ニューラルネットワークを用いた日本語解析の試み†

高橋直人†† 板橋秀一†††

日本語の形態素解析および係り受け解析は，一種の最適化問題，すなわち制約条件を最もよく満たすようにコスト関数の値を最小とする問題，と考えることが可能であり，それゆえに相互結合型ニューラルネットで解くことが可能である．この場合，上記のコスト関数はニューラルネットワークのエネルギ関数として表現される．本稿では日本語の形態素解析用と係り受け解析用の2種類のニューラルネットの構成，および計算機上でシミュレーションを行った結果について述べる．形態素解析用ネットワークは，分かち書きされていない日本語文を単語単位に切り分ける．複数通りの切り分け方が可能な場合は，文節数最少法と同等の評価関数に基づいて最尤候補を決定し，それを出力とする．係り受け解析ネットワークは文節列を入力とし，それら文節間の係り受け関係を決定する．この際に文法的制約のみでなく，文節間の意味的制約も考慮されるような配慮がなされている．意味的制約は学習用例文中の文節間係り受け頻度に基づいて決定された．どちらの解析においても，パラメータを調整することによって約95%の成功率を得ることができた．入力文の長さが増加しても，ネットワークが収束するまでのステップ数はあまり増加しなかった．

1. はじめに

自然言語文に含まれている曖昧性を解消するためには，統語的解析のみならず，意味的・文脈的・語用論的解析が必要である．人間はこの困難な問題を極めて効率良く行っている．一方，自然言語に関する知識やヒューリスティックスをいかに表現するかという問題にはまだ明確な回答が得られておらず，したがって，計算機で自然言語を解析する場合にはしらみつぶし的な方法をとらざるを得ないことが多い．すなわち，文法的に可能なすべての解析木を生成し，それらのなかから与えられた評価関数を最大にするものを選択するのである．評価関数には意味的・文脈的・語用論的妥当性を反映するように構成されたものが選ばれる．このような文法的に可能なすべての解析木の妥当性を評価する方法には，大きな計算コストが必要とされる．

解析対象言語が日本語の場合，形態素解析とは文を単語単位に分割することを指す．日本語文では，通常，単語と単語の間に空白を置かないので，同一の文を単語単位に分割する方法が複数通り存在し得る．特に文が仮名のみあるいはローマ字のみで書かれている場合は分割方法が多数存在するため，組合せ的爆発が生じ得る．

一方，日本語の統語的構造は，文節間の係り受け関係で表現することができる．一般的に，文法的制約のみでは係り受け関係を一意に決定することはできない．ある2文節間の係り受け関係の成立しやすさは，それら2文節に含まれている単語間の意味的関係によって影響を受ける．

形態素解析・係り受け解析の両方とも一種の最適化問題，すなわち与えられた制約条件を最もよく満たすようにコスト関数の値を最小とする問題，と考えることができる．この場合，上記のコスト関数はニューラルネットワークのエネルギ関数として表現される．HopfieldとTankは，ある種の最適化問題は相互結合型のニューラルネットを用いることで極めて効率良く解けることを示した[2),3)]．ニューラルネットを自然言語解析に応用した例は幾つか見られるが[4)~6),8),9),14)]，筆者らの研究[10)~13)]以外で日本語解析を最適化問題とみなし，それを相互結合型のニューラルネットで解いたという例は多くない．自然言語解析をニューラルネット上で実行する方法が確立されれば，近年開発されるようになったニューラルネット・チップと組み合わせることで極めて高速の自然言語解析エンジンの作成が可能となろう．

本稿では日本語文を相互結合型ニューラルネットで解析する方法，およびそのニューラルネットを計算機上でシミュレートした実験結果について述べる．以下，第2章で相互結合型ニューラルネットの一般的な振舞いおよびエネルギ関数について簡単に述べる．第3章では形態素解析に用いたネットワークとその実験結果について，また第4章では係り受け解析に用いたネットワークとその実験結果についてそれぞれ説明する．最後に第5章でまとめと今後の展望について述べる．

† An Experiment of Japanese Sentence Analysis with Neural Networks by NAOTO TAKAHASHI (Doctoral Program in Engineering, University of Tsukuba) and SHUICHI ITAHASHI (Institute of Information Sciences and Electronics, University of Tsukuba).

†† 筑波大学工学研究科
††† 筑波大学電子・情報工学系

2. 相互結合型ニューラルネット

相互結合型ネットワーク内の各ユニット u_i が

$$i_i=\sum_{j=1}^{n} w_{ij}u_j-\theta_i \qquad (1)$$

に対し

$$i_i=\begin{cases} >0 \text{ then } u_i=1 \\ =0 \text{ then 変化せず} \\ <0 \text{ then } u_i=0 \end{cases} \qquad (2)$$

に従って，非同期に状態遷移を行うものとする．ここで w_{ij} は u_i から u_j への結合の強さ，θ_i は u_i のしきい値である．この時 w が，$w_{ij}=w_{ji}$ かつ $w_{ii}=0$ を満たすならば，このネットワークの**エネルギ関数**

$$E=-\frac{1}{2}\sum_{i=1}^{n}\sum_{j\neq i} w_{ij}u_iu_j+\sum_{i=1}^{n}\theta_iu_i \qquad (3)$$

は時間とともに単調減少し，定常状態に収束することが証明できる[1),2)]．したがって相互結合型ネットワークを用いてある問題を解くためには，そのネットワークが解に達したときにエネルギが最低になるように結合の強さおよびしきい値を決定すればよい．

ここで問題となるのは式(3)がエネルギの最小値ではなくて，極小値に収束する可能性のあることである．この場合は最適解を得ることができない．これを避けるための方法の一つに，ネットワークの動作を確率的にすることがあげられる．Boltzmann Machine は式(2)のネットワークの動作を確率的にしたもので，各ユニットは次のような確率でその状態を1にする．

$$p(u_i=1)=\frac{1}{1+\exp(-i_i/T)}. \qquad (4)$$

式(4)中の T はネットワークの温度と呼ばれる正の数である．T が0の極限で，Boltzmann Machine は式(2)のネットワークと同じ動作をする．エネルギの差を強調するためには温度が低い方がよいが，その場合はエネルギの局所的なくぼみにはまりやすくなり，最小値付近に到達するまで長い時間がかかる．これを避けるために，初めのうちは温度を高くしておき，状態変化をさせながら徐々に温度を下げる**疑似焼きなまし**の技術を使う．焼きなましのスケジュールは経験的によいものを採用する．

3. 形態素解析

形態素解析ネットワークの各ユニットは，文の構成要素となり得る各単語に対応する．例として「かれがくるまでまつ」という文字列からなる文があり，この

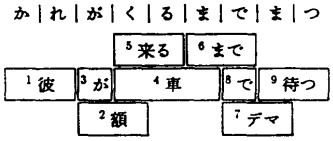

図 1 入力文を単語列に分解する方法の例
Fig. 1 An example of dividing a sentence into words.

文に含まれ得る単語としては以下のものがあるとしよう（**図1**）．

u_1：彼(代名詞)　u_4：車(名詞)　u_7：デマ(名詞)
u_2：額(名詞)　u_5：来る(動詞)　u_8：で(助詞)
u_3：が(助詞)　u_6：まで(助詞)　u_9：待つ(動詞)

この場合は互いに異なる9単語が存在するので，ネットワーク内のユニット数は u_1〜u_9 の9となる．ここでの目的はネットワークの活動が収束した状態で値1をとっているユニットに対応する単語のみが文の構成要素となるようにユニット間の重みおよび各ユニットのしきい値を調節することである．簡単のため，以下では「ユニット u_i で表される単語」を単に「単語 u_i」と表記する．

文を正しい単語列に分割するためには，以下の制約を満たすようなエネルギ関数を作成する必要がある．

1. 入力文中の各文字が，ちょうど1回ずつ用いられるような単語の組合せが選択されたときに最小値をとる
2. 隣接する単語の組合せがすべて文法的に接続可能な場合に最小値をとる
3. 選択された単語列に含まれる自立語の数が少ないほど小さな値をとる

最初の2条件は文法的なものである．最後の条件は，文節数最少法に類似した制限条件を与えるためのものである．上の 1〜3 を表す関数をそれぞれ E_C, E_G, E_I とした場合，形態素解析ネットワークのエネルギ関数はこれらの一次結合

$$E=pE_C+qE_G+rE_I \qquad (5)$$

となる．p, q, r は正の定数である．以下では E_C, E_G, E_I について少し詳しく説明する．

3.1 文字列の分解

n をネットワーク中のユニットの総数，N を入力文の文字数とする．このとき

$$C_{ik}=\begin{cases} 1 & \text{単語 } u_i \text{ が入力文中で } k \\ & \text{番目の文字を含む場合} \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

と定義される $n\times N$ 行列 C を考える．図1の例だと

$$C=\begin{bmatrix}1&1&0&0&0&0&0&0&0\\0&0&1&1&0&0&0&0&0\\0&0&1&0&0&0&0&0&0\\0&0&0&1&1&1&0&0&0\\0&0&0&1&1&0&0&0&0\\0&0&0&0&0&1&1&0&0\\0&0&0&0&0&0&1&1&0\\0&0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1&1\end{bmatrix} \quad (7)$$

となる．この C を用いると，入力文中の各文字がちょうど1回ずつ使用されるような単語の組合せが選択された時に最小値0をとる関数は

$$E_C=\sum_{k=1}^{N}\left(\sum_{i=1}^{n}C_{ik}u_i-1\right)^2 \quad (8)$$

と表すことができる．

3.2 文法的接続可能性

式(8)を用いれば，選択された単語を連結して得られる文字列を入力文と同じものにすることができる．しかし，字面が同一であるからといって，選択された単語列が必ずしも文法的に正しい接続になっているとは限らない．

すべての隣接した単語間の接続が文法的に正しい場合に最小値0をとる式は，以下のように記述することができる．

$$E_G=\sum_{i=1}^{n}\sum_{j\neq i}G_{ij}u_iu_j \quad (9)$$

ただし

$$G_{ij}=\begin{cases}1 & \text{単語 } u_i \text{ と単語 } u_j \text{ が入力文中で隣接しており，かつ文法的に接続可能でない場合}\\ 0 & \text{otherwise}\end{cases} \quad (10)$$

である．

例えば「代名詞＋名詞」という接続は文法的でないとする．すると図1の u_1「彼」と u_2「額」は，文中で隣接しておりかつ文法的に接続可能でないので $G_{12}=G_{21}=1$ となる．

3.3 自立語数

分ち書きされていない日本語文を単語に分解する際の基準はいろいろ提案されているが，ここでは文節数最少法に類似した**自立語数最少法**を採用する．すなわち，ある文を分解する方法が複数通りある場合は，そのなかでもっとも自立語数の少ない分解方法を選択するものとする．今回の実験では1文節に含まれる自立語は一つだけとしたので，複合名詞が現れる場合を除けば自立語数最少法と文節数最少法は同じ結果を与える．

まず I_i を次のように定義する．

$$I_i=\begin{cases}1 & \text{単語 } u_i \text{ が自立語の場合}\\ 0 & \text{その他．}\end{cases} \quad (11)$$

この I_i を用いると，式

$$E_I=\sum_{i=1}^{n}I_iu_i \quad (12)$$

は，選択されたユニット中に自立語が少ないときほど小さな値をとるようになる．

3.4 形態素解析ネットワークのエネルギ関数

以上で E_C, E_G, E_I が決定された．これより式(5)と式(3)を比較すると，形態素解析ネットワークにおけるユニット間の結合の強さ w_{ij} および各ユニットのしきい値 θ_i は以下のようになることがわかる．

$$\begin{cases}w_{ij}=-2\left(p\sum_{k=1}^{N}C_{ik}C_{jk}+qG_{ij}\right)\\ \theta_i=-p\sum_{k=1}^{N}C_{ik}+rI_i.\end{cases} \quad (13)$$

3.5 実　　験

本節では，上で述べたネットワークのシミュレーションプログラムを作成し，実際に文を解析させてみた結果について述べる．計算機は Sun-4/330，言語は Allegro Common Lisp を用いた．入力文は「新明解国語辞典第二版（磁気テープ版）」[15] の語義説明文のうち，植物に関する百科事典的記述を抜き出し，平仮名べた書きに改めたものを用いた．平仮名べた書きに改めたのは同音異義語が含まれる可能性を意図的に高くし，それに対してネットワークがどの程度の曖昧性解消能力を持つか見るためである．

まず予備実験として本実験における解析対象文の約1/4にあたる50文を用い，ネットワーク中の各定数をさまざまに変化させて収束の様子を観察した．その結果，式(13)の係数としては

$$p=4,\ q=2,\ r=3. \quad (14)$$

という値の組合せが比較的良い結果を与えることがわかった．

また，時刻 t のときのネットワークの温度 $T(t)$ は次式で与えられるものとした．

$$T(t)=\frac{T_0}{1+t/\tau} \quad (15)$$

ただし

$$T_0=5,\ \tau=10 \quad (16)$$

である．この T_0 および τ も予備実験によって経験的に得られたものである．

なお，ユニットの初期状態はすべて0とし，状態遷移は非同期に行うこととした．

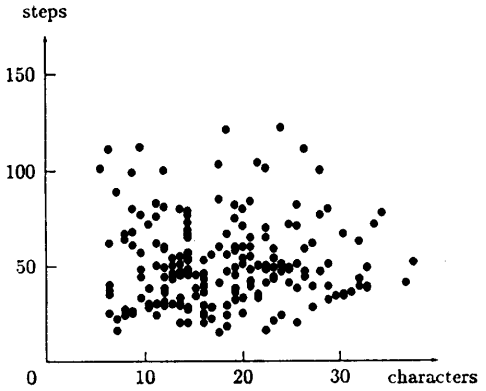

図 2 入力文中の文字数と形態素解析ネットワークが収束するまでのステップ数との関係
Fig. 2 Relation between the number of characters contained in input sentences and the steps required to make the network converge.

予備実験で得られた各定数を用いて本実験を行った．本実験の結果を**図 2** に示す．ユニットの状態遷移は非同期かつランダムに生じるので，同一条件で追試を行った場合でも実験結果の細部は異なる可能性がある．ここではすべてのユニットが平均して1回ずつ発火するのに必要な時間を1ステップとしている．なお，約 200 の入力文に対する解析成功率は約 95% であった．

一般に，入力文が長くなるとその中に含まれ得る単語数（すなわちネットワーク内のユニット数）は急激に増加する．しかし図2を見る限り，そのような場合にもネットワークが収束するまでのステップ数はさほど影響を受けないように見える．このように解析時間が入力文の長さにあまり影響を受けないという点は，逐次型の自然言語解析システムにはない大きな特徴といえよう．

実験を行った結果，解析途中で一度エネルギの極小状態に落ちいった場合は，ネットワークが収束するまでのステップ数が増加する傾向にあることがわかった．このことは最終的に正しい解が得られた場合にもそうでない場合にも成り立つ．エネルギのくぼみ（極小状態）から脱出するためには一度エネルギが増加する方向へ状態遷移を行う必要がある．くぼみが深い場合，すなわちエネルギ障壁が高い場合はそのような状態遷移を行う確率は小さく，したがってそのくぼみの周辺を長時間「うろつく」ことになるのがこの原因と考えられる．図2中でステップ数が 100 を越えているものは，ほとんどが一度エネルギ極小状態にとらえられたものである．

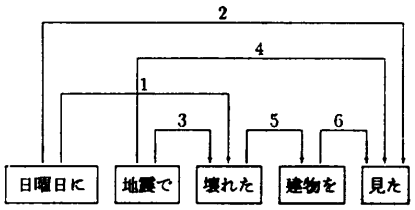

図 3 文節間係り受け関係の例 (1)
Fig. 3 An example of modification relation of a sentence (1).

4. 係り受け解析

文節間係り受けの条件としては以下の四つが考えられる[7]．

R1. 最後の文節以外の文節は，自分よりも後方にある文節のどれか一つに係る．
R2. 係り受け関係は互いに交差しない．
R3. 文節中の語の意味によって係り受けの成立しやすさが影響を受ける．
R4. 位置的に近い文節間ほど係り受けが成立しやすい．

R1 より係り受けの総数は構文にかかわらず

(文節数) − 1 (17)

となることがわかる．R2 は非交差条件として有名である．R3 は，例えば**図 3** のような文においては係り受け関係4よりも係り受け関係3の方が成立しやすい，ということを表している．R4 は運用論的な制約であり，統語的あるいは意味的なものではない．

係り受け解析ネットワークがなすべきことは，すべての可能な係り受け関係の中から必要十分な数の関係のみを選択することである．最終的に選択された係り受け関係は上の R1 から R4 を満たしている必要がある．

係り受け解析ネットワークにおいて各ユニットに割り当てられるのは文節そのものではなく，成立する可能性のある**係り受け関係**である．図3の文の場合には全部で6個のユニットが存在することになる．ネットワークがエネルギ最小状態に収束したとき，値1をとっているユニットが最終的に選択された係り受け関係を表しているものとする．例えば図3の文には解となり得る状態が全部で3通りあるが，そのなかの一つ

unit	u_1	u_2	u_3	u_4	u_5	u_6
value	1	0	1	0	1	1

は，「日曜日に」と「地震で」とが「壊れた」に，「壊れた」が「建物を」に，そして「建物を」が「見た」

に係っている状態を表している*.

一方，状態

unit	u_1	u_2	u_3	u_4	u_5	u_6
value	1	0	0	1	1	1

は係り受け関係1と係り受け関係4とが交差するため，解とはなり得ない．

R1～R4を表す関数をそれぞれ E_1～E_4 としたとき，係り受け解析ネットワークのエネルギ関数は

$$E=aE_1+bE_2+cE_3+dE_4 \tag{18}$$

（ただし a, b, c, d は正の定数）で表される．以下に E_1～E_4 の定義を与える．

4.1 係り受け関係の数

最終的に値1をとるべきユニットの数は，式(17)より（文節数−1）となる．n をユニットの総数，m を最終的に1となるべきユニットの数とすると，望ましい個数のユニットが1となったときに最小値をとる関数は以下のように表すことができる．

$$E_1=\left(\sum_{i=1}^{n} u_i-m\right)^2. \tag{19}$$

4.2 相互排他性

各文節の係り先はただ一つである．したがって，係り元が同一の文節であるような係り受け関係が複数個ある場合は，それらのうちの一つだけが値1をとるようにしなければならない．また，非交差条件より，互いに交差するような係り受け関係の両方が1となることは許されない．

このような係り受け関係間の相互排他性を表すために，**相互排他行列** X を導入する．X は $n\times n$ 行列（n は成立する可能性のある係り受け関係の総数）であり，その (i, j) 要素の値は

$$X_{ij}=\begin{cases}1 & \text{係り受け関係 } i \text{ と係り受け関係 } j \text{ が相互排他的である場合}\\ 0 & \text{otherwise}\end{cases}$$

である．相互排他行列は明らかに対称行列になる．例えば図3の文に対する相互排他行列は

$$X=\begin{bmatrix}0&1&0&1&0&0\\1&0&0&0&0&0\\0&0&0&1&0&0\\1&0&1&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{bmatrix}$$

となる．

この相互排他行列 X を用いると，値1をとるユニット中のどの二つも相互排他的でない場合に最小値をとる式は

$$E_2=\sum_{i=1}^{n}\sum_{j=1}^{n} X_{ij}u_iu_j \tag{20}$$

と表せる．

4.3 文節間の意味的関係

ある2文節間に係り受け関係が成立するか否かは，係り側の文節と受け側の文節の文法的関係および意味的関係に依存する．係り受け解析ネットワークにおける各ユニットは，もともと文法的に成立し得る係り受け関係のみを表しているので，ここでは文節間の意味的関係を係り受けの成立しやすさに反映させる手段を考えればよい．i を係り受け関係とし，Y_i を i という係り受け関係が意味的に成立しやすいほど小さい値をとるものとする．図3の例では「地震で見た」という係り受けよりも「地震で壊れた」という係り受けの方が意味的尤度が高いと考えられるので，$Y_3<Y_4$ となるであろう．

この Y_i を用いると，全体として意味的尤度が高い係り受け関係が多く選ばれるほど小さな値をとるような式は

$$E_3=\sum_{i=1}^{n} Y_iu_i \tag{21}$$

と表すことができる．

この Y_i をいかに設定するかは重要な問題である．今回の実験では約500の学習用例文から得られた**単語間係り受け共起頻度**に，式(22)で示す操作を施したものをこの Y_i として用いた．ただし単語間係り受け共起頻度とは，

> 文節Aが文節Bに係っているとき，文節A中の自立語と文節B中の自立語の単語間係り受け共起頻度を1増やす

という操作を学習用例文全体にわたって施した結果である．

$$Y_i=\begin{cases}0 \cdots \xi_i=0 \text{ のとき}\\ -1 \cdots \xi_{\max}=\xi_{\min} \text{ のとき}\\ -\dfrac{1}{2}\left(\dfrac{\xi_i-\xi_{\min}}{\xi_{\max}-\xi_{\min}}+1\right)\\ \quad \cdots \text{その他.}\end{cases} \tag{22}$$

ただし ξ_i は i 番目のユニットによって示される係り受け関係の単語間係り受け共起頻度であり，

$\xi_{\max}=\max\{\xi_i \mid i=1, \cdots, n\}$

$\xi_{\min}=\min\{\xi_i>0 \mid i=1, \cdots, n\}$

---

* 他の2状態は

unit	u_1	u_2	u_3	u_4	u_5	u_6
value	0	1	1	0	1	1

および

unit	u_1	u_2	u_3	u_4	u_5	u_6
value	0	1	0	1	1	1

である．

である．

式(22)は，1) 単語間係り受け共起頻度の符号を反転し，さらに，2) その絶対値が [0, 1] の間におさまるような変形を行う．符号を反転するのは，単語間係り受け共起頻度は意味的尤度が高いほど大きな値をとるのに対し，Y_i は意味的尤度が高いほど小さい値をとる必要があるためである．

また絶対値を一定の範囲におさめるのは以下の理由による．入力文が異なればその中で用いられる単語も異なる．用いられる単語が異なれば単語間係り受け共起頻度のとる値の範囲も変化する．入力文ごとに意味的尤度を表す値の範囲が変化するのでは安定したパラメータを得ることが困難であるので，どのような入力文に対しても一定の範囲内におさまるようにする必要がある．

式(22)は 0 以外の値をとる単語間係り受け共起頻度を [−1.0, −0.5] におさまるよう正規化する．ただし，単語間係り受け共起頻度が 0 のときはそのままで変化させない．これは学習用例文中で一度も係り受け関係を生じなかった単語の組合せをそうでないものから差別化するためである．この操作により，特に入力文中の単語間係り受け共起頻度の範囲が広い場合の解析成功率が向上した．

4.4 文節間の距離

日本語では一般的に，文節間の距離が小さいほど係り受けが生じやすいという傾向がある．係り受け関係 i の係り元を k_i 番目の文節，係り先を l_i 番目の文節とするとき，文節間の距離が短い係り受け関係が選ばれるほど小さな値を取る式は

$$E_4=\sum_{i=1}^{n} Z_i u_i \tag{23}$$

と表すことができる．ただし，

$$Z_i=l_i-k_i \tag{24}$$

とする．

4.5 係り受け解析ネットワークのエネルギ関数

以上で E_1〜E_4 が決定された．これより式(18)と式(3)の係数を比較すると，ユニット間の結合の強さ w_{ij} および各ユニットのしきい値 θ_i は

$$\begin{cases} w_{ij}=-2(a+bX_{ij}) \\ \theta_i=(1-2m)a+cY_i+dZ_i \end{cases} \tag{25}$$

とすればよいことがわかる．

4.6 実　　験

形態素解析の場合と同様，本実験の約 1/4 の入力データを使って予備実験を行い，ネットワーク内の定数を手作業で決定した．

まず，式(18)の a, b, c, d は以下のように決定された．

$$a=6,\ b=5,\ c=15,\ d=1. \tag{26}$$

時刻 t のときのネットワークの温度 $T(t)$ は次の式で与えられるものとした．

$$T(t)=\frac{T_0}{1+t/\tau}. \tag{27}$$

ただし

$$T_0=10,\ \tau=10 \tag{28}$$

である．

また，ユニットの初期状態はすべて 0 とし，状態遷移は非同期に行った．計算機環境は形態素解析の場合と同一である．

解析対象の文としては，単語間係り受け共起頻度の学習用例文中に現れた単語のみから構成されているものを選んだ．解析対象文・学習用例文はともに文献15) 中の植物に関する語義説明文を用いた．形態素解析の場合と異なり，平仮名への変換は行っていない．今回の実験では解析対象文の一部をそのまま学習用例文として用いたが，このように学習用例文中の係り受け関係と解析対象文中の係り受け関係とが類似した傾向のものである場合は，入力した 200 文のうち 95% 以上の文において正しい係り受け関係が得られることがわかった．解析した文全体の約 5% は，エネルギの極小値に捕まって脱出できなかった．ユニットの状態遷移は非同期かつランダムに生じるので，同じ文を解析しても異なる結果が導かれる場合もある．

本実験で係り受け解析ネットワークが安定状態に収束するまでに必要とされたステップ数を図 4 に示す．ここではすべてのユニットが平均して 1 回ずつ発火するのに要する時間を 1 ステップとしている．前述のと

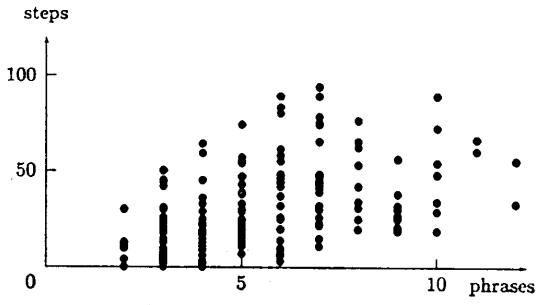

図 4 入力文中の文節数と係り受け解析ネットワークが収束するまでのステップ数との関係
Fig. 4 Relation between the number of bunsetu phrases and the steps required to make the network converge.

おり状態遷移は非同期かつランダムに生じるので，同じ文を解析しても収束までのステップ数が同一になるとは限らないが，全体の傾向はいつまでも図4と大差なかった．図4からは文節数と収束までのステップ数の間に強い相関関係があるとは言えないが，文節数の増加に比してステップ数はそれほど増加しないように見える．

図4の中には，文節数＝2，すなわち係り受け関係が一つしかないにもかかわらず収束までに数十ステップを必要とした例が見られる．これはネットワークの温度が高いときには比較的高い確率でエネルギの増加方向への状態遷移が生じるためである．ネットワークの初期温度 T_0 を小さくとればこのようなむだな状態遷移は生じにくくなるが，そのようにするとエネルギ極小状態からの脱出が困難になるため，正しい解析結果を得るためにはより一層ゆっくりと焼きなましを行わなければならなくなる．全体としてある程度以内のステップ数でネットワークを収束させるためには，多少むだな状態変移があってもやむを得ないと言えよう．

式(25)のネットワークに式(26)のパラメータを与えた場合，エネルギ最小の状態が常に文法的に正しい係り受けの組を表すとは限らない．例えば，**図5**に示すような文節数＝3，係り受け数＝3の入力があったとする．また，このときの Y_i は，

$$Y_1=-1,\ Y_2=0,\ Y_3=-1 \tag{29}$$

で与えられるものとする．このときエネルギ最小となるのは

$$u_1=1,\ u_2=0,\ u_3=1 \tag{30}$$

という状態であるが，これは文法的に正しい係り受けの組ではない．

にもかかわらず，図4の実験結果においては文法的に誤った解析結果となったものは一つもなかった．これは，学習用例文と解析用入力文が同一分野であるために，尤度の高い係り受けの組合せを選択することがそのまま文法的にも正しい組合せを選択することにつながるためと考えられる．また予備実験の結果得られた定数（式(26)）を見ると，係り受けの尤度を表す関数の係数である c が，文法的正当性を表す関数の係数である a および b よりも大きくなっているが，これも同じ理由と考えられる．

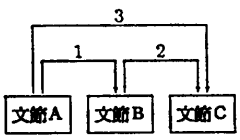

図5 文節間係り受け関係の例 (2)
Fig. 5 An example of modification relation of a sentence (2).

単語間係り受け共起頻度をまったく考慮しなかった場合，すなわち式(26)の c を0に変更し，

$$a=6,\ b=5,\ c=0,\ d=1. \tag{31}$$

とした場合の解析成功率は 44% に留まった．誤った解析結果のほとんどでは，位置的にもっとも近い文節が係り先として選ばれていた．これは，単語間係り受け共起頻度という意味的な制約がなくなり，位置的に近い係り受けほど成立しやすいという運用論的な制約のみが残っている状態から容易に予測される結果と一致する．なお，この場合も文法的に誤った解析結果は得られなかったが，これは文法的正当性を表す関数の係数 a および b が，係り受けの距離を反映する関数の係数 d に比して大きいためと考えられる．

5. おわりに

相互結合型のニューラルネットを用いて，日本語の形態素解析および係り受け解析を行った．係り受け解析では単語間の意味的整合性を表す関数が解析結果に大きな影響を与えた．どちらの解析においてもパラメータを調整することによって約 95% の成功率を得ることができた．入力文の長さが増加しても，ネットワークが収束するまでのステップ数はあまり増加しなかった．

ニューラルネットを利用した自然言語解析システムは，従来のシステムに比べてハードウェア化が容易であると思われる．また上に述べたように，入力文の長さが長くなっても，解析に要する時間はそれほど増加しない．したがって特定の分野における比較的複雑な文を専門に解析する自然言語解析エンジン等としての利用が考えられる．分野を限定すれば単語間の意味的関係を表す関数も比較的容易に作成できよう．

今後の課題としては，解析率の向上，形態素解析ネットワークと係り受け解析ネットワークの統合，一般の CFG を解析するネットワークの作成等が残されている．また，エネルギ関数の定数決定に Boltzmann Machine の学習アルゴリズムを適用することの可能性についても考察を進める予定である．

参考文献

1) 麻生英樹：ニューラルネットワーク情報処理，産業図書 (1988).
2) Hopfield, J.J.: Neurons with Graded Re-

sponse Have Collective Computational Properties like those of Two-state Neurons, *Proceedings of the National Academy of Science USA 81*, pp. 3088-3092 (1984).
3) Hopfield, J. J. and Tank, D. W.: Neural Computation of Decisions in Optimization Problems, *Biological Cybernetics*, Vol. 52, pp. 141-152 (1985).
4) 木村和広，鈴岡 節，伊藤悦雄，天野真家：神経回路網の連想機能を用いたかな漢字変換システム—ニューロワープロの実験試作—，第4回人工知能学会全国大会，9-3, pp. 301-304 (1990).
5) 森 辰則，中川裕志：Connectionist Model による構文解析モデル，情報処理学会論文誌，Vol. 30, No. 4, pp. 447-456 (1989).
6) 村瀬 功，中川聖一：ボルツマンマシンによる文節ラティスの係り受け解析，第38回情報処理学会全国大会論文集，5 E-8, pp. 378-379 (1989).
7) 尾関和彦：最適文節列を選択するための多段決定アルゴリズム，電子通信学会技術研究報告，SP 86-32, pp. 41-48 (1986).
8) 奥村明俊，山端 潔，村木一至：ニューラルネットワークによる日本語係り受け構造の学習，第4回人工知能学会全国大会，11-5, pp. 353-356 (1990).
9) 田村 淳，安西祐一郎：Connectionist Model を用いた自然言語処理システム，情報処理学会論文誌，Vol. 28, No. 2, pp. 202-210 (1987).
10) 高橋直人，板橋秀一：相互結合型ニューラルネットワークによる日本語の係り受け解析，第40回情報処理学会全国大会論文集，4 F-7, pp. 464-465 (1990).
11) Takahashi, N. and Itahashi, S.: Japanese Sentence Analysis Utilizing Mutually Connected Neural Network, *Proceedings of PRICAI '90*, pp. 257-262 (1990).
12) 高橋直人，板橋秀一：ニューラルネットによる日本語形態素・係り受け解析，情報処理学会研究会報告，90-NL-80 (1990).
13) 高橋直人，板橋秀一，平井有三：日本語形態素解析用ニューラルネットワークについて，第42回情報処理学会全国大会論文集，1 C-2 (1991).
14) Waltz, D. L. and Pollack, J. B.: Massively Parallel Parsing: A Strongly Interactive Model of Natural Language Interpretation, *Cognitive Science*, Vol. 9, pp. 51-74 (1985).
15) 金田一京助，金田一春彦，見坊豪紀，柴田 武，山田忠雄(編)：新明解国語辞典第二版（磁気テープ版），三省堂 (1974).

(平成3年2月18日受付)
(平成3年6月13日採録)

高橋 直人（正会員）

1987年筑波大学第三学群情報学類卒業．同年筑波大学大学院博士課程工学研究科に入学，現在に至る．自然言語処理に関する研究に従事．人工知能学会会員．

板橋 秀一（正会員）

昭和39年東北大学工学部通信工学科卒業．昭和45年同大学院(博)電気及通信工学専攻退学．同年東北大学電気通信研究所助手．昭和47年電子技術総合研究所入所．昭和49年同所主任研究官．昭和52年ストックホルム王立工科大学客員研究員．昭和57年筑波大学電子・情報工学系助教授．現在同教授．工学博士．音声・画像・自然言語処理の研究に従事．電子情報通信学会，日本音響学会，人工知能学会，日本認知科学会，IEEE，アメリカ音響学会各会員．