

結合価パターンを用いた音声認識 Speech Recognition using Valency Pattern

村上仁一* 佐伯 亮 徳久 雅人 池原 悟
Jin'ichi Murakami Ryou Saeki Masato Tokuhisa Satoru Ikehara

1 はじめに

言語モデルには、大きくわけて文法規則的な統語情報と、確率的な統計情報がある。現在の音声認識では、一般には言語モデルに単語 N -gram が用いられている [1]。しかし、単語 N -gram 言語モデルは、言語の隣接統計情報であり統語的な情報を含んでいない。

音声認識のための言語モデルにおいて、統語的な情報として文脈自由文法を用いたもの [2] や、ネットワーク文法を用いたもの [3] がある。しかし、これらのシステムで使用された言語モデルは、語彙やルール数が少ないため、言語の統語情報がどれくらい有効であるか不明であった。

本研究では、統語的な情報として結合価パターンを使用する。結合価パターンとは格要素と用言に対する制約効果を記述したものである。この結合パターンとして岩波出版から出版されている日本語語彙大系 [4] を利用する。この語彙大系は、単語総数 40 万、単語の結合価パターン数約 14000、名詞の分類 12 段の 2710 種類に分類されている規模の大きな辞書である。

本報告では、始めに言語モデルに bigram を用いた音声認識プログラムを使用して複数の候補文を出力する。次に、結合価パターンを用いて候補文を選出する。最後に、結合価パターンを用いることによる累積文正解率の変化を調査する。

2 結合価パターンによる候補文の削減

2.1 結合価パターン

結合価パターンは、名詞の持つ意味属性により用言と格要素(名詞+格助詞)の意味的共起関係を体系化したものである。結合価パターンの使用により名詞と動詞の意味的な誤りが検出可能である。本研究では、結合価パターンとして日本語語彙大系 [1] の「構文意味辞書」のパターンを使用する。表 1 に結合価パターンの例を示す。結合価パターンでは、格要素として一般名詞意味属性(図 1)が、用言として動詞が用いられている。

表 1: 結合価パターンの例 (括弧内は意味属性)

$N1(3 \text{ 主体})$ が $N2(1236 \text{ 人間活動 } 2064 \text{ 変動})$ に 動く
$N1(962 \text{ 機械})$ が $N2(2335 \text{ 電気})$ で 動く

2.2 一般名詞意味属性体系

一般名詞意味属性とは、対象と単語の関係を対象の捉え方に注目して分類したものである。一般名詞意味属性では約 40 万語の単語を約 2700 種類の意味属性に分類されている。日本語のシソーラスとしては、規模・語彙数・意味分解能において非常に規模の大きな辞書である。

図 1 に意味属性体系の構造の一部を記す。この意味属性体系には、下位の属性は上位の属性に付随するといった特性が

*鳥取大学工学部知能情報学科

ある。例えば、「地域」という意味属性には「土地」、「世界」などの属性が含まれているが、「地域」の上位属性の「場所」には「地域」だけでなく、「施設」、「自然」といった属性に含まれる単語すべてが内包されている。

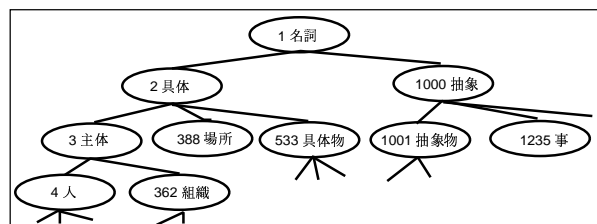


図 1: 一般意味属性体系の一部

2.3 結合価パターンによる候補文の選択

音声入力文に「私は彼の高校を受験する。」が入力されると仮定する。そして、言語モデルに単語 bigram を用いた音声認識プログラムにおいて、以下の 2 文が出力されたと仮定する。

第 1 位候補文 私からの高校と受験する。

第 2 位候補文 私は彼の高校を受験する。

2 つの候補文で用いられている用言「受験する」の結合価パターンを表 3 に示す。(括弧内は名詞意味属性)

表 2: 「受験する」の結合価パターン

①	$N1(4 \text{ 人})$ [は/が] $N2(362 \text{ 組織})$ を 受験する
②	$N1(4 \text{ 人})$ [は/が] $N2(1426 \text{ 試験})$ を 受験する
③	$N1(4 \text{ 人})$ [は/が] 受験する

上記のパターン内で、第 1 位候補文を満たすパターンは存在しないので、第 1 位候補文は非文と判断し削除する。第 2 位候補文を満たすパターン①が存在するため、第 2 位候補文を第 1 位候補文とする。

3 実験

3.1 実験手順

本論文における実験手順を以下に示す。

1. 音声入力文として単文を収録する。
 2. 単語 bigram を利用した音声認識プログラムを用いて、音声入力文に対して第 16 位までの候補文を出力する。
 3. 各位候補文に対して結合価パターンを用いて候補選択を行う。
 4. 累積文正解率を調査する。
- 以下、各々について説明する。

3.2 音声入力文

音声入力文として、IPAL 日本語名詞、日本語動詞の例文 [7] に記載されている単文を利用する。IPAL には日本語において使用頻度が高い文が収集されている。実験は男性話者 1 名により 70 文で行う。

3.3 音声認識プログラム

音声認識プログラムには、当研究室にある tree-trace 型の音声認識プログラムを用いる。音素 hmm の学習は、HTK [2] を用いる。ATR-Aset5240 単語男性話者 10 名の発話データを用いて不特定話者モデルを作成し、音声入力文と同一話者の ATR-Bset と同一の発話文 503 文を用いて連結学習し音素モデルを作成する。単語 bigram は毎日新聞 1993 年度 1 年分と音声入力文 70 文を用いて作成する。

3.4 評価方法

本研究は、結合価パターン適用前と適用後の累積文正解率を比べ評価する。文正解率は、評価文と正解文の単語の並びが同一のものを正解とする。また、評価は正解文が候補文の第 16 位までに含まれるものを用いる。

3.5 実験結果

3.5.1 累積文正解率

実験の結果、結合価パターンを適用することで 70 文中 23 文において累積文正解率が向上した。また第 1 位文正解率は 50% (35 文) から 67% (47 文) に向上した。

表 3 に結合価パターンを用いる前と、結合価パターンを用いた後の場合の累積文正解率を示す。

表 3: 累積文正解率

順位	結合価パターン無	結合価パターン有
第 1 位	50% (35/70)	67% (47/70)
~ 2	71% (50/70)	81% (57/70)
~ 4	84% (59/70)	93% (65/70)
~ 8	94% (66/70)	99% (69/70)

3.5.2 出力例

表 4 に、音声入力文が「母が私に新しい本をくれた」の候補文を示す。

表 4: 出力例

- | | |
|---|---------------|
| ① | 歯が私に新しいのんをくれた |
| ② | 歯が私のに新しいほど遅れた |
| ③ | 母が私に新しい方をくれた |
| ④ | 母が私に新しい本をくれた |

第 1 位候補文は用言「くれる」の結合価パターンが全て当てはまらないため削除される。第 2 位候補文は用言「遅れる」の結合価パターンが出力文に全く当てはまらないため削除される。第 3 位候補文、第 4 位候補文は用言「くれる」のパターンが存在する。

したがって、結合価パターンを適用することで、第 1 位、2 位候補文は削除され、正解文である第 4 位候補文が第 2 位候補に昇格する。

4 考察

4.1 文末の名詞

実験結果から結合価パターンを適用することで 70 文中 23 文において累積文正解率が向上した。この 23 文を調査した

ところ、8 文が文末に名詞が来ることで累積文正解率が向上した。例を以下に示す。

< 音声入力文 > 彼女は英語を始めた

第 1 位候補文 彼女は英国は自宅

このような例は、結合価パターンを用いなくても、文末が用言で終わる規則を用いることで、容易に削除できる。

本研究で利用した音声認識プログラムの言語モデルには単語 bigram を用い、単語 bigram の計算には新聞記事 1 年分 [6] を利用した。新聞記事には見出が多く含まれる。したがって文が名詞で終わる場合も多い。そのため、音声認識の出力文として名詞で終る文がある。音声入力文が単文である場合、単語 bigram を計算するテキストデータも単文が必要と考えている。

4.2 未知語

出力結果を調査すると、候補文の中に未知語としてパターンがあたるものが存在した。表 4 の第 1 位候補文がこの例に相当する。第 1 位候補文中の「のん」は意味的に不明な語句で未知語と見なせる。しかし、結合価パターンはこのような未知語が存在していてもパターンを適用できるため非文にならない。

今後、未知語を考慮した結合価パターンを適用することにより、累積文正解率はさらに向上すると考えられる。

4.3 複雑な文への応用

今回の実験データの中において、結合価パターンに当てはまらない音声入力文は存在しなかった。しかし、結合価パターンは全てのパターンを網羅していない。そのため、音声入力文の用言に対するパターンがない場合は正解文であっても削除してしまう場合があると考えられる。

また今回の実験は、単純な単文であるため、結合価パターンの有効性を示されたと考えられる。より複雑な重文や複文の場合の結合価パターンの有効性を調査する必要がある。

5 まとめ

本研究では、統語的な情報である結合価パターンが認識精度向上にどの程度役立つかを検証した。実験の結果、結合価パターンを適用することで 70 文中 23 文において累積文正解率が向上し、第 1 位正解率は 50% から 67% に向上した。これらの結果、結合価パターンは認識精度向上に有効であることが示された。

今後の課題として、より複雑な重文や複文の場合の結合価パターンの有効性を調査する必要がある。

参考文献

- [1] 村上 仁一, 荒木 哲朗, 池原 悟, “日本文音節入力に対して 2 重マルコフ連鎖モデルを用いた漢字仮名交じり文節候補の抽出精度”, 電子通信学会論文誌, Vol. j75-D-2, No. 1, pp11-20, (Jan 1992).
- [2] 北 研二, 川端 豪, 斎藤 博昭, “HMM 音韻認識と拡張 LR 構文解析法を用いた連続音声認識”, 情報処理学会論文誌, Vol. 31, No. 3, pp. 472-480. 1990.
- [3] Lowerre, B.T, “The HARPY Speech Recognition System”, PHD Thesis CMU, April, 1976.
- [4] 池原 悟, 宮崎 正弘, 白井 諭, 他, “日本語語彙大系”, 岩波書店 (1997).
- [5] “Hidden Markov Model Toolkit”, <http://htk.eng.cam.ac.uk>.
- [6] “毎日新聞記事 (1993.1~1993.12)”, 毎日新聞社.
- [7] “計算機用日本語名詞辞書 IPAL 解説編”, 情報処理振興事業協会技術センター, 1996. [<http://www.ipa.go.jp/STC/NIHONGO/IPAL/ipal.html>]