

## 課題番号14-03

最長一致法と分割数最小法で形態素への分割結果が異なる例を示す。

例えば、

「全国都道府県議会議長席」

を最長一致法で解析すると

「都道府県議」という単語が辞書にあった場合は

「全国 都道府県議 会議 長 席」

上記のように5つに分割されるような解析結果になる。

続いて、分割数最小法で解析すると、

「全国 都道府県 議会 議長席」

のように、4つに分割されるような解析結果となる。

この文を、形態素解析エンジンMeCabを用いて解析すると

以下のような出力結果を得る。

```
全国 名詞,一般,*,*,*,*,*全国,ゼンコク,ゼンコク
都道府県 名詞,一般,*,*,*,*,*都道府県,トドウフケン,トドーフケン
議会 名詞,一般,*,*,*,*,*議会,ギカイ,ギカイ
議長 名詞,一般,*,*,*,*,*議長,ギチョウ,ギチョー
席 名詞,一般,*,*,*,*,*席,セキ,セキ
EOS
```

予想に反して、「議長席」の部分は「議長」と「席」に分割されてしまった。これは、議長席という単語がMeCabの使用する辞書になかったためと考えられる。

MeCabの制約付き解析機能オプション(-p)を利用し、都道府県議を名詞と指定して解析すると、以下のようになる。

```
%> mecab -p
全国
都道府県議      名詞
会議長席
EOS
全国 名詞,一般,*,*,*,*,*全国,ゼンコク,ゼンコク
都道府県議      名詞,一般,*,*,*,*,*
会 名詞,接尾,一般,*,*,*,*,*会,カイ,カイ
議長 名詞,一般,*,*,*,*,*議長,ギチョウ,ギチョー
席 名詞,一般,*,*,*,*,*席,セキ,セキ
```

また、議長席を名詞と指定した場合は以下のようになる。

```
全国都道府県議会
議長席      名詞
```

EOS

全国 名詞,一般,\*,\*,\*,\*,\*全国,ゼンコク,ゼンコク

都道府県 名詞,一般,\*,\*,\*,\*,\*都道府県,トドウフケン,トドーフケン

議会 名詞,一般,\*,\*,\*,\*,\*議会,ギカイ,ギカイ

議長席 名詞,一般,\*,\*,\*,\*,\*

EOS

## 参考文献

- 中川裕志."自然言語処理入門「ここではきものをぬぐこと」".<https://taku910.github.io/mecab/>
- mwSoft.MeCabのコマンドライン引数一覧とその実行  
例.2009/09/28.[http://www.mwsoft.jp/programming/munou/mecab\\_command.html#partial](http://www.mwsoft.jp/programming/munou/mecab_command.html#partial)
- 京都大学情報学研究科."MeCab: Yet Another Part-of-Speech and Morphological Analyzer".<https://taku910.github.io/mecab/>