

```

# vim: fileencoding=utf-8
import MeCab
import wordAnalysis

m = MeCab.Tagger ("-Ochasen -l1")
print m.parse ("今日もしないとね")
print m.parseNBest(2, "今日もしないとね。今日も。")#2つ分の解析が出力され
れます。

# taggerをもう一つ作成して分ち書きをリストに格納。
t = MeCab.Tagger("-Owakati")
sent = t.parse("今日はするよね。今日は。今日今日")
print sent

sentList = sent.split(" ")#空白で区切ってリスト構造に変換

print sent.count("今日")#2が出力されます。今日が2個ありますので。

#pythonではlistやdictionaryをprintする際、日本語があると正しく出力できな
い。
#print sentListは故に駄目。
for i in range(len(sentList)):
    print sentList[i]
#単語の共起頻度を文章から計算する。
#第1引数を文章、第2に一つ目の単語,第3に二つ目の単語を指定する。
co_oc = wordAnalysis.co_occurrence("今日は今日とて。今日は。", "今日",
"は")
print co_oc
#事前に重複単語を削除したリストを作成
NoSameWordList = wordAnalysis.trim_redundancy(sentList)
#全単語について共起頻度を調べる。
for i in xrange(0,len(NoSameWordList)-1,2):
    co_oc = wordAnalysis.co_occurrence(sent,NoSameWordList
[i],NoSameWordList[i+1])
    print NoSameWordList[i],NoSameWordList[i+1],co_oc
#co2のテスト
for i in xrange(0,len(NoSameWordList)-1,2):
    co_oc = wordAnalysis.co_occurrence2(sent,NoSameWordList

```

```
[i],NoSameWordList[i+1])
```

```
    print NoSameWordList[i],NoSameWordList[i+1],co_oc
```