

推論システムの日本語処理に適した文章解析機能の開発

中村 勝則[†], 小林 翔子[†], 大野 知佳[†], 溝端 佳奈[†], 西岡 知夏[†]

武庫川女子大学 情報メディア学科[†]

1. はじめに

近年は、日本語、英語をはじめとする自然言語でユーザとインタラクションするシステムが普及しはじめている。また、自然言語インタラクションを実現するためのソフトウェアツールやプログラムライブラリも一般に利用できる形（OSS などの形態）で増えてきており、人工知能系アプリケーションプログラムを開発するための環境も整ってきている。

本研究は、日本語の文章の特徴を抽出する機能の開発に関するものであり、本発表では、文章の形態素解析と係り受け構造の解析の結果から「主部」「述部」といった語の機能を特定するための基本機能の開発について報告する。

2. 研究の動機と目的

本研究では、ユーザから入力された自然言語（日本語）の文章をコマンドとして受け付け、それに対して自然言語で応答するシステムの構築を目指している。そのためには、入力された日本語の文章構造を統計的に特徴化（語の出現頻度に基づく特徴化など）するだけでは不十分であり、格解析に基づいた「語の役割」を特定することが必要となる。このような処理は、自然言語処理（NLP）の分野の中でも特に「自然言語理解」（NLU）と呼ばれるものであり、NLUのためのソフトウェア製品が既にいくつか（IBM 製 Watson API など）存在している。また、自然言語を述語論理の式に変換する研究（参考文献 1, 2

参照）も行われている。ただし、商用の API サービスにおいては十分な機能が公開されていない状況であるのと、学術研究の成果がソフトウェアライブラリとして十分に整備されていない状況があるので、現在のところ NLU のための API が手軽に利用できるとは言い難い。幸いにして、日本語形態素解析、係り受け構造の解析、格解析に関しては、相当に高精度で高機能なソフトウェアツール（JUMAN, KNP：京都大学 黒橋・河原研究室）が公開されており、それらツールの出力結果を利用することで、比較的簡単に日本語文章の要素の特徴化をするための機能が構築できる。

本研究の目的は、一階述語論理に基づく処理系（Prolog）の上に構築した推論アプリケーションとユーザが自然言語でインタラクションするために必要な機能の構築であり、そのためには、自然言語と一階述語論理の式を相互に変換する機能が求められる。現在のところ、日本語の文章を論理式に変換するための前処理として、係り受けのある文章の要素を特徴化する機能が概ね完成したのでここに報告する。

3. 構築したシステム

テキスト（文字情報）として入力した日本語文章を JUMAN で形態素解析し、その結果を KNP によって解析（係り受け構造と格構造）したのから JSON 形式の特徴データを生成する機能を開発した。例えば ”私は大阪の書店で京都観光の本

処理結果 1

S-ID:1 KNP:4.2-e3fb1e34 DATE:2018/12/16 SCORE:-17.83613

```

      私は———| <体言>
大阪の———| <体言>
      書店で———| <体言>
京都———| <体言>
      観光の———| <体言><格解析結果:ノ/->
      本を———| <体言>
      買いました。 <用言:動><格解析結果:ガ/私;ヲ/本;ニ/-;デ/書店;時間/->
    
```

EOS

A Development of Japanese Language Processing Function for Logical Reasoning System

[†] Katsunori Nakamura, Shoko Kobayashi, Tomoka Ohno, Kana Mizobata, Chika Nishioka
Mukogawa Women's University, Dept.of Informatics and Mediology

処理結果 2

```
[{'apdx': [{'text': 'は',
            'wcl': '助詞',
            'wc2': '副助詞'}],
  'feature': {'cp': 'は',
              'type': 'subject'},
  'text': '私',
  'wcl': '名詞',
  'wc2': '普通名詞'},
{'apdx': [{'text': 'で',
            'wcl': '助詞',
            'wc2': '格助詞'}],
  'feature': {'cp': 'で',
              'type': ['cp',
                       'で']}],
  'mod': [{'apdx': [{'text': 'の',
                    'wcl': '助詞',
                    'wc2': '接続助詞'}],
            'feature': {'cp': 'の',
                       'type': 'adj'},
            'text': '大阪',
            'wcl': '名詞',
            'wc2': '地名'}],
  'text': '書店',
  'wcl': '名詞',
  'wc2': '普通名詞'},
```

(右に続く→)

を買いしました。”という文章をJUMANとKNPで解析すると前ページに示す**処理結果 1**のようになる。

JUMANとKNPの機能はWebサイト

<http://reed.kuee.kyoto-u.ac.jp/nl-resource/cgi-bin/knp.cg>

でもデモンストレーションとして公開されており、同様の処理が確認できる。

KNPでは、プログラム起動時に`-tab`オプションを与えることで解析結果を詳細な情報を持つデータレコードとして出力することができ、その出力をJSON形式のデータ構造に変換するためのPython用モジュール**afterKNP**を開発した。先に挙げた日本語の例文の解析結果をafterKNPに与えてJSONデータにしたものが上に示す**処理結果 2**である。このデータ構造は基本的には語の情報を要素とするリストであり、各要素の`feature`キーの値としてその語の役割を示唆する情報(品詞、時制、格による特徴)を保持している。

4. 発展

現在のところafterKNPが完成に近い状態にあるが、語の特徴情報から述語論理の式を生成する機能に関しては開発中である。これについて

```
{'apdx': [{'text': 'を',
            'wcl': '助詞',
            'wc2': '格助詞'}],
  'feature': {'cp': 'を',
              'type': 'object'},
  'mod': [{'apdx': [{'text': 'の',
                    'wcl': '助詞',
                    'wc2': '接続助詞'}],
            'feature': {'cp': 'の',
                       'type': 'adj'},
            'mod': [{'feature': {'type': 'adj'},
                    'text': '京都',
                    'wcl': '名詞',
                    'wc2': '地名'}],
            'text': '観光',
            'wcl': '名詞',
            'wc2': 'サ変名詞'}],
  'text': '本',
  'wcl': '名詞',
  'wc2': '普通名詞'},
{'apdx': [{'text': 'ました',
            'wcl': '接尾辞',
            'wc2': '動詞性接尾辞'}],
  'text': '.',
  'wcl': '特殊',
  'wc2': '句点'}],
  'feature': {'affm': True,
              'tense': -1,
              'type': 'verb'},
  'text': '買い',
  'wcl': '動詞',
  'wc2': '*}]}
```

は状況演算(Situation Calculus)に基づいた推論機能をProlog上に開発して、システムとユーザのインタラクションの脈絡を用いた推論を行い、記号接地の精度向上を図りながら、語の意味と役割を特定する機能を開発する予定である。

5. 参考文献

1. 稲田和明, 松林優一郎, 井之上直也, 乾健太郎, “効率的な推論処理のための日本語文の論理式変換に向けて”, 言語処理学会 第19回年次大会 発表論文集, 2013年3月
2. 高柳俊祐, 上條敦史, 石川勉, “日本語文から拡張型述語論理式への自動変換ツール: CONV”, 人工知能学会論文誌 27 卷 5 号 B, pp. 271-280, 2012

※ afterKNP を公開している。(下記)

http://www.k-techlabo.org/blog2/?page_id=1222

