

自己相関連想記憶の想起過程とその改良

正 員 森田 昌彦[†] 正 員 吉澤 修治[†] 正 員 中野 馨[†]

Analysis and Improvement of the Dynamics of Autocorrelation Associative Memory

Masahiko MORITA[†], Shuji YOSHIZAWA[†] and Kaoru NAKANO[†], Members

あらまし 自己相関型の連想記憶は、神経回路による連想記憶の中で最も単純かつ基本的なものである。しかし、その想起のダイナミックスに関する性質には理解しにくいものがあり、理論的解析も十分とは言えない。本論文は、数値実験を通じて連想記憶の想起過程について考察し、ダイナミックスの改良によって回路網の力学的性質が大幅に改善されることを示すものである。また、その具体的な方法として「部分反転法」を提案する。これは、回路網の状態が遷移するたびに一部の細胞の出力を反転させる（離散時間の場合）、あるいは従来と異なった形の出力関数を用いる（連続時間の場合）という方法である。これにより、想起特性が大きく向上し、記憶容量も増加する。このことは、同時に連想記憶の力学構造を理解する手掛りとなるものであり、最適化問題への応用等を考える上でも重要な示唆を与える。

1. ま え が き

自己相関行列による連想記憶は、神経回路による連想記憶モデルの中で最も基本的なものであり、歴史的にも最初に提案された⁽¹⁾。また、その性質について以前から解析がなされてきたが^{(2),(3)}、最近ではスピングラスとの類似性も一因となり、記憶容量や平衡状態の分布といった研究が盛んである⁽⁴⁾⁻⁽⁶⁾。

しかしながら、連想記憶神経回路網の力学的性質を明らかにするという点から見たとき、これらの解析は決して十分なものではない。むしろ、この問題に対する理解はまだ不足していると言える。特に、Amariら⁽⁷⁾によって報告されている想起の際の興味深い現象は、一見奇妙であるだけでなく、連想記憶の能力を低下させる原因ともなっている。

本研究の主な目的は、この現象を中心に自己相関連想記憶の想起過程に対して検討を加え、ダイナミックスの改良によって連想記憶モデルとしてより優れた性質をもつ力学系を実現することにある。同時に、それを通じて連想記憶の力学について考察し、理解を深めることを目指している。

以下では、まず従来モデルについて、これまでに知られている性質を簡単にまとめた後、想起過程における問題点を明らかにし、その原因を推察する。次いで、具体的な改善策として「部分反転法」を提案し、この方法により実際に連想能力が大幅に向上することを示す。最後に、実験結果などをもとに、連想記憶のダイナミックスに関していくつかの考察を行う。

2. 連想記憶のダイナミックス

2.1 自己相関連想記憶

本論文では、主として次のような最も単純なタイプの連想記憶モデルを扱う。

まず、 n 個の細胞からなる神経回路網を考え、 j 番目の細胞から i 番目の細胞への結合荷重を w_{ij} とする。また、時刻 t における i 番目の細胞の出力を $x_i(t)$ 、回路網の状態を

$$X_t = (x_1(t), x_2(t), \dots, x_n(t))'$$

で表す（' は転置を意味する）。

ここで、 x_i が 1, -1 の 2 値をとり、離散時間で同期的に動作するモデルを用いると、系のダイナミックスは

$$x_i(t+1) = \operatorname{sgn}\left(\sum_{j=1}^n w_{ij} x_j(t)\right) \tag{1}$$

または行列 $W = \{w_{ij}\}$ を用いて

[†] 東京大学工学部計数工学科, 東京都
Faculty of Engineering, The University of Tokyo, Tokyo, 113
Japan

$$X_{t+1} = \text{sgn}(WX_t) \quad (2)$$

という形で記述される。但し $\text{sgn}(u)$ は $u > 0$ のとき 1, $u \leq 0$ のとき -1 をとる関数 (ベクトルに対しては成分ごとに作用する) である。

この回路網に記録させるパターンを S^1, S^2, \dots, S^m としたとき、自己相関連想記憶では、

$$W = \frac{1}{n} \sum_{\mu=1}^m S^\mu S^{\mu'} = \frac{1}{n} \Sigma \Sigma' \quad (3)$$

を荷重行列として用いる。ここで Σ は $n \times m$ の行列 (S^1, S^2, \dots, S^m) である。また、 $S^\mu = (s_1^\mu, s_2^\mu, \dots, s_n^\mu)$ は系のとり得る 2^n 個の状態の中からランダムに選ばれ、近似的に

$$\frac{1}{n} \sum_{i=1}^n s_i^\mu s_i^\nu = 0 \quad (\mu \neq \nu) \quad (4)$$

が成り立っているものとする。

想起の際は、想起入力 X_0 を回路網の初期状態として与え、式(2)に従って次々に状態を更新していく。状態が遷移する度にエネルギー⁽⁸⁾

$$E = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} x_i x_j \quad (5)$$

が減少していくため、十分に時間が経つと状態 X_t は初期状態に応じて定まるある平衡状態 $\bar{X}(X_0)$ に到達する[†]。このとき、 X_0 が記憶パターンの一つ S にある程度似ていれば、

$$\bar{X}(X_0) = S \quad (6)$$

となることが期待される。逆に、式(6)が成り立つような X_0 の集合を S の引込み領域 (basin of attraction) と呼び、 $B(S)$ で表す。

このほか、以下では次の2種類の連想記憶モデルにも触れる。

一つはいわゆる一般逆行列を用いた連想記憶⁽⁹⁾ (以下、一般逆行列モデルと呼ぶ) で、自己相関行列を用いる代わりに

$$W = \Sigma (\Sigma' \Sigma)^{-1} \Sigma' \quad (7)$$

によって荷重行列を構成する (相関学習の代わりに直交学習⁽²⁾ によって w_{ij} を求める) 点だけが上述のモデルと異なる。

もう一つは、層状の神経回路を用いた連想記憶⁽¹⁰⁾ (以下、層状モデルと呼ぶ) である。そのダイナミクスは

$$X_{t+1} = \text{sgn}(W_t X_t) \quad (8)$$

$$W_t = \frac{1}{n} \sum_{\mu=1}^m S_{t+1}^\mu S_t^{\mu'} \quad (9)$$

で表される。 S_t^μ がすべて独立であり、時刻ごとに別々の荷重行列 W_t を用いる点が自己相関モデルと異なっている。

2.2 平衡点の分布

今、細胞数に対する記憶パターンの割合 (以下、記憶率と呼ぶ)

$$r = \frac{m}{n}$$

を一定したまま、 $n \rightarrow \infty$ にした場合を考える。よく知られているように、このとき S が平衡状態である確率は0である^{(2),(3),(7)}。つまり、式(6)はどんな初期状態 X_0 に対しても成り立たない。にもかかわらず上述の回路が連想記憶モデルとして動作するのは、 r や X_0 が適当な範囲にあれば、 $\bar{X}(X_0)$ は S に非常に近いものになるからである⁽⁵⁾。

そこで、平衡点がどう分布しているかという点にまず興味をもたれるが、これに関して次のことが解析的に知られている⁽⁶⁾。

まず、二つの状態 X^1 と X^2 の間の距離 d を正規化したハミング距離

$$d(X^1, X^2) = \frac{1}{2n} \sum_{i=1}^n |x_i^1 - x_i^2| \quad (10)$$

で定義する。また、ある記憶パターン S に注目したとき、そこからの距離を $\rho = d(X, S)$ で表すことにする。このとき、ほとんどすべての状態が $\rho \approx 0.5$ のところに集中していることに注意されたい。

平衡点は $\rho = 0.5$ のところに最も多く分布するが、 ρ が小さくなるにつれてその数は急速に減少し、あるところで0になる。しかし、記憶率 r がある値 r_e (≈ 0.11) 以下のとき、更に ρ が小さいところに多数の平衡点が集中的に存在する領域がある。すなわち、平衡点の存在域は、 S と無相関 ($\rho = 0.5$) の状態を中心とする領域と、 S の近傍 ($\rho < 0.005$) の領域の二つに分かれている。

この S 近傍の領域は、 r が大きくなるにつれて S から遠ざかり、 $r > r_e$ ではもう一つの領域とつながってしまう。また、一般に r が大きいときほど、 S から最も近くにある平衡点までの距離が大きくなる。

なお、ここでは単に平衡点と呼んだが、この解析は正確にはエネルギー極小点の分布を扱っている。荷重行列 W の対角成分 w_{ii} が0のとき両者は一致する (そのためしばしば $w_{ii} = 0$ とされる) が、 $w_{ii} > 0$ のときにはエネルギーが極小でなくてもこう配が十分小さければ平衡状態となる。式(3)によって W を構成した場合、

[†] 厳密には、これは回路が非同期的に動作する場合に成り立つことで、同期的に動作する場合には周期2で振動することがある。

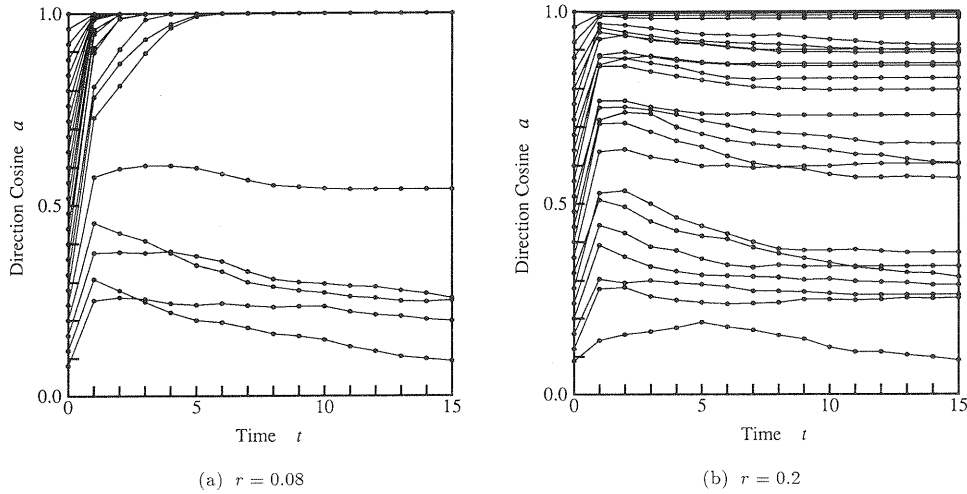


図1 方向余弦の時間変化
Fig. 1 Temporal change in direction cosines.

$w_{ii}=r$ であるから、 r が小さければ対角成分は平衡点の分布やダイナミクスにほとんど影響を及ぼさない。また、 r が大きくなるにつれて $w_{ii}=0$ とした場合との違いが現れてくるが、これは力学系の本質的な構造の違いを反映したものとは言えない。

2.3 想起過程

2.2に述べた平衡点の分布は、回路網の力学を考える上で重要であるが、これだけでは想起のダイナミクスに関する情報はあまり得られない。例えば、自己相関連想記憶では $r=r_c \approx 0.15$ が「記憶容量」とであるとされているが^{(4),(8)}、この値は r_c と一致しない。また、 $r > r_c$ の場合にも、 S のある程度近傍(例えば $r=0.2$ のときで $\rho < 0.03$)にはエネルギー極小点が存在するから、想起がうまくいかないのは S の周りに平衡状態がないためではない。

このことからわかるように、想起のダイナミクスは、最も近い平衡点に向かって状態が遷移していく、というような単純なものではない。従って、それぞれの平衡点がどのような引込み領域をもっているかを知る必要があるが、これは解析的にはもちろん、数値実験でも求めるのは困難である。そこで、ある初期状態 X_0 を選んだときに、 X_t が時間と共にどのように遷移していくかを見ることにする。

まず、記憶パターン S^1, S^2, \dots, S^m およびその反転パターン $-S^1, -S^2, \dots, -S^m$ の中で最も想起入力 X_0 と近いものを S とする。 S からの距離 ρ の代わりに、ここでは X と S の間の方向余弦(direction cosine)

$$a = \frac{1}{n} \sum_{i=1}^n x_i s_i \quad (11)$$

を考え、時刻 t での方向余弦を a_t とする。 $a=1-2\rho$ であるから、 a_t が大きいほど X_t が S に近いことを意味する。

この a の時間変化については既に興味深い現象が報告されているが⁽⁷⁾、本研究と関係が深いのでここで触れておく。

図1は、初期状態 X_0 と S との距離を変えつつ[†]、 a_t をプロットしていったグラフである。(a)は $r=0.08$ 、(b)は $r=0.20$ の場合で、いずれも $n=1000$ で実験した。この実験その他の結果から、以下のことが言える。

(1) $r=0.08$ の場合、 a_0 がある値 a_c より大きければ、比較的短い時間で a_t はほぼ1になる。すなわち、 X_t は S の極めて近傍にある平衡状態の一つ \hat{S} に達する。しかし、 $a_0 < a_c$ のときには a_t はある程度までしか大きくなり、 X_t は S から離れたところにある平衡点にトラップされてしまう。また、しばしば a_t はある程度大きくなった後に減少する。

(2) S の近傍の平衡点 \hat{S} の引込み領域全体をもって S の引込み領域 $B(S)$ とみなすことにすれば、 $B(S)$ はある意味で図2に模式的に示すような形をしていると考えられる⁽⁷⁾。図で内側に切れ込んでいる部分は n 次元空間のうちのごく一部を占めるに過ぎないが、 $X_0 \in B(S)$ のとき X_t はこの切れ込みへ入っていく。

[†] 正確には、 S からの距離が ρ である状態の中から一つをランダムに選んで X_0 とする。

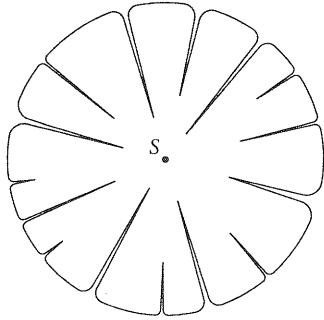


図2 Sの引込み領域の模式図
Fig. 2 Schematic basin of attraction $B(S)$.

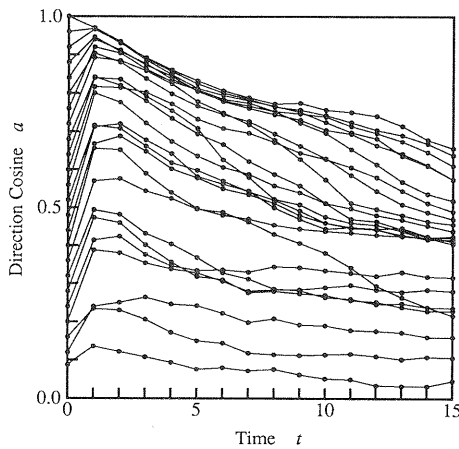
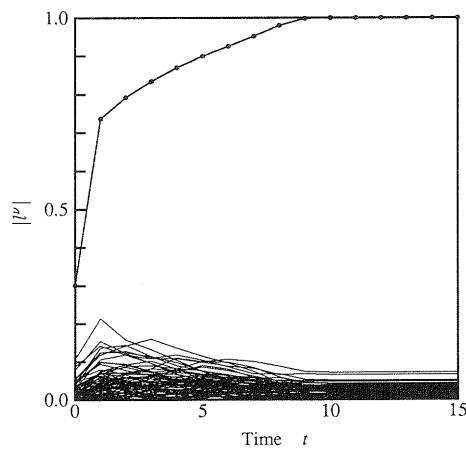
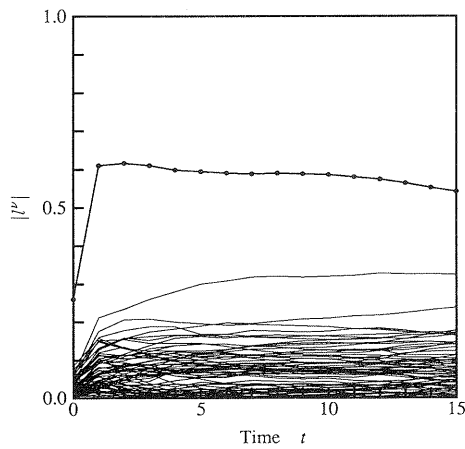


図3 $w_{ii}=0$ の場合の想起過程 ($r=0.2$)
Fig. 3 Recalling process of the network with $w_{ii}=0$, $r=0.2$.



(a) $a_0 = 0.30$



(b) $a_0 = 0.26$

図4 $|l^\nu|$ ($\nu=2, 3, \dots, m$) の推移 ($r=0.08$)
Fig. 4 Transition of $|l^\nu|$ with $r=0.08$.

(3) a_0 に関する臨界値 a_c は記憶率 r に依存し, r が大きいほど大きくなる (図 15 参照). しかし, r が r_c を超えるあたりからしきい現象がはっきりとしなくなり, $r > r_c$ で完全になくなる.

(4) $r > r_c$ のとき, a_0 がいくら大きくても X_t が単調に S に近づいていくことはない. ほとんどの場合, a は一度大きくなってから小さくなる. この傾向は w_{ii} を 0 にしてみるといっそう明確になり, $X_0=S$ としたときでさえ X_t は S から遠く離れていってしまう (図 3).

なお, 層状モデルでも同様な現象が見られるが^{(10),(11)}, いくつかの点で自己相関モデルの場合と異なる. 特に, 次の 2 点が後の議論と関連して重要である.

(1) $r \approx 0.27$ までしきい現象がある. この値を記憶容量 r_i とみなすことができるが, r_i は自己相関モデルの容量 r_c に比べてかなり大きい.

(2) $a_0 < a_c$ のとき, a_t は $t > 1$ で指数関数的に小さくなっていき, 最終的にはほぼ 0 となる. また $r > r_i$ のときには, 初期値に関係なく a_t は 0 に近づく.

2.4 考 察

以上のような性質に対して, 統計神経力学の手法を用いて理論的解析が試みられているが^{(7),(11)}, 自己相関モデルについてはまだ満足すべき結果が得られていない. そこで, ここでは想起の過程を少し別の観点から眺めてみる.

まず, S 以外の記憶パターン S^ν と X との方向余弦を考え,

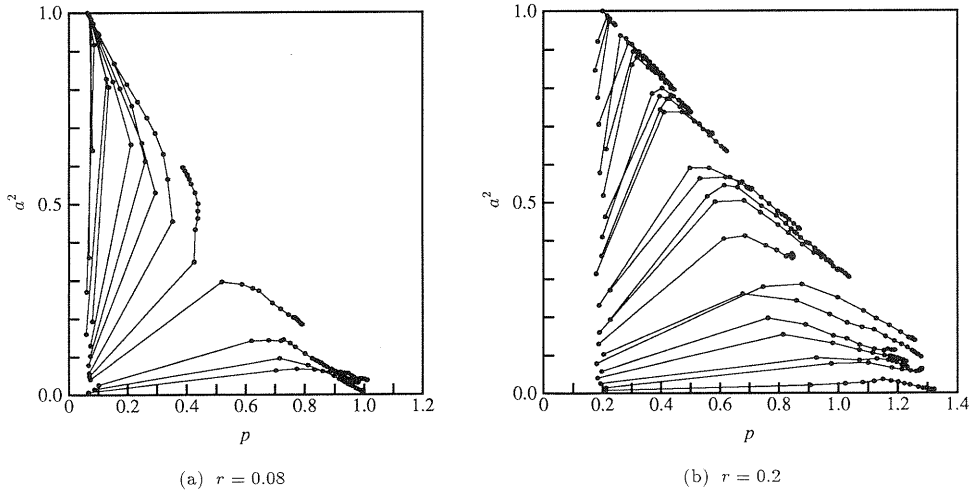


図5 状態遷移に伴う p と a^2 の変化
Fig. 5 Change in p and a^2 with state transitions.

$$l^\nu = \frac{1}{n} \sum_{i=1}^n x_i s_i^\nu \quad (12)$$

とおく。今、 $S = S^1$ としても一般性を失わないので、以下では $\nu = 2, 3, \dots, m$ であるとする。

図4に、 $|l^\nu|$ および a (太線) の遷移の様子を示す。(a)は正しく想起できた場合で、 $|l^\nu|$ は $t=1$ で大きくなった後は減少する傾向にある。これに対して(b)は正しい想起ができなかった場合であり、 $|l^\nu|$ は $t > 1$ でも減少せず、むしろ全体的に増加する傾向にあることがわかる。

次に、このような現象を理解するため、 l^ν の2乗和

$$p = \sum_{\nu=2}^m (l^\nu)^2 \quad (13)$$

を考える。 X_0 をランダムに選んだとき、 l^ν ($\nu = 2, 3, \dots, m$) はそれぞれ平均0、分散 $1/n$ の独立な正規分布に従うから、 $p_0 \approx r$ である (p の添字は時刻を表す)。また、 X_t が S に十分近いときにも $p_t \approx r$ が成り立つ。更に、 $a_1 > a_0$ が言える⁽⁷⁾ のと同様に $p_1 > p_0$ となることが示されるから、正しい想起ができるためには p_t は $t > 2$ で小さくなる必要があることがわかる。

ところが、

$$\begin{aligned} a^2 + p &= \frac{1}{n^2} \sum_{\mu=1}^m (X' S^\mu)^2 \\ &= \frac{1}{n^2} X' \left(\sum_{\mu=1}^m S^\mu S^{\mu'} \right) X \\ &= \frac{1}{n} X' W X \end{aligned}$$

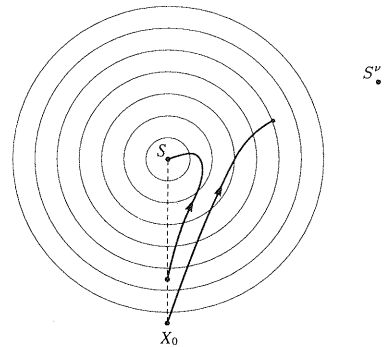


図6 従来の想起過程の模式的表現
Fig. 6 Schematic representation of the conventional recalling process.

$$= -\frac{2}{n} E \quad (14)$$

であるから、 p_t が減少する方向は一般にはエネルギーが増える方向となる。従って、 a^2 だけが大きく増加する領域(そのような領域は S の近傍に限られる)に X_t がある場合を除き、 p は状態遷移の度に増大する。従って、いったん p_t が大きくなり過ぎてしまうと、想起は非常に困難になると考えられる。

実際に a^2 と p とを状態変数として、 X_t の遷移の様子を描いたグラフが図5である。(a)は $r=0.08$ の場合で、前述のしきい現象が明確に現れている。この図からわかるように、 p の値がある程度以上大きくなれば、 $p_t \rightarrow r$ となって想起がうまくいく。逆に p があまり大きくなってしまうと、言い換えれば X_t の S へ

の近づき方に比べてエネルギーの減少が大き過ぎると、正しい想起ができない。また(b)は $r=0.2$ の場合であるが、 k_i は増加する一方で減少しないことがわかる。

以上の結果をまとめるならば、 X_i は X_0 から S へ「まっすぐ」に向かうのではなく、他の記憶パターン S^v に引っ張られて「斜め」に近づいていく、ということができよう(図6)。もちろん実際の状態空間は n 次元であるから、この図式にはかなり無理がある。しかし、現状ではこのような直観的な理解に頼らざるを得ないことも事実であろう。

3. 想起ダイナミックスの改良

2.3 で述べたような現象は、仮に式(4)が厳密に成り立っていても、やはりある程度生じる。すなわち、記憶パターン同士が完全には直交していないことだけが、想起がうまくいかない原因ではない。従って、一般逆行列モデルのように、ある意味で記憶パターンを直交化して記憶しても、 X_i が S 以外のエネルギー極小状態にトラップされるという問題は避けられない。

一方、自己相関モデルの記憶容量 r_c が層状モデルの容量 r_i の半分程度しかないことは、 $r > r_c$ のとき正しい想起ができないのは荷重行列 W から記憶情報が失われるためではない、ということを示唆している。つまり、前章で議論したような想起過程の特性が、回路網の能力を低下させる大きな原因となっているのである。

従って、想起のダイナミックスの問題は、荷重行列や学習法の改良の問題とはある程度切り放して考えることが必要である。また、連想記憶能力を向上させるには、従来のように単に W を改善するだけでは限界があると予想される。

そこで、ここでは W はそのままにしてダイナミックスを改良するというアプローチをとることにする。これは、優れた連想能力を得る手段であると同時に、連想記憶の力学を調べる一つの手段でもある。

もちろんダイナミックスを変えるとといっても無制限ではなく、この種の連想記憶の長所は残しておく必要がある。つまり、各細胞(素子)は独立かつ並列的に動作しなくてはならないし、一様性も保たれた方がよい。また、想起入力 X_0 以外に、 a_0 の値など特別な情報を与えるべきではない。

このような条件を満たしつつ、従来よりも優れた連想記憶能力を実現する方法の一つが、以下に述べる「部分反転法 (partial reverse method)」である。

3.1 部分反転法

回路網の状態がエネルギー極小値にトラップされるのを防ぐ方法としてよく用いられるのが、ノイズ(熱的揺らぎ)を加えるという方法である。もし S の引込み領域が本当に図2のような形をしているのならば、適当なノイズを加えることによって X を引込み領域の内部に飛び込ませることが可能なように思われる。しかし、ただランダムなノイズを加えてもうまくいかないことは、図5や図6などからも明らかであろう。

しかしながら、同時にこれらの図は、ランダムではなく ρ を減らす方向に X を動かしてやればよいことを示唆している。すなわち、状態が遷移するたび、 S からの距離はあまり変わらないようにしつつ、その他パターン S^v との距離を遠ざける(相関を減らす)方向に戻す、ということを繰り返せば、 X_i は S の十分近傍まで近づいていくと期待される(図7)。

これが部分反転法の基本的な考え方である。 X をある方向に動かすことは特定の細胞の出力を反転させることであるから、どのような細胞の出力を反転させるかが大切であるが、 $x_i = s_i$ が成り立っている可能性が低く、反転によって全体のエネルギーがなるべく大きくなるような x_i を選ばばよいと考えられる。

その具体的アルゴリズムは次に示すとおりである。これが唯一の方法ではないし、最適であるという保証もないが、単純な割に大きな効果が得られる。

3.2 アルゴリズム

まず、想起の1ステップをフェーズIとフェーズIIとに分ける。前者は従来と同様の想起を行うフェーズで、後者が出力の一部を反転するフェーズである。

具体的には、フェーズIで入力の荷重和

$$u_i = \sum_{j=1}^n w_{ij} x_j(t) \quad (15)$$

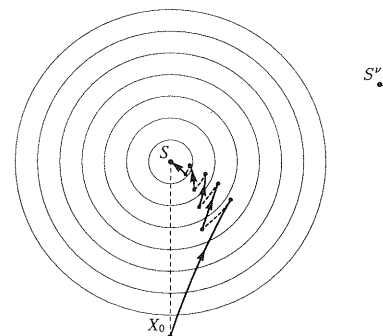


図7 部分反転法の考え方

Fig. 7 Illustration of the partial reverse method.

を求める。式の上では必要ないが、ここでいったん

$$x_i = \text{sgn}(u_i) \quad (16)$$

によって X を更新するものとすれば、フェーズ I でのダイナミクスは従来と等価になる。

フェーズ II では、まず式(16)で求められた u_i のうち、 $|u_j| > h$ ($h > 0$) であるような x_j についての荷重和を求め、その値を v とする。式で表すと

$$v_i = \sum_{j=1}^n w_{ij} \phi(u_j) \quad (17)$$

となる。但し、

$$\phi(u) = \begin{cases} -1 & (u < -h) \\ 0 & (-h \leq u \leq h) \\ 1 & (u > h) \end{cases} \quad (18)$$

である。

そうした上で、次式によって改めて X_{i+1} を求める (λ は正の定数)。

$$x_i(t+1) = \text{sgn}(u_i - \lambda v_i) \quad (19)$$

これにより、フェーズ I でいったん求めた x_i のうち、

$$0 < \frac{u_i}{v_i} < \lambda \quad (20)$$

が成り立つものの値が反転することになる[†]。

以下、フェーズ I と II の操作を交互に行って X を順次更新していき、平衡状態に達するまで繰り返す。但し、平衡状態に必ず落ち着くという保証はない。

このダイナミクスにおいて、パラメータ h および λ の選び方が重要である。理論的な最適値は求められていないが、数値実験によれば λ は 2.7 前後、 h は 1.8~2.0 程度にするのがよい。また、記憶率 r がわかっているのであれば

$$h = 1 + r + k\sqrt{r} \quad (21)$$

に従って h を変える方がよい ($k \approx 2$)。

$X = S$ のとき、 $s_i u_i$ の値は平均 $1+r$ 、分散 r の正規分布に従うから、パラメータを適切に選ぶと、 X_i が S に十分近づいたとき $|u_i| > h$ となる細胞の割合は非常に小さくなる。従って、正しく想起できたときにはフェーズ II での反転がほとんど起こらない。また、その逆が常に成り立つという保証は得られていないが、数値実験の結果から見た限りでは、平衡状態で反転が起こらなければ正しい想起ができたものと判断してよいと思われる。

3.3 連続時間モデル

例えば神経回路網をアナログ回路で実現する際、連

[†] x_i が 0 をとることを許すならば、反転する代わりに値を 0 にしてもほぼ同じ効果が得られる。

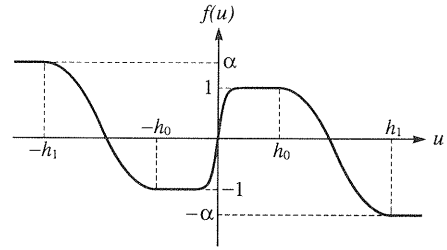


図8 連続時間部分反転法の実出力関数 $f_r(u)$

Fig. 8 Output function $f_r(u)$ for the continuous-time partial reverse method.

続時間-連続情報の細胞モデル⁽²⁾をよく用いる。連続時間の自己相関連想記憶の場合、そのダイナミクスは

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j=1}^n w_{ij} x_j \quad (22)$$

$$x_i = f(u_i) \quad (23)$$

で表される。ここで $f(u)$ は非線形の実出力関数であり、通常

$$f_s(u) = \frac{1 + e^{-cu}}{1 - e^{-cu}} \quad (24)$$

のようなシグモイド形の関数を用いる (c は正の定数)。

このようなモデルに対して、前節のアルゴリズムをそのままの形で適用することはできない。しかし、 f として図8に示すような形の関数 $f_r(u)$ を用いれば同様な効果があると考えられる。但し、最終的な出力結果は $x_i = \text{sgn}(u_i)$ で与えられるものとする。 f_r の形やパラメータはそれほど厳密でなくてもよいが、以下の数値実験では $\alpha = 1.0$ 、 $h_0 = 0.2$ 、 $h_1 = 1.0$ としている。

連続時間の部分反転法を用いると、二つのフェーズを考える必要がないだけでなく、後述のように離散時間の場合よりも優れた特性が得られる。

4. 数値実験および考察

部分反転法を用いることによって想起特性が実際にどう改善されるかを調べるため、数値実験を行った。

まず、図9に離散時間の部分反転法を用いた場合の方向余弦 a の推移を示す。部分反転を行う以外はすべて従来法の場合(図1)と同じ条件で実験し、パラメータ λ は 2.7、 h は式(21)に従って $r = 0.08$ のとき 1.65、 $r = 0.2$ のとき 2.09 とした。

この図から、以下の点で連想記憶能力が向上していると言える。

(1) 臨界方向余弦 ac が小さくなる。すなわち、 S の引込み領域 $B(S)$ が拡大する。

(2) 従来法で想起可能な場合においても、 X_i が平

平衡点 \hat{S} に至るまでに要する時間(ステップ数)が短くなる。また、グラフからは読みとりにくいですが、より S に近いところにある \hat{S} に達する(想起誤差が小さくなる)場合が多い。

(3) 記憶容量が増加し、 $r > r_c$ でも想起が可能となる。 r の上限はパラメータ等にも依存するためはっきりしないが、層状モデルの容量 r_i とほぼ一致する。

次いで a および $|l^\nu|$ ($\nu=2, 3, \dots, m$) の推移を調べた結果を図 10 に示す。(a)は図 4 (b)と同じ初期状態を与えた場合であるが、 $t > 1$ で $|l^\nu|$ はすばやく減少し、4 ス

テップで想起が完了している。また(b)は a_0 がほぼ臨界値 a_c に等しく、想起に成功するまで長い時間を要するような場合である。 $|l^\nu|$ の増加が抑えられ、 a との差が徐々に大きくなっていく様子が読み取れるが、これは一種の競合現象とみなすこともできよう。

図 11 は、図 5 と同様 a^2 と p を座標軸にとり、各フェーズの直後の状態をプロットしたグラフである。実線はフェーズ I での状態遷移、点線はフェーズ II での遷移を表す。この図から、図 7 に示した図式とは違い、むしろ部分反転によって X は S に近づくことがわ

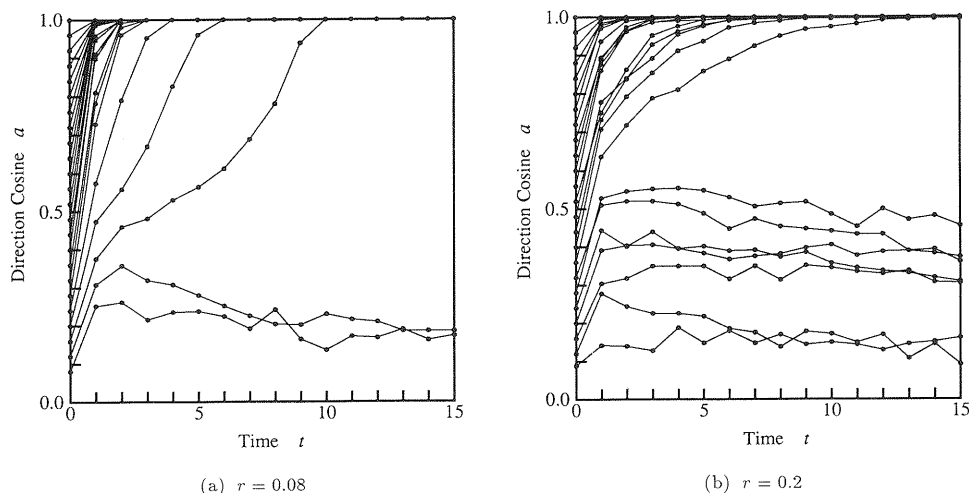


図 9 部分反転法(離散時間モデル)の想起過程
Fig. 9 Temporal change in direction cosines: discrete-time partial reverse method.

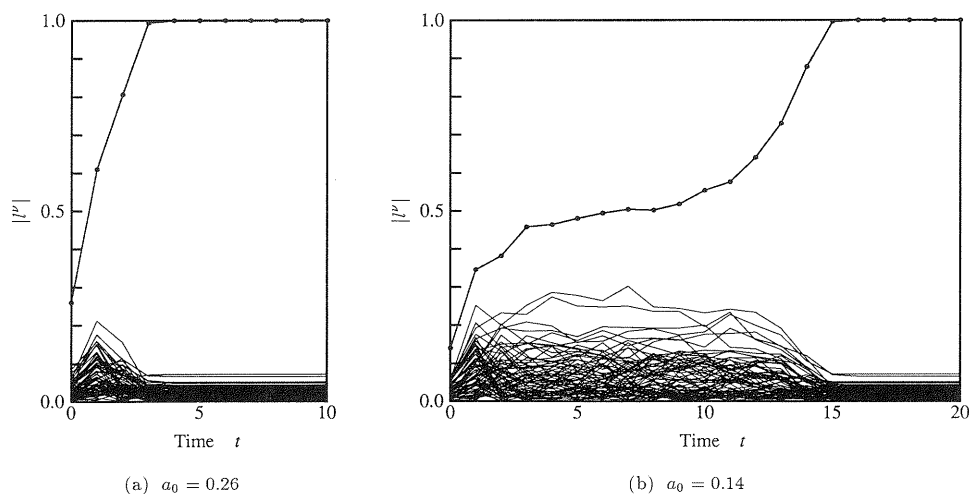


図 10 $|l^\nu|$ の推移 ($r=0.08$)
Fig. 10 Transition of $|l^\nu|$ with $r=0.08$.

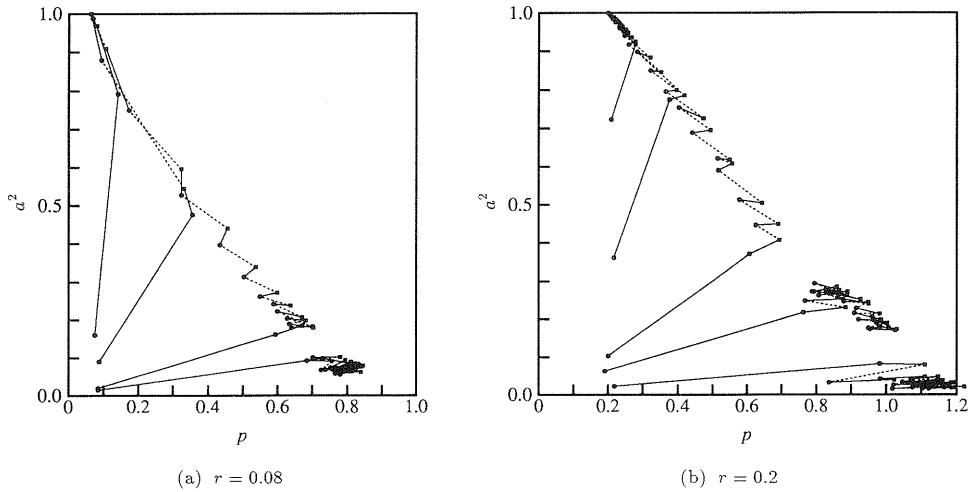


図 11 部分反転法による回路網の状態遷移
Fig. 11 State transitions of the network by the partial reverse method.

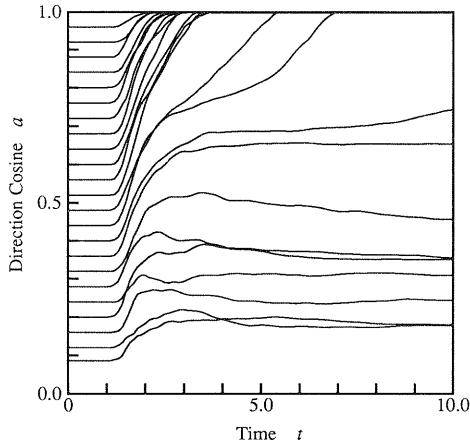


図 12 連続時間部分反転法の想起過程 ($r=0.32$)
Fig. 12 Transition of a : continuous-time partial reverse method with $r=0.32$.

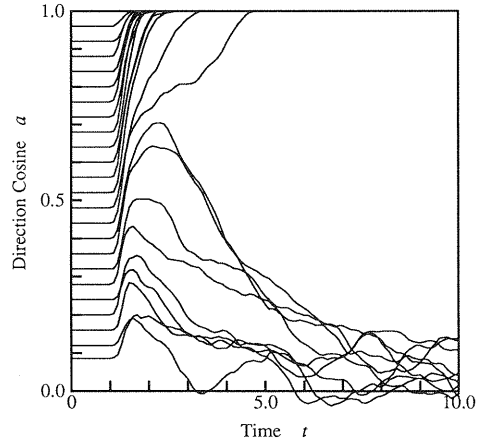


図 13 $w_{ii}=0$ のときの想起の様子
Fig. 13 Behavior of the recalling process with $w_{ii}=0$.

かる。このことは、この系の力学構造や従来のダイナミックスの問題点を考える上で示唆的である。

次に、連続時間モデルを用いた場合の実験結果を示す。

図 12 は $r=0.32$ で f_r を出力関数として用いた際の方向余弦 a の時間変化を示したものである。ここで、 a は

$$a = \frac{1}{n} \sum_{i=1}^n f_s(u_i) s_i \quad (25)$$

と定義した。また、横軸(時間)の単位は時定数 τ である。

$r=0.32$ というのは r_i よりも大きく、離散時間モデ

ルでは想起ができない記憶率であるが、 a_0 が約 0.42 以上のとき想起に成功し、想起誤差もほとんどないことがわかる。また、比較のため W の対角成分 w_{ii} を 0 にしてみたところ、しきい現象がより明確に現れた(図 13)。想起に要する時間が短くなるだけでなく、想起がうまくできない場合に a_i が急速に小さくなり、いつまでも定常状態に達しない点が興味深い。このことは、この系が記憶パターンの近傍以外に平衡状態をもたないことを示唆している。

ここでよく考えてみると、図 12 や図 13 の結果はやや不思議に思える。従来のダイナミックスでは、 S が不安定状態であるだけでなく、 r が大きいとき S の近

傍にもエネルギー極小点が存在しないのであった。にもかかわらず，出力関数を f_r とするだけで a_i がほぼ1に達し，しかもそこで平衡状態になる，というのはなぜであろうか。

2.の結果も考え合わせると， $r > r_c$ のとき S は一種の鞍点になっているものと推測される。つまり，エネルギー関数 E は S の周辺で図14に模式的に示すような形をしており（図の細線はエネルギーの等高線を表す），大部分の領域で E のこう配は S の側に向かって下がっているが，中心部に達するエネルギーの谷があり，その底では外側に向かって E が小さくなっていると考えられる。但し，谷底での傾斜は緩やかなので， w_{ii} が大きいと X はそこで止まる。また，谷に近い部分ほど $|u_i$

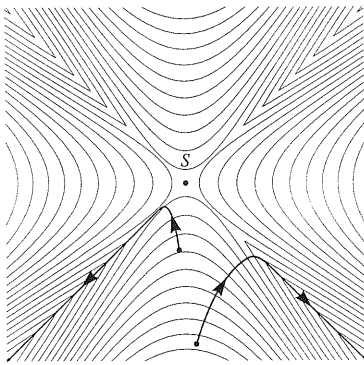


図14 S 周辺でのエネルギー関数 ($r > r_c$)
Fig. 14 Energy function around S.

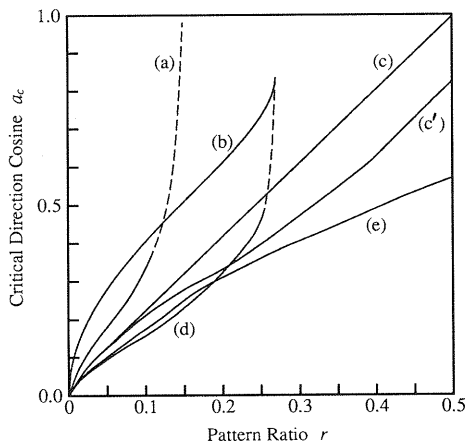


図15 臨界方向余弦の変化
Fig. 15 Critical direction cosines of (a) autocorrelation model, (b) layered model, (c) generalized inverse model, (c') generalized inverse model with $w_{ii}=0$, (d) discrete-time partial reverse method, (e) continuous-time partial reverse method.

の分散が大きく，強い入力を受ける細胞があるためにエネルギーが低くなっている。

結局， S の引込み領域は「切れ目」の入った「たらい (basin)」であり， $|u_i|$ の分散を大きくしない力を加えることによって切れ目がふさがり， $r > r_c$ でもたらいとして機能する，と解釈することができよう。このような解釈が妥当かどうか疑問ではあるが，今のところ種々の実験結果と矛盾しない。

最後に， r を変えたときの a_c の変化の様子を図15に示す。図の曲線は，それぞれ(a)自己相関モデル，(b)層状モデル，(c)一般逆行列モデル，(d)離散時間の部分反転法，(e)連続時間の部分反転法の場合を表す。また，(c')は一般逆行列モデルで w_{ii} を0にした場合である。(b)は理論的に求めたもの⁽¹⁰⁾であるが，それ以外は $n=1000$ での実験により求めたので若干の誤差がある（図の破線部分は推定値）。

(e)の方が(c')よりも小さな a_c をもつという事実は，自己相関モデルの場合，荷重行列の改良よりもダイナミックスの改良の方が有効であることを物語っている。

ところで，一般逆行列モデルに部分反転法を適用すると，より優れた連想記憶能力が得られる（この際，パラメータまたは反転のアルゴリズムを若干変更する必要がある）。しかし， a_c に関しては(d)または(e)とほとんど同程度にしか改善されない。このことは， S の近傍の力学構造を改良しても，全体的な構造はそれほど変化しないことを意味するものと考えられる。

5. む す び

自己相関連想記憶のダイナミックスについて考察し，その改良によって連想記憶能力を向上させる方法を示した。この方法を用いれば，学習法や回路の基本的構造を変えることなしに，記憶パターンの引込み領域を拡大し，記憶容量を場加させることができる。

このことは，既に議論したように連想記憶の力学的構造を理解する上で役立つだけでなく，直接的にもいくつかの意義をもつ。

まず，この種の連想記憶が工学的に有用かどうかという議論は別にして，少なくとも従来モデルの長所を残したまま，より優れた特性を得ることができる。特に，正しく想起できたかどうか判別できる点は重要である。また， n および r があまり大きくない場合には，ランダムな初期値からかなり高い確率で S のどれかを想起できる（例えば， $n=500$ ， $r=0.08$ の場合で確率0.5以上，従来法では 10^{-6} 以下）から， s_i^{μ} がランダ

ムという条件のもとで W から Σ を求めるという、一種の逆問題を解くのに利用できるかもしれない。

更に、部分反転法の考え方は、ある種の最適化問題を神経回路網で解こうとする際に応用できる可能性がある。実際、荷重行列 W が式(7)で与えられた場合を考えると、 $X = S^{\mu}$ のときエネルギーが最小である⁽¹²⁾から、部分反転法によってエネルギー最小解を得る可能性が大幅に高められることになる。但し、その他の問題(例えば巡回セールスマン問題)に対して、この方法をそのまま適用してもうまくいくとは思えない。というのは、そうした問題を解く神経回路網の力学構造がほとんどわかっていないからである。

しかしながら、少なくとも従来の単純なダイナミックスが最適とは限らないことを、本研究の結果は指し示している。また、荷重行列や学習法の改良だけでは限界があることから、神経回路網による情報処理の長所を残しつつダイナミックスを改良することが期待される。例えば、 $f_{\mu}(u)$ のパラメータを時間的に変化させるといった方法は、従来のアニーリング法よりも有望かもしれない。

いずれにせよ、神経回路網の力学についてもっと理解を深めることが必要であり、ダイナミックスを変えろというアプローチはそのための手段としても重要な意味をもつものと思われる。

そのほか今後の課題としては、連想記憶の想起過程や部分反転法に関する数理解析を進めることが挙げられる。また、本論文では触れなかったが、脳における記憶および想起のメカニズムを考えたとき、本研究で得られた結果には興味深いものがある。これを念頭においたモデルを構成することも今後の重要な課題の一つである。

謝辞 本研究に関して貴重な御指摘を下された東京大学工学部甘利俊一教授に感謝します。また、図面の作成や数値実験の一部を手伝ってくれた大学院の田中宏一良、北田宏両氏および学部学生の鈴木竜自、高友吾両氏に謝意を表します。

文 献

- (1) 中野 肇：“アソシエイトロンとその応用—連想記憶装置に関する研究”，信学インホメーション理論研資，IT69-27 (昭44)。
- (2) 甘利俊一：“神経回路網の数理”，産業図書 (昭53)。
- (3) 上坂吉則，尾関和彦：“連想記憶の二，三の性質”，信学論(D)，55-D，5，pp. 323-330 (昭47-05)。
- (4) D. J. Amit, H. Gutfreund and H. Sompolinsky：“Storing infinite numbers of patterns in a spin-glass model of

neural networks”，Phys. Rev. Lett., 55, pp. 1530-1533 (1985)。

- (5) R. J. McEliece, E. C. Posner, E. R. Rodemich and S. S. Venkatesh：“The capacity of the Hopfield associative memory”，IEEE Trans. Inf. Theory, IT-33, pp. 461-482 (1987)。
- (6) E. Gardner：“Structure of metastable states in the Hopfield model”，J. Phys., A19, pp. L1047-L1052 (1986)。
- (7) S. Amari and K. Maginu：“Statistical neurodynamics of associative memory”，Neural Networks, 1, pp. 63-73 (1988)。
- (8) J. J. Hopfield：“Neural networks and physical systems with emergent collective computational abilities”，Proc. Natl. Acad. Sci. USA, 79, pp. 2554-2558 (1982)。
- (9) T. Kohonen：“Self-Organization and Associative Memory (2nd ed.)”，Springer, New York (1988)。
- (10) R. Meir and E. Domany：“Exact solution of a layered neural network model”，Phys. Rev. Lett., 59, pp. 359-362 (1987)。
- (11) 齊藤公博：“連想記憶のダイナミクス”，東京大学工学部卒業論文 (昭63)。
- (12) M. Cottrell：“Stability and attractivity in associative memory networks”，Biol. Cybern., 58, pp. 129-139 (1988)。

(平成元年6月26日受付，8月31日再受付)

森田 昌彦



昭61東大・工・計数卒。昭63同大学院修士課程了。現在、同大学院博士課程に在学中。生体情報工学，特に神経回路網の研究に従事。

吉澤 修治



昭37東大・工・応物卒。昭39同大学院修士課程了。同年日立製作所中研勤務。昭41東京大学工学部勤務。現在同学部計数学科助教授。バイオサイバネティクス，生体工学，非線形力学の研究に関心をもつ。IEEE, INSS, 計測自動制御学会，日本ME学会，神経科学協会各会員。

中野 肇



昭35東大・工・応物卒。同年(株)東京芝浦電気入社。中央研究所勤務。昭41東大・工・計数助手。昭47工博。現在、東大・工・計数助教授。脳機能のモデル化の研究に従事。