

平成16年度 情報工学コース卒業研究報告要旨

外山 研究室	氏 名	國 安 結
卒業研究題目	対訳コーパスを用いた法令文口語化に関する研究	
<p>国際取引の円滑化や法整備支援を目的に、近年、法令文翻訳に対するニーズが増加している。しかし、法令文には専門用語や独特の言い回しが多く含まれており、翻訳作業において一般の辞書が使用できない場合もある。また、法令の翻訳作業は現在専門家が人手で行っているが、コストがかかる作業になっている。</p> <p>さて、我が国では片仮名・文語体で書かれた法令文を平仮名・口語体に改める「法令文の口語化」が現在進められている。第二次世界大戦前に制定された法令は片仮名・文語体で書かれており、国の法律では約100件が今日でもそのままの形で残されている。そのような法令を改正する場合、原則として同じ文体で記述することになっているため、現在でも片仮名・文語体の法令文に改正されることがある。この作業には通常の場合法令作成以上に労力がかかることは言うまでも無い。さらに、前述した法令文翻訳のニーズの増加に伴い、片仮名・文語体で書かれた法令文を外国語に翻訳する要求もある。その際、直接外国語に翻訳するのは困難であるため、まず口語化し、その上で外国語に翻訳することが考えられる。以上の理由から、法令文の口語化は必要であると考えられるが、この作業も外国語への翻訳の場合と同様にコストがかかる作業である。</p> <p>そこで本研究では、法令文口語化の支援を目的に、対訳コーパスを用いた法令文自動口語化を行った。なお、対訳コーパスには片仮名・文語体で書かれた旧刑法とそれを口語化した新刑法を用いた。まず始めに、計算機を用いて文語 - 口語間の法令文対訳辞書を自動的に作成した。その手法は、コーパス中に3回以上出現する両言語の単語間で類似度を求め、その値が最も高いものを対訳として抽出するというものである。類似度にはDice係数を使用した。その結果、対訳ペア806個からなる対訳辞書を自動的に作成することができた。この辞書の精度（正しい対訳ペアが得られた割合）は69.5%であった。また、得られた辞書の精度を向上させるために、対訳の語頭や語尾に余分に登場する文字の削除を行うことによって、辞書の修正を図った。その結果、精度は73.3%に上昇した。</p> <p>次に、作成された対訳辞書を用いて、旧刑法の自動口語化を行った。具体的には、片仮名・文語体を平仮名・口語体に変換しても文法的な構造や語順はほぼ同じであることに着目し、辞書に掲載されている片仮名・文語体の単語が旧刑法に出現した場合、その対訳である平仮名・口語体の単語に置き換えることにより自動口語化を実現した。この結果と新刑法を比較した結果、自動口語化の精度（正しく口語化された形態素の割合）は71.5%だった。さらに、「辞書は文語の文字数の長い順に適用する」、「一度書き換えたものに対して書き換えは行わない」という、辞書の適用方法に関する2つの規則を追加した結果、精度は81.9%に上昇した。</p> <p>また、旧民法に対しても同様の方法を適用したところ、最終的に75.7%の精度で旧民法の自動口語化を行うことができた。以上より、刑法以外の法令文でも本手法が有効であることが確認できた。</p>		