

1対多の連合を記憶する相関型連想記憶モデルの解析

川村 正樹[†] 岡田 真人^{††} 平井 有三^{†††}

Analysis of a Correlation-Type Associative Memory with One-to-Many Associations

Masaki KAWAMURA[†], Masato OKADA^{††}, and Yuzo HIRAI^{†††}

あらまし 1対多の連合を記憶する相関型連想記憶モデルを議論する。連想記憶モデル HASP は相互想起型連想記憶回路と相互抑制回路からなる1対多の連合を扱うモデルである。我々はこれまでこの二つの回路を分離して、絶対容量に関して議論してきた。しかしながら、想起過程に関しては連合ベクトルと回路の状態との統計的な相関のため二つ回路を分離して議論することができない。そこで、本論文では HASP 型の最も簡単な相関型連想記憶モデルを提案し、このモデルの想起過程を議論する。このモデルは HASP と同様に相互想起型連想記憶回路と自己想起型連想記憶回路から構成される。1対多の連合を扱うために、複数の連合ベクトル間の対称性を破る外部入力を導入する。外部入力を導入する際に次の点を考える。1. 外部入力は想起される連合ベクトルに近いものから、無相関な入力までの様々なものを考慮する。2. 外部入力は相互想起部または自己想起部のいずれに加えればよいかを議論する。想起過程の性質がこれらの点にどのように依存するかを統計神経力学と計算機シミュレーションにより調べた。その結果、外部入力によって連合ベクトルが想起できることを示し、自己想起部に外部入力を加えた場合の方がより連合ベクトルを想起しやすいことがわかった。

キーワード 連想記憶, 1対多の連合, 統計神経力学, HASP

1. まえがき

これまでに解析されてきた多くの相互想起型連想記憶モデル [1], [2] は、その取り扱いの簡単さのために主に1対1の連合を扱っていた。しかしながら、明らかに人間は一つの“キー”項目に対して多くの“連合”項目を対応づけて連想記憶することができる。従って、より人間らしい連想をする神経回路モデルの構築のためには、1対多の連合を記憶することができるモデルを考察する必要がある。ここで、相互想起型連想記憶モデルに、一つのキーベクトルに対して複数個の連合ベクトルを学習させる場合を考える。このとき、キーベクトルを入力すると、対応する複数個の連合ベクトルの重なりに相当する混合ベクトルが出力される。連

合ベクトルの各成分がそれぞれ確率 $1/2$ で ± 1 の値をとる場合、生じた混合ベクトルは各連合ベクトルと等距離にある。そのため、記憶した連合ベクトルの一つを取り出すためには、連合ベクトル間の対称性を破る機構が必要である。

連想記憶モデル HASP [3] は1対多の連合を扱ったモデルであり、複数のキー項目に連合した項目の積集合や和集合などを検索できる。HASP は相互想起型連想記憶回路と自己想起を行う相互抑制回路の二つの回路から構成されている。相互想起部では1対多の連合を記憶している。自己想起部には連合ベクトル間の対称性を破る機構が含まれている。また、相互想起部の出力にクロストークノイズが含まれている場合、自己想起部はそれを除去する能力をもつ。これまでに、HASP の特性を利用して、足し算過程やプライミング効果などの記憶に関する認知心理学的現象が説明されている [4]~[7]。川村と平井 [8] は HASP の二つの回路を別々に議論することによって、クロストークノイズを評価し、絶対容量の解析を行った。しかしながら、想起過程のダイナミックスの解析においては、連合ベクトルと回路の状態との間の統計的な相関を考慮する

[†] 筑波大学工学研究科, つくば市
Doctoral Program in Engineering, University of Tsukuba,
Tsukuba-shi, 305-8573 Japan

^{††} 科学技術振興事業団 ERATO 川人学習動態脳プロジェクト, 京都府
Kawato Dynamic Brain Project, ERATO, Japan Science and
Technology Corporation, Kyoto-fu, 619-0288 Japan

^{†††} 筑波大学電子・情報工学系, つくば市
Institute of Information Sciences and Electronics, University of
Tsukuba, Tsukuba-shi, 305-8573 Japan

必要がある。二つの回路を別々に議論するとこの相関が取り扱えない。そのため、クロストークノイズを定量的に評価できず、想起過程のダイナミクスを議論できない。

本論文の目的は HASP の二つの回路を統一的に議論し、1 対多の連合を行う HASP の想起過程を議論することである。本論文は以下のように構成される。まず HASP 型の構成をもつ、相互想起型連想記憶回路と自己想起型連想記憶回路の二つの回路からなる最も簡単なモデル [9] を提案する。二つの回路を全体として解析することにより、連合ベクトルと回路の状態との間の相関を統計神経力学 [10], [11] を用いて考慮し、想起過程でのクロストークノイズを評価する [9]。次に複数の連合ベクトル間の対称性を破るために、どのような方法を用いる必要があるかを議論する。ここでは、連合ベクトルの対称性を破るためにバイアスとして外部入力を導入する。外部入力は記憶検索における文脈情報に相当する [5]。外部入力を導入する際に、次の点を議論する必要がある。まず、相互想起部または自己想起部のいずれの回路に外部入力を加えた方が連合ベクトルを想起しやすいかを検討する。次に、外部入力と連合ベクトルとの間の相関の大きさが、想起のしやすさにどのように影響するかを検討する。ここでは、外部入力としては想起する連合ベクトルに近いものから無相関な入力までの様々なものを考える。以上の点について、統計神経力学と計算機シミュレーションを用いて調べる。

2. モデル

本モデルは図 1 に示すように、相互想起型連想記憶回路と自己想起型連想記憶回路から構成される。説明の都合上、入力 y が相互想起部へ入力される時刻を $t = -1$ とし、相互想起部の出力が自己想起部へ入力される時刻を $t = 0$ とする。

2.1 相互想起部

今、 p 個の M 次元のキーベクトル $\eta^\mu = (\eta_1^\mu, \dots, \eta_M^\mu)^T, \mu = 1, 2, \dots, p$ に対して、それぞれ k 個の N 次元の連合ベクトル $\xi^{\mu, \kappa} = (\xi_1^{\mu, \kappa}, \dots, \xi_N^{\mu, \kappa})^T, \kappa = 1, 2, \dots, k$ を連合し、1 対 k の連合を記憶する場合を考える。 T は転置を表す。各ベクトルの成分 $\eta_i^\mu, \xi_i^{\mu, \kappa}$ は以下の確率で独立に生成する、

$$\text{Prob}[\eta_i^\mu = \pm 1] = \frac{1}{2}, \quad (1)$$

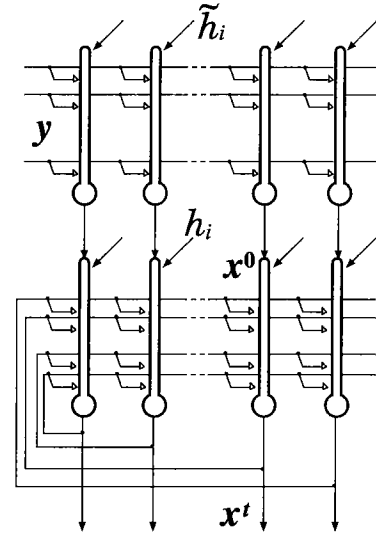


図 1 モデルの構成。上段が相互想起型連想記憶回路、下段が自己想起型連想記憶回路である。

Fig.1 A schematic diagram of model: An upper network operates as hetero-associative memory, while a lower network as auto-associative one.

$$\text{Prob}[\xi_i^{\mu, \kappa} = \pm 1] = \frac{1}{2}. \quad (2)$$

図 1 の相互想起部は N 個のニューロンで構成される。相互想起部への入力 y はキーベクトル η^μ と同じ M 次元のベクトル $y = (y_1, \dots, y_M)^T$ である。ここでは $M = \beta N, N \rightarrow \infty, \beta \sim O(1)$ の極限を考える。連合ベクトルの記憶数 pk は $pk \sim O(N)$ を超えることはできない [12]。ここでは $p \sim O(N), k \sim O(1)$ の場合を中心に議論する。 $pk = \alpha N$ とおき、 α を記憶率と呼ぶ。相互想起部の第 i 番目のニューロンの出力 x_i^0 は、

$$x_i^0 = \text{sgn} \left(\sum_{j=1}^M \tilde{J}_{ij} y_j + \tilde{h}_i \right), \quad (3)$$

で表される。ここで、 \tilde{J}_{ij} は入力の第 j 成分から相互想起部の第 i 番目のニューロンへのシナプス結合であり、関連学習を用いて、

$$\tilde{J}_{ij} = \frac{1}{N} \sum_{\mu=1}^p \sum_{\kappa=1}^k \xi_i^{\mu, \kappa} \eta_j^\mu, \quad (4)$$

で与えられる。また、 \tilde{h}_i は連合ベクトル間の対称性を破る外部入力であり、2.3 で説明する。

いま、第 1 番目のキーベクトル η^1 に対応する k 個の連合ベクトルのうち、第 1 番目の連合ベクトル

$\xi^{1,1}$ を想起する場合について考える。相互想起部への入力 \mathbf{y} とキーベクトル $\boldsymbol{\eta}^1$ とのオーバラップ (方向余弦) を、

$$\tilde{m} = \frac{1}{M} \sum_{j=1}^M \eta_j^1 y_j, \quad (5)$$

と定義する。これより、式 (3) は、

$$x_i^0 = \text{sgn} \left(\sum_{\kappa=1}^k \beta \tilde{m} \xi_i^{1,\kappa} + \tilde{z}_i + \tilde{h}_i \right), \quad (6)$$

と表すことができる。但し、 \tilde{z}_i は、

$$\tilde{z}_i \equiv \frac{1}{N} \sum_{j=1}^M \sum_{\mu=1}^p \sum_{\kappa=1}^k \xi_i^{\mu,\kappa} \eta_j^\mu y_j, \quad (7)$$

であり、時刻 $t = -1$ におけるクロストークノイズを表す。

2.2 自己想起部

図 1 の自己想起部は N 個のニューロンが相互に結合している。時刻 $t \geq 0$ における第 i 番目のニューロンの状態 x_i^t のダイナミクスは、

$$x_i^{t+1} = \text{sgn} \left(\sum_{j \neq i}^N J_{ij} x_j^t + h_i \delta_{t,0} \right), \quad (8)$$

で表される。ここで、 J_{ij} は第 j 番目のニューロンから第 i 番目のニューロンへのシナプス結合を表し、相関学習を用いて、

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \sum_{\kappa=1}^k \xi_i^{\mu,\kappa} \xi_j^{\mu,\kappa}, \quad (9)$$

で与えられる。また、 h_i は連合ベクトル間の対称性を破る外部入力であり、2.3 で説明する。また、 $\delta_{t,0}$ は $t = 0$ のときに 1 を、 $t \neq 0$ のときに 0 をとるクロネッカーデルタである。本論文では外部入力は時刻 $t = 0$ にのみ存在し、時刻 $t > 0$ においては存在しない場合を議論する。従って、外部入力 h_i は自己想起部の平衡状態での安定性に影響を与えない。

時刻 $t \geq 0$ における自己想起部のニューロンの状態 $\mathbf{x}^t = (x_1^t, \dots, x_N^t)^T$ と連合ベクトル $\boldsymbol{\xi}^{1,\kappa}$ とのオーバラップを、

$$m_i^\kappa = \frac{1}{N} \sum_{i=1}^N \xi_i^{1,\kappa} x_i^t, \quad (10)$$

と定義する。これより、式 (8) は、

$$x_i^{t+1} = \text{sgn} \left(\sum_{\kappa=1}^k m_i^\kappa \xi_i^{1,\kappa} + z_i^t + h_i \delta_{t,0} \right), \quad (11)$$

と表される。但し、 z_i^t は、

$$z_i^t \equiv \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu=1}^p \sum_{\kappa=1}^k \xi_i^{\mu,\kappa} \xi_j^{\mu,\kappa} x_j^t, \quad (12)$$

であり、時刻 t におけるクロストークノイズを表す。

2.3 外部入力

本節では、式 (3) と式 (8) の \tilde{h}_i と h_i について定性的に説明する。簡単のためにキーベクトルの個数 p を $p \sim O(1)$ とする。 $k \sim O(1)$ なので、この場合は記憶率 $\alpha = 0$ に対応する。また一つのキーベクトルに対応する連合ベクトルの個数 k を $k = 3$ とする。キーベクトルと連合ベクトルの次元は $M = N (\beta = 1.0)$ とする。

相互想起部への入力 \mathbf{y} が第 1 番目のキーベクトル $\boldsymbol{\eta}^1$ であったとする ($\tilde{m} = 1.0$)。この場合、自己想起部への入力、すなわち相互想起部の出力 x_i^0 は式 (6) より、

$$x_i^0 = \text{sgn} \left(\xi_i^{1,1} + \xi_i^{1,2} + \xi_i^{1,3} + \tilde{h}_i \right), \quad (13)$$

となる。もし外部入力 \tilde{h}_i がなければ、相互想起部の出力 x_i^0 は、キーベクトル $\boldsymbol{\eta}^1$ に対応する $k = 3$ 個の連合ベクトルの和の符号をとった混合ベクトル $\boldsymbol{\zeta}^1 = (\zeta_1^1, \dots, \zeta_N^1)^T$,

$$\zeta_i^1 = \text{sgn} \left(\xi_i^{1,1} + \xi_i^{1,2} + \xi_i^{1,3} \right), \quad (14)$$

となる。

次に、この混合ベクトル $\boldsymbol{\zeta}^1$ が自己想起部の初期状態になった場合 ($\mathbf{x}^0 = \boldsymbol{\zeta}^1$) を考えてみよう。次の時刻 ($t = 1$) の自己想起回路の状態は、

$$x_i^1 = \text{sgn} \left(\frac{1}{2} \xi_i^{1,1} + \frac{1}{2} \xi_i^{1,2} + \frac{1}{2} \xi_i^{1,3} + h_i \right), \quad (15)$$

となる。ここで、式 (15) の $1/2$ は回路の状態 \mathbf{x}^0 と連合ベクトル $\boldsymbol{\xi}^{1,\kappa}$, $\kappa = 1, 2, 3$ とのオーバラップ m_0^κ である ($m_0^\kappa = 1/2$)。このとき $h_i = 0$ であれば、 $\mathbf{x}^t = \boldsymbol{\zeta}^1$, $t > 1$ となる。

これらの議論からわかるように、 \tilde{h}_i や h_i の外部入力がなければ、連合の結果として混合ベクトル $\boldsymbol{\zeta}^1$ が得られ、想起は失敗に終る。出力が混合ベクトルでは

なく、記憶した連合ベクトルになるためには、 \tilde{h}_i や h_i の外部入力が必要である。本論文では外部入力を加える時刻の違いにより、次の二つのモデルを考えた。
モデル 1. 相互想起部に時刻 $t = -1$ のとき入力する場合、

$$\text{Prob}[\tilde{h}_i = \pm 1] = \frac{1 \pm a\xi_i^{1,1}}{2}, \quad h_i = 0. \quad (16)$$

モデル 2. 自己想起部に時刻 $t = 0$ のとき入力する場合、

$$\tilde{h}_i = 0, \quad \text{Prob}[h_i = \pm 1] = \frac{1 \pm a\xi_i^{1,1}}{2}. \quad (17)$$

もちろん、 \tilde{h}_i と h_i の両方をともに用いることができる。この場合の定性的な性質は上述の両極端な例の結果から容易に推測できると考えられる。また、時刻 $t > 0$ においても、自己想起部へ外部入力を加えることも考えられる。ここでは議論しないが、本論文の解析手法をこの場合に拡張することは容易であり、その結果も今回の解析結果から容易に推測できると考えられる。

式 (16) と式 (17) から連合ベクトル $\xi_i^{1,1}$ と外部入力 \tilde{h}_i, h_i の相関は、

$$E[\xi_i^{1,1}\tilde{h}_i] = E[\xi_i^{1,1}h_i] = a, \quad (18)$$

となる。本論文では、 a を連合ベクトルと外部入力の類似度と呼ぶ。 a の値の範囲は $0 \leq a \leq 1$ を考える。 $a = 0$ のとき外部入力 \tilde{h}_i, h_i はランダムに ± 1 の値をとり、 $\xi_i^{1,1}$ とは統計的に無相関な入力となる。 $a = 1$ のときは $\xi_i^{1,1}$ に等しくなる。これらのことから、 $a = 0$ では連合ベクトル $\xi_i^{1,1}$ を想起することは困難であるが、 $a = 1$ では容易であることが予想される。従って、 $\xi_i^{1,1}$ を想起できる類似度 a の臨界値 a_c が存在すると予想される。また、この外部入力を加える回路の違いによって、 $\xi_i^{1,1}$ が想起できる臨界値 a_c は異なると予想される。従って、本論文では二つのモデル 1 とモデル 2 の臨界値 a_c を理論的に議論し、どちらの方式が優れているかを議論する。

3. 結 果

3.1 クロストークノイズの分布

まず、相互想起部と自己想起部の二つの回路を統一的に解析しなければならないことを示す。簡単のために、外部入力は $\tilde{h}_i = h_i = 0$ とし、1 対 1 の連合

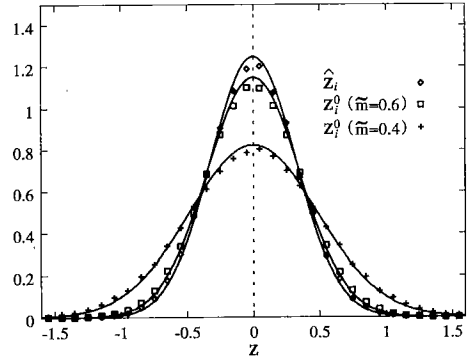


図 2 クロストークノイズの分布. $M = N = 500$, $\alpha = 0.1$
 Fig.2 The distribution of crosstalk noise: $M = N = 500$, $\alpha = 0.1$.

($k = 1$) の場合について考える。ここでは、自己想起部の初期状態に対応する時刻 $t = 0$ におけるクロストークノイズの分布を評価する。自己想起部に $\xi^{1,1}$ と相関をもち、他の連合ベクトルと無相関な状態 \hat{x} が入力されたとする。すなわち、相互想起部がなく、自己想起部のみで連合ベクトル $\xi^{1,1}$ を想起する場合を考える。この場合、他の連合ベクトルからの寄与を表すクロストークノイズ \hat{z}_i ,

$$\hat{z}_i = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=1}^{\alpha N} \xi_i^{\mu,1} \xi_j^{\mu,1} \hat{x}_j, \quad (19)$$

は平均が 0 であり、分散が記憶率 α であるガウス分布に従う [10]。図 2 に計算機シミュレーションと理論で得た記憶率が $\alpha = 0.1$ である場合のクロストークノイズ \hat{z}_i の分布を示す。但し、計算機シミュレーションにおいては、 $M = N = 500$ である。

次に、相互想起部に初期オーバーラップ \tilde{m} の入力が与えられ、相互想起部の出力が自己想起部の初期状態として与えられた場合を考える。この場合、自己想起部の初期状態 x^0 はすべての連合ベクトル $\xi^{\mu,1}$ に依存しているので、その相関を考慮してクロストークノイズを評価する必要がある。そのため、本モデルの巨視的状态方程式を統計神経学 [10], [11] を用いて導出した (付録 1)。図 2 にこの巨視的状态方程式を用いて求めたクロストークノイズ z_i^0 の分布を示す。式 (A.4) より、クロストークノイズ z_i^0 は平均が 0 であり、分散が σ_0^2 である正規分布に従う。式 (5) の初期オーバーラップは $\tilde{m} = 0.4, 0.6$ とした。各点は計算機シミュレーションによる結果であり、実線は理論値であ

る。理論と計算機シミュレーションの結果はよく一致している。図2からわかるように、時刻 $t=0$ でのクロストークノイズ z_i^0 は初期オーバーラップ \tilde{m} に強く依存することがわかる。 $\tilde{m}=0.4$ に比べ、 $\tilde{m}=0.6$ のときのクロストークノイズの分布は、クロストークノイズ \hat{z}_i の分布に非常に類似している。これは $\tilde{m} \rightarrow 1$ ではより顕著に現れる。従って、これまでの絶対容量の解析[8]において、二つの回路を分離して、クロストークノイズの分布を $z_i^0 \approx \hat{z}_i$ とみなして議論することはよい近似であったことを示している。一方、想起過程のダイナミクスを解析する場合においては、 \tilde{m} が小さい値の場合も考慮する必要があるため、分離して議論できないことを示している。ここで重要な点は、相互想起部に入力を加えた場合のクロストークノイズ z_i^0 の分散が、自己想起部に入力を加えた場合のクロストークノイズ \hat{z}_i の分散よりも大きくなるということだけではなく、相互想起部への入力ベクトルのオーバーラップ \tilde{m} にも依存するという点である。この効果は二つの回路を統一的に議論することにより、初めて得られる。

3.2 1対多の連合の想起過程

さて次に1対多の連合を記憶した場合を考えよう。ここでは、 $k=3$ の場合を考える。外部入力を加える時刻の違いによる影響を調べるためにモデル1とモデル2の場合を議論する。図3に記憶率 $\alpha=0.09$ における連合ベクトル $\xi^{1,1}$ の想起過程を示す。相互想起部へはキーベクトル η^1 そのものを入力した ($\mathbf{y}=\eta^1$)。横軸が時間を示し、縦軸が式(10)のオーバーラップ m_i^1 を示す。時刻 $t=-1$ のオーバーラップは式(5)の入力 \mathbf{y} とキーベクトル η^1 のオーバーラップ $\tilde{m}=1.0$ を表す。グラフは下から順に類似度 a が0.1刻みに $a=0.0, 0.1, \dots, 1.0$ となっている。図3(a), (b)は計算機シミュレーションから得られた想起過程の一例である。ニューロン数は $N=M=1000$ ($\beta=1.0$) とした。図3(c)~(f)は理論から得られた想起過程である。近似の次数が $n=1, 4$ の場合について示した。

図3からわかるように、類似度 a がある値 a_c 以上であれば、連合ベクトルが想起され、それ以下の場合には混合ベクトル ζ^1 に収束するか、その他の平衡状態(偽記憶)に収束する。すなわち外部入力を用いることにより、混合ベクトル ζ^1 や偽記憶の想起を避けて、連合ベクトルを想起することができる。臨界値 a_c はモデル1の場合、図3(a)より、計算機シミュレーションでは $a_c > 0.6$ である。図3(c), (e)より、1次

近似の理論 ($n=1$) では $a_c > 0.3$ 、4次近似の理論 ($n=4$) では $a_c > 0.6$ である。また、モデル2の場合、図3(b)より、計算機シミュレーションでは $a_c > 0.3$ である。図3(d), (f)より、1次近似の理論 ($n=1$) では $a_c > 0.1$ 、4次近似の理論 ($n=4$) では $a_c > 0.3$ である。これらの結果より、高次近似の理論はほぼ定量的に計算機シミュレーションの結果を表していると言える。

3.3 モデルの比較

ここで、外部入力によってモデル1とモデル2のどちらの場合がより連合ベクトルを想起しやすいかを検討しよう。ここでは、各ベクトルの次元が $M=N$ ($\beta=1.0$) であるとする。式(A-1)と式(A-4)より、時刻 $t=-1$ と $t=0$ でのクロストークノイズの分散を比較すると、 $\sigma_0^2 > \bar{\sigma}^2$ である。クロストークノイズの点から考えると、クロストークノイズの分散が小さいほど、外部入力の影響を強く受ける。そのため、外部入力を時刻 $t=-1$ に加えたモデル1の方が連合ベクトル $\xi^{1,1}$ が想起されやすくなると思われる。しかしながら、図3より、臨界値 a_c はモデル2の方が小さく、想起能力が高くなることが示唆された。このことは、連合ベクトル $\xi^{1,1}$ のオーバーラップ \tilde{m}, m_0^1 と外部入力の類似度 a を比較することで説明できる。式(13)と式(15)からわかるように、オーバーラップ \tilde{m}, m_0^1 と類似度 a との比は、 $\tilde{m} > m_0^1$ より外部入力を時刻 $t=0$ に加えた場合の方が時刻 $t=-1$ に加えた場合よりも小さくなっている。すなわち、モデル2ではいったんオーバーラップが時刻 $t=0$ で小さくなっているため、外部入力の影響を大きく受け、臨界値 a_c が小さくなると考えられる。

そこで、臨界値 a_c の記憶率 α への依存性を調べた。図4のグラフは横軸が記憶率 α を示し、縦軸が類似度 a であり、連合ベクトル $\xi^{1,1}$ を想起できる臨界値 a_c を示している。臨界値 a_c 以上では $\xi^{1,1}$ が想起され、それ以下では想起は失敗に終わる。図4(a), (b)はそれぞれモデル1とモデル2の場合を示している。相互想起部へはキーベクトルそのものを入力した ($\mathbf{y}=\eta^1, \tilde{m}=1.0$)。計算機シミュレーションは $M=N=1000$ ($\beta=1.0$) を用いた。図のエラーバーは20回試行した際の a_c の標準偏差を表す。理論は次数 $n=1, 2, 3, 4$ のそれぞれの場合について臨界値 a_c を求めたものである。近似の次数が高くなるにつれて、計算機シミュレーションによる結果に漸近することがわかる。図4において $\alpha \approx 0.03$ 付近で、 a_c が一度

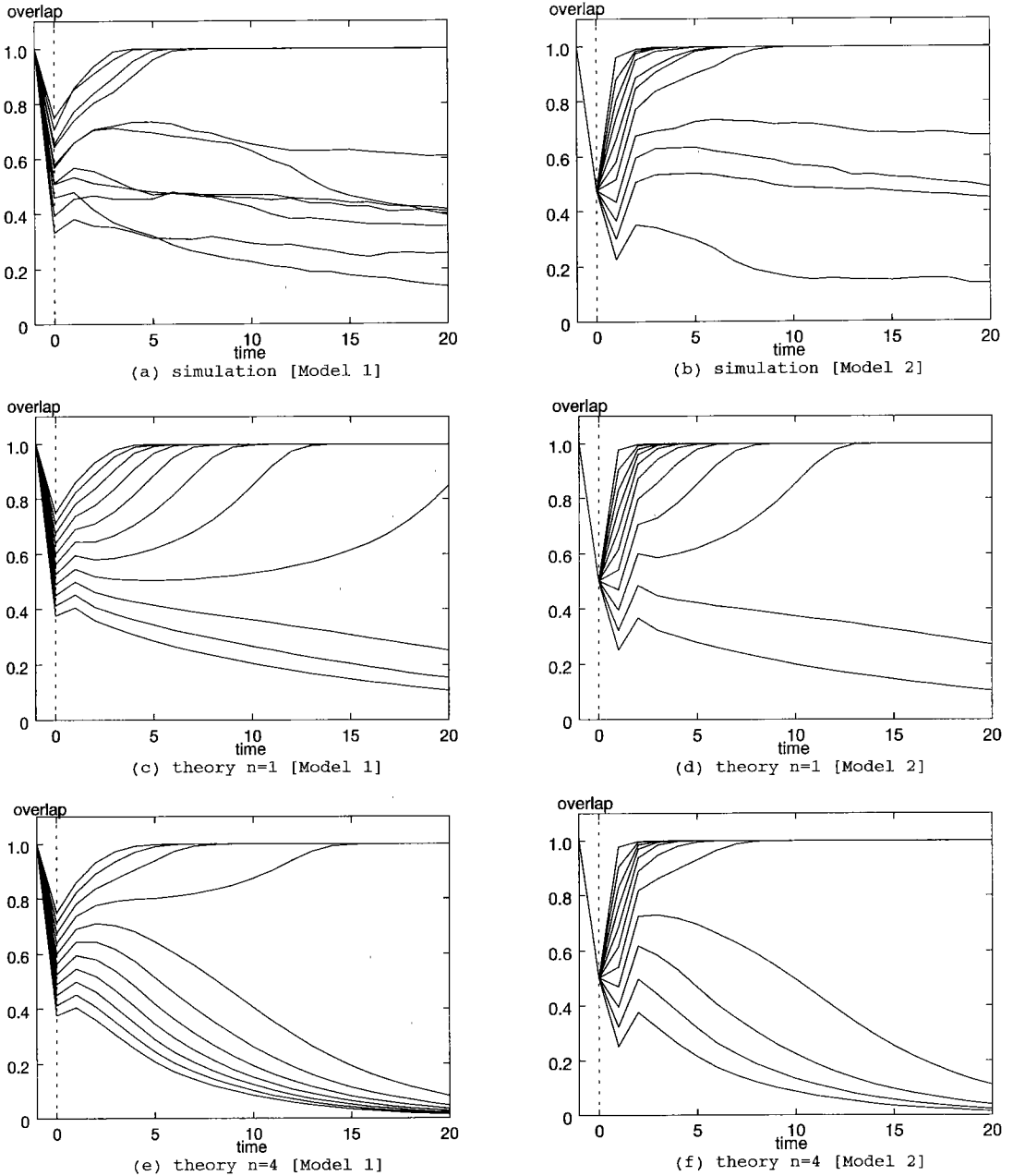


図3 オーバラップ \tilde{m}, m_t^1 の時間変化。記憶率 $\alpha = 0.09$, 1対3の連合 ($k = 3$) の場合。(a), (b) は計算機シミュレーションの結果。ニューロン数が $M = N = 1000 (\beta = 1.0)$ である。(c)~(f) は理論値であり、それぞれ次数が $n = 1, 4$ の場合である。(a), (c), (e) はモデル1の場合を、(b), (d), (f) はモデル2の場合を示している。各曲線は下から類似度 $a = 0.0, 0.1, \dots, 1.0$ の順にオーバラップを示す。

Fig. 3 Overlaps \tilde{m} and m_t^1 : In the case of loading rate $\alpha = 0.09$ and one-to-three associations ($k = 3$). (a), (b) Results of the computer simulations, where $N = 1000$. (c)~(f) correspond to 1st- and 4th-order approximations, respectively. (a), (c), (e) show the result for model 1. (b), (d), (f) for model 2. The curves represent the overlaps in order of similarity $a = 0.0, 0.1, \dots, 1.0$ from the bottom.

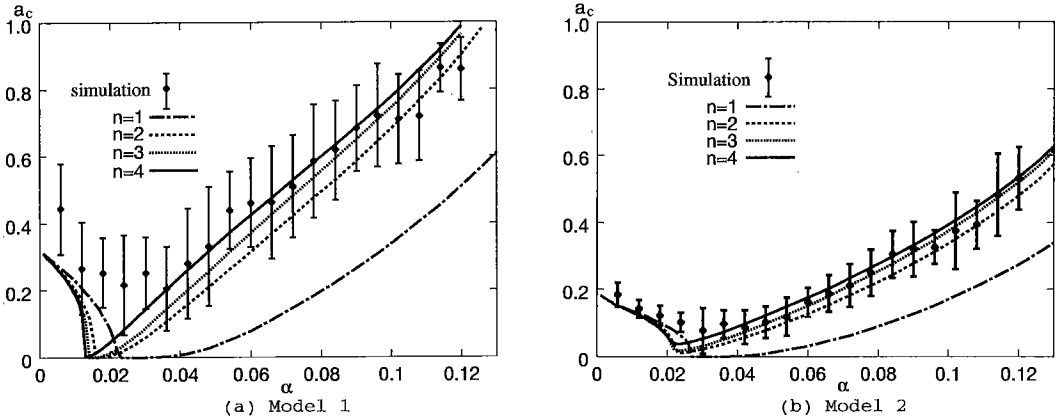


図4 類似度 a の臨界値 a_c . 横軸は記憶率 α , 縦軸は類似度の臨界値 a_c である. 臨界値 a_c 以上では $\xi^{1,1}$ を想起することができる. (a)モデル1:時刻 $t = -1$ に外部入力を加えた場合. (b)モデル2:時刻 $t = 0$ に外部入力を加えた場合. それぞれ, 計算機シミュレーションの結果, 20回平均の値を示す. エラーバーは標準偏差を示す. 素子数は $M = N = 1000 (\beta = 1.0)$. 理論値は次数 $n = 1, 2, 3, 4$ の場合である.

Fig.4 The critical value a_c of similarity a : The abscissa is time t , the ordinate is critical similarity a_c . The associative vector $\xi^{1,1}$ can be recalled if a is more than the critical value a_c . (a) Model 1: An external input is added at $t = -1$, (b) Model 2: An external input is added at $t = 0$, the results of computer simulation show 20 times average of a_c , where $M = N = 1000 (\beta = 1.0)$ and the results of the theory show the case of 1st, 2nd, 3rd, and 4th order approximations, respectively.

小さくなり, その後増加する点が興味深い. これは混合ベクトル $\zeta_i^1 = \text{sgn}(\xi_i^{1,1} + \xi_i^{1,2} + \xi_i^{1,3})$ が $\alpha = 0.030$ で不安定になるからである [13]. 図4より, モデル2の方がモデル1よりも小さい臨界値 a_c をもつため, より連合ベクトル $\xi^{1,1}$ を想起しやすいと言える.

これまでに議論したように, どちらのモデルが連合ベクトルを想起しやすいかは式(4)と式(9)の結合荷重の決め方に依存する. 上記の結果は, 学習過程における式(4)と式(9)の増分を共通の $1/N$ とした場合に成り立つ. ここで不自然ではあるが, 自己想起部と相互想起部で異なった増分を用いると仮定してみよう. この場合は, ここで得られた結果をそのまま適用することができないことは明らかである. しかしながら, このような場合であっても本論文で提案した解析手法を用いて, 仮定に応じた結果を得ることができる.

4. むすび

連想記憶モデル HASP は認知心理学的現象を説明することができるモデルである. 本論文では, HASP の想起過程を議論するために, 最も簡単な相互想起部と自己想起部からなる HASP 型の連想記憶モデルを提案した. 想起過程のダイナミックスを解析する場合

においては, 回路の状態と連合ベクトルとの間の統計的な相関を考慮しなければならない. 特に HASP 型の連想記憶モデルの場合, 相互想起部の出力が自己想起部へ入力されるため, 自己想起部での初期状態が問題となる. そこで, 時刻 $t = 0$ におけるクロストークノイズの分布を解析した. 計算機シミュレーションより時刻 $t = 0$ におけるクロストークノイズの分布は, 自己想起部のみで連合ベクトルを想起する場合と相互想起部を含む場合とは明らかに相違があった. これは相互想起部を含む場合では, 時刻 $t = 0$ におけるクロストークノイズの分布が初期オーバーラップ \tilde{m} に依存するからである. この効果を理論的に解析するために統計神経力学を用いて, 巨視的状态方程式を導出した. その結果, 理論で予測した結果と計算機シミュレーションによる結果はよく一致し, 理論が相互想起部を含め統一的に扱えることが確かめられた.

次に1対多の連合を行う場合について解析した. 外部入力がない場合では, 出力は混合ベクトルになってしまう. そこで, 本論文では各回路に外部入力に加え, その有効性を検討した. 外部入力は連合ベクトルに類似したものから無相関なものまでを考慮し, 連合ベクトルとの相関を類似度とした. その結果, 連合ベクトル

ルが想起できる類似度には臨界値が存在し、臨界値以上では外部入力によって連合ベクトルが想起できることがわかった。また、外部入力を与える回路の違いによりモデル1とモデル2の二つの場合を考えた。計算機シミュレーションと理論による結果より、モデル2ではオーバーラップがいったん時刻 $t = 0$ で小さくなるため外部入力の影響を強く受けることがわかった。従って、モデル2の方がより小さな臨界値をもち、連合ベクトルを想起しやすいという結果を得た。この結果は3.3で述べたように、学習過程での増分が $1/N$ である場合に成り立ち、増分が異なればこの結果は変わり得る。

今回は相互想起部の入力の次元 M と自己想起部のニューロン数 N を同じにした場合 ($\beta = M/N = 1.0$) を議論したが、外部入力の影響はオーバーラップ \tilde{m} と β に依存するので、今後、 β の値を変え、シナプス結合数 M とニューロン数 N の関係について検討する必要がある。また、入力 y が不完全な場合 ($\tilde{m} < 1.0$) についても考慮する必要がある。

文 献

[1] K. Nakano, "Associatron—A model of associative memory," IEEE Trans. Sys. Man Cyb., vol.SMC-2, no.3, pp.380–388, 1972.
 [2] T. Kohonen, "Correlation matrix memories," IEEE Trans. Comput., vol.C-21, no.4, pp.353–359, 1972.
 [3] Y. Hirai, "A Model of Human Associative Processor (HASP)," IEEE Trans. Sys. Man. Cyb., vol.SMC-13, no.5, pp.851–857, 1983.
 [4] 平井有三, "連想記憶モデル HASP による手続き的知識の表現," 信学論 (D-II), vol.J69-D-II, no.11, pp.1743–1753, 1986.
 [5] 平井有三, 上條憲一, "連想記憶モデル HASP の心理学的妥当性について," 信学論 (D-II), vol.J70-D-II, no.1, pp.175–184, 1987.
 [6] Y. Hirai and Q. Ma, "Modeling the process of problem-solving by associative networks capable of improving the performance," Biol. Cybern., 59, pp.353–365, 1988.
 [7] Q. Ma and Y. Hirai, "Modeling the acquisition of counting with an associative network," Biol. Cybern., 61, pp.271–278, 1989.
 [8] 川村正樹, 平井有三, "連想記憶回路 HASP の記憶容量の解析," 信学論 (D-II), vol.J78-D-II, no.11, pp.1692–1700, 1995.
 [9] 川村正樹, 岡田真人, 平井有三, "連想記憶モデル HASP の統計神経力学," 日本神経回路学会第7回全国大会論文集, pp.257–258, Sept. 1996.
 [10] S. Amari and K. Maginu, "Statical neurodynamics of associative memory," Neural Networks, vol.1, pp.63–73, 1988.
 [11] M. Okada, "A hierarchy of macrodynamical equations

for associative memory," Neural Networks, vol.8, no.6, pp.833–838, 1995.

[12] M. Okada, "Notions of Associative memory and sparse coding," Neural Networks, vol.9, no.8, pp.1429–1458, 1996.
 [13] D.J. Amit, H. Gutfreund, and H. Sompolinsky, "Information storage in neural networks with low levels of activity," Phys. Rev. A, vol.35, no.6, pp.2293–2303, 1987.

付 録

1. 巨視的状態方程式

想起過程のダイナミックスを解析するために、本モデルの巨視的状態方程式を統計神経力学 [10], [11] を用いて導出する。この理論では、回路全体の状態をいくつかの巨視的状態変数を用いて記述する。また、想起過程は巨視的状態方程式で表される。Okada [11] は Amari-Maginu 理論 [10] を発展させ、想起過程のダイナミックスをほぼ定量的に説明することに成功した。この理論は階層的な構造をもち、 $n - 1$ 時刻離れた時刻までのクロストークノイズの時間相関を考慮している。 $n = 1$ の場合 Amari-Maginu 理論に一致する。本論文では $n \geq 2$ の巨視的状態方程式を高次近似の理論と呼ぶ。

1.1 相互想起部

相互想起部の出力に関して、巨視的状態方程式を導出する。式 (7) より、時刻 $t = -1$ でのクロストークノイズの平均は 0 であり、その分散は、

$$\tilde{\sigma}^2 = E [\tilde{z}_i^2] = \alpha\beta, \tag{A.1}$$

である。ここで、 α は $\alpha = pk/N$ である。時刻 $t \geq 0$ でのクロストークノイズは、統計神経力学的手法 [10], [11] に基づき平均が 0 で、分散が σ_t^2 である正規分布に従うものと仮定する。

時刻 $t = 0$ での巨視的状態方程式は、

$$\tilde{Y}(z) \equiv \text{sgn} \left(\sum_{\kappa=1}^k \beta \tilde{m} \xi^{1,\kappa} + \tilde{\sigma}z + \tilde{h} \right), \tag{A.2}$$

と定義すると、

$$m_0^\kappa = \int Dz \left\langle \xi^{1,\kappa} \tilde{Y}(z) \right\rangle_{(\xi)}, \tag{A.3}$$

$$\sigma_0^2 = \alpha + \tilde{\sigma}^2 U_0^2, \tag{A.4}$$

$$U_0 = \frac{1}{\tilde{\sigma}} \int Dz z \left\langle \tilde{Y}(z) \right\rangle_{(\xi)}, \tag{A.5}$$

と表される。ここで、 $\langle \cdot \rangle_{(\xi)}$ は $\xi^{1,1}, \xi^{1,2}, \dots, \xi^{1,k}$ に

関する平均を表す。\$Dz\$ は \$(2\pi)^{-1/2} \exp(-z^2/2)dz\$ と定義する。

1.2 自己想起部

次に、自己想起部に関して巨視の状態方程式を導出する。時刻 \$t \ge 0\$ での \$n\$ 次近似の巨視的状态方程式[11]は、

$$Y(z) \equiv \text{sgn} \left(\sum_{\kappa=1}^k m_t^\kappa \xi^{1,\kappa} + \sigma_t z + h\delta_{t,0} \right), \quad (\text{A.6})$$

と定義すると、

$$m_{t+1}^\kappa = \int Dz \left\langle \xi^{1,\kappa} Y(z) \right\rangle_{(\xi)}, \quad (\text{A.7})$$

$$\sigma_{t+1}^2 = \alpha + U_{t+1}^2 \sigma_t^2 + 2\alpha \sum_{\tau=t-n+1}^t q_{t+1,\tau} \prod_{r=\tau+1}^{t+1} U_r, \quad (\text{A.8})$$

$$U_{t+1} = \frac{1}{\sigma_t} \int Dz z \left\langle Y(z) \right\rangle_{(\xi)}, \quad (\text{A.9})$$

と表される。但し、\$\sigma_{-1} \equiv \tilde{\sigma}\$ である。

ここで、\$q_{t+1,\tau}\$ は時刻 \$t+1\$ と \$\tau\$ の間の状態の相関であり、

$$q_{t,\tau} \equiv E \left[x_t^t x_\tau^\tau \right], \quad (\text{A.10})$$

と定義する。\$q_{t+1,\tau}\$ は、

$$q_{t+1,\tau} = \int Dc \int Da \int Db \left\langle F_t(a,c) F_{\tau-1}(b,c) \right\rangle_{(\xi)}, \quad (\text{A.11})$$

$$F_t(z,c) \equiv \text{sgn} \left(\sum_{\kappa=1}^k \xi^{1,\kappa} m_t^\kappa + \sigma_t(d_0 z + d_1 c) + h\delta_{t,0} \right),$$

で求めることができる。ここで、

$$d_0 = \sqrt{1 - \frac{C_{t,\tau-1}}{\sigma_t \sigma_{\tau-1}}}, \quad d_1 = \sqrt{\frac{C_{t,\tau-1}}{\sigma_t \sigma_{\tau-1}}}, \quad (\text{A.12})$$

であり、\$C_{t,\tau-1}\$ は時刻 \$t\$ と \$\tau-1\$ (\$\tau \ge 1\$) の間のクロストークノイズの相関を表す。\$C_{t,\tau-1}\$ は次の式で表される、

$$C_{t,\tau-1} = 0, \quad (\tau = t - n + 1, n \ge 1), \quad (\text{A.13})$$

$$C_{t,\tau-1} = \alpha q_{t,\tau-1} + U_t C_{t-1,\tau-1}, \quad (\tau = t - n + 2, n \ge 2), \quad (\text{A.14})$$

$$C_{t,\tau-1} = \alpha q_{t,\tau-1} + U_t U_{\tau-1} C_{t-1,\tau-2} + \alpha \sum_{\eta=\tau-n+1}^{\tau-2} q_{t,\eta} \prod_{r=\eta+1}^{\tau-1} U_r + \alpha \sum_{\eta=t-n+1}^{t-1} q_{\eta,\tau-1} \prod_{r=\eta+1}^t U_r, \quad (t - n + 3 \le \tau \le t, n \ge 3). \quad (\text{A.15})$$

但し、時刻 \$t = -1\$ との相関は \$q_{t+1,-1} = 0\$ である。
(平成9年6月16日受付, 10月3日再受付)



川村 正樹 (学生員)

平6筑波大・第三・情報卒。平8同大大学院修士課程了。現在同大学院博士課程在学中。ニューラルネットワーク、記憶に関する研究に従事。日本神経回路学会会員。



岡田 真人 (正員)

昭60阪市大・理・物理卒。昭62阪大大学院前期課程了。同年三菱電機(株)入社。平3阪大・基礎工・生化工助手。平8年10月科学技術振興事業団・川人学習動態脳プロジェクト研究員。博士(理学)。物性物理、半導体レーザ、神経回路モデル、計算論的神経科学の研究・開発に従事。平6,8神経回路学会研究賞受賞。



平井 有三 (正員)

昭45慶大・工・電気卒。昭50同大大学院博士課程了。同年富士通株式会社入社。昭53より筑波大学電子・情報工学系勤務。現在同大教授。工博。ニューラルネットワーク関係の研究に従事。昭54本学会学術奨励賞受賞。情報処理学会、日本神経回路学会、IEEE、INNS各会員。