

日本語述語項構造解析における マイクロブログに対する文分割手法の比較

述語項構造解析とは？

述語項構造解析は自然言語処理の一種であり、自然言語処理とは、人間が普段使う言語である自然言語を計算機で扱えるようにするための様々な処理を指す。言語を計算機で扱えることから自然言語処理の応用先は多岐に渡る。

述語項構造解析自体は、文中の事態を表す述語とそれに関わりのある名詞である項を項の種類と共に見つけ出す技術である。

マイクロブログと自然言語処理

Twitterなどを始めとする、非常に短い文章を投稿するマイクロブログはインターネットの普及に伴い生まれた言語を扱う新たなサービスの種類の一つである。

マイクロブログのような新たな言語資源は自然言語処理の研究が十分に進んでおらず、既存手法の有用性の検証や文の特性に合った新たな処理の考案が望まれている。

研究概要

この研究は、マイクロブログの投稿に対して述語項構造解析を行うための前処理の一つであるマイクロブログの一投稿を文の集合に変換する手法に着目したものである。

実際には、既存研究で提案されたいくつかの文分割手法をそれぞれ適応することで述語項構造解析の精度が変化するかを検証している。