

ニューラルコンピューティングを用いた 現代いろは歌の作成*

4T-8

北端美紀 吉池紀子 武藤佳恭†
慶應義塾大学 政策・メディア研究科‡

1 はじめに

平安時代に作られたいろは歌は、同じ仮名が重複せずに「ん」の字を除く47文字で表現された仏教的な教訓歌である。以下の歌が代表的である。

「いろはにほへと ちりぬるを わかよたれそ つねならむ うゐのおくやま けふこえて あさきゆめみし めいもせず」

この他にも、いくつかの「いろは歌」が存在することが知られている。

“あ～ん”までのすべての平仮名を使うという条件のもとで歌を作る、というのは一見単純そうであるがなかなか難しい。その理由として、私達の日常話す言葉の中には1つの単語で同じ仮名を何回も使ってしまうものが多いことや、濁音・破裂音などを含む単語が多いことなどが挙げられる。このような条件を満たすような歌を作成するのは、難解な組み合わせ最適化問題として捉えることができる。本論文ではニューラルコンピューティングを用いた「現代いろは歌」の作成方法とそのシミュレーション結果を紹介する。

シミュレーションでは、あらかじめ語群として任意の964語の単語を与え、その中から本論文で提案するニューラルネットワークにいろは歌の条件を満たすフレーズを選ばせることに成功した。

2 ニューラル表現

データとして与えられた語群の一語一語に対し、その単語を使用するか使用しないかを定めることで問題を表現するために、単語数の長さをもつ1次元配列の正規ニューロン (V_0, \dots, V_z , z は語群の単語数)

*A method of making "Iroha-Uta" by using neural networks

†Miki Kitabata Noriko Yoshiike Yoshiyasu Takefuji

‡Graduate School of Media and Governance Keio University

を用意する。単語 n を使用することを $V_n = 1$, 使用しないことを $V_n = 0$ とすると、任意のいろは歌は0か1を要素にもつ1次元配列のセルオートマトンの状態のひとつとして表される。

語群から選択的に単語を選びだすときの制約条件は、平仮名46文字をすべて使うこと、それぞれの仮名が重複しないことである。この制約条件を満たすように z 個のニューロンの状態を変化させていくニューラルネットワークの動作式は次のようになる。

$$\frac{dU_i}{dt} = -A \times \sum_{c=0}^{cnum_i} \left(\sum_{a=0}^{wordnum} \sum_{b=0}^{cnum_a} V'_{abd_{ic}} - 1 \right) + B \times hill \quad (1)$$

ただし、ニューロン V_n は式(2)で表されるマッカロックピッツニューロンで、ニューロン V_n の状態値 U_n は式(3)で表される1次のオイラー法で変化させるものとする。

$$\begin{cases} V_n = 1 & \text{if } U_n \geq 0 \\ V_n = 0 & \text{otherwise} \end{cases} \quad (2)$$

$$U(t+1) = dU(t+1) + U(t) \quad (3)$$

V'_{ijk} は、単語 i が選択されていて、かつ単語 i の j 文字目がある仮名 k に対応するとき $V'_{ijk} = 1$, それ以外のとき $V'_{ijk} = 0$ となる V_i の関数である (図1)。 $cnum_i$ は単語 i の文字数、 d_{ic} は単語 i の c 文字目の仮名の種類、 $wordnum$ は語群の単語数である。仮名の種類は“あ:0, い:1, ..., ん:45” というふうに番号づけされている。

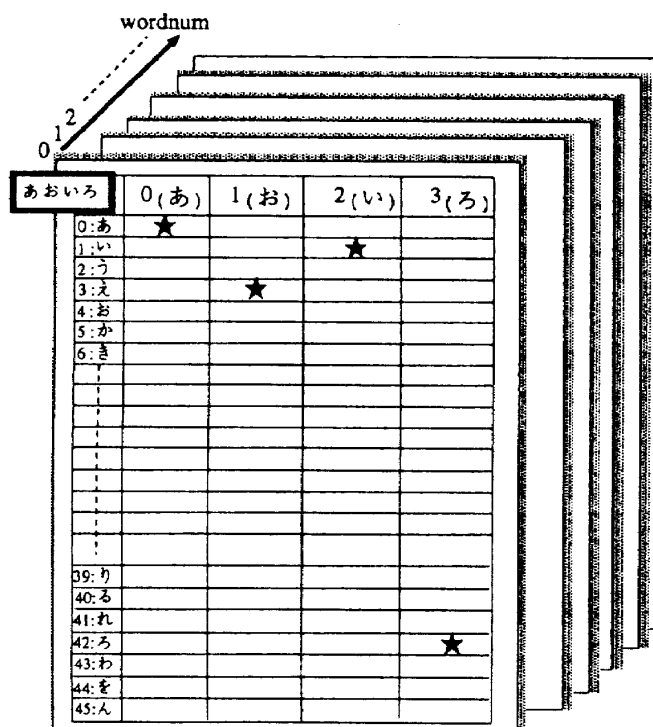


図1: V' の発火の様子 (★は発火の状態を表す)

$hill(x)$ は局所解から抜け出す働きをする関数であり、式(4)で与える。¹

$$\begin{cases} hill(x) = 1 & \text{if } x < cnum \\ hill(x) = 0 & \text{otherwise} \end{cases} \quad (4)$$

3 シミュレーション結果

シミュレーションに用いた語群は、任意に選んだ文章を文節ごとに区切ったものと、単語毎に区切ったものの2つを合計964語用意した。

実際のシミュレーションでは文字数の短い単語が多くなってしまったため、次のように動作式を正規化して文字数の長い単語を優先して選ばせた。さらに収束率をあげるため、50ステップごとに $A = A + 1, B = B/5$ という演算を行ってパラメータの調整を行った。

$$\frac{dU_i}{dt} = -A \times \sum_{c=0}^{cnum_i} \left(\sum_{a=0}^{wordnum} \sum_{b=0}^{cnum[a]} V'_{abkanai,c} \right) \times \frac{1}{cnum_i} - 1 + B \times hill \times cnum_i \quad (5)$$

¹ここでの局所解とは、ニューロンの1次元セルオートマトンの状態が解ではないデッドロックに陥った状態を指す。

シミュレーション結果:

「そんなぬすまれたほうへつきひとのけむりにわかあめおをみせよふゆはてもえるいちやさくらしろねこ」

この単語を品詞ごとに分類し、文法規則に当てはめて並びかえをすれば、歌をつくることが可能である。今回はデータの入力を文節の単位で行なって、人間の手で適当に並び変えを行った。シミュレーション結果で得られた単語を並べ替えてできた歌を以下に示す。

歌1:

しろねこおをみせ
にわかあめけむり
ぬすまれたつきひとのほうへ
ふゆはて
さくらもえるよ
そんないちや

4 おわりに

歌1は次のように解釈することができる。

「しろねこが歩いている。
にわか雨が降りだし、月を隠した。
冬も終り桜もえるようにさいている。
そんな一夜の風景。」

制約式は意味を考えて文節を選ぶように作られていないが、単語を後で並べ替えることによって意味の通るような歌を作ることができた。いろは歌作成の問題には、歌の読み手が勝手に意味を考えて状況を想像する楽しさがある。ニューラルネットワークで文節の並び方の制約を決めて、自動的に文章らしく並び換えをすることは今後の課題としたい。

参考文献

- [1] Yoshiyasu Takefuji "Neural Network Parallel Computing" Kluwer Academic Publishers, 1992
- [2] 稲垣足穂, 「一千一秒物語」, 新潮社, 1969
- [3] 長尾真編, 「自然言語処理」, 岩波書店, 1996