工学科卒業。1985 年同大学院理工学研究科修士課程修了。同年 (株) 三菱総合研究所入社。1986 年東京工業大学大学院博士課程入学。現在、東京工業大学情報理工学院教授。博士 (工学)。専門は自然言語処理、計算言語学。情報処理学会, 人工知能学会, 言語処理学会, 計量国語学会, Association for Computational Linguistics, ACM SIGIR, Cognitive Science Society, International Cognitive Linguistics Association 各会員。

(2018 年 11 月 1 日 受付)

(2019 年 3 月 5 日 再受付)

(2019 年 4 月 12 日 採録)