

簡単な日本語を1つの単語に言い換える

C0119175 関 大輝

1. はじめに

日本語には様々な言い換えが存在する。同じ意味を表す異なる表現を言い換えと言う。

私が目指すものは、どのような日本語でも自由に言い換えができるサービスである。現在、難しい文章を日本在住の外国人にも理解できるような易しい文章にすることは可能である。また、文章内で使用している単語を抽出し、その単語に近い意味の単語の候補を複数表示することも可能である。しかし、文章をまとめるときや、物事を一言で表したいときに、文を1つの単語に言い換えるというサービスは、未だ見つかっていない。そのため、私は日本語の文を1つの単語に言い換える仕組みを考える。

一致率を計算する手法として、ゲシュタルトパターンマッチング、レーベンシュタイン距離法、ジャロ・ウィンクラー距離法がある。この3つの手法のうち、全ての手法を比較し、適しているものを使用する。

この作業を行うことで、文を1つの単語に言い換える技術の実現に近づくことができると考えた。

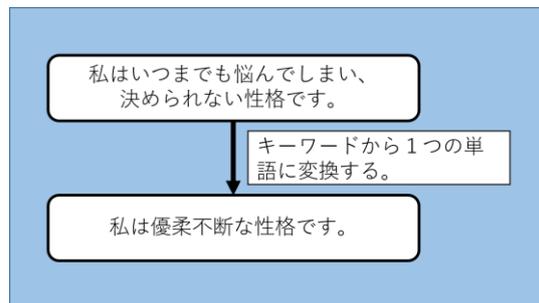


図1 実験のイメージ図

2. 関連研究

2-1. 伝えるウェブについて

伝えるウェブは、語彙の制限、分ち書き、短い文章への分割、振り仮名の追加などの処理を機械学習によるAIが自動で行う国内初のサービスである。伝えるウェブは、防災の目的で開発され、日本在住の外国人にも理解できるような基礎的な日本語を使用し、言い換えを可能としている。直訳タイプの機械翻訳を用いて、原文1つ1つの言葉を忠実に置き換えている。[1]

文章に出てくる単語を分かりやすい表現に言い換えることはできるが、文を言い換えることはできないのが問題である。

2-2. 言語能力に合わせた文への言い換えに関する研究

この研究は、子供や外国人などが難しい言い回しの文章を理解できるように言い換えるというものである。

サ変名詞と2字の漢字熟語を言い換えるの対象としている。言い換えをするアルゴリズムとして、対象文の入力、「茶釜」による解析、サ変名詞・漢字熟語検出、言い換えるべきか判定、単語の置き換え、言い換え文の出力の順に言い換え作業が行われている。[2]

3. 研究の概要

本研究では、日本語で書かれた文を、その文と同じ意味になるような単語に言い換えをするというを行う。また、二字熟語への言い換えのみを対象とする。

まず、二字熟語の意味を形態素解析する。その中から、キーワードとなる言葉を抽出し、文との一致率から1つの単語を導き出す。

4. 研究計画

	8月	9月	10月	11月	12月	1月	2月
単語のデータ収集	■	■					
アルゴリズム作成			■	■			
卒業論文作成					■	■	
卒業論文提出						■	
卒業論文最終発表							■

5. 進捗状況

- ・関連研究の調査
- ・言い換え方法の模索
- ・データ収集方法の模索

6. おわりに

文章をまとめることや物事を簡潔に説明することを目的とし、日本語の文を1つの単語に言い換えるシステムの概要と、構築手順について述べた。

参考文献

- [1]アルファサード株式会社, “伝えるウェブ | やさしい日本語で情報発信”, <https://tsutaeru.cloud/>
- [2]三戸部 矩倫, 横山 晶一: “言語能力に合わせた平易な文への言い換えシステム”, [情報処理学会第69回全国大会](#), 6Q-7, pp. 2-435~2-436 (2022).