自己連想記憶の想起失敗とロジスティック写像の関係

1J－1

浅川　芳元　　長橋　宏
東京工業大学 大学院総合理工学研究科

1．はじめに

パターンをヘッブ則により埋め込んだホップフィールド型ニューラルネットワーク（以下，HNN）において，エネルギー極小値との対応により，自己連想記憶が可能であることが知られている．しかし，実際の人の記憶想起においては，同一のパターンに対して必ずしも想起が成功するとは限らず，最終的に思い出せない不完全な記憶想起過程が存在する．

本研究では HNN の数理モデルを用いて，各ニューロンのエネルギー，発火確率の時間変化量を考え，ロジスティック写像との関数形としての相似性より，想起失敗となる現象についての考察を与える．

2．HNN の数理モデル

発火状態を＋1，非発火状態を－1で表現すると，i 番目のニューロンの状態 $S_i(t)$ に対するダイナミックスは，次のように表される．

$$S_i(t+1) = sign\left[\sum_{j\neq i}\frac{J_{ij}\left(S_j(t)+1\right)}{2}-\theta_i\right] \qquad (1)$$

ここで J_{ij} はヘッブ則により定められたシナプス結合強度，θ_i は閾値である．

$$\theta_i = \frac{1}{2}\sum_{j\neq i}J_{ij} \qquad (2)$$

この閾値を(2)のように定め，符号関数の引数を2倍することにより，(1)のダイナミックスは次のように簡略化される．

$$S_i(t+1) = sign\left[\sum_{j\neq i}J_{ij}S_j(t)\right] \qquad (3)$$

(3)のダイナミックスによるN個のニューロンからなる，HNN の系全体のエネルギーは次式で定義される．

$$E(t) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}J_{ij}S_i(t)S_j(t) = \sum_{i=1}^{N}\varepsilon_i(t) \qquad (4)$$

$$\varepsilon_i(t) = -\frac{1}{2}S_i(t)\sum_{j=1}^{N}J_{ij}S_j(t) = -\frac{1}{2}S_i(t)h_i(t) \qquad (5)$$

$$h_i(t) = \sum_{j=1}^{N}J_{ij}S_j(t) \qquad (6)$$

ε_i，h_i はそれぞれニューロン i のエネルギー，入力総和を表す．各ニューロンのエネルギーと統計力学の手法におけるボルツマン分布を用いて，各ニューロンの発火確率を定義すると次式のようなシグモイド関数が導かれる．

$$W_i(h_i(t)) = \frac{1}{1+\exp(-\beta h_i(t))} \qquad (7)$$

β は系のノイズTの逆数で定義される定数とする．

3．ロジスティック写像

カオスを生ずるシンプルな写像として，ロジスティック写像が知られている．[3]

$$x(n+1) = ax(n)(1-x(n)) \qquad (8)$$

ロジスティック写像ではパラメーター a の値が，臨界値 3 以上で安定固定点の解から周期解を生じ、 $a^{*}>3.5699\cdots$ でカオス状態を生ずる．

4．ロジスティック写像的なニューロンの発火確率

ニューロン i における発火確率は，時間の関数である入力総和の合成関数となっているので，その時間変化を考えるために合成微分を行なうと，時刻 t における各ニューロンの状態，および発火確率より，時刻 $t+1$ の発火確率を表すことができる．

$$W_i(h_i(t+1)) = W_i(h_i(t))\left[1+\beta\{1-W_i(h_i(t))\}\lambda_i(t)\right]$$
$$= \beta\lambda_i(t)W_i(h_i(t))\left[\left(\frac{1}{\beta\lambda_i(t)}+1\right)-W_i(h_i(t))\right] \qquad (9)$$

ここで発火確率 W_i を Z_i で次のように変数変換する．

$$Z_i(t) = \frac{\lambda_i(t)}{T+\lambda_i(t)}W_i(h_i(t)) \qquad (10)$$

Z_i は系のノイズが 0 の時に W_i と一致する変数である．

The Relationship between Failure to Retrieve the Memory by Self Associative Memory and Logistic Mapping
Yoshiyuki Asakawa : asakawa@n.isl.titech.ac.jp , Hiroshi Nagahashi : longb@n.isl.titech.ac.jp
Tokyo Institute of Technology, Interdisciplinary Graduate School of Science and Engineering

この Z_i を用いて(9)を表すと次のようなロジスティック写像と相似の時系列関数が得られる．

$$Z_i(t+1)=\left(1+\frac{\lambda_i(t)}{T}\right)Z_i(t)\left(1-Z_i(t)\right) \tag{11}$$

ロジスティック写像のパラメーターとの比較より，ロジスティック写像の臨界値を用いて，次のような関係式が得られる．

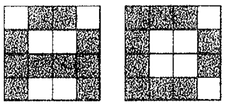
$$\lambda_i(t)>2T \tag{12}$$

$$\lambda_i(t)>\left(a^*-1\right)T=2.5699\cdots\times T \tag{13}$$

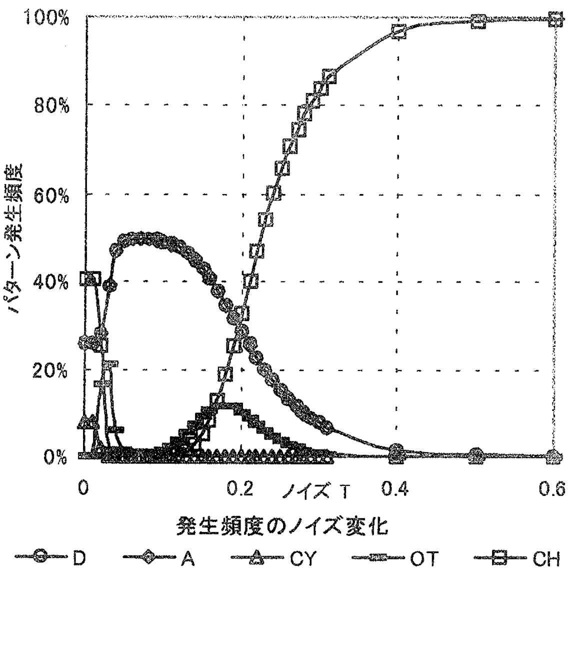
ニューロン i の入力総和の時間変化量が，系のノイズの2倍以上になると(12)より発火確率は周期的状態になり，また2.5699…倍以上になると(13)より発火確率はカオス状態になると考えられる．

5．数値実験

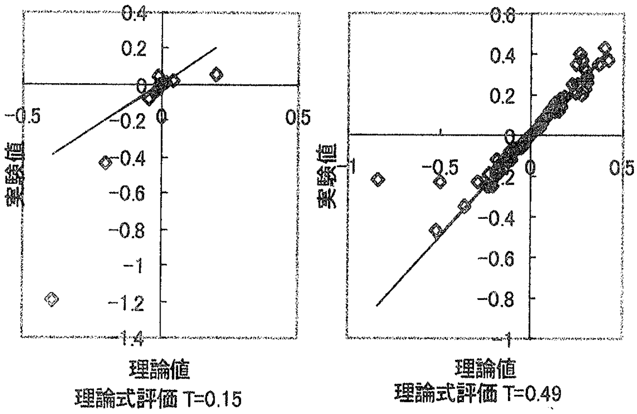
16個のニューロンから構成されるHNNに、ヘッブ則により以下のような2つのパターンA、Dを記憶させ、それらの想起について数値実験を行った．



2^{16} 通りの初期状態すべてに対し、系のノイズを0から1まで変化させ、更新回数46～50回の5回を観測対象とし、埋め込んだパターンA、D（反転も含める)、2周期状態(CY)、他の固定パターン(OT)、4周期以上およびカオス状態(CY)の計5つに分類し、それぞれの相対的発生頻度を表したグラフを以下に示す．



また乱数により設定した初期状態に対して，状態の更新を行ない，得られた時系列データを元に，発火確率の差による時間変化量を縦軸（実験値)，(9)式により得られる理論値を横軸としてプロットしたグラフを以下に示す．



5．考察

ニューロンの発火確率をエネルギーより導いたので，逆の操作をすることで発火確率からエネルギーを算出できる．そして発火確率が安定な状態ではなく，周期的，カオス的となるならば，エネルギーも同様の振る舞いをするものと考えられる．自己連想記憶ではエネルギー極小値に記憶パターンを対応させることで，想起を行なっているため，各ニューロンのエネルギー状態が周期的，カオス的となることで，安定的なエネルギー状態をとり続けることはできなくなり，このような場合に記憶の想起失敗が起こるものと考えられる．

6．まとめ

ホップフィールド型ニューラルネットワークにおいて，その系のノイズを各時刻における各ニューロンへの入力総和との関係より，ロジスティック写像と同様なダイナミックスが存在することを示した．

この研究では，系のノイズが定常的である場合のみに適応されるが，実際には状態更新に伴い，ノイズも変化していくものと考えられる．しかしニューラルネットワークの系におけるノイズには，熱的，電磁気的，情報量的なさまざまな要因が考えられ，それらの定式化は，今後の大きな課題となっていくものと思われる．

参考文献

[1] H.Nishimori:『ニューラルネットワークの統計力学』（丸善株式会社，1995）
[2] G.Takeda:『脳と力学系』（講談社，1997）
[3] K.Aihara:『カオス －カオス理論の基礎と応用－』（サイエンス社，1990）