

キリル文字モンゴル語形態素解析器の開発

—マイナー言語に対する言語処理基盤開発—

村脇 有吾 (京都大学大学院情報学研究科)

形態素解析器をキリル文字モンゴル語に対して開発した。形態素解析器とは、入力として文を受け取り、それを最小の単位である形態素の列に分解して出力するプログラムである。計算機は、人が読み書きする言語、自然言語のデータをただの文字の並びとしか認識していない。計算機に言語を扱わせるためには、文字の並びが構造を持つことを教えなければならない。形態素解析は、そのもっとも基礎的な処理である。

```
c:/home/nurawaki/esp/qar $
c:/home/nurawaki/esp/qar $ rhino main.js 入力文
Веб програмчлал болон XML, веб үйлчилгээнд ашигладаг.
+ веб
веб веб noun - э <plural:-yya> <ja:ウェブ>
+ програмчлал
програмчлал програмчлал noun - а <ja:プログラミング>
+ болон
болон болон conjunction - о <ja:~と;~及び>
@1 + болон
@1 бол бол verb - о <ja:なる>
@1 он н verbal suffix converb -
+ XML undefined
+ ,
, , special comma - <ja:, >
+ веб
веб веб noun - э <plural:-yya> <ja:ウェブ>
+ үйлчилгээнд
үйлчилгээн үйлчилгээ noun - э <hiddenN> <ja:サービス;奉仕/ほうし> <plural:-yya> <hiddenNsurfaced>
д д nominal suffix dative-locative - <inflectionFunction:dativeLocativeC
aseInflection> <inflectionContext:LibNoun> <ja:~に;~で>
+ ашигладаг
ашигла ашигла verb - а <ja:利用する/リようする;取り扱う/とりあつかう>
даг даг verbal suffix verbal noun а <inflectionFunction:fourInflection>
<inflectionContext:LibUtility> <variants:аар;дэр;аор;аер>
+ .
. . special period - <ja:。>
EOS
```

解析結果