

# 先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	自然言語処理学 (渡辺 太郎 (教授))		
学籍番号	1911117	提出日	令和 3年 1月 25日
学生氏名	澤田 悠治		
論文題目	Coordination Boundary Identification without Labeled Data for Elliptical Entity Disambiguation 並列構造の教師なし範囲同定手法を用いた複合化された固有表現の曖昧性解消		
要旨			
<p>固有表現認識は、科学技術論文などのテキストで記述される専門用語(固有表現)を自動的に抽出するタスクである。これまでの固有表現認識は、連続した範囲にある固有表現のみが評価対象とされており、並列構造によって省略を伴う複合化された固有表現を個別に抽出できないという問題があった。本研究では、近年の自然言語処理タスクで広く使用されている学習済みの言語モデルを用いて、教師情報を用いずに並列する句の範囲を同定し、複合化された固有表現を正規化する手法を提案する。GENIA Treebank とGENIA Term annotation を用いた評価実験では、教師情報を使用した先行研究と同等の解析性能を示し、提案手法によって固有表現認識の精度が向上することが確認された。</p>			