

## 照応性判定を含む名詞句照応解析の実験と分析

飯田龍 乾健太郎 松本裕治

本稿では、我々が提案した照応解析手法 [9] が照応性判定に必要な先行文脈の情報と局所文脈の情報を効果的に併用できていることについて説明する。さらに、我々の照応解析モデルは Soon ら [21] や Ng ら [20] の従来の学習に基づく照応解析モデルの利点をすべて継承しながらも、この既存手法の欠点を克服している点について議論する。提案手法を評価するために日本語名詞句照応解析の実験を行い、従来の学習に基づく手法の性能を改善できたことを報告する。また、提案手法を用いた解析の結果を人手で分析し、今後の方向性を論じる。

### Noun Anaphora Resolution Combining Anaphoricity Determination and Antecedent Identification: Experiments and Analysis

Ryu Iida Kentaro Inui Yuji Matsumoto

We discuss how to combine the anaphoricity determination process with the antecedent identification process in anaphora resolution. In doing so, we provide existing models, such as Soon et al. [21] and Ng and Cardie [20], and present a new model which effectively incorporates clues obtained from preceding contextual information together with those from local information of a given target noun phrase. We conducted experiments on resolving Japanese anaphora with noun phrases. The results show that the proposed model outperforms earlier learning-based approaches. We manually analyze major error sources, and discuss remaining problems and future directions.

#### 1 はじめに

文章中の同一指示対象を同定する照応解析は、機械翻訳や情報抽出、質問応答など、さまざまな言語処理アプリケーションで必須の処理である。これまでの照応解析の手法はおおきく理論指向の規則作成に基づく手法 [1, 3, 12, 16, 17, 23] とタグ付きコーパスを利用した学習手法 [4, 7, 15, 14, 19, 21, 22, 25] に分類できる。機械学習を利用した手法は主に表層的な手がかりを用いた単純なものが多いが、それでも規則作成に基づく手法と同程度かもしくはそれ以上の成果をあげており、照応解析に必要な手がかりをうまく導入することでさらなる精度向上を目指すことができると考えられる。

照応解析の処理は、おおきく照応性判定と先行詞同定の二つの処理に分解できる。照応性判定は、文章中の名詞句が先行文脈に対となる先行詞を持つ照応詞か、もしくはそれ以外(非照応詞)であるかを分類するタスクである。また、先行詞同定は、照応性判定で検出した照応詞に対して先行詞を同定する処理である。

初期の照応解析に関する研究 [1, 6, 12] は統語的な特徴から照応詞と判断できる代名詞や指示詞のみ

奈良先端科学技術大学院大学 情報科学研究科  
Graduate School of Information Science, Nara Institute of Science and Technology

を研究の対象としている。つまり、文章中のどの要素が照応詞となるかはあらかじめ与えられた上で先行詞同定処理の精度の向上を目指すことが照応解析の主要な目的のように考えられてきた。

しかし、以下の二つの理由により、照応性判定処理も視野に入れた照応解析処理にも研究者の関心が集まってきている [2, 18, 20, 24]。

- 照応性判定の手がかりとして定冠詞を利用できる英語などの言語においても、照応性を判定することはそれほど簡単な問題ではない。
- 照応解析全体の精度は照応性判定の結果に依存する。

もちろん、日本語のように定冠詞を手がかりとして利用できない言語の場合は、照応性判定はさらに重要な問題となる。

これまでの照応性判定に関する研究から以下の二つの点が明らかになってきた。

- 照応性判定の重要な手がかりの一つは先行詞候補を探索することによって得ることができる。なぜなら、適当な先行詞候補が先行文脈に存在することが照応詞候補が照応詞と判断される必須条件となるためである。
- 照応性を効果的に判定するためには、照応詞の振舞いだけでなく非照応詞の振舞いについても学習する必要がある。

	SM	CM	TM
先行文脈情報が利用可能	✓		✓
非照応詞を訓練事例に利用可能		✓	
照応性を判定する	✓	✓	
訓練事例の正負例に偏りが無い			✓

SM:探索型モデル, CM:分類型モデル, TM:トーナメントモデル

表 1: 先行研究の各手法の利点

しかし, 2 節で述べるように, これらの二つの情報を効果的に併用する解析モデルはこれまでのところ報告されていない. そこで, 本稿では, 文献 [9] で提案した照応解析モデルがこれら二種類の手がかりを併用しながらも, 既存手法の問題点を克服していることについて説明し, このモデルと既存のモデルを詳細に比較した結果について報告する.

2 節では機械学習に基づく照応解析手法の先行研究をまとめ, 3 節では既存手法で利用されている照応解析に有益な情報をどのように併用したかについて説明する. 4 節で日本語の名詞句照応解析の評価実験を行った結果について報告する. この実験では, 機械学習に基づく照応解析手法の先行研究である Soon らの探索型の手法 [21] と Ng らの分類型の手法 [18] の 2 種の手法と比較することで提案手法の性能を評価する. 5 節で提案手法を利用して得られた解析結果のうち, 誤り事例について分析した結果を報告する. 最後に 6 節でまとめる.

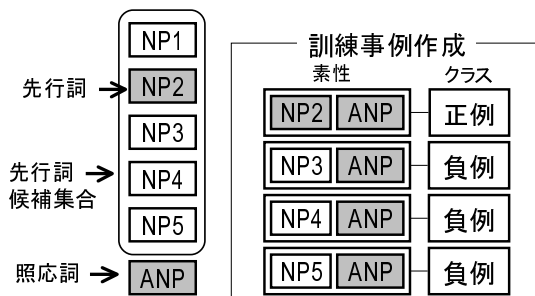
## 2 先行研究

従来の機械学習に基づく照応解析手法はおおきく探索型手法と分類型手法に分類できる. この節では表 1 に示すような観点で先行研究の利点と欠点をまとめる.

### 2.1 探索型モデル

探索型手法では, 文章中の任意の名詞句 (照応詞候補) に対して先行文脈に先行詞となる候補が存在するか否かを探索することにより間接的に照応性判定を行う. つまり, もし先行詞となる適切な候補がみつければ, 照応詞候補は照応詞と判定され, それ以外の場合には非照応詞と判定される. 探索型の代表的な手法である Soon ら [21] の手法では, 照応詞候補に対して先行文脈中の各名詞句をそれぞれ先行詞候補とみなし, 照応詞候補に近い先行詞候補から順に先行詞候補と照応詞候補の対が照応関係にあるか否かの分類問題を解く (図 1 を参照). この手法では, 先行文脈に適切な先行詞候補を持つか否かを調査するため, 照応詞候補の照応性を判定する際に先行文脈情報を利用できるという利点を持つ.

一方, 探索型手法は, 非照応詞の振舞いを学習するように設計されていないという問題があり, 解析する際に非照応詞を適切に棄却できるとは限らない. 例えば, 図 1 に示した Soon らのモデル [21] では,



図の左側は照応詞 ANP に対して 5 つの先行詞候補 ( $NP_1, \dots, NP_5$ ) が出現している状況を表している.  $NP_2$  を ANP の先行詞とする. この状況において, Soon らのモデルでは照応詞と (照応詞から最も近い) 先行詞の対  $ANP-NP_2$  を正例, 照応詞と他の (先行詞と照応詞の間の) 各名詞句の対  $ANP-NP_3, ANP-NP_4, ANP-NP_5$  を負例として学習する. 解析の際には, 照応詞候補から先行文脈に向かって, 先行詞候補となる名詞句の各々について, その対が照応関係なるか否かを分類していく. Soon らのモデルでは照応関係にあると分類された照応詞候補を照応詞に決定し, 同定された名詞句を先行詞として出力する. もしどの対も照応関係にならないと分類された場合には照応詞候補は非照応詞として出力される. このモデルの詳細な説明については文献 [21] を参考にされたい.

図 1: Soon らの探索型モデル

訓練事例を作成する際に照応詞 ANP に対して最も近い先行詞との対  $NP_2-ANP$  を正例に, また照応詞と先行詞の間の名詞句それぞれと照応詞の対 ( $NP_3-ANP, NP_4-ANP, NP_5-ANP$ ) を負例とする. そのため, 非照応詞の事例は訓練時に参照されることがない. この事例作成方法は Ng ら [19] や Yang ら [25] の手法でも採用されているため同様に問題となる.

さらに, 探索型手法では, 上述の訓練事例の作成方法のために, 一つの正例に対して多数の負例を作成することになる. このような正負例の偏りも精度向上のためのおおきな障害となる.

### 2.2 分類先行探索型モデル

探索型手法に対し, 分類型手法 [18, 20] では, 照応性判定の問題を先行詞同定の処理と切り離して考える. この手法では, 照応詞の情報に加え, 探索型手法で利用できていない非照応詞の情報を用いて照応性判定の分類器を作成できるという利点がある. ここでは分類型手法の一例として Ng [18] の手法を図 2 を用いて説明しよう. Ng の手法では以下の手順で照応解析を行う.

- (1) 対象とする名詞句 ( $TNP$ ) の照応性を判定し, 照応性判定のスコア (確信度) が閾値  $\theta_{ana}$  を下回る場合は非照応詞に分類する.
- (2) 非照応詞に分類されなかった名詞句に対し, 探索型モデルを利用して先行詞を同定する.
- (3) (2) で同定した先行詞のうち最も先行詞らしいと判定された名詞句 ( $NP_2$ ) の持つ先行詞らしさのスコアが  $\theta_{ant}$  以上の場合は  $TNP-NP_2$  の対を照応関係として出力する.  $\theta_{ant}$  より低い場合は  $TNP$  を非照応詞に分類する.

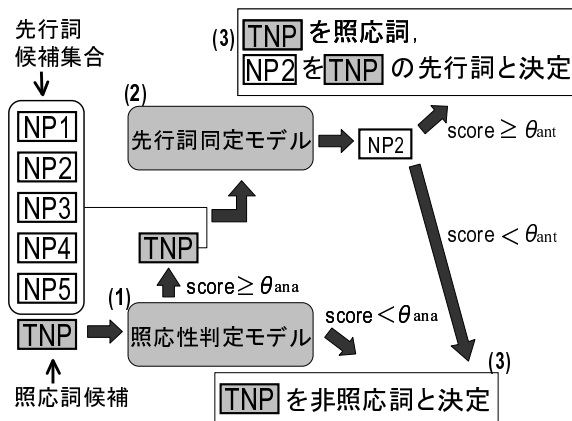


図 2: Ng[18] の分類先行探索型モデル

この手法では最初に照応性を判定し、次に先行詞を探索型モデルを利用して解析するため、以後このモデルを分類先行探索型モデルと呼ぶ。

Ngら [18] の評価実験によると、分類先行探索型モデルの二つの閾値をうまく推定してやることで、探索型モデルと比べて照応解析全体の精度が向上したと報告されている。

分類先行探索型モデルでは、非照応詞を学習に利用することで探索型手法を改善できているように見える。しかし、(1) の時には先行詞の情報を参照できないので、探索型手法で利用できていた先行文脈の情報が利用できないという欠点が存在する。この問題に対して、Ngら [20] は照応詞候補と語彙のもしくは意味的な観点で類似する先行詞候補が先行文脈に存在するか否かという情報を表 2 に示すような素性として照応性判定に導入している。ただし、明示的に先行詞を同定してるわけではないので、表 2 の素性が利用できない場合は先行文脈の情報を参照できない。さらに先行詞同定に探索型モデルを利用するため、訓練事例の正負例に偏りが生じるという問題は解決されないまま残る。

### 2.3 トーナメントモデル

先行詞同定の処理では、我々が提案したトーナメントモデル [7] が利用可能である。このモデルでは、照応詞に対して先行詞候補となる名詞句の間で勝ち抜き戦を行うことで先行詞を同定する。照応詞 ANP に対して先行文脈に 5 つの先行詞候補 ( $NP_1, \dots, NP_5$ ) が出現している状況を描いた図 3 を使ってトーナメントモデルを用いた先行詞同定の処理を説明しよう。このモデルでは照応詞から文章の先頭に向かって勝ち抜き戦を行い、最初の比較では、図 3(1) に示すように、最も照応詞に近い 2 つの候補  $NP_5$  と  $NP_4$  を比較し、分類器はより先行詞らしい名詞句を選択する。以降の比較では、1 つ前の比較において勝ち残った(より先行詞らしいと判

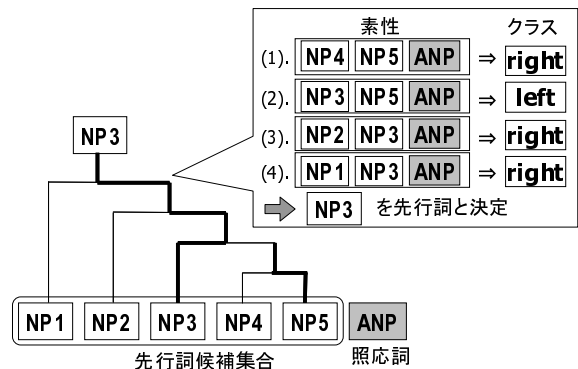


図 3: トーナメントモデル

定された) 候補と新たな先行詞候補との比較を行う。例えば、 $NP_5$  と  $NP_4$  の比較で  $NP_5$  が勝ったとすると、次は  $NP_5$  と新たな候補  $NP_3$  を比較する(図 3(2))。この処理を繰り返す、最後の比較では、文章の先頭に最も近い先行詞候補との比較を行い、勝ち残った候補を与えられた照応詞に対する最尤先行詞候補と決定する。

このモデルは 2.1 の探索型モデルと比較していくつかの利点がある。まず、先行詞候補間で先行詞らしさを比較するため、センタリング理論 [5] で導入されている先行詞になりやすい談話要素 (center) の遷移を明示的に学習できるという利点がある。また、Soonらや Ngら の探索型モデルとは異なり、二つの先行詞候補のうちどちらがより先行詞らしいかを学習するように設計されているため、探索型モデルで問題となる正例と負例の偏りを自然に解消できる。

ただし、トーナメントモデルは、対象とする一つの照応詞に対して必ず一つの先行詞を同定するため、照応性判定の能力を持たない。そのため、そのままでは名詞句照応解析に適用できないという問題が残っていた。

### 3 探索先行分類型モデル

ここで、分類先行探索型モデルとは異った方法で先行文脈の情報を照応性判定に利用するモデルを考える。この手法では先行詞同定のモデルと照応性判定のモデルを仮定し、各照応詞候補に対して次の 2 段階の処理で照応性を判定する。

1. 照応詞候補  $NP_i$  に対して先行詞同定モデルを用いて先行詞候補集合から最も先行詞らしい候補 (最尤先行詞候補)  $AC$  を同定する。
2.  $AC-NP_i$  の対が照応関係にあるか否かを分類する。もし  $AC-NP_i$  が照応関係にあると分類された場合は  $NP_i$  は照応詞と決定される。そうでなければ、 $NP_i$  は非照応詞と判断される。

この 2 段階の処理は分類先行探索型モデルと対照的であるため、提案モデルを探索先行分類型モデルと

素性の種類	素性名	詳細
Lexical	STR_MATCH	それぞれの限定詞を取り除いた場合、照応詞候補 $NP_j$ と同じ文字列となる先行詞候補 $NP_i$ が $NP_j$ の先行文脈に存在するなら Y. それ以外は N.
	HEAD_MATCH	照応詞候補 $NP_j$ と同じ主辞を持つ先行詞候補 $NP_i$ が $NP_j$ の先行文脈に存在するなら Y. それ以外は N.
Semantic	ALIAS	照応詞候補 $NP_j$ が縮約形となるような先行詞候補 $NP_i$ が $NP_j$ の先行文脈に存在する (もしくは $NP_i$ が $NP_j$ の縮約形となる) なら Y. それ以外は N.
	SUBCLASS	照応詞候補 $NP_j$ が WordNet の上位下位関係となるような先行詞候補 $NP_i$ が $NP_j$ の先行文脈に存在するなら Y. それ以外は N.

表 2:  $Ng$  らが分類型モデルで利用した先行文脈の素性.  $NP_i$  が先行する名詞句を表し,  $NP_j$  は照応詞を表す. 素性は対象となっている  $NP_j$  に対してその性質を満たすか (Yes) 満たさないか (No) の 2 値をとる.

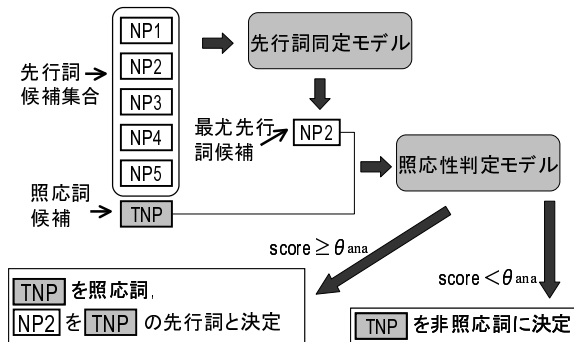


図 4: 探索先行分類型モデルによる照応解析処理呼ぶ.

提案手法では, 照応性判定のために分類型手法のモデルを拡張し, 照応詞候補と最尤先行詞候補の対を用いて照応性の分類問題を解く. 図 4 を使って照応解析処理の全体像を説明しよう. 図 4 では対象となる名詞句  $TNP$  に対して先行文脈に 5 つの先行詞候補 ( $NP_1, \dots, NP_5$ ) が出現している状況を仮定している. この状況で解析モデルは, まず  $TNP$  に対して先行詞候補集合から先行詞同定モデルを用い最尤先行詞候補を選択する. ここでは仮に  $NP_2$  が最尤先行詞候補として選ばれたとすると, 次に解析モデルは  $NP_2$ - $TNP$  が照応関係にあるか否かを分類する. もしモデルが対を照応関係にあると分類した場合には  $TNP$  を照応詞, 最尤先行詞候補  $NP_2$  を先行詞として出力する. そうでない場合は  $TNP$  を非照応詞と判断する.

探索先行分類型モデルで使用する照応性判定モデルを作成するために, 照応詞を正例, 非照応詞を負例とし, 以下の方法で訓練事例を作成する. 正例については, 訓練用コーパスに出現する各照応詞に対して先行詞との対を正例集合に加える. 図 5 の上部では, 照応詞である  $ANP$  とその先行詞である  $NP_4$  の対を正例集合に加えている. 一方, 負例については, 非照応詞に対して先行文脈の先行詞集合から先行詞同定モデルを用いて最尤先行詞候補を決定する. この最尤先行詞候補と非照応詞の対を負例集合に追加する. 図 5 の下部の例では, 非照応詞である  $NANP$  に対し先行文脈に出現している先行詞候補 ( $NP_1, \dots, NP_5$ ) の中から最尤先行詞  $NP_3$  を

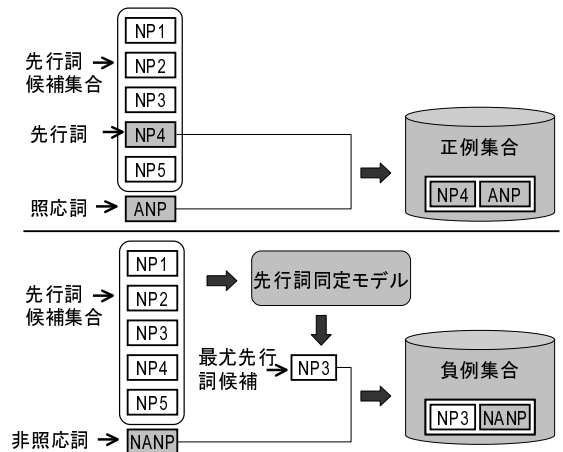


図 5: 照応性判定モデルのための訓練事例作成

決定し,  $NP_3$ - $NANP$  を負例集合に追加する. この手続きにより, 非照応詞に対し先行文脈の情報を明示的に加えて最終的な照応性判定の分類器を作成することができる.

提案モデルは一見すると分類先行探索型モデルと大きな違いはないように見えるかもしれない. しかし, 探索先行分類型モデルは以下に示す 3 つの点において表 1 に示した既存手法の利点を全て継承しており, 分類先行探索型モデルの欠点を克服している.

第 1 に, 提案モデルは探索型モデルの特徴をうまく継承しており, 照応性を判定する際に最尤先行詞候補を利用できるという利点がある. 先行詞同定の処理で選ばれた最尤先行詞候補は照応性判定に有益な情報を提供することが期待できる. もし最尤先行詞候補が真の先行詞でない場合は, 照応詞候補は前方文脈中に先行詞を持たないかもしれない. この点で, 分類先行探索型モデルが単純に文字列一致情報を素性として利用しているのに比べ, 提案モデルはより効果的に先行文脈の手がかりを利用できているといえる.

第 2 に, 分類先行探索型モデルから継承した利点として, 提案モデルでは照応詞の事例に加えて非照応詞の事例を訓練事例として利用できる. これにより, 非照応詞のふるまいを明示的に学習することが可能になる.

第3に、分類先行探索型モデルでは2段階の処理とともに照応性判定の処理が必要であるのに対し、提案モデルでは2段階の各処理が照応性判定と先行詞同定のそれぞれの処理に厳密に対応する。つまり、先行詞同定の処理では与えられた照応詞候補に対して先行詞を同定するだけでよい。このような厳密な対応付けにより、提案モデルではトーナメントモデルを先行詞同定に利用できるという利点が生まれる。これに対して、分類先行探索型モデルでは最初の処理で完全に照応性を判定することが困難であり、後の処理でも先行詞を同定するとともに照応性を判定する必要がある。このため、分類先行探索型モデルではトーナメントモデルが利用できない。4.4で述べるように、この違いが提案モデルと分類先行探索型モデルの精度におおきく影響する。

#### 4 評価実験

提案する探索先行分類型モデルの有効性を調査するために、日本語名詞句照応解析の評価実験を行い、先行研究の探索型モデルや分類先行探索型モデルと比較を行った。

##### 4.1 比較する3つのモデルの実装

3つのモデルを比較するために先行研究のモデルを実装する必要がある。まず、探索型モデルについてはNgら[19]の提案するモデルを実装した。このモデルでは、図1で示したSoonら[21]のモデルとは若干異なり、解析の際に、照応詞候補と先行文脈中の各先行詞候補との対がそれぞれ照応関係となるか否かを分類し、最も先行詞らしいと判定された候補を先行詞として決定する。ただし、どの対も確信度があらかじめ与えられた閾値 $\theta_{ana}$ を下回るときには照応詞候補を非照応詞に分類する。

また、分類先行探索型モデルには2.2で示したNgら[18]のモデルを実装した。ただし、今回の実験では二つの閾値は訓練事例から推定せずに入手で設定することで、このモデルの精度の上限値を調査する。

提案する探索先行分類型モデルについては、トーナメントモデルを導入することでどの程度精度に影響が出るかを調査するために、先行詞同定の処理に探索型モデルを利用したモデルとトーナメントモデルを利用したモデルの二種類のモデルを実装した。

##### 4.2 素性

学習には、以下の4種の素性を導入した<sup>1</sup>。

- *TNP*: 照応詞候補に関する語彙、統語、意味(名詞の意味属性)、位置情報に関する素性。
- *ANT*: (i) 先行詞候補に関する語彙、統語、意味(名詞の意味属性)、位置情報、(ii) 照応詞候補

<sup>1</sup>素性の詳細については文献[8]を参考にされたい。

	SM	CSM		SCM	
		照応性判定	先行詞同定	先行詞同定	照応性判定
<i>TNP</i>	✓	✓	✓	✓	✓
<i>ANT</i>	✓		✓	✓	✓
<i>ANT-SET</i>		✓			
<i>ANT-ANT</i>				✓	

表3: 各モデルで利用可能な素性。SM: 探索型モデル、CSM: 分類先行探索型モデル、SCM: 探索先行分類型モデル。

と先行詞候補の関係から抽出可能な情報(例えば、意味的な整合性や照応詞候補と先行詞候補の間の距離など)に関する素性。

- *ANT-SET*: 照応詞候補と先行詞候補集合の関係から抽出可能な情報(例えば、照応詞候補と文字列一致する先行詞候補が先行文脈に存在するか否か)に関する素性。
- *ANT-ANT*: 先行詞候補間の情報(例えば、二つの候補間の距離)に関する素性。

分類先行探索型モデルでは照応詞候補と先行詞候補集合から*ANT-SET*素性を抽出可能であるが、あらかじめ先行詞同定を行っていないため、照応性判定の処理では*ANT*素性を利用することができない。また、探索型モデルや分類先行探索型モデルでは*ANT*素性は利用可能だが、先行詞同定の処理では照応詞候補と先行詞候補の対に関して照応関係にあるか否かを分類するため、*ANT-ANT*素性を利用できない。これに対し、提案手法のモデルではトーナメントモデルを導入可能なので、*ANT-ANT*素性を利用できる。

実験では、茶釜[13]とCaboCha[11]を用い形態素解析、固有表現タグ付与、係り受け解析を行い、すべての素性は自動的に抽出した。

##### 4.3 評価事例

評価実験のために日本語新聞記事90記事に照応関係のタグを付与し、照応関係タグ付きコーパスを作成した<sup>2</sup>。このコーパスは884の照応詞と6,591の非照応詞、あわせて7,475の名詞句を含み、照応詞については先行詞がどの名詞句に相当するかのタグも付与されている。実験では、このコーパスを記事単位で分割し、10分割交差検定を行った。

##### 4.4 実験結果

今回の実験の目的は3つの解析モデルの照応解析の精度を比較することである。そこで照応関係を正しく同定できた場合を正解とし、再現率、精度を以下の式を使って求める。

$$\text{再現率} = \frac{\text{照応詞関係を正しく同定できた数}}{\text{照応詞の総数}}$$

<sup>2</sup>[http://cl.naist.jp/~ryu-i/coreference\\_tag.html](http://cl.naist.jp/~ryu-i/coreference_tag.html) を参照。

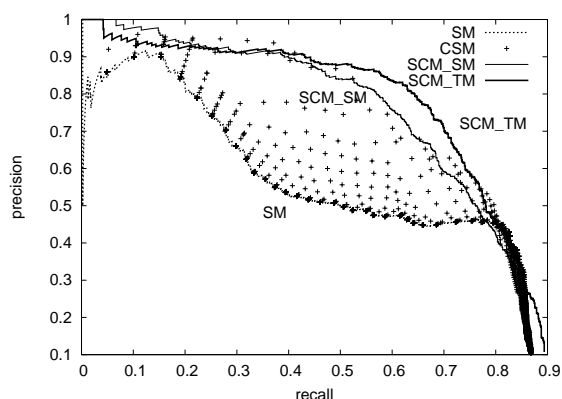


図 6: 名詞句照応解析の再現率-精度曲線

$$\text{精度} = \frac{\text{照応詞関係を正しく同定できた数}}{\text{システムが検出した照応詞の総数}}$$

解析のスコア（確信度）には入力の素性ベクトルと分類器が出力する分離平面との距離を利用した。探索型モデルや提案モデルでは各事例のスコアが  $\theta_{ana}$  以上の場合に照応詞候補を照応詞に分類する。また、分類先行探索型モデルでは 2 つの閾値  $\theta_{ant}$  と  $\theta_{ana}$  を設定してやる必要がある。今回の実験ではこれらの閾値を訓練事例から求めるのではなく、人手で動かしてやることで、それぞれのモデルの精度の上限値を見る。閾値を動かして再現率-精度曲線を描いた結果を図 6 に示す。

まず、先行詞同定に探索型モデルを利用した探索先行分類型モデル (SCM\_SM) と分類先行探索型モデル (CSM) を比較することで、明示的に先行詞を同定して照応性判定を行うことがどの程度精度向上に貢献するかを調査する。図 6 の二つのモデルの結果を見ると、各再現率での SCM\_SM の精度は CSM の上限値に至っており、このことから最尤先行詞候補がより良い先行文脈情報の近似として利用できていることがわかる。これに加え、CSM ではあらかじめ 2 つ閾値を設定してやる必要があるのに対し、SCM\_SM では 1 つの閾値のみを設定してやるだけでよい。このように解析モデルの設計についても提案モデルが優れていることがわかる。

次に、先行詞同定にトーナメントモデルを利用した探索先行分類型モデル (SCM\_TM) と SCM\_SM を比較する。まず、先行詞同定処理単独で見ると、トーナメントモデルの方が探索型モデルより高い精度で先行詞を同定している (表 4)。照応性判定まで含めた全体の処理の精度についても、SCM\_TM が SCM\_SM の精度を上回っている。この結果より、提案モデルでは、精度良く先行詞を同定できるトーナメントモデルの性能を生かしたモデル化ができていくことがわかる。

	探索型モデル	トーナメントモデル
精度	86.9% (768/884)	89.4% (790/884)

表 4: 先行詞同定の実験結果

誤りの原因	頻度
(a) 高品質な意味属性が必要	26
(b) 一つの名詞句が複数の実体を指す	13
(c) 先行詞ではない名詞句と照応詞の文字列が (一部、完全) 一致する	11
(d) 前方文脈の照応関係の情報が必要	10
(e) 代名詞の意味属性が学習できていない	6
(f) タグ付与誤り	6
(g) 問題設定の修正 (GPE)	3
(h) その他	27

表 5: 先行詞同定の誤り分析

## 5 誤り分析

提案する探索先行分類型モデルの解析結果のうち、解析を誤った事例を (i) 先行詞同定と (ii) 照応性判定の 2 つの観点で分析した。

### 5.1 先行詞同定の誤り分析

表 4 に示した先行詞同定の結果のうち、解析を誤った 94 事例をいくつかの誤りの原因に人手で分類した。分類結果を表 5 に示す<sup>3</sup>。表 5 のうち、最も多い誤りの原因は意味属性に関するもので、照応詞と同定した先行詞候補の間に意味的な不整合があると人手で判断したものである (a)。この誤りでは、例えば照応詞「読者」に対して「日本」が先行詞候補として同定されるなど、二つの名詞句の間に意味的な整合性がない場合が多い。そのため、日本語語彙大系 [10] のような既存の言語資源の情報を制約として加えることで照応詞候補と整合性のない先行詞候補をあらかじめ棄却することも考えられる。

また別の頻出した問題として複数実体を指す名詞句に関する誤りがある。名詞句の中には「両首脳」や「二人」のように一つの名詞句で複数の実体を指すことができる表現があり、このような名詞句が複数実体を指している場合には、先行詞を一つだけ出力する既存のモデルでは対応できない (b)。この問題に対応するには、まず複数実体を指す可能性のある表現をあらかじめ保持しておき、その表現にあてはまる場合には複数個の先行詞を同定することが考えられる。この処理を行うことで、どの程度問題に対処可能かについては今後調査を進めたい。

また、今回の解析モデルでは解析誤りの伝播を防ぐために、前方文脈の解析結果を利用していないが、その情報を利用しなければ解析が不可能な場合がある (d)。ここでいう前方文脈の解析結果はおおきく二種類に分類できる。一つはこれまでに述べてきた照応解析の結果が必要となる場合であり、例え

<sup>3</sup>一つの事例を複数のカテゴリに分類しているため、合計すると調査した事例の総数 94 事例より多くなる。

ば、先行詞候補が「それ」のような代名詞の場合には、代名詞の先行詞の情報を導入することによって意味の不整合についての素性が有効にはたらく可能性がある。もう一つが先行詞候補と間接照応の関係にある名詞句の情報が必要となる場合であり、例えば、照応詞候補と先行詞候補がそれぞれ「前年」という名詞句である場合にはそれぞれの名詞句が何年の前年なのかという間接照応的な情報が必要となる。名詞句照応解析結果の効果的な導入方法や間接照応の解析についても今後の課題となる。

さらに、少数ではあるが、どのような実体を指しているかを厳密に区別して照応関係のタグを付与していることも解析の際の問題となっている(表 5(g))。例えば、次の文章では複数回出現する「ロシア側」という表現がそれぞれ「ロシア政府」と「ロシア軍」を指しているが、この違いを厳密に区別してタグ付与したとしても、現状ではそれら二つの「ロシア側」の違いを特徴付けることはできず、「ロシア側<sub>1</sub>」の先行詞として「ロシア側<sub>2</sub>」を同定してしまう。

ロシア南部チェチェン共和国の首都グロズヌイに進攻したロシア軍<sub>1</sub>は二日、ドゥダエフ政権部隊の激しい抵抗で大量の被害を出し、首都攻略の第一次作戦は事実上失敗した模様だ。ロシア側<sub>2</sub>は支援部隊の投入を発表し、三日も首都空爆を行った。一方、ドゥダエフ大統領は二日夜、多数のロシア兵を捕虜にしたと発表。これを盾にロシア政府<sub>2</sub>に改めて即時停戦を求めた。グロズヌイからの報道によると三日、大統領官邸の北西一・五キロの鉄道駅付近でロシア軍部隊とチェチェン側部隊が衝突したが、ロシア側<sub>1</sub>は中心部への進撃を阻まれて苦戦。

このような二つの実体の差分をどのくらい考慮すべきかは、照応解析の解析結果を利用する応用分野に依存するが、現状の応用分野を見るかぎりこのような区別を必要とする分野は多くはない。逆に、Automatic Content Extraction (ACE)<sup>4</sup>の仕様で採用されているように、政府や国民を区別せず、それらを包含するような地政学的実体 (Geo-Political Entity; GPE) とみなすことによって作業員間のタグ付与の揺れを抑えることができ、対象とする問題を単純化できるかもしれない。この GPE の採択も含め照応関係タグの仕様についても再度検討したい。

## 5.2 照応性判定の誤り分析

次に照応性判定を誤った事例について調査した結果をまとめる。非照応詞を照応詞と誤った事例のうち、4.4で導入した確信度が高い100事例を分析した結果を表6に示す<sup>5</sup>。また、照応詞に対して適切に先行詞を同定できたにもかかわらず、非照応詞に

誤りの原因	頻度
(a) 非照応詞と先行詞候補が文字列一致	46
(b) タグ付与誤り	23
(c) 非照応詞と先行詞候補が別実体を指す	19
(d) 対象とする名詞句が総称名詞	15
(e) 高品質な意味属性が必要	14
(f) その他	27

表 6: 非照応詞の照応性判定の誤り

誤りの原因	頻度
(a) 照応詞と先行詞候補が文字列一致	86
(b) 対象名詞句単体の情報では実体と対応付けられない	23
(c) 前方文脈の照応関係の情報が必要	20
(d) タグ付与誤り	5
(e) その他	17

表 7: 照応詞の照応性判定の誤り

分類された事例についても確信度が高い100事例を分析した(表7)。

まず、表6と表7の(a)を比較するとわかるように、照応詞に関する多くの誤り事例では、対象となる照応詞と同じ文字列を持つ先行詞候補を最尤先行詞候補として同定しており、非照応詞についても同様の傾向にある。この原因を調査したところ、照応性判定モデルが文字列一致に関する素性に過剰に重みを付与して分類していることがわかった。このような不適切な重み付けが起こる原因として、今回の実験では、固有表現と普通名詞、代名詞をすべてをまとめて学習していることに問題があると考えられる。固有表現では文字列一致の情報が有効にはたらくため、その素性に大きな重みが付くが、これを普通名詞や代名詞についての分類に利用すると誤った結果を導くことになる。名詞句の種類毎に分けて学習させることも考えられるが、代名詞などは出現頻度が少ないため、学習事例をいかに効率的に作成するかについても今後の課題となる。

また、(a)と関連して、二つの名詞句が別の実体を指している場合であっても、同じ文字列を含んでいるために照応関係として同定される事例も多い(表6(c))。下の例では、「ロシア兵<sub>1</sub>」はすでに死亡したロシア兵を指しているが、「ロシア兵<sub>2</sub>」は捕虜となっているロシア兵を指しており、それぞれが指している実体は異なる。このような複数実体をどのような情報を用い区別するかも課題として残る。

... またロシア政府は、チェチェン側が塩素ガス、青酸ガスなど化学兵器を使用したためロシア兵<sub>1</sub>八人と一般市民が死亡したと述べた。チェチェン外務省はこれを否定している。ドゥダエフ大統領は二日夜、声明を発表し、ロシア側がさらに攻撃を続ければ、捕虜となったロシア兵<sub>2</sub>を「いつでも処刑できる」と言明した。

<sup>4</sup><http://www ldc.upenn.edu/Projects/ACE/>

<sup>5</sup>表5と同様に表6や表7でも一つの事例を複数のカテゴリに分類しているため、合計すると100事例を越える。

## 6 おわりに

本稿では、探索型モデルで利用されている先行文脈の情報と分類型モデルで利用されている非照応詞の情報を我々が提案した探索先行分類型モデルが効果的に併用していることについて説明し、名詞句照応解析の評価実験を通じて提案モデルの有効性を示した。さらに解析を誤った事例を先行詞同定と照応性判定の二つの観点から分析し、今後の方向性を示した。今後は、5節に示した分析結果のうち、特に前方文脈の照応解析の結果をどのように導入するかについて検討する。さらに名詞句照応だけではなくゼロ照応の問題も同時に考えた場合に、二つの照応の問題をどのような順序で解析し、その解析結果をどのように利用するのが最適かを考える文章全体の解釈の最適化問題に取り組みたい。

## 参考文献

- [1] Baldwin, B.: *CogNIAC: A Discourse Processing Engine*, PhD Thesis, Department of Computer and Information Sciences, University of Pennsylvania (1995).
- [2] Bean, D. L. and Riloff, E.: Corpus-based Identification of Non-Anaphoric Noun Phrases, *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 373–380 (1999).
- [3] Brennan, S. E., Friedman, M. W. and Pollard, C.: A Centering Approach to Pronouns, *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 155–162 (1987).
- [4] Ge, N., Hale, J. and Charniak, E.: A Statistical Approach to Anaphora Resolution, *Proceedings of the 6th Workshop on Very Large Corpora*, pp. 161–170 (1998).
- [5] Grosz, B. J., Joshi, A. K. and Weinstein, S.: Centering: A framework for modeling the local coherence of discourse, *Computational Linguistics*, Vol. 21, No. 2, pp. 203–226 (1995).
- [6] Hobbs, J.: Resolving Pronoun References, *Lingua*, Vol. 44, pp. 311–338 (1978).
- [7] 飯田龍, 乾健太郎, 松本裕治: 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*, Vol. 45, No. 3, pp. 906–918 (2004).
- [8] 飯田龍, 乾健太郎, 松本裕治, 関根聡: 最尤先行詞候補を用いた日本語名詞句同一指示解析, *情報処理学会論文誌*, Vol. 46, No. 3, pp. 831–844 (2005).
- [9] 飯田龍, 乾健太郎, 松本裕治: 先行文脈と局所文脈を併用した照応性判定モデルの学習, *言語処理学会 第11回年次大会 発表論文集*, pp. 1048–1051 (2005).
- [10] 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).
- [11] 工藤拓, 松本裕治: Support Vector Machine を用いた Chunk 同定, *自然言語処理*, Vol. 9, No. 5, pp. 3–21 (2002).
- [12] Lappin, S. and Leass, H. J.: An Algorithm for Pronominal Anaphora Resolution, *Computational Linguistics*, Vol. 20, No. 4, pp. 535–561 (1994).
- [13] 松本裕治, 北内啓, 平野善隆, 松田寛, 高岡一馬, 浅原正幸: 形態素解析システム『茶釜』version 2.3.3 使用説明書, 奈良先端科学技術大学院大学 (2003).
- [14] McCallum, A. and Wellner, B.: Object Consolidation by Graph Partitioning with a Conditionally Trained Distance Metric, *Proceedings of the KDD-2003 Workshop on Data Cleaning, Record Linkage, and Object Consolidation*, pp. 19–24 (2003).
- [15] McCarthy, J. F. and Lehnert, W. G.: Using Decision Trees for Coreference Resolution, *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1050–1055 (1995).
- [16] Mitkov, R.: Factors in anaphora resolution: they are not the only things that matter. A case study based on two different approaches, *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics (ACL) and the 8th Conference of the European Chapter of the Association for Computational Linguistics (EACL) Workshop on Operational Factors in Practical, Robust Anaphora Resolution* (1997).
- [17] 中岩浩巳, 池原悟: 語用論的・意味論的制約を用いた日本語ゼロ代名詞の文内照応解析, *自然言語処理*, Vol. 3, No. 4, pp. 49–65 (1996).
- [18] Ng, V.: Learning Noun Phrase Anaphoricity to Improve Coreference Resolution: Issues in Representation and Optimization, *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 152–159 (2004).
- [19] Ng, V. and Cardie, C.: Improving Machine Learning Approaches to Coreference Resolution, *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 104–111 (2002a).
- [20] Ng, V. and Cardie, C.: Identifying Anaphoric and Non-Anaphoric Noun Phrases to Improve Coreference Resolution, *Proceedings of the 19th International Conference on Computational Linguistics (COLING)*, pp. 730–736 (2002b).
- [21] Soon, W. M., Ng, H. T. and Lim, D. C. Y.: A Machine Learning Approach to Coreference Resolution of Noun Phrases, *Computational Linguistics*, Vol. 27, No. 4, pp. 521–544 (2001).
- [22] Strube, M. and Müller, C.: A Machine Learning Approach to Pronoun Resolution in Spoken Dialogue, *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 168–175 (2003).
- [23] 田村浩二, 奥村学: センター理論による日本語談話の省略解析, *情報処理学会研究会報告 (自然言語処理研究会)*, Vol. 107, No. 12, pp. 91–96 (1995).
- [24] Uryupina, O.: High-precision Identification of Discourse New and Unique Noun Phrases, *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL) Student Research Workshop*, pp. 80–86 (2003).
- [25] Yang, X., Zhou, G., Su, J. and Tan, C. L.: Coreference Resolution Using Competition Learning Approach, *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 176–183 (2003).