

1 E-3

日本語アクセント生成システム

水野孝司・高田正之・小谷善行・西村恕彦
(東京農工大学 工学部 電子情報工学科)

1. 概要

入力された日本語分かち文に対して、アクセント法則(1)を用いて高・低のアクセントを生成するシステムを構築した。アクセントを決定するパラメータには様々なものがあるが、本研究ではそのうちの特に重要な品詞、拍数、活用形に関するアクセント法則をPrologをもちいて記述した。

2. アクセント法則

2-1. 用語の説明

アクセントの法則に使われる用語の説明を以下に記す。

- ・アクセント型
アクセントにはある種のパターンがあり、それに名前をつけたものが型である。型は大きく分けて平板型と起伏型があり、起伏型には頭高型、中高型、尾高型がある。
- ・名詞類、形容詞類、動詞類
意味的な要素を考慮した、単語の並びの品詞的な分類。例えば「私」は名詞類に属し、「私らしい」は形容詞類に属する。
- ・アクセントリスト
アクセントを、1(高)・0(低)の数字のリストで表したもの。例えば「わたし」は[0, 1, 1]となる。
- ・アクセント節
アクセントに関連したまとまりのある単語の並び。アクセント節内の個々の単語は、各々その前後の単語のアクセントに影響を及ぼし合う。

2-2. アクセント節

このシステムで取り入れた、アクセント節を構成する語のつながりを表1に示す。

表1. アクセント節を構成する品詞のつながり

		後部の品詞				
		名	動	形容	副・接・感	助 助動
前	名詞	○				○ ○
部	動詞		○			○ ○
の	形容詞					○ ○
品	副・接・感					○ ○
詞	助詞					○ ○
(類)	助動詞					○ ○

3. システム

3-1. システム構成

システムの構成を図1に示す。システム全体を制御するのが制御部で、入力される日本語分かち文に対し、①～④の過程で文のアクセントを生成する。

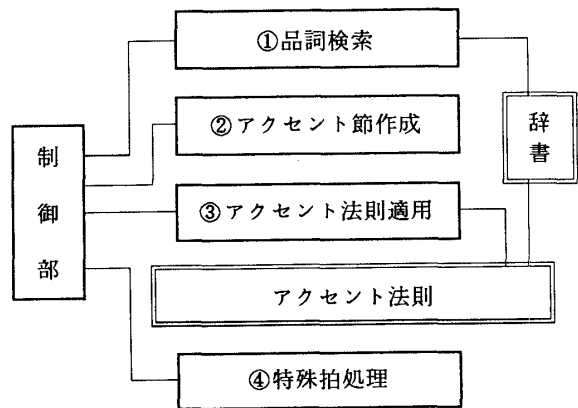


図1. システム構成図

3-2. システムの流れ

図1の①～④の説明を以下に記す。

① 単語の品詞情報の検索

入力された日本語分かち文を、品詞名つきのリストに変換する。

入力
[原稿, を, 読み, 直す]
出力
[[名詞(原稿), 助詞(を), 動詞(読み), 動詞(直す)]]

② アクセント節の作成

名詞、形容詞、動詞、副詞などを先頭にした節に分ける。ただし、名詞+名詞+名詞、動詞+動詞など、名詞のつながりと動詞のつながりは、例えば下の入力の場合、[[動詞(読み), 動詞(直す)]] のようにまとめる。

入力
[[名詞(原稿), 助詞(を), 動詞(読み), 動詞(直す)]]
出力
[[[[名詞(原稿)], 助詞(を)], [[動詞(読み), 動詞(直す)]]]]

③アクセントの生成

②で分けられた節に対して、アクセント法則を用いてアクセントを生成する。[[名詞(原稿)], 助詞(を)]の場合、
 [0, 1, 1, 1, 1] / 平板型
 (ゲ, ン, コ, ウ, ヲ)
 というアクセントが生成される。

④特殊拍の処理

③で生成されたアクセントに、引き音(-)、撥音(ん)のアクセントの処理をする。
 [0, 1, 1, 1, 1] → [1, 1, 1, 1, 1]
 (ゲ, ン, コ, ウ, ヲ)

制御部により、①から④の過程を経て、入力文に対し次のような文全体のアクセントを生成した。
 [1, 1, 1, 1, 1, 0, 1, 1, 1, 0]
 (ゲ, ン, コ, ウ, ヲ, ヨ, ミ, ナ, オ, ス)

4. 辞書

アクセントを生成する際に必要な、単語に関する情報を記す。

- ・品詞
 アクセントの法則がすべて「名詞のXXのアクセントは・・・」というように、まず最初に品詞で区別されているため、単語の品詞は絶対に必要なものである。
- ・読み
 単語の拍数の算出と特殊拍の処理に使われる。
- ・活用形
 動詞、形容詞のアクセントを生成するのに使われる。
- ・アクセント型
 アクセントの型は、助詞、助動詞以外の場合に使われる。もし単語の型が登録されていない場合、その単語の品詞におけるもっとも多いパターンとして処理するようにしてある。例えば名詞の場合にはアクセントの型が登録されていない場合には平板型として処理される。したがって、平板型の名詞についてはアクセントの型を辞書に登録する必要はない。

5. アクセント法則のプログラミング

5-1. アクセント節におけるアクセント生成の流れ

アクセントは基本的に、アクセント節の最後部と、前部のアクセントを見て生成される。図2は、アクセント節においてアクセントがどのように生成されるかを示したものである。

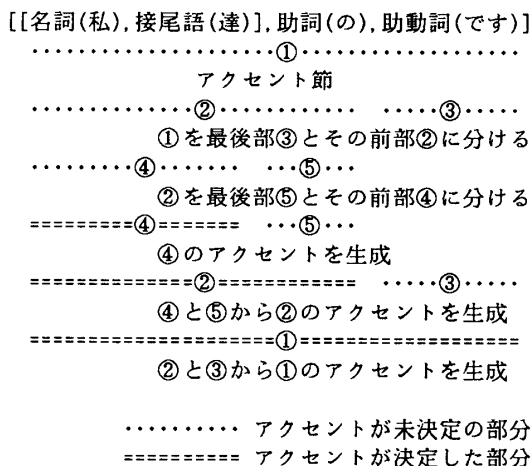


図2. アクセント節におけるアクセント生成の流れ

5-2. アクセント法則のProlog記述

アクセントの法則は一部を除いて、用語をパラメータとしてアクセントを生成する、というものである。これを達成するために、処理系を考えた結果、細かい数値演算は必要とせず、IF~THENルールの際にマッチングを多用する、という理由から本研究では処理言語としてPrologを使用した。

アクセント法則はすべて次のような、Prologでいう頭部を持ち、その後続く本体に法則が記述されている。

```

accent(S, L, K):-
    SがS1の場合,
    Sの前部のアクセントを生成する,
    これと後部からSのアクセントを生成.
    
```

```

accent(S, L, K):-
    SがS2の場合,
    Sのアクセントを生成.
    
```

```

S . . . . . アクセント節
L . . . . . アクセントリスト
K . . . . . アクセント型
    
```

図3. アクセント法則のProlog記述

S1は、Sを最後部とその前部に分けて、前部のアクセントを生成し、その結果からSのアクセントを生成するようなアクセント節である。図2の①や②がこれにあてはまり、例えば②の場合だとS1は「最後部が助詞(の)で、前部が名詞類である」ということになる。これに対しS2は、図2の④のような、他のアクセント法則を適用しないアクセント節である。

アクセント生成には、単語、品詞(類)、活用形、拍数、アクセント型、さらにはS1のように前部のアクセントリスト、アクセント型などが必要になる。

アクセント法則は、Prologの実行順序を利用し、語の組合せの複雑なものから単純なものへ、またパラメータの多いものから少ないものへと格納する。例えば<~+助詞+助動詞>の法則は<~+助動詞>や<~+助詞>などの法則より上に格納し、<動詞+動詞(する)>の法則は<動詞+動詞>の法則より上に格納する。こうすることにより、入力文に対しよりきめ細かいアクセント法則を対応させることができる。

6. むすび

アクセントの法則を5章のような方法で組み立てることにより、表1の語のつながりについてはすべてアクセント法則を組み立てることができた。3-2章の②では、このアクセント法則に対応するアクセント節しか作成しないので、構文に誤りのある文でも品詞と読みが辞書に登録してあればアクセントが生成できる。

使用処理系

K-PROLOG (SONY・NWS-830)

参考文献

- (1) 金田一春彦監修 秋永一枝編
 「明解日本語アクセント辞典」第二版
 三省堂、1988