

## システム情報工学研究科修士論文概要

年 度	平成 25 年度	学位名	修士( 工学 )
専 攻	知能機能システム	専攻	著者氏名 豊田 樹生
指導教員氏名 宇津呂 武仁			
論文題目  パテントファミリーを用いた専門用語訳語推定における対訳文対非抽出部分の利用			
論文概要 <p>特許文書の翻訳は、他国への特許申請や特許文書の言語横断検索などといったサービスにおいて必要不可欠なものである。この特許文書翻訳の過程において、専門用語の対訳辞書は重要な情報源である。本論文では、日米パテントファミリーを情報源として、専門用語対訳辞書を生成する手法を提案する。従来より、日米パテントファミリーの対応特許文書中において、「背景」および「実施例」の部分の日英対訳文対の対応付けを行い、これを情報源として専門用語の対訳辞書を生成する手法が提案されている。例えば、NTCIR-7 特許翻訳タスクにおいて配布された日英 180 万件の対訳特許文を用いて、対訳特許文からの専門用語対訳対獲得を行った研究がある。この研究では、句に基づく統計的機械翻訳モデルを用いることにより、対訳特許文から学習されたフレーズテーブル、要素合成法、Support Vector Machines (SVMs)を用いることによって、専門用語対訳対獲得を行った。</p> <p>しかし、上述の方式では、対訳文対が抽出される部分は、「背景」及び「実施例」全体の約 30% であり、約 70%は利用されていなかった。そこで、本論文では、「背景」および「実施例」のうちの残りの 70%の部分を言語資源として、既存の対訳辞書を用いて専門用語の訳語推定を行う方式の有効性を実証する。特に、人手で作成された辞書である英辞郎及びその部分対訳対辞書に加えて、全体の約 30%を訓練例として学習したフレーズテーブルを併用して要素合成法を適用し、専門用語の訳語推定を行う方式、および、フレーズテーブルのみを用いて要素合成法を適用し、専門用語の訳語推定を行う方式を提案する。また、要素合成法を適用する際には、訳語候補が相手言語側特許文書中に存在するか否かの検証を行うことにより、高精度な訳語推定を実現する。</p> <p>また、このとき、Moses と呼ばれる統計的機械翻訳モデルのツールキットを用いて専門用語の訳語推定を行った場合と提案手法の要素合成法との比較実験を行う。その結果、比較実験において、提案手法の要素合成法により、Moses で訳語推定を行った場合よりも高い適合率・再現率を達成することができた。</p>			
審査日	平成 26 年	1 月	30 日
審査員	(大学名 職名)	(学位)	(氏名)
主査	筑波大学 教授	博士(工学)	宇津呂 武仁
副査	筑波大学 准教授	博士(工学)	古賀 弘樹
副査	筑波大学 准教授	博士(工学)	矢野 博明