# 日本語照応解析における深層格推定に基づいた先行詞の同定

伊藤清晃　寺岡丈博
拓殖大学大学院工学研究科
20m301@st.takushoku-u.ac.jp　tteraoka@cs.takushoku-u.ac.jp

## 概要

日本語の照応解析において，先行詞の同定は文意を解釈する必要があるため容易ではない．照応され易い語として，一般的に主語や目的語が挙げられ，深層格に対応づけると動作主格と対象格が多い．そこで，本研究では動詞，名詞，助詞の関係から推定した深層格を用いて先行詞を同定する手法を提案した．先行研究と同様に，NAIST テキストコーパスにおける先行詞の同定精度を評価した結果，本手法の有用性を確認することができた．

## 1　背景と目的

省略された項をはじめ代名詞や指示詞を照応詞と呼び，照応解析はそれらと照応関係にある先行詞を同定する手法である．本研究における照応関係とは，同じ内容や対象を示す異なる 2 語間に成立する関係を指す．例えば，「太郎はパンを買った．そして，それを食べた.」においては，1 つ目の文の「太郎」が先行詞，2 つ目の文の「それ」が照応詞であり，両者の間に照応関係が成立する．また，「太郎は疲れていました．そして眠ってしまいました.」においては，2 つ目の文の「太郎」が先行詞であり，2 つ目の文の「眠ってしまいました」の前に省略された代名詞（ゼロ代名詞）が照応詞となり両者の間に照応関係が成立する．ゼロ代名詞による照応関係を特にゼロ照応と呼ぶ．特にゼロ照応では，照応詞が省略されているものに対してどこに省略部分があるか，何が省略されているかを明らかにする必要がある．とりわけ，人間にとって当たり前である事柄が省略される傾向にあるため，このような表現が含まれる文は理解が困難である．

照応解析は，機械翻訳や対話システムなどの自然言語処理の応用研究において，文脈を理解するために必要とされる技術である．日本語において，一般的に言い換えられる語や省略され易い語として，主語や目的語が挙げられる．これらの語を述語と項の意味的な関係（深層格）に当てはめると動作主格と対象格が多い傾向がある．また，照応解析における従来の研究では，大規模な辞書やシソーラスを用いて学習した人工ニューラルネットワーク（ANN）や学習モデルを利用した手法がよくみられる．そこでは，特徴量に，格フレームや格要素，動詞と名詞と助詞の関係性を用いている．これらの研究では，位置関係や格助詞などによって決まる格（表層格）を利用していることが多い．

人は先行詞を見つける際に，文章と知識を照らし合わせて，雑多な手がかりを利用している．その中でも，先行詞を同定しやすくするには，表層格と深層格を組み合わせることが重要だと考える．そこで，表層格と深層格を組み合わせるために，動詞連想概念辞書 [1] と述語項構造シソーラスを利用する．本研究では，照応解析における先行詞の同定において，深層格と表層格を組み合わせて利用する手法の提案と，その有用性を検証することを目的とする．

## 2　関連研究

照応解析の研究として，笹野らの研究 [2] や西念らの研究 [3]，山城らの研究 [4]，河崎らの研究 [5] が挙げられる．笹野らの研究では，比較的小規模の照応関係タグ付きコーパスから獲得した構文的な手がかり，および，大規模なタグなしコーパスから獲得した意味クラスといった語彙的手がかりを素性として用いた対数線形モデルを用いて先行詞を同定している．そして，が格，を格，に格のみを対象としており，本手法と異なり，他の格を考慮していない．また，格フレーム辞書から素性を作成しているため，本手法とは異なり，先行詞と深層格の関係を考慮していない．

西念らの研究では，先行詞候補に対して，概念類似度情報や距離情報などの 素性値を取得し，各素性値の正解先行詞としての尤度を求め，ナイーブベイズ法による尤度スコアを求めることにより，先行詞を同定している．西念らは，先行詞の深層格につい

て考慮されているが，本手法では動詞と名詞，助詞の関係から ANN で，先行詞候補の深層格を推定し，先行詞を同定する．

山城らの研究では，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ている．述語や格要素の分散表現をはじめとした素性を使用している．山城らは，格フレームにより述語と格要素の関係を利用することで，先行詞を同定している．対して，本手法では，先行詞候補を深層格により絞り込み，動詞連想概念辞書や述語項構造シソーラスから得られる動詞と名詞の関係や名詞と深層格の関係を用いて，先行詞を同定する．

河崎らの研究では，ANN を使用して深層格を用いた照応解析の手法を提案している．河崎らは，ANN による先行詞の誤割り当てを防ぐために名詞クラスと代名詞の作成と導入を行なった．また，使用した ANN は動詞項シソーラスを使用しており，入力として，動詞と先行詞候補，助詞を入力する．河崎らの手法では，同じ名詞クラスに分類された名詞を区別することが十分ではなかった．対して，本研究では単語分散表現を利用する事で，この問題の解決を試みた．

## 3 提案手法

図 1 は，提案手法の処理手順を表している．はじめに，文章から名詞を抽出し，文間距離によって先行詞候補の名詞を決定する．このとき，先行詞候補に含める文は対象の文に加えて，対象文の前の 3 文とする．これは，文中のゼロ代名詞が前の 3 文内に 80 %の割合で出現することを踏まえて決定している．

次に，学習した人工ニューラルネットワーク（ANN）を用いて，先行詞候補の名詞の深層格を求める．ANN の入力には，動詞と名詞と助詞の特徴をそれぞれ数値化したものを入力する．動詞の特徴量として，述語項構造シソーラスにある大分類 2 に基づいて動詞クラスを 53 クラスに分け，one-hot ベクトルを作成する．名詞の特徴量として，分類語彙表 [6] の中項目に基づいて名詞クラスを 51 クラスに分け，1 つの名詞が複数のクラスに割り当てることができるように，ダミー変数を作成する．助詞の特徴量として，述語項構造シソーラスにある表層格の助詞が 36 種あり，これらの助詞から one-hot ベクトルを作成する．

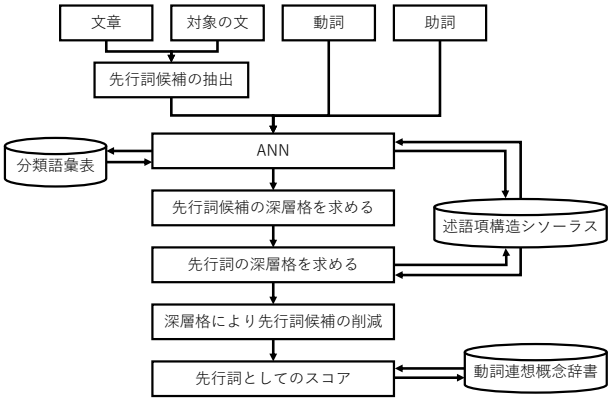


**図 1** 提案手法の処理手順

この ANN は複数の深層格が割り当てられるように，全ての深層格に対して割り当てられる確率を算出可能なニューラルネットワークを用いる．割り当てる深層格は述語項構造シソーラスの意味役割から 32 種類にしている．ANN の出力の中で，最も高い出力値の深層格をその先行詞候補の名詞の深層格とする．そしてその出力値を深層格スコアとする．

さらに，動詞と助詞を手がかりに，述語項構造シソーラスから先行詞が取り得る深層格を抽出する．先行詞候補の名詞が割り当てられた深層格が，抽出した深層格と一致しているものを先行詞候補として残す．残った先行詞候補の名詞の深層格と動詞連想概念辞書の連想課題を対応させる．この動詞連想概念辞書（VACD）は, 連想実験のデータから構築された辞書で, 実際に人間が言葉に対して連想した内容をまとめたものである [1]. これを用いることで人間が普段, 文章を読む際に言葉に対して連想する情報をコンピュータが利用することができる．

対応させた連想課題と動詞に対応する連想語を動詞連想概念辞書から抽出する．そして，抽出した全ての連想語と先行詞候補の名詞の類似度をそれぞれ計算し，得た類似度の最大値を連想スコアとする．また，京都大学格フレーム（CF）[7] と共起概念ベース（共起 GB）から動詞に対応する語を抽出し，先行詞候補の名詞と類似度をそれぞれ計算して得た類似度を動詞スコア，共起スコアとする．これらのスコアを組み合わせて先行詞の同定を行う．

## 4 評価実験

### 4.1 実験設定

NAIST テキストコーパス（NTC）を利用して，評価用データを作成する．評価対象は，条件を満たす

**表 1** 実験 1 の結果

| | ANN | CF | VACD | 共起 GB |
|---|---|---|---|---|
| top1 | 0.185 ( 37 / 200 ) | **0.235** ( 47 / 200 ) | **0.265** ( 53 / 200 ) | 0.090 ( 18 / 200 ) |
| top3 | 0.505 ( 101 / 200 ) | 0.525 ( 105 / 200 ) | 0.520 ( 104 / 200 ) | 0.255 ( 51 / 200 ) |
| top5 | 0.625 ( 125 / 200 ) | 0.670 ( 134 / 200 ) | 0.655 ( 131 / 200 ) | 0.415 ( 83 / 200 ) |
| top10 | 0.865 ( 173 / 200 ) | 0.865 ( 173 / 200 ) | 0.880 ( 176 / 200 ) | 0.620 ( 124 / 200 ) |

**表 2** 実験 2 の結果

| | ベースライン | 先行研究 | 提案手法 | | |
|---|---|---|---|---|---|
| | ANN | ANN | ANN | AV | ACVK |
| top1 | 0.185 ( 37 / 200 ) | 0.190 ( 38 / 200 ) | **0.265** ( 53 / 200 ) | **0.345** ( 69 / 200 ) | **0.265** ( 53 / 200 ) |
| top3 | 0.505 ( 101 / 200 ) | 0.500 ( 100 / 200 ) | 0.550 ( 110 / 200 ) | 0.610 ( 122 / 200 ) | 0.615 ( 123 / 200 ) |
| top5 | 0.625 ( 125 / 200 ) | 0.640 ( 128 / 200 ) | 0.680 ( 136 / 200 ) | 0.720 ( 144 / 200 ) | 0.750 ( 150 / 200 ) |
| top10 | 0.865 ( 173 / 200 ) | 0.820 ( 164 / 200 ) | 0.820 ( 164 / 200 ) | 0.840 ( 168 / 200 ) | 0.855 ( 171 / 200 ) |

事例のみとする．1 つ目は，NTC の述語項構造・共参照タグにおいて，述語と項の関係がゼロ照応の関係であることである．これは，照応に対する先行詞の同定の精度を評価するためである．2 つ目は，先行詞が文章中に存在することである．これは，文章外に照応詞の照応先が存在する外界照応を，本手法では考慮していないためである．3 つ目は，「する」以外の動詞で述語項構造シソーラスに含まれている動詞であることである．これは，本手法において，「勉強する」といった「名詞＋する」からなる動詞は考慮していないためである．また，未知語も考慮していない．

作成した評価データに対して，先行詞の同定を行い，提案手法の評価を N 位正解率で行う．実験は，各スコア単独で先行詞の同定を行うもの（実験 1）と深層格推定による先行詞候補の絞り込みを行わない ANN 単体の手法（ベースライン）と先行研究をもとに作成した手法，深層格推定による先行詞候補の絞り込みを行なった ANN 単体の手法，各スコアを組み合わせて先行詞の同定を行うもの（実験 2）からなる．

## 4.2 結果と考察

表 1 では，ANN と 3 つの言語資源から作成したスコアを使用して先行詞の同定を行った結果をまとめている．1 位正解率では，格フレーム（CF）が 0.235，動詞連想概念辞書（VACD）が 0.265 で 2 割を上回る正解率となった．対して，ANN は 1 位正解率が，0.185 と 2 割を下回った．また，共起 GB は，0.090 と 1 割を下回る結果となった．これにより，各スコア単体では，先行詞を 1 つに決めることは，難しいことが確認できる．

対して 10 位正解率をみると，ANN，CF，VACD は 8 割を上回っている．また，共起 GB は 6 割を超えている．これは，各スコア単体では先行詞の候補から先行詞を上位 10 候補までに，6 割から 8 割ほどの精度で絞り込むことができていると考えられる．また，共起 GB の正解率をみると，先行詞の同定において，共起 GB 単体では効果が低いことが考えられる．

そして，3 位正解率と 5 位正解率をみると，共起 GB 以外のスコアは，5 割を上回っており，特に CF と VACD は ANN の正解率を超えている．これらの結果から先行詞の同定において，各スコア単体でも少なからず効果がみられた．そして，これらのスコアを組み合わせることにより，更なる精度向上の見込みがあると考えられる．

表 2 では，先行詞の同定において，ベースラインと先行研究の手法，提案手法の結果をまとめている．ベースラインには，深層格推定による先行詞候補の絞り込みを行わずに ANN のみを用いた手法を用いた．提案手法として，深層格推定による先行詞候補の絞り込みを行なった手法（ANN），深層格推定による先行詞候補の絞り込みと ANN と動詞連想概念辞書（VACD）を組み合わせた手法（AV），深層格推定による先行詞候補の絞り込みと ANN と格フレーム（CF），VACD，共起 GB を組み合わせた手法（ACVK）の 3 つを比較する．

ベースラインと提案手法を比較すると，どの提案手法もベースラインの正解率を上回っている．深層格推定による先行詞候補の絞り込みが，先行詞の同定における正解率の向上に効果があることが確認で

きる．

先行研究の手法と提案手法の ANN を比較すると，1 位正解率において提案手法の ANN の方が 15 事例多く正解している．3 位正解率以降は，2 つの手法の正解数にほとんど差がみられなかった．原因として，ANN の違いが挙げられる．先行研究では，動詞と名詞の単語分散表現と助詞の one-hot ベクトルを入力していることに対して，提案手法では，動詞クラスと名詞クラス，助詞の one-hot ベクトルを入力している．提案手法では，動詞と名詞について，意味や属性によりクラス分けを行なっているため，名詞が取り得る深層格の傾向が先行研究より学習されているものだと考えられる．

3 つの提案手法をそれぞれ比較すると，AV と ACVK の正解率が，ANN の正解率を上回っている．ANN 単体より，複数のスコアを組み合わせた手法の方が正解率は高くなるが AV の方が ACVK より 1 位正解率が高くなっていた．つまり，本研究で使用した 3 つのスコアのうち動詞連想概念辞書から得られるスコアが先行詞の同定において，最も効果が高いことがわかる．つまり，動詞連想概念辞書は先行詞の同定に直接的な効果が期待できる．対して，格フレームと共起 GB からそれぞれ得られるスコアは，先行詞の同定に効果があまりみられなかった．つまり，これらの情報は先行詞の同定に直接的な効果が期待できない．

本研究では先行詞の同定に使用するスコア以外の要素の影響がないようにしているため，先行詞候補はスコアの高さによってのみ決定する．そして，提案手法も 3 位正解率は，5 割を超えており 10 位正解率になると 8 割を超えていることからスコアのみでの限界がこの正解率であると考えられる．

## 5 まとめ

本研究では，照応解析における先行詞の同定において，深層格と表層格を組み合わせて利用する手法を提案した．ANN により，先行詞の候補を絞り込み，候補の名詞に対して，動詞連想概念辞書と格フレーム，共起 GB からそれぞれスコアを求めて組み合わせることにより先行詞を同定する．先行詞の同定において，スコアごとの効果の比較，組み合わせによる正解率を比較した．スコア単体でみると，格フレームと動詞連想概念辞書は，先行詞の同定において，ANN より正解率が高いことが確認できた．また，深層格推定による先行詞候補の絞り込みの効果があることが確認できた．

ANN と動詞連想概念辞書を組み合わせた手法が提案手法の中で最も高い正解率を示した．この手法は，1 位正解率が 0.345 で，3 位正解率だと 0.610，10 位正解率までみると 0.840 であった．先行詞の同定において，深層格推定と動詞連想概念辞書の有用性が確認できた．また，この手法は先行詞の候補を 10 個程度までに絞り込むことが確認できたが，1 つに絞り込むことは，ほとんどできなかった．

今後の課題として，深層格推定による先行詞候補の絞り込みの精度向上と新たな先行詞らしさのスコア計算方法が考えられる．深層格推定による先行詞候補の絞り込みによって先行詞が省かれてしまう事例があったため，構文情報などを利用してより正確な深層格推定，絞り込みを行うこと必要だと考えられる．また，現在の手法は，候補を絞り込んだ後に先行詞を 1 つに同定することがほとんどできていない．先行詞の候補を 1 つに絞り込むために，現在考慮していない文中の他の名詞との関連性や構文情報などの情報を取り入れた先行詞らしさのスコアを考えることで，先行詞の同定精度の向上を目指すことが考えられる．

## 謝辞

本研究は JSPS 科研費 JP18K12434 の助成を受けたものです．

## 参考文献

[1] 寺岡丈博, 東中竜一郎, 岡本潤, 石崎俊. 単語間連想関係を用いた換喩表現の自動検出. 人工知能学会論文誌, Vol. 28, No. 3, pp. 335–346, 2013.

[2] 笹野遼平, 黒橋禎夫. 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, Vol. 52, No. 12, pp. 3328–3337, 2011.

[3] 西念星宝, 谷津元樹, 原田実. 語意類似度を用いた指示代名詞の照応解析システム AnasysD. 知能と情報, Vol. 31, No. 5, pp. 797–807, 2019.

[4] 山城颯太, 西川仁, 徳永健伸. 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. 自然言語処理, Vol. 26, No. 2, pp. 509–536, 2019.

[5] Takumi Kawasaki and Masaomi Kimura. A Novel Japanese Anaphora Resolution Method Using Deep Cases. In **International Symposium on Computer Science and Intelligent Controls (ISCSIC)**, pp. 129–134. IEEE, 2017.

[6] 国立国語研究所. 『分類語彙表増補改訂版データベース』(ver.1.0), 2004.

[7] 林部祐太, 河原大輔, 黒橋禎夫. 格パターンの多様性に頑健な日本語格フレーム構築. 研究報告自然言語処理 (NL), Vol. 2015, No. 14, pp. 1–8, 2015.