

ニューラルネットワークを用いた日本語係り受け解析

3K-1

清水浩行*

佐藤 秀樹†

立岡 章‡

林 達也*

*名古屋工業大学 電気情報工学科

†日本電装（株）

‡CICC

1 研究主旨

本研究は、広い意味で一種の最適化問題である日本語の文節間係り受け解析を、数学的な最適化問題—評価関数を最小化するような状態を検索する問題—として整理すること、更にその求解にニューラルネットの導入を試みることをその主旨とする。

2 係り受け解析

本研究でいう（日本語の）係り受け解析とは、文節単元に区切られた日本語文に対し、2項関係により意味的に成立し得る文節間の係り受け（図1の実線および破線矢印）の中から、文全体として成立するもの（図1の実線矢印）を選び出すことである。

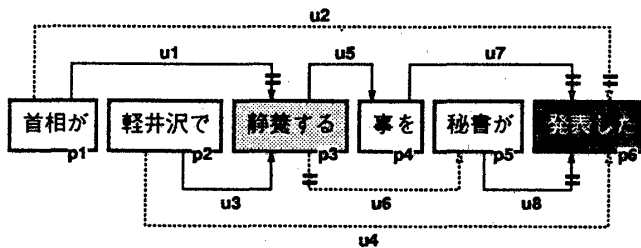


図1: 文節間係り受け関係の例

2.1 解釈のベクトル表現

係り受け線 i ($i = 1, 2, \dots, n$; $n \triangleq$ 2項関係により意味的に成立し得る係り受け関係の総数) を、それが選択されていれば値“1”を、そうでなければ値“0”をもつような変数 u_i で表す。このとき、文の係り受けをいかに解釈するかは n 次元の状態ベクトル $u = [u_1, u_2, \dots, u_n]$ を使って表現できる。

2.2 構文原理

本研究では、次の4つの構文原理を評価関数化して係り受け解析に利用した。

Dependency Analysis for Japanese Sentence
Using Mutual Connected Neural Networks
Hiroyuki SHIMIZU (NIT), Hideki SATO (NIPPONDENSO Co., Ltd.), Akira TATEOKA (CICC), Tatsuya HAYASHI (NIT)
Department of Electrical and Computer Engineering,
Nagoya Institute of Technology (NIT),
Gokiso-cho, Showa-ku, Nagoya, 466, Japan

1. 非交差性: 係り受けは互いに交差しない。
2. 係り先の（1文節による）占有性: 全ての文節は自分より後方のただ一つの文節に係る。
3. 卑近接続性: 距離的に近い文節間ほど係り受けが成立しやすい。
4. 格充足性: 用言が必要とする必須格が満たされる。

2.2.1 非交差性の評価関数 E_1

相互排他行列 X を用いて、次のように定義する。

$$E_1 = \sum_{i,j=1}^n X_{ij}u_iu_j \quad (1)$$

$$X_{ij} = \begin{cases} 1 & u_i \text{ と } u_j \text{ が交差している場合} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

2.2.2 係り先占有性の評価関数 E_2

係り元行列 Y を用いて、次のように定義する¹。

$$E_2 = \sum_{k=1}^m \left(\sum_{i=1}^n Y_{ki}u_i - 1 \right)^2 - 1 \quad (3)$$

$$Y_{ki} = \begin{cases} 1 & u_i \text{ の係り元が文節 } k \text{ である場合} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

2.2.3 卑近接続性の評価関数 E_3

文節間距離ベクトル Z を用いて、次のように定義する。

$$E_3 = \sum_{i=1}^n Z_iu_i \quad (5)$$

$$Z_i = (u_i \text{ の長さ}) \quad (\leq m-1) \quad (6)$$

2.2.4 格充足性の評価関数 E_4

必須格候補行列 P および必須格数ベクトル N を用いて、次のように定義する。

$$E_4 = \sum_{k=1}^m \left(\sum_{i=1}^n P_{ki}u_i - N_k \right)^2 \quad (7)$$

$$P_{ki} = \begin{cases} 1 & u_i = 1 \text{ で文節 } k \text{ に含まれる用言の} \\ & \text{必須格が一つ埋まる場合} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

$$N_k = \text{文節 } k \text{ に含まれる用言の欲する必須格数} \quad (9)$$

¹ $m \triangleq$ 文節数。

2.3 逐次検索による解析結果

以上、4つの評価関数に対して、次のようにその和をとることで、4つの原理の成立度の全てを反映するような評価関数が得られる² ($b > d \geq a \geq c = 1$)。

$$E = aE_1 + bE_2 + cE_3 + dE_4 \quad (10)$$

$$\begin{cases} b = c \{D_1 - (m-2)\} + 1 \\ a = \{c(D_1 - D_2) \text{div } 2\} + 1 \\ d = c(D_1 - D_2) + 1 \end{cases} \quad (11)$$

$$D_1 = \sum_{k=1}^m \left\{ \max_{1 \leq i \leq n} (Y_{ki} Z_i) \right\}, \quad D_2 = \sum_{k=1}^m \left\{ \min_{1 \leq i \leq n} (Y_{ki} Z_i) \right\}$$

この評価関数を用いて、実際にいくつかの例文について逐次検索により解析—評価関数の最小値検索—を行なった。その結果、本システムが完全な単文・複文に対しては、評価値最小状態として妥当な解釈を得られる、係り受け解析システムと成り得ることが実証できた。

一方、本システムでは重文や自明格の省略された文としては不完全なものについては、正確な解析が困難であり、これらを取り扱うには、より強力な解析システムが必要となることもわかった。

3 ニューラルネットを用いた最小値検索

3.1 相互結合型ニューラルネットワークと最適化問題

ネットワーク内の各ユニット u_i が、

$$I_i = \sum_j w_{ji} u_j - \theta_i \quad (12)$$

$$\text{遷移確率 } p(u_i = 1) = \frac{1}{1 + \exp(-I_i/T)} \quad (13)$$

に従って非同期的に状態遷移を行なう相互結合型ニューラルネットワークが、互いに等しい強さで結合—対称結合—し ($w_{ij} = w_{ji}$)、自身に対して直接、フィードバック結合をもたない ($w_{ii} = 0$) ならば、ネットワーク全体は次式で与えられるエネルギーと呼ばれるネットワークの状態の関数 E を、近似的に最小化するように動作する。

$$E(u) = -\frac{1}{2} \sum_{i,j} w_{ij} u_i u_j + \sum_i \theta_i u_i \quad (14)$$

ここで、 T は温度と呼ばれる正の定数で、はじめのうちはエネルギーの極小値からの脱出を図るため高温でネットワークを作動させ、しばらくして最小解付近に状態が近付いたら低温へと冷却してやることで、最小解へと収束させる。実際にはどの程度の速さで冷却してゆくか—クーリング・スケジュール—が大きな問題となる。

²div は整数除算の商を返す二項演算子。

3.2 係り受け解析用のネットワークの設計

ニューラルネットを使った係り受け解析では、各係り受け線にユニットを一つずつ割り当てる。

評価関数 ((10) 式) とネットワークのエネルギー式 (14) との対応をとれば、ネットワークの重み・閾値を与えるものとして次の2式が得られる。

$$w_{ij} = -2 \left\{ aX_{ij} + \sum_k (bY_{ki}Y_{kj} + dP_{ki}P_{kj}) \right\} \quad (15)$$

$$\theta_i = cZ_i - b - d \sum_k \{P_{ki}(2N_k - 1)\} \quad (16)$$

3.3 実験とその結果

本研究ではクーリング・スケジュールには、一般によく用いられる次の関数を用いた (T_0 = 初期温度)。ここに ΔT はインターバルと呼ばれる定数で、この値が大きいほどゆっくりとした冷却に、小さいほど急冷となる。

$$T(t) = \frac{T_0}{1 + t/\Delta T} \quad (17)$$

このスケジュールを用いて、実際にいくつかの例文に対して構成したネットワークでシミュレートを行なった結果、初期温度やインターバルを上手く調整することで (逐次検索によって事前に得られていた) 真の最小値へと系を取束させることができることがわかった。

図2は、最小解を得るのに初期温度に対して、インターバルをどこまで小さくとれるか、その限界値がある例文について実際に計測した様子である。

この図から初期温度を小さくとるほど、よりゆっくりとした冷却が必要になることが確認できる。また、初期温度が一定値以上でなければならないこともわかる。

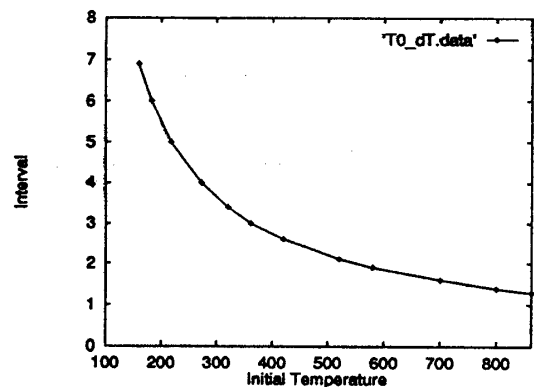


図2: 初期温度に対するインターバルの限界値

参考文献

- [1] 合原 一幸: “ニューラルコンピュータ 脳と神経に学ぶ”, 東京電機大学出版局 (1988)
- [2] 高橋 直人, 板橋 秀一: “ニューラルネットワークを用いた日本語解析の試み”, 情報処理学会論文誌 Vol. 32, No. 10, pp. 1330-1337 (1991)