

## 平成22年度 情報工学コース卒業研究報告要旨

外山 研究室	氏 名	小 林 隆 誠
卒業研究題目	構文パターンに基づく法令ターミノロジーの構築	
<p>近年、日本では、経済活動の国際化に伴い法令翻訳の需要が増大していることから、法務省により日本法令の英訳が進行している。しかし、現在、法令翻訳に利用可能な言語資源は対訳辞書のみである。そのため、訳者や翻訳担当省庁間での翻訳の質のばらつきは依然として問題となっている。</p> <p>一方で、法令翻訳に関しては欧州連合 (EU) が先進的である。EU では、23ヶ国語もの公用語が使われている。どの公用語でも運営できるように翻訳総局 (DGT) を有しており、すべての法令が23ヶ国語に翻訳されている。そして、翻訳業務の利便を図るために、様々な分野の翻訳用の言語資源が構築されている。その一つとして、DGTにより、EU内の各機関に既存のターミノロジーデータベースを統合してIATE(InterActive Terminology for Europe)が構築された。IATEは、EU公用語にラテン語を含む24ヶ国語での専門用語とその定義を分野ごとに収録したターミノロジーデータベースである。およそ840万語を収録し、用語の収録範囲は法学、金融、政治から農学、工学まで広範に及ぶ。</p> <p>ターミノロジーとは、いくつかの定義が存在するが、本研究では個々の主題領域における概念体系を表現する専門用語の集合と捉える。専門用語は概念を持ち、それは定義として自然言語で表現される。そして、その概念は概念の限定や結合を行うことで体系を生じ、概念に基づいた専門用語の体系が形成される。このように作られたターミノロジーは、ある分野の知識を整理するだけでなく、同じ概念を持つ複数の言語で表現された専門用語を収録することにより、その分野に対する翻訳資源として活用できる。</p> <p>そこで、本研究では、法令翻訳のための言語資源として、日本法令に対するターミノロジーの構築を行う。</p> <p>自然言語を対象に構文パターンから概念階層を抽出する研究は広く行われ、一定の成果を上げている。例えば、Miller(1990)らは、“<math>y</math> is a (kind of) <math>x</math>”という表現があるとき、<math>x</math>が<math>y</math>の上位語であるとしている。Hearst(1992)はさらに、“such <math>x</math> as <math>y</math>”などのパターンを指摘した。この手法は日本語でも適用できることも知られており、安藤(2003)らは「<math>y</math>などの<math>x</math>」などの構文パターンを指摘した。</p> <p>本研究では、これらの研究の成果を踏まえ、法令ターミノロジー構築に当たって概念階層を表す構文パターンの分析を行った。法令には定義規定と呼ばれる規定があり、その規定は定義語と定義文で構成されている。それに着目して構文パターンを発見し、定義語と定義文の抽出を行った。次に、得られた定義語と定義文のペアから概念階層の構築を行った。定義文から概念を得るために、抽出する定義文に対して観察に基づき、構文パターンを得た。</p> <p>例えば、文部省設置法第二条第一号の「「学術」とは、人文科学及び自然科学並びにそれらの応用の研究をいう。」という規定から、定義語「学術」と定義文「人文科学及び自然科学並びにそれらの応用の研究」が得られる。さらに、定義文から「学術」の上位語として「人文科学」、「自然科学」、「研究」が抽出できる。</p> <p>これらの獲得した構文パターンを基に、法令ターミノロジー構築システムを作成し、構築されたターミノロジーを評価した。</p>		

