

フリーの音声認識システムである Julius を用いて、簡単に音声対話システムを構築してみる。

1 Julius のインストール

まず、Julius 記述文法音声認識実行キットの Windows 版を <http://julius.sourceforge.jp/index.php?q=grammar-kit.html> からダウンロードする。Windows 版は zip アーカイブなので、適当なツールを用いて解凍する。

練習 1 展開したフォルダにある HOWTO.txt を読んで、「果物購入タスク」を試してみよ。安価なものでよいので、マイクを用いて試すことが望ましい。

練習 2 Julius のシステム構成と認識のしくみを調べよ。

2 音声認識用文法の理解

一般に、音声認識システムは、言語モデルで生成できる文だけしか認識できない。今回ダウンロードした「記述文法音声認識実行キット」では、「文法」と呼ばれる規則で言語モデルが定義される。

Julius では「文法」は、文法規則と辞書から構成される。文法フォルダに含まれる拡張子 .grammar のファイルが文法規則 (grammar) を記述したファイルである。拡張子 .voca のファイルが辞書ファイルである。この 2 つのファイルを理解すれば、文法から生成される文が理解できる。

「果物購入タスク」を例に、文法から生成される文を生成する方法を説明する。

「果物購入タスク」の辞書 fruit.voca を以下に示す。

```
% FRUIT
蜜柑          m i k a N
リンゴ        r i N g o
ぶどう        b u d o:
% NUM
0             z e r o
1             i c h i
2             n i:
3             s a N
4             y o N
5             g o:
6             r o k u
7             n a n a
8             h a c h i
9             k y u:
% KO
個            k o
% WO
を            o
% KUDASAI
ください     k u d a s a i
% NISHITE
にして      n i s h i t e
% DESU
```

```

です          d e s u
% NS_B
<s>           silB
% NS_E
</s>         silE

```

この辞書では、単語クラスごとに単語とその語がどのように発声されるかが定義されている。例えば、最初の部分では、FRUIT という単語クラスには、「蜜柑」「リンゴ」「ぶどう」という 3 語が属していることが定義されている。さらに、「蜜柑」という語は、m i k a N という 5 つの音素から構成されることが定義されている。(Julius における読みと音素記号の対応は、SampleGrammar フォルダの下にある type というフォルダにある type.voca を見よ。) ファイルの末尾の silB と silE は、それぞれ発話の開始前・終了後の無音部分に対応する音素記号である。

このファイルで定義される単語がどのようにつながって文になるのかを定義するのが .grammar ファイルである。「果物購入タスク」の .grammar ファイルを以下に示す。

```

S           :  NS_B FRUIT_N PLEASE NS_E
FRUIT_N    :  FRUIT
FRUIT_N    :  FRUIT NUM KO
PLEASE     :  WO KUDASAI
PLEASE     :  NISHITE KUDASAI
PLEASE     :  DESU

```

Julius の文法は文脈自由文法で書く。上記の式は、通常は以下のように書かれる。

```

S           NS_B FRUIT_N PLEASE NS_E
FRUIT_N    FRUIT
FRUIT_N    FRUIT NUM KO
PLEASE     WO KUDASAI
PLEASE     NISHITE KUDASAI
PLEASE     DESU

```

この文法を使って実際の文を生成する例を示す。

1. Julius の文法では、S は特殊な記号であり、文全体を意味する。1 番目の規則は、文全体の S という記号が、NS_B FRUIT_N PLEASE NS_E1 という記号列に置き換えられることをあらわす。
2. NS_B と NS_E は、文法規則ではこれ以上置き換えられない記号である。これらの記号は、上記の辞書を参照して単語に置き換える。NS_B は <s> に、NS_E は </s> に置き換えられる。これらは、それぞれ、Julius で文の最初、文の最後を示す特別な記号である。
3. FRUIT_N は、2 番目、3 番目の規則で置き換えることができる。ここでは、2 番目の規則で置き換えてみる。すると、FRUIT に置き換えられる。FRUIT は、文法規則ではこれ以上置き換えられない記号なので、単語に置き換える。辞書には、FRUIT に属する単語は 3 つある。このいずれに置き換えても構わない。
4. 引き続き、残った PLEASE を置き換える。
5. このような手順を繰り返し、全ての記号が単語に置き換えられるまで続ける。すると、例えば、「<s> 蜜柑を下さい </s>」というような文が生成される。

練習 3 「果物購入タスク」の文法を用いて、異なる文を 10 文生成せよ。また、その文を発話して認識させて、認識できるか確認せよ。