

## 論文審査の結果の要旨および担当者

報告番号	※ 甲 第	号
------	-------	---

氏 名 山腰 貴大

論 文 題 目

Coordination Analysis and Term Correction for Statutory Sentences using Machine Learning

(機械学習による法令文の並列構造解析及び用語校正)

論文審査担当者

主 査 名古屋大学教授 外山 勝彦

委 員 名古屋大学教授 武田 浩一

委 員 名古屋大学教授 松原 茂樹

委 員 名古屋大学准教授 小川 泰弘

山腰貴大君提出の学位論文「Coordination Analysis and Term Correction for Statutory Sentences using Machine Learning」(機械学習による法令文の並列構造解析及び用語校正)は、日本の法令文の解析および生成に関して、その構造や用語の特徴を利用した計算機支援の方法を扱った論文である。法令文は、その意味を正確に規定するために長く複雑になりがちであり、また、正しく解釈するためには法令文特有の構造や用語を理解する必要がある。そのため、一般的な自然言語処理技術をそのまま適用しても十分な処理性能は得られない。本論文は、特に法令文における並列構造解析と用語校正について新しい手法を提案し、実際の法令文を用いた評価を通じて、その有効性を示したものであり、全7章で構成されている。

第1章は序論であり、日本の法令文について、法制執務と呼ばれる実務上の記述規則を踏まえて説明した後、関連研究と比較して、本論文の位置付けを明らかにしている。

第2章では、法令文の並列構造について、一般の文とは異なる特有の構造がある点を説明するとともに、本論文で提案する手法の基礎となる言語モデルや分類器などの機械学習手法について概説している。

第3章では、法令文に対する並列構造解析について述べている。一般的な構文解析手法では法令文の並列構造を正しく解析できない点を明らかにした後、この問題を解決するために、法制執務における階層的並列構造規則を利用し、並列句の範囲を特定するためにニューラル言語モデルを用いる手法を提案している。この手法では、特に言語モデルの構築において、並列構造を示すタグが付与されていない法令文コーパスから文脈情報を学習している。

第4章では、法令文における法令用語の校正について述べている。法令用語には一般の文における用法とは異なる特有の使い分けがあるが、それを正しく判定する問題をテキストの選択肢付き穴埋め問題として位置付けている。その解決手法として、伝統的な機械学習であるランダムフォレストを用いる手法と、最新のニューラルネットワークに基づくBERTを用いる手法の二つを提案し、比較している。また、機械学習におけるデータの不均衡性について、本論文の対象には2種類のレベルがある点を指摘し、それぞれに対策を講じている。

第5章は、第4章におけるランダムフォレストを用いた手法をタイの法令文に適用した場合について述べ、この手法が日本以外の法令文にも適用可能であることを示している。その際、提案手法を単純に適用するのではなく、タイの法令用語の用法が制定年に依存するなどの特徴を捉え、日本の法令文が有しない素性の追加などにより、処理性能を向上させている。

第6章では、本研究と実世界におけるデータ循環の関係について考察し、本研究の成果について議論している。

第7章は、本論文を総括し、今後の課題と展望について述べている。

以上のように、本論文は、法令文という特有の規則をもつ言語領域に対して、自然言語処理技術をどのように適用すべきかを示した点で学術的な寄与があるだけでなく、近年リーガルテックと呼ばれるようになった立法・行政・司法の電子化やその関連産業への応用においても寄与が大きい。よって、本論文の提出者・山腰貴大君は博士(情報学)の学位を受けるのに十分な資格があるものと判定する。