

論文

ホップフィールドニューラルネットにおける
逐次学習パターン記銘法の検討

Examination of Iterative Pattern Storage Method in Hopfield Neural Network

道法 浩孝 (高知大学教育学部)

Hirotaka DOHO

Faculty of Education, Kochi University

ABSTRACT

The Hopfield associative memory model is significant because the memory function can theoretically be represented as a local-minimum stability problem in a dynamical system introducing the concept of energy function. Based on this property, the model has been applied as an engineering model not only to associative memory, but also to combinatorial optimization problems. In this paper, the focus is on the pattern storage process of this neural network, investigating the storage method, and analyzing the effect through computer simulation. As a result, a learning algorithm based on the iterative pattern storage method set all stored patterns into the local minimum state of energy function and was able to control the value adjusting a stability of memory. In addition, we verified that after completion of storage, all network states move to the state of the nearest local minimum or stored pattern in a few steps when initially in a non-stored pattern state, thus confirming the effectiveness of a pattern storing control model.

Key words : Hopfield Neural Network, Iterative Pattern Storage Method, Associative Memory, Energy Function, Stability of Memory

1. はじめに

脳は、自己組織化、環境適応、並列分散処理などの機能を備えており、認知・記憶・思考および精密性や正確性のみならず柔軟な情報処理などを実現している。このような機能は、従来の2値論理に基づくデジタルコンピュータでは実現困難であり、脳に学んだ新しい情報処理手段を確立するための様々な模索がなされている [1, 2, 3].

ニューロコンピューティングは、この脳に学ぶことでノイマン型コンピュータに代わる新しい情報処理機構の脳型コンピュータを構築しようとするものであり、その手法は、脳を構成する神経回路網のモデルを構築し、それによる様々な情報処理の可能性を数学的解析および計算機シミュレーションによって探るものである [4, 5, 6, 7, 8].

ところで、ニューロコンピューティングの基礎となるニューラルネットワークに関する研究は、1940年代のマカロック (McCulloch) とピッツ (Pitts) による形式ニューロンモデルの提案に始まる。これは、脳を構成するニューロンを多入力・1出力の非線形素子としてモデル化し、それらを接続してニューラルネットを構成したものである。以後、1950年代のローゼンブラット (Rosenblatt) が提案した学習認識のできるパーセプトロン、1980年代のラムエルハート (Rumelhart) らが提案した誤差逆伝搬法など多くの学習アルゴリズムやモデルの提案が行われ現在に至っている。

そのなかで、記憶機能を最も簡明な形でモデル化したニューラルネットワークとしてホップフィールド型ニューラルネット [9] が知られている。これは、先のマカロックとピッツによる形式ニューロンモデルの回路網 (ニュー

ラルネット)にエネルギー関数の概念を導入し、記憶すべき対象をそのエネルギー関数の局所安定状態に対応させるというものである。記憶という機能を力学系における極小値安定問題として理論化され得ることを示した点で、このモデルは意義深いものである。そして、この性質に着目して、連想記憶や組合せ最適化問題等の諸課題に適用され工学的モデルとして利用されてきた。

本研究では、このホップフィールド型ニューラルネットワークのパターン記憶過程に着目し、コンピュータシミュレーションを通してその検討・評価を行う。具体的には、パターン記憶の処方およびそれに含まれるパラメータの設定・調節の効果を、記憶成立後のネットワーク構造と入力情報に対するネットワークの振舞いから定量的に分析する。そして、モデルが適用される課題の条件に即したパターン記憶の制御方法の提案を行う。

2 モデリング

2.1 ニューロンモデル

脳における情報処理は、多数の他の神経細胞とシナプス経由で結合する個々の神経細胞の性質とその間の相互作用形式によって基本的に担われている。そして、神経細胞には、他の神経細胞からの信号電位の影響を総和し(線形加算性)、その総和が神経興奮のためのしきい値を越えなければ何事も起こらず、越えれば信号を出す(非線形しきい値性)という2つの特徴が確認されている[10]。このような特徴を持つ脳の神経細胞を多入力・1出力の非線形要素(ニューロンモデル)として図1のように数理的に形式化し(以降、ニューロンと呼ぶ)、そのニューロンを多数個結合させ、その相互作用によって情報処理を人工実現させようとするのがニューラルネットワークの考え方である。

時刻 $t+1$ における i 番目のニューロンの状態値 $X_i(t+1)$ は、入出力関数 f を通して次のように決定される。

$$X_i(t+1) = f\left(\sum_{j=1}^N w_{ij} X_j(t) - \theta_i\right) \quad (1)$$

ここで、 w_{ij} は j 番目のニューロンから i 番目のニューロンへの結合の強さ(以後、結合荷重と呼ぶ)、 θ_i は i 番目のニューロンのしきい値、 $X_j(t)$ は時刻 t における j 番目のニューロンの状態値、 N は全ニューロン数である。入出力関数 f は以下の式で定義される。

$$f(h_i) = \tanh\left(\frac{h_i}{2\varepsilon}\right) \quad (2)$$

(1)式では、

$$h_i(t) = \sum_{j=1}^N w_{ij} X_j(t) - \theta_i \quad (3)$$

ということになるが、ニューロンの状態値(出力値) X_i

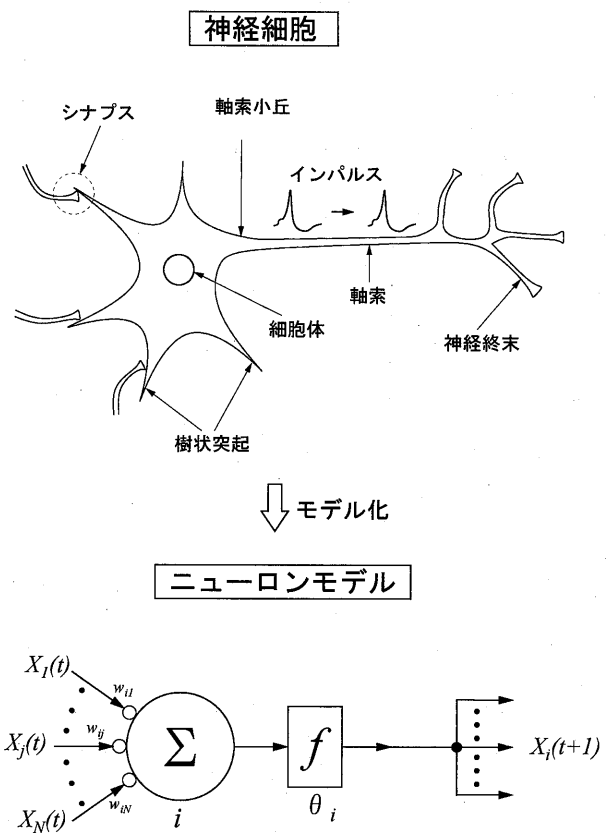


図1 神経細胞とその数理モデル

に対してこの h_i のことを以後、ニューロンの内部状態と呼ぶことにする。関数 f は S 字形をしており、 ε はその傾きのパラメータである。このとき X_i は $[-1, +1]$ の連続値をとるが、 $\varepsilon \rightarrow 0$ の極限で(2)式は

$$f(h_i) = \text{sgn}(h_i) \quad (4)$$

となる。 sgn は符号関数

$$\text{sgn}(x) = \begin{cases} +1 & (x > 0 \text{ のとき}) \\ -1 & (x \leq 0 \text{ のとき}) \end{cases}$$

で、このときニューロンの状態値 X_i は ± 1 の離散値となる。

2.2 静的連想記憶モデル

相互結合型ニューラルネットワークは、多数個のニューロン間の可能な結合をすべて認めるものであり、図示すると図2のように表される(ただし、自己結合は明示していない)。ニューロン i とニューロン j の間には、ニューロン j からニューロン i への結合荷重 w_{ij} とその逆の w_{ji} という双方向の結合が存在する。この両者の結合は一般には互いに違った値を持つ ($w_{ij} \neq w_{ji}$) が、両者が互いに等しい対称結合 ($w_{ij} = w_{ji}$) としたものが、ホップフィー

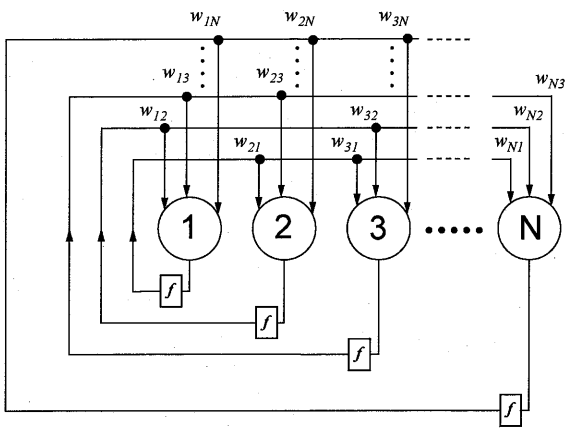


図2 相互結合型ニューラルネット

ルド (Hopfield)による静的連想記憶モデル[9]である。

今、結合が対称 ($w_{ij}=w_{ji}$) でかつ自己結合が正またはゼロ ($w_{ii} \geq 0$) という条件の下で、各ニューロン値を逐次的に (全ニューロンが一斉にではなく、1時刻に1つのニューロンだけ)、

$$X_i(t+1) = \text{sgn} \left(\sum_{j=1}^N w_{ij} X_j(t) - \theta_i \right) \quad (5)$$

にしたがって状態を更新するものとする。このとき、ネットワーク状態 $\{X_i\}$: $\mathbf{X} \equiv (X_1, X_2, \dots, X_N)$ の関数として

$$E(\mathbf{X}) = -\frac{1}{2} \sum_{ij} w_{ij} X_i X_j + \sum_{i=1}^N \theta_i X_i \quad (6)$$

を導入すると、この関数の値は以下に示すようにニューロンの状態遷移とともに時々刻々単調減少し、最終的には極小値に収束する。ここで、関数 E はネットワークのエネルギー関数と呼ばれている。

今、 i 番目のニューロンの値が時刻 t から $t+1$ の間に、 $X_i(t)$ から $X_i(t+1)$ へ変化したとすると、このときの $E(\mathbf{X})$ の変化量 $\Delta E(\mathbf{X}) = E(\mathbf{X}(t+1)) - E(\mathbf{X}(t))$ は、

$$\Delta E(\mathbf{X}) = -(X_i(t+1) - X_i(t)) h_i(t) \quad (7)$$

となる。したがって、(7)式の時刻 t における i 番目のニューロンの内部状態 h_i が0より大きい場合、 $X_i(t+1)$ は+1になり右辺の $(X_i(t+1) - X_i(t))$ の値は0または正で、右辺全体は0または負になる。一方、 h_i が0または負の場合、 $X_i(t+1)$ は-1になり右辺の $(X_i(t+1) - X_i(t))$ の値は0または負で、右辺全体は0または負になる。すなわち、常に $E(\mathbf{X}(t+1)) \leq E(\mathbf{X}(t))$ 、(等号成立は $\mathbf{X}(t+1) = \mathbf{X}(t)$ のとき) の関係が成立する。極小値に落ち着いたときのネットワークの状態 $\{X_i\}$ はもう変化せず ($\mathbf{X}(t+1) = \mathbf{X}(t)$)、局所安定状態となる。

また、各ニューロン値を(5)式にしたがって同期的に (全ニューロンを1時刻に一斉に) 更新する場合も、逐次的場合のエネルギー関数と同様の性質を持つ関数を導入でき、その極小値への収束を示すことができる。ただし、

この場合のネットワークの安定解は $\mathbf{X}(t+1) = \mathbf{X}(t-1)$ で与えられ、逐次更新の場合のホップフィールド解 $\mathbf{X}(t+1) = \mathbf{X}(t)$ に加えて、2つの異なるネットワーク状態間の周期2の周期 (振動) 解 ($\mathbf{X}(t+1) \neq \mathbf{X}(t)$) も存在し得る。

上記のようなネットワークの性質を脳における連想記憶の機能と対応させる。すなわち、連想記憶を構成する学習と想起の2つの過程を、エネルギー関数に基づいて以下のように捉える。連想記憶における学習過程は、ネットワークに記憶させたいパターン情報 $\{\xi_i^\mu\}$: $\xi^\mu \equiv (\xi_1^\mu, \xi_2^\mu, \dots, \xi_N^\mu)$ (以下、記銘パターンと呼ぶ) に対応するネットワーク状態がちょうど(6)式のエネルギー関数の谷 (すなわちネットワークの局所安定状態) に一致するように、結合荷重のセット $\{w_{ij}\}$ の値を調節し、エネルギー関数の形を決定することに相当する。ここでいうパターン情報とは、図2のネットワークを構成する各ニューロンの状態が+1のときは黒、-1のときは白と解釈した際の配列状態をさすが、図3の $N=12$ の場合の例のようにネットワーク状態を2次元の配列パターンに対応させることで文字表現等が可能となる。学習過程を図 (図4(a)) を用いて説明すると、はじめ無結合 ($\{w_{ij}\} = 0$) 状態のネットワーク ($E(\mathbf{X}) = 0$) が記銘パターン情報 (例として文字パターンH, I) に基づいて $\{w_{ij}\}$ を変更し、最終的に記銘パターンに対応するネットワーク状態がエネルギー関数の谷にくるように整えるというものである。

一方、想起過程は、学習過程によって形成されたエネルギー関数の曲面に沿って極小値へ山降りするプロセスとして捉えられる。つまり、想起過程においては $\{w_{ij}\}$ は固定されており、ネットワークが記銘パターンに対応

$$\{X_i(t)\} : \mathbf{X}(t) = (+1, -1, +1, +1, -1, +1, -1, -1, -1, +1, -1, +1)$$



(● : +1 ○ : -1)

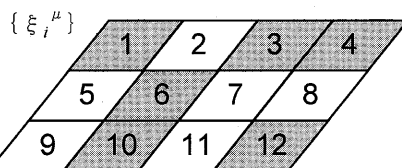
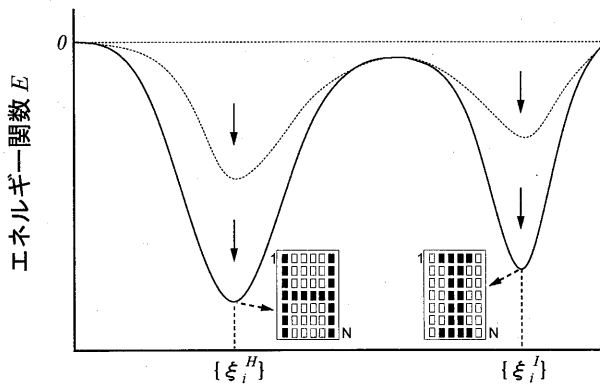
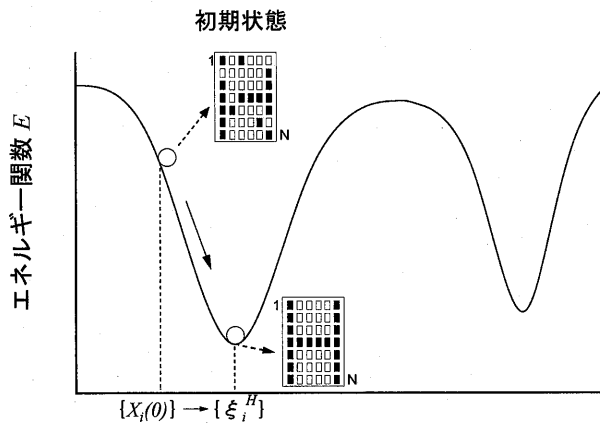


図3 ネットワーク状態とパターン情報の対応 ($N=12$ の場合)



(a) 学習過程: $\{w_{ij}\}$ 可変



ネットワーク状態 $\{X_i\}$

(b) 想起過程: $\{w_{ij}\}$ 固定

図4 学習と想起の過程の概念図

するエネルギー関数の極小状態へと変化する。図で示すと、ネットワークはある1つの状態(例としてパターンH)に対しノイズを持った状態から出発して最寄りの谷の記録パターン対応状態(パターンH)に至るといふものである。

2.3 静的記憶の逐次学習則

ネットワークに記憶させたいパターン $\{\xi_i^\mu\}$ がネットワークの局所安定状態(すなわち, $\{X_i(t+1)\} = \{X_i(t)\} = \{\xi_i^\mu\}$) に一致するための必要十分条件は、個々のニューロンにおける発展方程式(5)が

$$\begin{aligned} X_i(t+1) &= \text{sgn} \left(\sum_{j=1}^N w_{ij} X_j(t) \right) \\ &= \text{sgn} \left(\sum_{j=1}^N w_{ij} \xi_j^\mu \right) = \xi_i^\mu \end{aligned} \quad (8)$$

となることである(ただし、ニューロンのしきい値は全て $\{\theta_i\} = 0$ とする.)。これは ξ_i^μ と $\sum_{j=1}^N w_{ij} \xi_j^\mu$ が同符号、

すなわち両者の積(これを γ_i^μ と表す)が正:

$$\gamma_i^\mu \equiv \xi_i^\mu \sum_{j=1}^N w_{ij} \xi_j^\mu > 0 \quad (9)$$

という条件に帰着される。さらにこの(9)式を満足する十分条件として

$$\gamma_i^\mu \equiv \xi_i^\mu \sum_{j=1}^N w_{ij} \xi_j^\mu > \kappa (> 0) \quad (10)$$

なる設定が可能である。正のパラメータ κ の導入により、局所安定な $\{\xi_i^\mu\}$ の引力圏が形成され、ネットワーク状態 $\{X_i(t)\}$ が記録パターン $\{\xi_i^\mu\}$ から多少変位していても最終的に $\{\xi_i^\mu\}$ の状態をとるようになる。その意味で γ_i^μ を記憶の安定度、その下限 κ を記憶の引力圏パラメータと呼ぶことにする[4]。

ディーデリッヒ(Diederich)とオPPER(Opper)[11]は、与えられた全ての記録パターン($\mu = 1 \sim p$)に対して(9)式が成立するまで

$$w_{ij}^{\text{new}} = w_{ij}^{\text{old}} + \delta w_{ij}^\mu \quad \left(\delta w_{ij}^\mu = \frac{1}{N} \xi_i^\mu \xi_j^\mu \right) \quad (11)$$

によって結合荷重を徐々に変更する逐次型の学習則(以後、逐次学習則と呼ぶ)を提出し、その収束性を証明した。この方法を用いれば、通常の文字パターンのように互いに相関を持つ非直交なパターン群もネットワーク上で局所安定状態とすることが可能である(記録対象が直交パターン群の場合には、この学習則はよく知られたヘップ則[12]に一致)。図5はこの学習アルゴリズムのフローを示したものであるが、全パターンの提示に対して結合荷重の更新回数 M がゼロになった時点で学習は終了する。つまり、(9)式の安定度 γ_i^μ は(11)式の逐次学習ごとに

$$\begin{aligned} (\gamma_i^\mu)^{\text{new}} &= \xi_i^\mu \sum_{j=1}^N w_{ij}^{\text{new}} \xi_j^\mu \\ &= \xi_i^\mu \sum_{j=1}^N \left(w_{ij}^{\text{old}} + \frac{1}{N} \xi_i^\mu \xi_j^\mu \right) \xi_j^\mu \\ &= (\gamma_i^\mu)^{\text{old}} + 1 \end{aligned} \quad (12)$$

となり、着目する記録パターンに対する安定度が1ずつ増加することになる。ただ、他の記録パターンに対する安定度が同時に増加する保証はないので、学習アルゴリズムとしては与えられた全ての記録パターン($\mu = 1 \sim p$)に対して(9)式が成立するまで繰り返し(11)式を適用することになる。

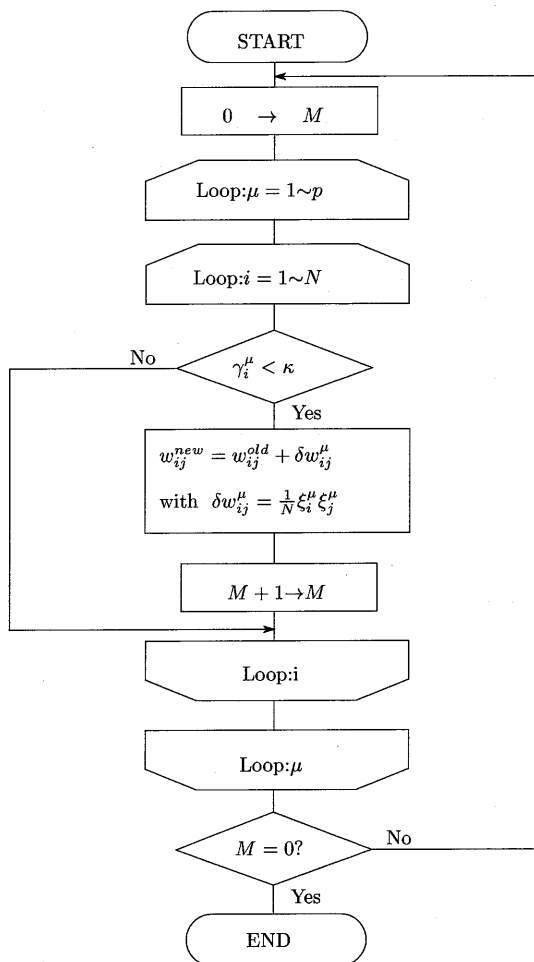


図5 逐次学習アルゴリズムのフローチャート

3. 静的記憶モデルの計算機シミュレーション

ここでは、前章で説明した連想記憶型ニューラルネットによる静的記憶モデルについて、計算機シミュレーションを通して実際に評価する。

シミュレーションでは、ネットワークを構成する全ニューロン数を 156 ($N=156$) とし、ニューロンのしきい値と入出力関数の傾きのパラメータを、 $\{\theta_i\} = 0$, $\varepsilon = 0.015$ (全ニューロン共通) と設定した。パターン情報としては、図6のように $12 \times 13 (=156)$ のユニット構成でラ

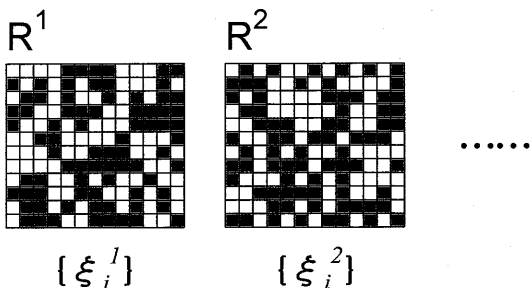


図6 シミュレーションで用いたパターン情報の例 ($\xi_i = 1$ が黒, $\xi_i = -1$ が白に対応)

ンダムパターンを20個 ($R^1 \sim R^{20}$) 用意し、その半数の $R^1 \sim R^{10}$ を記録パターンとした。これらのパターンはいずれも $\xi_i = 1$ と $\xi_i = -1$ のユニットが半々で、互いに直交に近いパターン群となっている。

図7は、 $R^1 \sim R^{10}$ を記録対象に 2.3 で説明した逐次学習則を $\kappa = 1$ という条件の下で実行したときの結果である。記録成立後の安定度 γ_i^μ ($156 \times 10 = 1560$ 個) の度数 (count) 分布 (区分の幅 $\Delta \gamma_i^\mu = 0.1$) は、10パターンすべてが $\gamma_i^\mu > 1$ を満たすまで逐次更新した結果、上限がほぼ2付近に及ぶものとなっている。

図8は、学習終了後の結合荷重 $\{w_{ij}\}$ に基づいて20パターンそれぞれのエネルギー値 E を求めたものである。ネットワークに記録させた $R^1 \sim R^{10}$ の10パターン (以降、既知パターンと呼ぶ) はいずれもほぼ -120 付近のエネルギー値に掘り下げられているが、記録対象としなかった $R^{11} \sim R^{20}$ の10パターン (以降、未知パターンと呼ぶ) は、0 付近のエネルギー値に留まっており、ほとんど掘り下げが見られない。各パターン状態 $\{\xi_i^\mu\}$ に対応するエネルギー値は、(6)式に(9)式を考慮すると、

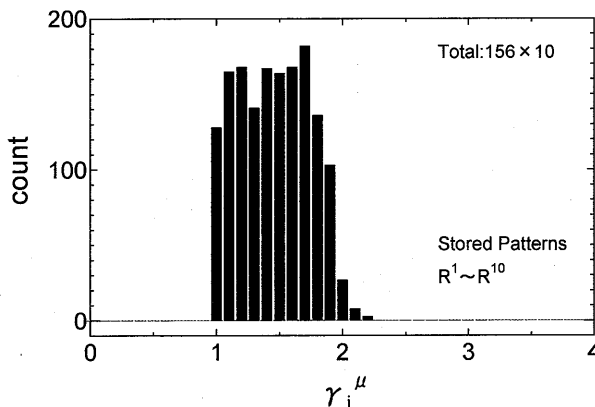


図7 10パターン ($R^1 \sim R^{10}$) 逐次学習後の安定度 (γ_i^μ) 度数分布 (学習条件: $\gamma_i^\mu > 1$)

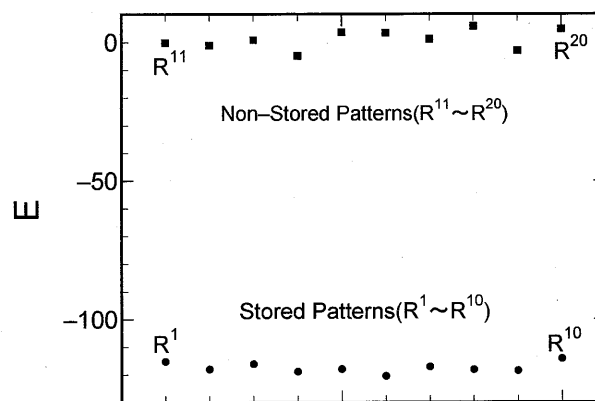


図8 記録パターン ($R^1 \sim R^{10}$) と非記録パターン ($R^{11} \sim R^{20}$) のエネルギー値 (学習条件: $\gamma_i^\mu > 1$)

$$\begin{aligned}
 E^{(\mu)} &= -\frac{1}{2} \sum_{ij} w_{ij} \xi_i^\mu \xi_j^\mu = -\frac{1}{2} \sum_{i=1}^N \xi_i^\mu \sum_{j=1}^N w_{ij} \xi_j^\mu \\
 &= -\frac{1}{2} \sum_{i=1}^N \gamma_i^\mu
 \end{aligned}
 \tag{13}$$

であり、これに図7より既知パターンの γ_i^μ の平均が概ね1.5であることを適用すると、

$$\begin{aligned}
 \overline{E^{(\mu)}} &\simeq -\frac{1}{2} \times N \times 1.5 = -\frac{1}{2} \times 156 \times 1.5 \\
 &= -117
 \end{aligned}
 \tag{14}$$

が得られる。

次に、このネットワークの想起過程について見ることにする。初期状態 $\{X_i(0)\}$ としていずれかの既知パターン ($\{\xi_i^\mu\}$, $\mu=1\sim 10$) から多少変位した状態を与え、(1)式による時間発展を調べたところ、いずれの場合もそれぞれの既知パターンへネットワーク状態が収束した。このことは、既知パターンが逐次学習によってエネルギー関数の局所安定状態に位置するに至っており、そのため先の図4での説明と同様の静的想起が実現していることを示すものである。

ところで、初期状態 $\{X_i(0)\}$ として未知パターン ($\{\xi_i^\mu\}$, $\mu=11\sim 20$) のいずれかに近い状態を与えた場合、ネットワークの時間発展はどのようになるだろうか。図9は、5つの未知パターン ($R^{12}, R^{14}, R^{16}, R^{18}, R^{20}$) を初期状態としたときの結果である。いずれの場合も、ネットワークは数ステップ (3~8ステップ) でエネルギー値が-100前後の局所安定な状態に遷移し、そのまま停留している。特に初期状態が R^{14}, R^{16}, R^{20} のパターン状態のときには、ネットワークはそれぞれ既知パターンの R^6, R^9, R^7 の状態に到達している。同様のシミュレーションを、上記以外の未知パターン ($R^{11}, R^{13}, R^{15}, R^{17}, R^{19}$) に対して行ったところ、 R^{11}, R^{13}, R^{17} のパターンを初期状態とした場合、それぞれ既知パターンの R^7, R^2, R^7 の状態に到達していた。

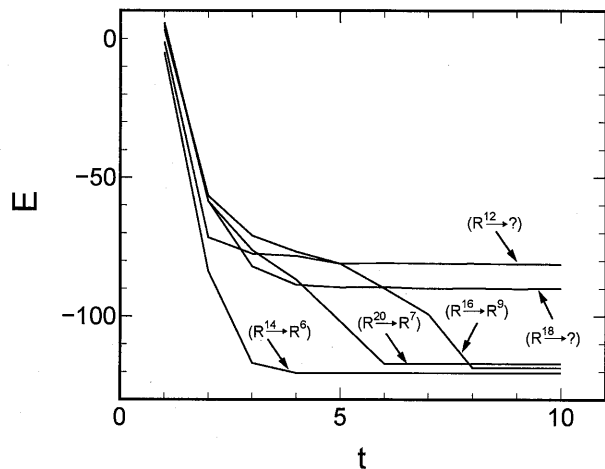


図9 未知パターン ($R^{12}, R^{14}, R^{16}, R^{18}, R^{20}$) を初期状態とする場合のネットワークの時間発展

図10は、逐次学習則のパラメータを変更して、上記のシミュレーションと同様に $R^1\sim R^{10}$ の10パターンをネットワークに記録させたときの結果である。ここでは、結合荷重更新の刻みを $1/N$ から $1/(5N)$ に変更して w_{ij} の更新を行い、 γ_i^μ に上限値を設定 ($1 \leq \gamma_i^\mu \leq 1.5$) している。 γ_i^μ の値を一定の範囲に制限し分散を小さくすることにより、記録パターンに対応するネットワーク状態のエネルギー関数値および各記録パターン状態が形成する引力圏がほぼ等しくなることが予想される。図11は、学習終了後の20パターンそれぞれのエネルギー関数値 E を求めたものである。10個の記録パターン状態に対応するエネルギー関数値は-122~-124の狭い領域に分布しており、学習方法改善の効果が表れている。

4. おわりに

本研究では、ホップフィールドによる連想記憶モデルにおけるパターン記録に対して、記録方法の検討とその効果の分析・評価をコンピュータシミュレーションを通して実施した。具体的には、学習アルゴリズムおよび記憶の安定度 γ_i^μ がパターン記録に与える影響を、記録成

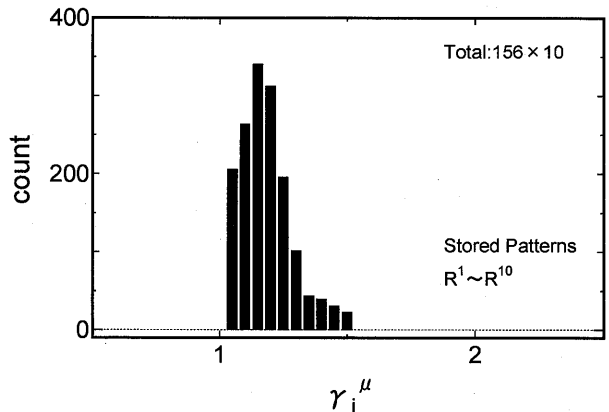


図10 10パターン ($R^1\sim R^{10}$) 逐次学習後の安定度 (γ_i^μ) 度数分布 (学習条件: $1 \leq \gamma_i^\mu \leq 1.5$)

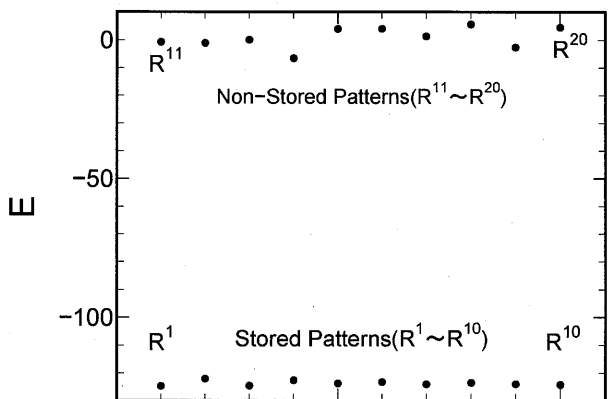


図11 記録パターン ($R^1\sim R^{10}$) と非記録パターン ($R^{11}\sim R^{20}$) のエネルギー値 (学習条件: $1 \leq \gamma_i^\mu \leq 1.5$)

立後のエネルギー関数および初期状態設定後のネットワーク状態の変化から検討・分析した。その結果、逐次学習則に基づく学習アルゴリズムにより、すべての記憶パターンがエネルギー関数の局所安定状態に位置するようネットワーク構造の設定がなされることが明らかになった。また、記憶パターンに相当する局所安定状態のエネルギー関数値は、記憶学習時の η を調節することによって制御可能であることが明らかになった。さらに、記憶成立後、初期状態として未知パターンを設定した場合、いずれのパターンに対しても数ステップで、最寄りの局所安定状態または記憶パターン状態に遷移することが確認され、パターン記憶の制御モデルとしての有効性が明らかになった。

参考文献

- [1] M.A.Arbib (ed.): The Handbook of Brain Theory and Neural Networks (2nd Edition), The MIT Press (2002)
- [2] 土井利忠・藤田雅博・下村秀樹 (編): 脳・身体性・ロボット, シュプリンガー・フェアラーク東京 (2005)
- [3] M. Minsky: The Emotion Machine-Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind-, Simon & Schuster (2007)
- [4] B.Müller and J.Reinhardt: Neural Networks An Introduction, Springer-Verlag (1990)
- [5] 中野馨: 人間情報工学-バイオニクスからロボットまで-, コロナ社 (1996)
- [6] 吉富康成: ニューラルネットワーク, 朝倉書店 (2002)
- [7] 甘利俊一: 神経回路網モデルとコネクショニズム (新装版), 東京大学出版会 (2008)
- [8] 渡辺桂吾 (編著): ニューラルネットワーク計算知能, 森北出版 (2006)
- [9] J.J.Hopfield: Neural networks and physical systems with emergent collective computational abilities; Proc. Natl. Acad. Sci. USA, Vol. 79, pp. 2554-2558(1982)
- [10] 松村道一: ニューロサイエンス入門, サイエンス社 (1995)
- [11] S.Diederich and M.Opper: Learning of Correlated Patterns in Spin-Glass Networks by Local Learning Rules; Phys.Rev.Lett., Vol. 58, pp. 949-952(1987)
- [12] D.Hebb: Organization of Behaviour, Wiley, New York(1949)