

## 問題のマッピングに k-out-of-n 規則を用いた確率的 ニューラルネットの収束特性に関する実験

3C-2

平岩伸次

中川 徹

北川 一

(豊田工業大学)

### 1 はじめに

ボルツマンマシンなどの確率的ニューラルネットを用いて組合せ最適化問題を解く研究が盛んに行なわれているが、これら確率的ニューラルネットでは以下のような問題点がある。

- パラメータ設定(特に焼き鈍し)が経験的
- 解への収束判定を並列に行なう手段が無い

本研究で用いるニューラルネットは、大域解において各ニューロンの入出力関係が一意的に定まるという”最適点”を有しており、この性質を用いることによる上記の問題点の克服を目的として、実験を試みた。

### 2 ネットワーク構成

本研究で用いるニューラルネットは”n個からなる集合の中からk個を選択(k-out-of-n)”という形の組合せ問題を対象とするものである<sup>[1][2]</sup>。この制約を満たす条件は、集合の各要素の出力  $O_i$  ( $= 0$  または  $1$ ) について

$$E = \left( \sum_{i=1}^n O_i - k \right)^2 + \sum_{i=1}^n O_i(1 - O_i)$$

が0となることであり、Hopfield型ネットワークで構成すると

$$T_{ij} = \begin{cases} -2 & (i \neq j) \\ 0 & (i = j) \end{cases}$$

$$I_i = 2k - 1$$

が成立する ( $T_{ij}$  はニューロン  $i$   $j$  間の結合係数、 $I_i$  はニューロン  $i$  のバイアス)。

さらに、このように構成されたネットワークでは、大域解において各ニューロンの入出力が

”出力=1 かつ 入力= $I_i$ ,” 或は

”出力=0 かつ 入力= $-I_i$ ,”

のいずれかを満たすことが判っている<sup>[3]</sup> ( $I_i$  はそのニューロンが属する集合の数)。

### 3 ニューロンモデル

ここでは出力  $O_i$  が離散値 (0 または 1) で、入力値に対し確率的に動作するニューロンを用いる。確率的ニューラルネットでは、温度  $T$  における平衡状態にてネットワークのエネルギー  $E$  がある確率分布  $P(E)$  に従うように、ニューロンの出力変化確率  $p(O_i = 1)$  を定める<sup>[4]</sup>。

本研究で用いる確率分布として、既に最適化問題への適用例のある Boltzmann 分布 (ボルツマンマシン)<sup>[5]</sup> 及び Cauchy 分布 (コーシーマシン)<sup>[6]</sup> を採用した。各々における確率分布式とニューロンの出力変化確率を下に示す ( $T$  は温度、 $u$  はニューロンの入力、 $C$  は定数)。

- Boltzmann 分布

$$P(E) = C \cdot \exp\left(-\frac{E}{T}\right)$$

$$p(O_i = 1) = \frac{1}{1 + \exp(-u/T)}$$

- Cauchy 分布

$$P(E) = \frac{T}{\pi(T^2 + E^2)}$$

$$p(O_i = 1) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \frac{u}{T}$$

### 4 実験方法

Sun-3 上に C 言語にて確率的ニューラルネットのシミュレータを作成し、収束実験を行なった。最適化問題としては n-Queen 問題を  $n = 12$  として用いた。n-Queen 問題をこのネットワークにマッピングする方法については<sup>[6]</sup>を参照されたい。

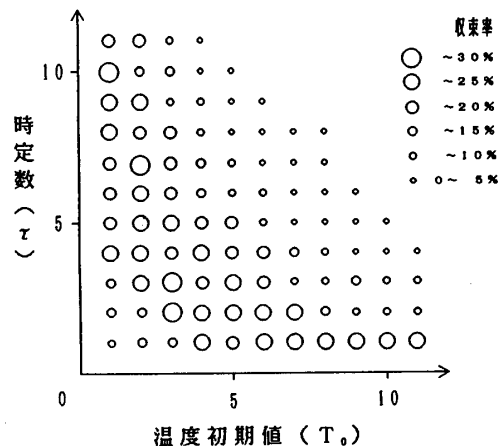


図1 コーシーマシンによる収束結果

実験条件を次に示す。

- 温度  $T$  の焼き鈍しスケジュール

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

( $T_0$ : 温度初期値  $\tau$ : 時定数  $t$ : ステップ数)

- 試行回数: 200 回
- 初期状態: 任意
- 1 試行当りの最大実行ステップ数: 400

但しニューロンの更新方式は逐次更新にて行い、全ニューロンが更新した時点で 1-step として計算を行なった。

収束判定は、ニューロンが全て最適点にいるかどうかを調べることにより行なっている(収束判定を並列に行なうことができる)。

## 5 結果及び考察

### 1. 焼き鈍しパラメータと収束率の関係

コーシーマシンを用い、初期温度  $T_0$ 、時定数  $\tau$  を変化させて収束率を測定した結果を図 1 に示す。 $T_0 \cdot \tau = 10$  の双曲線近辺で収束率が高くなっていることがわかる。測定した範囲での収束率の最大値は 30% であった。

ボルツマンマシンにて同様の測定を行なった結果、測定範囲内に於いて上述のような収束率の高まりは見られなかった(収束率の最大値は 19%)。

### 2. ステップ数と収束率の関係

ステップ数と収束率の関係を図 2 に示す。時定数  $\tau$  が大きいほど収束するまでのステップ数は大きくなるが、最終収束率も向上している。また  $\tau$  が大きいと最大実行ステップ数の制限のため、見かけ上は最終収束率がかえって低くなっており、これが図 2 で双曲線状に収束率ピークが現われた原因と考えられる。

$T_0 = 3, \tau = 3$  にて最大実行ステップを増加したところ約 2000 ステップで収束率が飽和状態に達し、最終収束率は 68% であった。

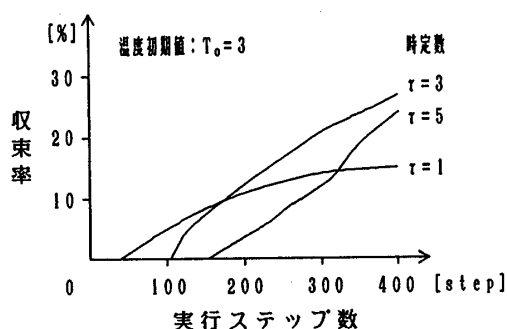


図 2 ステップ数と収束率の関係

### 3. 最適点情報を用いた場合

ニューロンが最適点にいるという情報を基に、収束速度の向上が図れないか試みた。具体的には、最適点にいるニューロンの温度を  $1/n$  倍することによる”部分的急冷

”を行なった。しかしその結果、急冷を行なわないものに対し収束率の向上は見られなかった(表 1)。

n	0.2	0.5	0.7	1.0
収束率 [%]	24	21	27	30

(コーシーマシン,  $T_0 = 3, \tau = 3$ )

表 1 部分的急冷による収束率の変化

## 6 おわりに

収束判定を並列に行うことのできる確率的ニューラルネットにより、 $n$ -クイーン問題を解くことができた。しかし最適点情報を用いた収束の高速化については効果が得られなかった。

この実験で得られた収束率は、同じ  $n$ -クイーン問題を扱った文献<sup>[5]</sup> に比べてかなり低い。しかし同じネットワーク構成において決定論的なニューロンモデルを採用し高い収束性能が得られていること<sup>[7]</sup> や、収束判定を並列に行える利点があることなどから、本研究で試みた最適点情報の利用について今後研究を進めてゆく価値はあると思われる。

## 7 参考文献

1. E.W.Page, G.A.Tagliarini, "Algorithm Development for Neural Networks," Proc. SPIE Symposium on Innovative Sci. and Tech., vol.880, pp.11-18, 1988
2. 中川, 平岩, 小早川, 北川, "SDNN: 厳密にデジタルの NN 計算モデルとその並列処理システムの提案," 信学技報 CPSY 89-42, pp.61-66, 1989
3. T.Nakagawa, E.W.Page, G.A.Tagliarini, "SDNN: A Computation Model for Strictly Digital Neural Networks and Its Applications," Proc. of 5th AAAIC '89, pp.1-9
4. 麻生, "ニューラルネットワーク情報処理," pp.27-33, 産業図書, 1988
5. 梶浦, 秋山, 安西, "ボルツマンマシンによる  $n$ -クイーン問題の解法," 情処学会第 38 回全国大会予稿集, pp.482-483, 1989
6. H.H.Szu, R.L.Hartley, "Nonconvex Optimization by Fast Simulated Annealing," PROC. of IEEE, vol.75, No.11, pp.1538-1540, 1987
7. 中村, 佐竹, 中川, 北川, "SDNN: 厳密にデジタルな NN による  $N$ -クイーン問題の解法," 情処学会第 40 回全国大会予稿集, 1990