

# 固有表現認識への応用を目指した 言語学的制約に基づく述語使用規則獲得手法の提案

寺本優香<sup>1</sup> Yuka Teramoto波多野賢治<sup>2</sup> Kenji Hatano<sup>1</sup> 同志社大学大学院文化情報学研究所<sup>2</sup> 同志社大学文化情報学部

## 1 はじめに

固有表現認識 (Name Entity Recognition) とは、文書内から固有表現 (Named Entity) と呼ばれる特定の定義に基づいた表現を抽出し、それらが属するカテゴリを判定するタスクで、情報検索の基礎技術で、一般に近年は分散表現に基づくニューラルネットワークが用いられ、高い精度を実現している [1].

一方で、ニューラルネットワークをはじめとした深層学習モデル全般の課題として判定プロセスの不明瞭さが指摘されており、近年では説明可能な AI (Explainable AI) の研究が提唱されている [2]. 固有表現認識といった自然言語処理においても、判定のプロセスを分析可能にすることはモデルの結果をユーザが利用・改善するうえで必要であると考えられる。しかし文法的な規則の情報を特徴として活用した固有表現認識のアプローチはほとんどない [3].

本研究では、言語学における規則を既存の固有表現認識モデルと併用することにより、1) モデルの解釈性向上 2) 文法的に整合性の取れない誤判別結果の修正を目指す。そのための前段階として、選択制限 (Selectional Restriction) と呼ばれる言語学的規則を表現するモデルを構築する。

## 2 先行研究

Dai らは、固有表現認識タスクの既存モデルにおいて、共参照 (Coreference) の情報を考慮した拡張モデルを提案している [3]. 共参照とは、同じ実体を表す語句が複数の文にまたがって存在する状態を指す。既存モデルに共参照情報処理層 (coreference layer) を加え、共通文脈上に同一文字列が異なるカテゴリの固有表現として認識される問題を解決した。また共参照関係であるために同じカテゴリを付与するというプロセスにより、判定根拠がより明確になっていると言える。

## 3 選択制限

本研究が選択制限を取り扱う理由を解説する。選択制限とは、言語学の生成文法理論において提唱された「使用されるの種類に応じて、主語や目的語となれる語に制限が存在する」という理論である [4].

**A** 太郎は書籍を読む

**B** 京都は生八つ橋を読む

選択制限の具体例を説明するために「読む」という述語について考える。「読む」が主語として取りうるのは動作が可能な主体であり、同様に目的語として取りうるのは読解可能な対象である。従って一般的な文脈では A は制約を満たしており、B のように地名や食品名が「読む」の主語、目的語となる可能性は低い。

表 1 選択制限に影響する名詞の型情報 [5]

<i>count or mass</i>
<i>places or portions of matter</i>
<i>kind or individual</i>
<i>abstract or concrete</i>
<i>eventualities or object</i>
<i>animate or non-animate</i>
<i>propositions or fact or eventualities</i>
<i>locations or objects</i>

## 4 提案手法

上に示した選択制限の例より、「読む」という語の主語にあたる固有表現のカテゴリは人名と地名では前者の確立が高い。このようにある語が選択制限を課すとき、制約の種類や強さを明らかにすることで、既存固有表現認識モデルの結果からより複雑な解釈を得たり、制約にそぐわない誤分類をルールベースで修正する。

表 1 は選択制限が課される条件として考慮すべき型情報で、素性 (feature) と呼ばれる。表 1 が示すように、可算名詞を表す *count* と非可算名詞を表す *mass* は互いに独立している。本研究ではある語を入力したとき、それらの制限制約に対し各素性がどの程度影響を持つのかを表す教師あり学習モデルを構築する。

## 5 おわりに

今後の課題として、学習データをどこから作成するかを検討が必要になる。ある名詞が表に示した型情報のどれを性質として含むのかを収集する必要がある。

## 参考文献

- [1] Krisztian Balog. *Entity-oriented search*. Springer, 2018.
- [2] Tim Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, Vol. 267, pp. 1–38, 2019.
- [3] Zeyu Dai, Hongliang Fei, and Ping Li. Coreference aware representation learning for neural named entity recognition. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), Macau*, 2019.
- [4] Noam Chomsky. *Aspects of the Theory of Syntax*, Vol. 11. MIT press, 2014.
- [5] Nicholas Asher. Selectional restrictions, types and categories. *Journal of Applied Logic*, Vol. 12, No. 1, pp. 75–87, 2014.