

5E-8

ボルツマンマシンによる 文節ラティスの係り受け解析

村瀬 功 , 中川聖一

豊橋技術科学大学 情報工学系

1. はじめに

文節単位に発声された音声でも、その認識は非常に難しく、通常は複数候補を出力し、言語処理によって最適な候補列を求める方法をとる。そのための構文解析アルゴリズムが種々提案されている^[1]。

本稿では、上述の構文解析の問題を一種の最適化問題^[2]とみなし、ボルツマンマシンを適用した結果について述べる。

2. 係り受け解析システム

2.1 ネットワークの構成

三層構造のネットワークを取り入れた。第一層は入力文節の候補をユニットに対応させた層、第二層は各文節各候補間のリンクを構成する層、第三層は係り受け規則の中の非交差の条件を取り入れる層、である。よって、第一層と第二層、第二層と第三層、が結合する。この様子を図1に示す。

入力文の文節数が n で、各文節の候補数が m であれば、第一層のユニット数は nm 、第二層のユニット数は $m^2 \cdot \sum_{i=1}^{n-1} i$ ($= m^2 \cdot n \cdot (n-1) / 2$)、第三層のユニット数は $\sum_{i=1}^{n-1} i$ ($= n \cdot (n-1) / 2$) である。

このネットワークの状態は、次式より更新される。

$$net_i = istr \cdot (\sum_j w_{ij} a_j + bias_i) + estr \cdot input_i \quad (1)$$

$$P_i = \frac{1}{1 + e^{-net_i/T}} \quad (2)$$

ここで、 w_{ij} はユニット i, j 間の重み、 net_i はユニット i への総入力、 $input_i$ は l サイクル前の net_i 、 $istr$ と $estr$ は現在の入力と l サイクル前の入力とのバランスをとるためのパラメータ、 P_i はユニット i の状態 l をとる確率である。

(1)(2)式で温度 T を下げながらネットワークのエネルギーを低くし、極値状態に落ち着かせる。

2.2 重み計算で用いるデータ

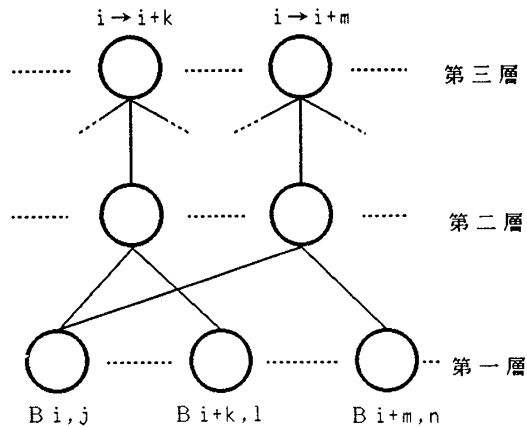
音声認識装置より出力される文節候補（先頭から n 番目に発声した文節の文節番号を n と呼ぶことにする）とそのスコアにより次に示すデータを計算する。

各文節各候補間の組合せから、係り受けデータ（係り受けが成立するならば5、成立してなかつ意味的にも通るならば10、成立しないならば-10）と、各文節当りのスコアを1~5の間に正規化したものを新しいスコアとする。

2.3 文節の分類

文節中から自立語を取り出し自立語の辞書を参照してその品詞、及び活用するものは活用形を決定する。

次に、文節中に付属語がある場合は付属語辞書を参照し、その品詞を決定する。この手続きからの出力より、その文節に係る側と受け側での両方の場合で表1に示すように分類する。これをすべての二文節の組合せにおいて係り受けが成立するかしないかを決定し、係り受けデータとする。



$B_{i,j}$ は、 i 番目の文節 j 番目の候補を示す。

○ はユニット、— はリンクを示す。

図1 ネットワークのリンク

表1 文節の分類

係り側の文節分類	受け側の文節条件
a 体言 + ϕ	1 体言を含む
b 条件・時を表す名詞	2 用言を含む
c 体言 + の	3 終止の用言、それに並立する用言を含む
d 体言 + 並立助詞	4 動詞を含む
e 体言 + 各助詞	5 形容詞を含む
f 体言 + 係助詞	
g 副詞	
h 連体詞	係り受け可能な組合せ
i 文の接続詞	
j 並立の接続詞	a→1, b→3, c→1
k 形容詞節 (連体)	d→1, e→2, f→3
l 形容詞 (連用)	g→2, h→1, i→3
m 動詞節 (連体)	j→1, k→1, l→4
n 動詞節 (連用)	l→5, m→1, n→4
o 動詞 + 接続助詞	o→4

2. 4 重み計算

排他的な重みをつけるユニット間は、第一層は各文節候補を示すユニット間（同一文節番号の候補文節集合のうち一つだけが活性化されるため）、第二層は同文節番号からの係り受けのリンクを示すユニット間（各文節は後続の一つの文節のみに係るため）、そして第三層は係り受けで交差しているユニット間である。この非交差の関係は次の手順で求めた。

第三層のユニットを X, Y とする。

$$flg1 = X \text{ の係り側の番号} - Y \text{ の係り側の番号} \quad (3)$$

$$flg2 = X \text{ の受け側の番号} - Y \text{ の受け側の番号} \quad (4)$$

もし flg1 と flg2 の符号が等しいならば、

$$flg3 = Y \text{ の受け側の番号} - Z \text{ の係り側の番号} \quad (5)$$

flg3 が正ならば、交差している。

第一層と第二層のユニット間では、次のように重みを設定した。第一層のユニット $B_{i,j}$ と $B_{i+k,1}$ 間のリンクを示す第二層のユニットとの重みは、その文節間の係り受けデータを Ka とすると、

$$A \cdot Ka - B \cdot (B_{i,j} \text{ のスコア} + B_{i+k,1} \text{ のスコア}) \quad (6)$$

とする。ただし A と B は定数。

第二層と第三層のユニット間の重みは、(6) 式の係り受けデータ Ka とスコアを用い、

$$C \cdot Ka - D \cdot (B_{i,j} \text{ のスコア} + B_{i+k,1} \text{ のスコア}) \quad (7)$$

とした。ただし C と D は定数。

また、ユニットに与える静的なバイアスは、第一層の各ユニットに $-E \cdot \text{スコア}$ 、第二、三層に全体を抑制するための負の値とした。ただし E は定数。

3. 解析結果

入力される文節候補間の全ての組合せにおいて係り受け解析を行い最適な文を全探索により計算して、このシステムからの結果と比較した。

このシステムに疑似的なデータを入力して解析を行

ったところ、ほぼ正しい解析結果が得られた。完全ではないのは、ボルツマンマシンが確率で各ユニットの状態を更新するため、いつでもまったく同じ解が得られるとは限らないからである。

例の結果を図2に示す。図2より、入力文節の各候補の中から「彼女の 白い帽子が 空に 飛んだ。」という候補が選出され、なおかつ「彼女の」と「白い」から「帽子が」に、「帽子が」と「空に」から「飛んだ」に係るという係り受けが、解析されたことが読み取れる。

次に、実際の音声認識装置から得られる文節候補列を入力して解析を行った。得られた解析結果は、上述の手法から得られた最適な文と比べて一致した結果が10%程であった。しかし残りも1つの文節の中の選ばれる候補が違う程度で、最適解とほぼ同一であった。

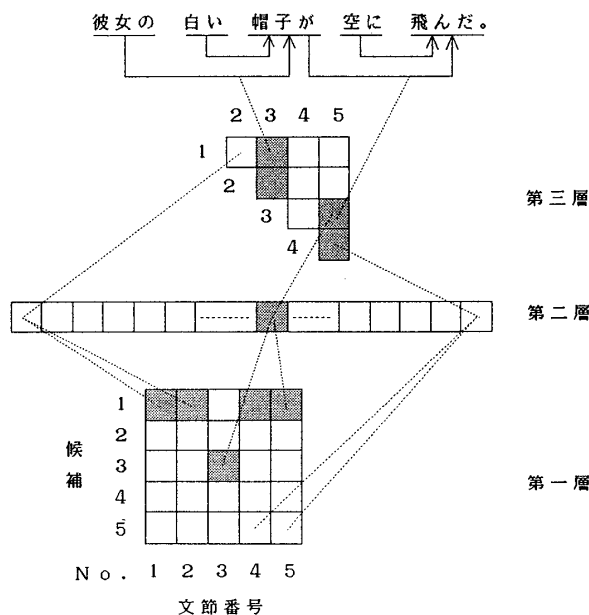
以前に示した通り、使用した係り受けデータは、意味情報を含んでおらず5または-10の値のみであるためにエネルギーの差があまりなく、最適解のときのネットワーク状態にたどり着けないために、文認識率が低かったと思われる。

4. おわりに

本稿では、係り受け解析の一手法として、問題を最適化問題に置換してボルツマンマシンで解く方法について述べた。しかし、このシステムも動き始めたばかりで完全なものとは言えない。3の結果のところでも触れたが、現在係り受けデータの中に意味情報が入っていないので、最適解に達することができないと思われる。したがって、これをデータの中に入れるためのソフトウェアの開発が必要である。まだ従来のアルゴリズム的な手法と比べて能力は低い、大規模な音声理解や、文としては間違っても大部分の文節は正しく認識できるから音声ワープロ・ディクテーションの場合などに、有効になると考えられる。

参考文献

- [1] 中川聖一「確率モデルによる音声認識」
(電子情報通信学会)
- [2] J. Hopfield and D. W. Tank : "Neural" Computation of Decisions in Optimization Problems,
Biol. Cybern. 52, pp141-152(1985)



(a) ネットワークの状態

	1	彼女の	白い	帽子が	空に	飛んだ
優候	2	彼女が	広い	胞子が	虎に	読んだ
先補	3	青の	におい	帽子が	きれいに	すてきだ
順番	4	彼を	きれい	牛が	上に	静かだ
位号	5	彼に	黒い	帽子を	山に	走った
		No. 1	No. 2	No. 3	No. 4	No. 5

文節番号

(b) 入力文節列

図2 疑似的なデータによる解析結果