

令和元年度 情報工学コース卒業研究報告要旨

武田 (浩) 研究室	氏 名	大 柳 慶 悟
卒業研究題目	自動構築された格フレームと日本語 FrameNet の対応づけ	
<p>自然言語処理では言語の意味を記号的に扱う手段として、フレーム知識と呼ばれる知識表現が利用されてきた。フレームとは、語の理解のために必要な背景知識であり、言語処理で使用されるフレーム知識は FrameNet や PropBank に代表されるような人手により整備されたフレーム知識と、京都大学格フレームなどのような大規模コーパスから自動構築されたフレーム知識の2つに大きく分けることができる。このうち前者は、各フレームがどのような意味役割をとるか等を記述したフレーム辞書とそれらがアノテーションされた文例から構成されることが一般的である。その整備には多くのコストが必要であり、実際にアノテーションされた動詞や文の数は限定的であるものの、フレームの粒度や意味役割が人手で決定されていることから、人間にとって直観的に理解しやすいという特長がある。一方、後者は、大規模コーパス中の文例で出現した述語についてその述語がとる格とその格の用例から自動構築されているため、フレームの粒度が人間にとって直観的であるとは限らず、また、各フレーム、格がどのような意味内容を持つかという情報は明示的には与えられていないものの、非常に高いカバレッジや大規模な語彙的情報を持つという特長がある。</p> <p>本研究では、これらの2つのタイプのフレーム知識が持つ特長を併せ持つフレーム知識、すなわち、人間にとって直観的なフレーム粒度、意味役割の情報を持ち、かつ、カバレッジの高いフレーム知識の獲得を目的とし、自動構築されたフレーム知識と人手により整備されたフレーム知識の自動的な対応づけに取り組んだ。具体的には、大規模日本語コーパスから自動構築されたフレーム知識である京大格フレームを、人手で整備された日本語のフレーム知識である日本語 FrameNet (以下 JFN) と対応づけることで、京大格フレームが持つ豊富な語彙的情報と、JFN が持つ意味情報を併せ持つフレーム知識の獲得を行った。</p> <p>京大格フレームは各述語の用法ごとに構築され、フレームごとにその用法において必要な格とその用例が記述されている。JFN は事象・概念ごとにフレームが構築され、それぞれの JFN フレームにはフレーム喚起語、関連する意味役割、意味役割の情報が付与された文例が記述されている。提案手法では述語が与えられたとき、その述語がとる各格フレームを、その述語を喚起語として含む JFN フレームの1つに対応づける。対応づけは格フレームの格に記述されている用例と対象となる JFN フレームの文例の類似度を用いて行った。このとき JFN の文例のうち、与えられた述語を含む文例に大きな重みを与えることで、関連性が高いと考えられる文例を重視した対応づけを行った。さらに、文例中で用いられた格を考慮することでフレーム間の対応づけの高精度化を目指した。</p> <p>提案手法の評価は対応先の候補となる JFN フレームが2つのみである格フレームに対して、専門家が正しい対応づけ情報を付与した19動詞119フレームのデータを用いて行った。さらに、各動詞の出現頻度がもっとも大きい格フレームに対して格レベルの対応づけ情報を付与したデータを用いて格の対応づけ精度の評価を行った。フレームの対応づけ実験の結果、提案手法では約79%の精度で正しいフレームに対応づけることができた。また、与えられた述語を含む文例を重視する場合とそうでない場合を比較した結果、与えられた述語を含む文例を重視することで精度が向上することが確認できた。一方、用例が伴う格を考慮する場合としない場合ではフレームの対応づけ精度に差は見られなかったが、格レベルの対応づけ精度は向上することが確認できた。</p>		