

1章 総論

概要

【本章の構成】

脳と神経モデル - 1 章

1-1 脳と数理モデル

(執筆: 甘利俊一) [2008 年 6 月 受領]

1-1-1 数理モデルについて

脳は生命がつくった巧妙な情報処理の装置である。それは神経細胞から成り、複雑な回路網を形成している。脳の情報処理の原理を解明するには、神経細胞、更にはその分子の仕組みを解明することも必要であるが、一方回路網に焦点を当て、多くの要素が共同して働く回路、そしてシステムとしての情報の仕組みを解明することが重要である。

脳の情報処理システムは進化の結果現れたものであり、単純な設計原理をもつわけではないから、その実情は極めて複雑であり、多岐に及んでいる。しかし、多数の要素を結合し、神経細胞のような動作をする素子の上で実現できる情報の仕組みには、基本的な可能性と制約がある。これが脳の情報原理であり、脳は進化の過程でこの原理に合う仕組みを見つけ出し、それを複雑な状況の上に実現した。

ここで状況を単純化し、原理を実現するのに必要な最小のモデルを作り、その動作原理を数理的な解析を通じて明らかにする方法が必要である。それは現実と比べて単純化してあるから、そのままでは複雑な脳の現実を説明できないであろう。しかし、この原理が組み合わせり、更に種々の現実的な要因を組み込んで脳になっていることを考えれば、これは脳の情報原理の解明に必要なことである。これが理論脳科学、更に純化した数理脳科学であり¹⁾、そのために数理モデルが必要になる。

本稿では、数理モデルの典型的なものとその成果を概括する。

1-1-2 神経細胞の数理モデル

神経細胞（ニューロン）は、脳を構成する基本素子である。その動作をモデル化することから始めよう。1 個のニューロンは、他のニューロンからの多数の情報を受け取り、これを総合し処理して出力を出す情報処理素子と考える。

入力信号を x_1, \dots, x_n とする。このとき、入力を受け取るシナプスという場所に効率があって、それを w_1, \dots, w_n とすれば、入力の総合は重み付線形和であり、ここから一定の閾値 h を引いたものを $u = \sum w_i x_i - h = \mathbf{w} \cdot \mathbf{x} - h$ としよう。 \mathbf{x}, \mathbf{w} は入力及び効率をベクトルにまとめたものである。出力は u の非線形関数であり、 $z = \varphi(u)$ となる。

入出力信号は、現実には電気パルスである。このとき、 x_i や z は 0, 1 の 2 値をとる。場合によっては、解析を容易にするために、1, -1 の 2 値をとるモデルもある。しかし、発火率符号化といって、情報は個々のパルスにあるのではなく、一定時間に現れるパルスの個数、つまりパルス頻度に乗っている場合がある。このとき、信号 z は大きさを正規化して、0 と 1 の間の実数をとると考えてよい。出力関数 $\varphi(u)$ は、単調増大で、0 から 1 までの値をとるシグモイド関数となる。離散の場合は、これが 0, 1 もしくは 1, -1 の 2 値をとる特別の関数である。

もう少し現実的なモデルは、時間の要素を取り入れて、ニューロンの平均膜電位と呼ぶ時間の関数 $u(t)$ が、微分方程式

$$\frac{du}{dt} = -u + \sum w_i x_i(t) - h \quad (1\cdot1)$$

で変化するものである。これは、リーキー積分モデルと呼ばれ、 $u(t)$ が一定値に達するとニューロンはパルスを出し、 u は元の値に戻る。頻度符号化ならば、このときの出力関数 $z = \varphi(u)$ が、不応期の特性などを考慮に入れて計算できる。

もっと現実にあったモデルとしては、膜のチャネルの特性を入れた Hodgkin-Huxley モデルなどがある。

さて、このモデルで、ニューロンの担う情報をパルス頻度ではなくて、パルスのタイミングが担うという根拠もある。この場合は、時間で平均をとる頻度符号化は用いず、上記のモデルで個々のパルスの時間を考えることになる。このとき、多くのニューロンのパルスの同期や、ニューロン集団が作り出す種々の振動の上でのパルスのタイミング（位相）などに興味が集まる。

ニューロンのモデルとして、更にシナプスの効率 w_i の変化に関する法則を考える。多くは、Hebb の提唱したモデルを一般化したものであり、ある学習関数 r を用いて

$$\frac{dw_i(t)}{dt} = -w_i + cr x_i(t) \quad (1\cdot2)$$

のようにその法則が書ける¹⁾。ここでは、教師信号が外部から来る教師ありの学習と、教師信号なしで環境の情報の構造を取り込む自己組織学習がある。これらは頻度符号化のもとの表現であるが、パルス符号化では、より詳細な時間構造を取り入れて、入力信号と出力信号の微妙な時間差に応じてシナプス効率が増えたり減ったりするモデルがある²⁾。

1-1-3 順結合層状回路

最も単純なモデルは、多数の共通入力信号が神経集団に入り、各ニューロンがそれを処理して出力を次の集団に伝える一方方向性の情報処理モデルである。簡単のため、 x_1, \dots, x_n が入力信号、それを受けて情報変換をする k 個の中間層ニューロンがあり、更にその出力 $\varphi(w_i \cdot x)$ 、 $i = 1, \dots, k$ を受けて情報を総合する最後の出力ニューロンからなるとする。出力ニューロンの 1 個を取り出し、それは線形ニューロンだとすると、その入出力関数は

$$z = f(x, \theta) \sum v_i \varphi(w_i \cdot x) \quad (1\cdot3)$$

のように書ける。ここに、 w_i は中間層（隠れ層とも呼ぶ）の i 番目のニューロンのシナプスの重みベクトル、 v_i は最終出力ニューロンの重みである。ここで、可変のパラメータをまとめて、ベクトル $\theta = (w_1, \dots, w_k, v_1, \dots, v_k)$ とした。

この関数の特徴を調べよう。任意の関数 $g(x)$ が与えられたとき、 k が十分大きければ、パラメータ θ をうまく選べばこれを $f(x, \theta)$ で十分によく近似できる。これを、層状回路網の万能性という。次にこの関数をよく見ると、基底関数 $\varphi(w_i \cdot x)$ を用いた関数 $g(x)$ の展開を、 k 次で打ち切った形をしていることが分かる。このとき、フーリエ変換、ウェーブレット変換、スプライン展開などと違う点は、基底関数が可変で、学習によって変更できる点である。つまり、神経回路網は可変基底による展開を実現する。このとき、ある条件の下でこの展開は「次元の呪」をかるうじて免れて、入力の次元が極めて大きくても、展開がうまく働くこ

とが証明できる³⁾。

1-1-4 回帰神経回路網

神経回路網は、情報を一方的に伝えるだけではない。各ニューロンは他のニューロンからの情報をフィードバック信号として受け取る。つまり、出力信号が入力へ戻ってくる。このとき、離散時間で表せば、時間 t での各ニューロンの状態を $x(t) = (x_1(t), \dots, x_n(t))$ とすれば、次の時間での状態は $x(t+1) = f\{x(t), \theta\}$ のように書ける。ここで、ベクトル関数 $f(x, \theta)$ は、ニューロンを一つの層にまとめたときの入出力関係を表し、 θ はすべてのパラメータである。

連続時間モデルでは、 $u(t)$ を各ニューロンの状態として状態の発展は微分方程式で

$$\frac{du(t)}{dt} = -u + f(x, \theta) + s(t) \quad (1.4)$$

のように書ける。 s は外部からの入力である。

回路の力学的な時間発展にはいろいろなものが考えられる。入力信号が時間とともに変化するならば、それはその変化に追従する。また、入力が消失しても時間発展を続ける。そこには、いくつかの安定平衡状態をもつ多安定回路網、振動現象を示す振動回路、更にカオス的な振る舞いをし、擬似安定的な状態を遍歴するカオス遍歴回路などが知られている⁴⁾。これはどれも情報処理の原型要素としての役割を果たしている。

更に構造をもった回路網を調べてみよう。例えば、ニューロンが一次元の場に一樣に並び、近隣のニューロンから出力信号を受け取る一樣な回路網を考えよう。これを神経場と呼ぶ。その方程式は

$$\frac{\partial u(\xi, t)}{\partial t} = -u(\xi, t) + \int w(\xi - \xi') \varphi[u(\xi', t)] d\xi' + s(\xi, t) \quad (1.5)$$

のように書ける。ここに、 $u(\xi, t)$ は場所 ξ にあるニューロンの時間 t での平均膜電位、 $w(\xi - \xi')$ は場所 ξ' から場所 ξ のニューロンに入る入力の結合の強さで、通常近所のニューロンに対しては正の、少し離れたニューロンに対しては負の値をとる結合関数である。また、 $s(\xi, t)$ は外部から場所 ξ のニューロンに入る入力の強さである。

神経場は、入力がなくなったときでも、場の一部に興奮状態を残しそれを保持する作業記憶の役割を果たすことから、情報処理の基本要素の一つとして注目を集めている¹⁾⁵⁾。これは、心理現象、神経内部の情報の競合、ロボットのナビゲーションなどに用いられる。場を二次元にし、またニューロンの層を抑制性をも含む多層にするなど、いろいろなモデルへと発展している。

1-1-5 統計神経力学

神経回路網の動作は、その結合の特性に応じて多様である。しかし、マクロな視点に立てば、その結合は一定の統計的なパターンに従うと見られる場合も多い。このとき、特定の確率法則を指定して、現実の回路はこの分布から生み出された典型的なものであるとし、その情報処理の特性を調べることができる。これを統計神経力学と呼ぶ¹⁾。

まず、1層の層状神経回路網から始めよう。簡単のため n 入力 n 出力の層状回路で、入出力が ± 1 の離散値をとる回路を考える。入出力関係は $z_i = \text{sgn}(\sum w_{ij}x_j)$ 、 sgn は符号関数である。ここで、結合の値 w_{ij} が独立同一のある確率分布に従うものとする。このとき、入出力情報のマクロな値として、全体の活動度 $X = (1/n) \sum x_i$ 、 $Z = (1/n) \sum z_i$ を用いると、活動度の変換法則 $Z = F(X)$ が得られる。ここで変換の関数 F は、結合 w_{ij} の平均と分散のみによって定まる。

更に、二つの入力 x と x' の間の距離を $D_X = (1/2n) \sum |x_i - x'_i|$ と置く。このとき、入力で D_X だけ離れた二つの状態が、変換後における z 及び z' ではその距離 D_Z がどのくらいはなれるかという、距離の分離法則 $D_Z = G(D_X)$ が得られる。ランダム結合の場合、これは原点において微係数 ∞ をもつ関数で、入力側でのほんの小さな差異を出力側で大きく拡大するので、差異の検出に有力である。

層状回路をフィードバックで結べば、マクロな状態が変化するモデルとなる。このとき、単純なランダム回路は多安定回路になる。興奮性と抑制性の素子をもつ回路では、振動現象が起こる。更にミクロな状態の間の距離の時間発展を調べることもできる。このとき、ミクロな状態遷移のグラフは、その次数の分布が代数的になる「スモールワールドネットワーク」の例になっている¹⁾。

ここまでは、確率分布で結合の重みがすべて独立（同一でなくてよい）の分布に従う例である。この場合でも、マクロな方程式が成立する基礎を探っていくと、難問に行き当たる。統計神経力学の基礎が必要である¹⁾。

もう一つの統計神経力学の例として、連想記憶がある。まず、層状の（フィードバック結合のない）、相互相関による連想の回路を考えよう。話を簡単にするために、入出力はいずれも n 次元とし、 p 個の入出力の組 $(x(k), y(k))$ 、 $k = 1, \dots, p$ を考える。このとき、入力として $x(k)$ もしくはそれに類似した入力を与えると、出力として $y(k)$ を出すような回路を作ることが要請される。

統計神経力学では、信号の成分 $x_i(k), y_i(k)$ はすべて 1 または -1 から成り、それらが等確率で独立に出たものと仮定する。そして、この入出力を用いて回路の重みを $w_{ij} = (1/p) \sum y_i(k)x_j(k)$ と設定する。

ここで前と同じに、入力信号が $x(k)$ から少しずれた $x'(k)$ だったときに、出力 $y(k), y'(k)$ はどのくらいずれるかを計算してみよう。すると、以前のランダムで独立な結合のときと違い、 $D_Y = G(D_X)$ は D_X が小さいときに微係数が極めて小さい関数となる。これは、誤差により入力が大きくずれても、回路の出力はその誤差を打ち消してしまうことを意味している。連想の安定性が保証される。

回路にフィードバックを付けると、これは自己相関による連想記憶のモデルとなる。そのとき、想起のダイナミクスは $x(t+1) = \text{sgn}Wx(t)$ と書ける。

連想記憶の想起のダイナミクスは極めて複雑である。これは、じつはマクロな状態遷移の方程式、すなわち第一段から第二段への方程式がそのままでは使えないことが分かるからである。これは先に述べた統計神経力学の基礎の問題である。甘利 馬被⁶⁾ は想起の力学方程式を提出したが、これは近似である。岡田⁷⁾ はこの考えを発展させ、更に良い高次の近似を与えた。

1-1-6 学習神経回路 多層パーセプトロン

1 個のニューロンの能力は限定されている．その学習能力は，パーセプトロンの学習定理で明らかにされている．次は，万能の能力のある層状回路，多層パーセプトロンの学習能力を調べる必要がある．いま，回路の入出力関係 $z = f(x, \theta) = \sum v_i \varphi(w_i \cdot x)$ を考え，学習用の教師信号として，入出力の組 $(x_1, y_1), \dots, (x_k, y_k)$ が逐次与えられるものとしよう．現在の回路のパラメータの値を θ_t とし，教師の与える入出力 (x_t, y_t) をもとにこれを更新する．入力 x_t をこの回路に与えれば，回路の出力は z_t であるから，これを教師の答 y_t と比較して，2乗誤差は $l(x_t, y_t, \theta_t) = (1/2)|y_t - f(x_t, \theta_t)|^2$ と書ける．教師信号にはガウス雑音が混ざっているとすれば，この損失の総和を最小にするのが，最尤推定であり，学習は逐次最尤推定と考えることができる．

最急降下学習法は $\theta_{t+1} = \theta_t - \eta \nabla l(x_t, y_t, \theta_t)$ のようになる⁸⁾． ∇ は θ による偏微分（グラジエント）である．ここで，誤差の微分を実行すると，パラメータを更新するために誤差信号が回路を逆順に伝播していくように見える．そのため，この学習法は誤差逆伝播学習法と呼ばれている⁹⁾．

しかし，その収束が極めて遅いこと，誤差関数の極小解が多数あることから，この方法をそのまま使うには問題もある．収束の遅さは，多層回路の統計モデルが特異点を含むこと，すなわち対称性由来するところで Fisher 情報行列が縮退することに起因することが明らかになった¹⁰⁾¹¹⁾．これを防ぐには，情報幾何の立場に立って，自然勾配学習法，すなわち， G を Fisher 情報行列として， $\theta_{t+1} = \theta_t - \eta G^{-1} \nabla l$ を用いればよい．なお，情報行列を求めその逆転を行うのは容易でないから，情報行列の逆行列を逐次的にデータから求める，適応自然勾配法が使われる¹²⁾．

また，極小解が多数ある問題は，種々の解を総合するブースティングと呼ぶ機械学習の方法が開発されている．ここでも，情報幾何が有用である¹³⁾．

1-1-7 強化学習

教師あり学習では，各ニューロンにそのときにどのような動作をすべきであるかを示す学習信号が与えられる．これは，直接に与えられる場合もあるし，誤差逆伝播学習のように間接的に全体の誤差信号から計算によって与えられる場合もある．しかし，現実の我々の動作は，多くの場合 1 回で決まるのではなくて，多段の決定を行う．すなわち，ある状況から出発してある動作を行い，それによって状態が変化し，そのうえで次の動作を決定する．こうした一連の過程の後で成否が決まり報酬が与えられる．これをもとに学習するには，途中の段階で何をすべきであったのかを推論しなければならない．これは多段決定過程と呼ばれ，計算論的には極めて複雑である．脳はこれをどのように実現するのかを示す，強化学習と呼ばれる枠組みがある．

1-1-8 自己組織神経回路網

神経回路網は，外界と相互作用する中で自己の入出力を学習する．これにはどう行動すべきかを指示する教師信号が必要である．しかし，外界からは単に多数の信号が入力するだけで，これをもとに外界の信号の構造を学習し，これらの信号に対して効率の良い処理を可能にするように備える学習もある．これは神経回路網の自己組織化である．

自己組織化には、例えば外界の典型的な信号に反応するように自己を特化する細胞（特徴検出細胞）の形成、信号の主成分を検出する学習、更に信号のトポロジカルな構造を配列して神経場に埋め込む学習などがある。トポロジカルマップの形成に関しては、Kohonenのモデルが有名であるが、これは別稿に譲ることにしよう。

1-1-9 スパイク系列の統計解析と情報幾何

1個のニューロンは、スパイクを発生する。しかし、そのタイミングは確率的な揺らぎをともなう。パルス発生頻度を一定として、この過程がランダムすなわちポアソン過程であるならば、パルスの個数と分散は同じになる。実験では、これはニューロンによっても異なり、同じではない。そこでどのような法則に従うのが議論される。

篠本らは、巧妙なアイデアのもとに、発火頻度が一定ではなくて時間変化する場合にも使える指標として L_V を提案している¹⁴⁾。

パルス感覚の分布がガンマ分布に従うと仮定し、かつパルス頻度は時間可変で勝手に変わってよいとしたときに、パルスの規則性を測る不変な指標として S_I が得られる。これは、情報幾何に基づいてセミパラメトリックモデルを解析した結果である¹⁵⁾。

一つのニューロンではなくて、二つ以上のニューロンがパルスを発生するときには、その同期が問題になる。各ニューロンの発火については、それぞれの頻度と相関が問題になるのである。特に、ニューロン集団の同期は相関にかかわっている。この場合、 n 個のニューロンの発火頻度と相関とを統計的な意味で直交する座標に分解することが求められる。

情報幾何はこの問題に良い答えを与える。いま、 n 個のニューロンの同時確率分布を $p(x_1, \dots, x_n)$ としよう。このとき、個々のニューロンの発火率は E を期待値として $\eta_i = E[x_i]$ で測ればよい。次は相関であるが、例えば 2 個のニューロンの発火の分散や相関係数を指標にすると、これは発火率と直交しない。そこで、確率分布の対数をとって、これを展開した係数を考える。このとき、係数は各次数のニューロン間の相関を表す。しかも、これは η_i と直交する座標を与えるので、パルスの集団時系列の解析に便利な基盤を与える¹⁶⁾。

参考文献

- 1) 甘利俊一, “神経回路網の数理,” 産業図書, 1978.
- 2) S. Song, K. D. Miller, and L. F. Abbott, “Competitive Hebbian learning through spike-timing-dependent synaptic plasticity,” *Nature Neuroscience*, 3, pp.919–926, 2000.
- 3) A. R. Barron, “Universal approximation bounds for superpositions of a sigmoidal function,” *IEEE Trans. on Information Theory*, 39, pp.930–945, 1993.
- 4) 津田一郎, “カオスの脳観,” サイエンス社, 1990.
- 5) S. Coombes, “Waves and bumps in neural field theories,” *Biological Cybernetics*, 93, pp.91–108, 2005.
- 6) S. Amari and K. Maginu, “Statistical neurodynamics of associative memory,” *Neural Networks*, vol.1, no.1, pp.63–73, 1988.
- 7) M. Okada, “A hierarchy of macrodynamical equations for associative memory,” *Neural Networks*, 8, pp.833–838, 1995.
- 8) S. Amari, “Theory of Adaptive Pattern Classifiers,” *IEEE Trans.*, EC-16, no.3, pp.299–307, 1967
- 9) D. E. Rumelhart and D. E. McClelland, “Parallel Distributed Processing,” vols.I, II. MIT Press, 1986.
- 10) S. Amari, H. Park, and T. Ozeki, “Singularities Affect Dynamics of Learning in Neuromanifolds,” *Neural Computation*, 18, pp.1007–1065, 2006

- 11) 福水健次, 栗木 哲, “特異モデルの統計学,” 岩波書店, 2004.
- 12) S. Amari, “Natural Gradient Works Efficiently in Learning,” *Neural Computation*, vol.10, no.2, pp.251–276, 1998.
- 13) 村田 昇, “パターン認識と学習の統計学 推定量を組み合わせる— バギングとブースティング,” 岩波書店, 2003.
- 14) S. Shinomoto, S. Yutaka, and J. Tanji, “Differences in spiking patterns among cortical neurons,” *Neural Computation*, 15, pp.2823–2842, 2003.
- 15) K. Miura, M. Okada, and S. Amari, “Estimating Spiking Irregularities Under Changing Environments,” *Neural Computation*, 18, pp.2359–2386, 2006.
- 16) S. Amari, “Information Geometry on Hierarchy of Probability Distributions”, *IEEE Transactions on Information Theory*, vol.47, no.5, pp.1701–1711, 2001.

脳と神経モデル - 1 章

1-2 脳と工学モデル

(執筆者：田中宏和，川人光男)[2008 年 5 月受領]

脳を工学の観点から理解しようとする試みには長い歴史がある。脳は 18 世紀には水力機に，19 世紀には機械式計算機に，20 世紀初頭には電話線網に喩えられたことがある。しかし，情報処理という観点からの比較が可能になったのは，デジタルコンピュータが開発された 20 世紀後半からであろう。脳に代表される生物の情報処理機構とコンピュータに代表される工学の情報処理機構の間には，どのような素子を用いるかという実装的な相違と問題をどのように解くかというアルゴリズム的な相違がある。具体的には，脳は神経細胞というミリ秒オーダーの素子を並列的に用いるのに対し，コンピュータはナノ秒オーダーの素子を直列的に用いている。こういった相違にもかかわらず，解くべき計算問題は同じ物理法則のもとで感覚処理・運動制御を行うことであるから，脳で行われている計算を工学理論やコンピュータ科学で発展した枠組みを用いて理解することが意味をもつ¹⁾²⁾。

脳が解くべき計算問題には，(1) 感覚入力と運動出力間の適切な入出力関係を学習すること，(2) 入力の統計性に基づいて自己組織的に分類すること，(3) 感覚入力から外界の状態を推定すること，そして (4) 身体を適切に制御すること，といったものがあるだろう。本稿では，上記の計算問題に対して，関数近似理論，特徴抽出理論，推定理論，制御理論がどのように解法を与えるかについて解説する。特に近年の流れとして，外界の非一意性や感覚入力，運動出力における不確定性を考慮した統計的手法が注目されてきている点を強調する。これは奇しくもロボティクスで近年強調されている確率的アプローチとも合致する³⁾。

1-2-1 関数近似理論

脳において，感覚系で得られた生の入力を認識もしくは運動制御に使いやすい出力に変換する必要がある。一般にこの感覚運動変換は解析的に簡潔に記述できるものではなく，また外界や身体の変化に対応して適応的でなくてはならない。このような入出力関係を記述するのに工学分野で発展した関数近似理論が有効である。ここでは代表的な関数近似法であるニューラルネットワーク，動径基底関数，モジュラーネットワークがどのように脳のモデル化に用いられてきたかを概観する。更に，入出力の不確実性まで考慮に入れることができるガウス過程と呼ばれる確率的な関数近似法を説明し，脳のモデル化に応用できる可能性を示唆する。

(1) パーセプトロン

1950 年代後半に Rosenblatt らによりパーセプトロンが提唱されてから，ニューラルネットによる関数近似・パターン認識の分野が開かれた⁴⁾。神経細胞の反応曲線に似せた基底関数を用いて，その「重み」すなわち結合定数を調節することで関数を近似する方法は，ニューラルネットワークによる近似と呼ばれる。この方法では複雑で未知の規則に従う入出力関係をトレーニングセットから学習するという，以後の関数近似理論で使われる枠組みを提唱したことで評価されよう。理論的には学習則が単純なことに加えて必ず学習が収束することから，応用の観点からは手書き文字認識などのパターン認識に使えることから，研究者の興味をひいた。しかし，パーセプトロンにより実現できる入出力関係は線形分離可能なものに限られることが示されてから，ニューラルネットの研究は一時下火になる。

(2) 多層ニューラルネットワーク

ニューラルネットワークモデルが再び注目を浴びるのは、バックプロパゲーションという多層ニューラルネットワーク学習則が再発見された 1980 年代半ばである。出力層で得られた誤差を出力層から入力層に後ろ向きに伝播させて中間層の結合定数を学習することから、この名が付けられた。多層にすることでどのような非線形変換も実現できることが示され、解析的手法では扱いにくい多様な実用的問題に適用された(図 1・1A)。特に有名な例には、英語の綴りからその発音への変換を示した NetTalk がある⁵⁾。言語学者によって見出された明示的だが複雑な発音規則が比較的単純なニューラルネットワークで陰に実現できること、またその学習過程が幼児が言葉を学ぶ様子に似ていることなどから注目された。

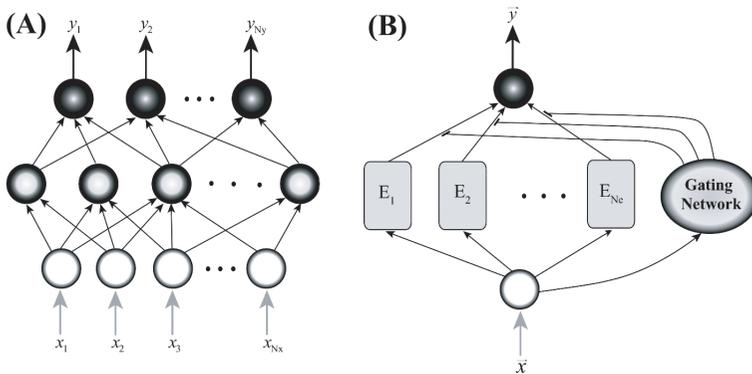


図 1・1 ニューラルネットワークの概念図。(A) 多層ニューラルネットワーク: ネットワークへの入力 $\{x_i\}$ は、隠れ層を経て出力 $\{y_i\}$ を生成する。(B) Mixture-of-Experts ネットワーク: 複数のエキスパートモジュール (E_i) が入力を受け取りそれぞれの出力を生成するが、全体としての出力はゲーティングネットワークにより現在の文脈に沿うように重み付けされて決定される。図の簡略化のため、入力と出力をベクトルとしてまとめているが、多層ニューラルネットワークと同様に多次元の入出力関係を記述できる。

(3) 動径基底関数

多層ニューラルネットワークではトレーニングに時間がかかること、また新たに学ぶ事例が過去のトレーニングに障害を引き起こすこと(干渉)が問題点として知られていた。これらの問題点は、ニューラルネットワークの基底関数がシグモイド関数のように大域的であること、ネットワークが多層であることに起因している。動径基底関数ネットワークは、ガウス関数といった局所的な関数を用いることでこれらの問題を回避している。直感的に言えば多数の「こぶ」を適当に重み付けて、非線形関数を近似する方法である。動径基底関数ネットワークは視覚認知に適用され、ワイヤフレーム状の物体を視点によらずに認識できることが示された。一方、代表的な視点からの物体画像が与えられた際の物体認識・分別課題をサルに学習させたところ、動径基底関数ネットワークから期待されるように、側頭連合野から代表的な視点に反応する神経活動が見出された⁶⁾。この結果から、視点によらない物体認識は代表的な視点からの 2 次元画像を内挿することで実現されていると考えられている。

(4) モジュラーネットワーク

局所化のアイデアを更に推し進めて、ネットワークを単純だが複数のモジュールから構成するモジュラーネットワークも提案されている。有名な例に Mixture-of-Experts ネットワークがある(図 1・1B)⁷⁾。そこではエキスパートと呼ばれる複数のモジュールからの出力を状況に応じて重み付けし、ネットワーク全体からの出力とする。同じネットワーク構造でもモジュールごとの重み付けを変えられるので、状況に応じた出力や学習が可能になる。

(5) ガウス過程

上述した関数近似法はトレーニングセットからパラメータ(結合定数)を学習することで関数を近似するので、パラメトリックな関数近似法と呼ばれる。また、これらの関数近似法において、入出力が決定論的である。近年ガウス過程と呼ばれるノンパラメトリックで確率的な関数近似法が機械学習の分野で提案されて注目を集めている(図 1・2)⁸⁾。

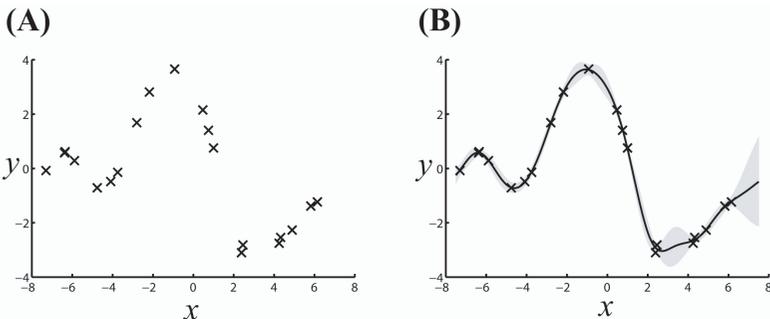


図 1-2 ガウス過程による関数近似の概念図。(A) トレーニングセット (x_i, y_i) 。(B) ガウス過程による出力の期待値(実線)と標準偏差(影)。新たな入力 x に対して出力の期待値は $\hat{y} = \sum_{i=\text{datapoints}} K(x, x_i) \cdot y_i$ と内挿される。ここで、 $K(x, x') = E[y(x)y(x')]$ は出力間の相関関数と定義する。

出力の「滑らかさ」である相関関数が与えられているとして、その相関関数に基づいてトレーニングセットを重み付けて和をとることで未知の入力に対する出力を計算する。特筆すべきは出力の分散、すなわちどれくらい確からしいかを記述できることである。トレーニングセットが密に得られるところの近傍では確信をもって出力を計算することができるが、トレーニングセットが与えられていないところでは出力は大きな不定性をともなう。ヒト心理物理実験において、関数近似の不定性がどのような役割を果たしているかは調べられていない。ヒトの学習過程に不定性がどのような影響を及ぼすかを、ガウス過程を用いてモデル化することは一つの方向であろう。

1-2-2 特徴抽出理論

前節で述べた関数近似のためには学習のためのトレーニングセットが必要であり、そのためにある入力に対応して何らかの出力がラベルされていないといけない。現実にはある入力に対して出力が明示的に与えられていることは稀であり、入力信号の規則性・統計性のみを

用いて分類もしくは特徴抽出する必要がある．ここでは脳のモデルとかかわりの深い独立成分分析とスパースコーディングについて概説する．

(1) 独立成分分析

脳にとって独立な情報源を特定することは生存にとって不可欠である．例えば、様々な音が飛び交っている中、天敵の鳴き声を抽出できなければ生き残れない．一方、複数のセンサからの入力を与えられる場合にそれぞれのセンサが単一の信号源からの信号を純粹に拾っていることは稀で、大抵の場合複数の信号が交じり合っている．独立成分分析において、問題は混合入力がどのように交じり合っているかを知らずして、元の信号を再構成することである．仮定するのは、個々の信号源が独立であるということのみである．

この問題は一見不可能に思えるが、独立成分分析における一つのアプローチとして、統計学における中心極限定理の「逆」を解くことで独立な成分を見出すことができる．中心極限定理によれば、複数の独立な信号の平均をとったものは元の信号に比べてよりガウス分布に近いものになっている．センサで得られた入力から元の信号を復元したければ、よりガウス分布から離れるように入力信号を組み合わせればよい．脳科学への応用としては、モデリングというよりは脳信号とりわけ機能イメージングの解析があげられる⁹⁾．

(2) スパースコーディング

自然界から感覚系が受ける信号は決してデタラメではなく、背後に潜む物理法則に由来する構造をもっている．例えば、網膜上の自然画像においては、隣り合う二点の輝度は決して独立ではなく高い相関をもっていて冗長な、言い換えれば無駄の多い表現になっている．感覚系での外界の表現が外界の冗長性を取り除いた効率の良いものになっていると考えて、ある自然画像の表現は少数の神経細胞活動から成り立っているとするスパースコーディングが提唱されている¹⁰⁾．

このアイディアは、独立な成分を見つけるという点で独立成分分析と同じであり、入力である自然画像の統計性から基底を決めるという点で適応フィルタともみなすことができる．元の自然画像をなるべく少ない基底を用いて再現できるようにトレーニングすると、第一次視野で知られている受容野に似た局在化して方向性をもつバンドパス状の基底が得られた．スパースコーディングは視野だけでなく聴覚野にも適用され聴覚細胞の反応性を再現することが示されている．

1-2-3 推定理論

感覚入力から外界で何が起きているかを推し量る問題を推定問題と呼ぶ．脳が解くべき推定問題には二種類の不定性がともない、一般には解が唯一に決まらない不良設定問題である．一つには外界の不定性がある．これはある感覚表現に対応する外界の状態が唯一に決まらないことに由来する．例えば、網膜上の2次元画像に対応し得る3次元構造は無数に存在する．もう一つには感覚系・運動系には必ず神経細胞由来のノイズがあることである．感覚ニューロンの活動はノイズをともなうため、神経活動だけから外界の情報を忠実に再現することはできない．したがって、推定問題を解くためには、外界及び神経細胞ノイズの先験的知識が不可欠である．最適推定理論はこのような推定問題を系統的に解く枠組みを与え、また近年

ヒトが行う感覚処理は最適推定で説明できることが示されてきた．ここでは最適推定理論，特にベイズ理論やカルマンフィルタが脳のモデルにどのように用いられたかを鳥瞰する．

(1) 最適推定理論

ベイズ理論は，視覚や触覚などの異なるモダリティからの情報を統合したり，感覚入力を先験的知識を基に解釈したりするための方法を与える．感覚系における神経活動データ D が与えられたとして外界が S である確率は事後確率と呼び $P(S|D)$ と書くことにする．外界 S が与えられたとき神経活動データ D の確率 $P(D|S)$ を尤度と呼び，そもそも外界 S がどれくらいありふれているかの事前知識 $P(S)$ を先験的確率と呼ぶ．尤度は神経細胞の反応性に関する知識を，先験的確率は外界に対する知識をそれぞれ表している．ベイズの定理では事後確率が尤度に先験的確率を掛けたものに比例していて， $P(S|D) \propto P(D|S) \cdot P(S)$ である．つまり，神経データがどれくらい確実に外界を反映しているかと外界に対する事前知識がどれくらい確実かの積になる．

脳がベイズ理論的に最適推定していることを支持する多数の心理物理実験結果が得られている．例としては，視覚刺激のコントラストに依存する速度知覚が説明されている¹¹⁾．コントラストが大きいときは尤度が支配的になるので，ヒトはほぼ実際の速度を報告する．一方，コントラストが小さいときには先験的知識が支配的になる．「たいてい物体はゆっくりと動く」という先験的知識を仮定すると，ヒトは速度を過小評価する心理物理の結果を説明できる．このようにしてベイズ理論は，先験的知識と観測データを統合して最適に推定する枠組みを与える．

(2) カルマンフィルタ

ベイズ理論の応用であるカルマンフィルタは，時々刻々変化する外界の状態を部分的でかつノイズの入った観測から推定する枠組みを与える．先験的知識として外界の時間発展を記述する運動方程式が与えられているとして，現在状態の推定量から少しだけ未来の状態を運動方程式を解くことで予測することができる．その予測された状態と実際に得られた観測量をベイズ定理に従い重み付け和をとれば，最適な推定量が計算できる．脳モデルへのカルマンフィルタの応用例として階層的な視覚認知モデルがあげられる（図 1・3）¹²⁾¹³⁾．

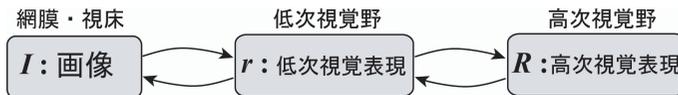


図 1・3 階層的視覚モデルの概念図．ここでは模式的に網膜もしくは視床，低次視覚野と高次視覚野を示してある．高次層では一つ下の低次層の活動を予測し，後ろ向き結合（図中左向きの矢印）によりその予測を低次層に伝える．その活動予測は実際の活動と比較され，その誤差は前向き結合（右向きの矢印）を通して高次層に伝えられる．

これらのモデルでは高次層は低次層の活動を予測すると仮定し，高次層から低次層への後ろ向き結合は高次層によって予測された低次層の活動を，低次層から高次層への前向き結合

はその予測活動と実際の活動の差を伝えているとする。この双方向計算のため、神経細胞の反応特性は低次層からの前向き結合のみで決まるのではなく、高次層から低次層への後ろ向き結合にも依存することが示唆される。これらのモデルでは、小数回の緩和計算による陰影からの構造復元や、endstopping 細胞といった第一次視覚野神経細胞の性質を再現することなどが示された。

1-2-4 制御理論

ヒトの行う身体運動は試行ごとにランダムになされるのではなく、不変的な特徴を示すことが知られている。これは脳の中で何らかの評価基準を最適化することで、身体運動を計画・実行していることを示唆している。ここでは制御理論がどのようにヒト身体制御の理解に適用されてきたかについて概説する。

(1) 最適制御

水平平面内での上腕到達運動において、端点の位置にかかわらず手先は直線の軌跡を描き、速度形状は釣鐘型を示すことが知られている。到達運動で見出された滑らかな手先軌道は、手先のキネマティクスもしくはダイナミクスに関する何らかの評価関数を最適化している結果と期待される¹⁴⁾。いままで提案された評価関数は大別して二つのクラスに分けられる。一つのクラスでは運動の滑らかさを要請するもので、例として手先軌道の滑らかさを要請する躍度最小モデル、その力学的拡張としてトルクの滑らかさを要請するトルク変化最小モデルがあげられる。もう一つのクラスでは、運動指令の大きさに比例するノイズ（信号依存ノイズ）の影響を最小にして標的からの誤差を最小化する分散最小モデルがあげられる。分散最小モデルではタスクを達成するという制約の下で運動指令をなるべく抑えることで、運動指令に付随するノイズの影響を最小化する。運動指令が抑えられることから、無用に大きな力を必要とする滑らかでない軌道よりも力の抜けた滑らかな軌道が期待される。

これらの計算論的モデルは2点間の滑らかな軌道だけではなく、経由点運動、 $2/3$ 乗べき則、Fittsの法則などといった到達運動の特徴を統一的に説明する。

(2) 内部逆モデル

到達運動や眼球運動といった速い運動では、遅い感覚入力に頼るフィードバック制御のみでは運動が不安定で振動的になる。したがって、望ましい軌道から運動指令を前もって準備しておく必要があり、この制御をフィードフォワード制御と呼ぶ。実現すべき理想軌道がタスク空間もしくは関節角空間で与えられたとして、それを実現するような運動指令に変換することをロボティクスの用語で逆ダイナミクス変換と呼ぶ。脳は素早い運動を実現するために逆ダイナミクス変換を用いてフィードフォワード制御を行っていると考えられており、その変換は内部逆モデルと呼ばれる¹⁵⁾。電気生理・脳イメージング研究から、眼球・上腕運動及び道具学習の内部逆モデルの計算は小脳で行われている証拠が示されている。

(3) フィードバック誤差学習

内部逆モデルの目的は望ましい身体運動を生成するような運動指令を出力することであるから、逆モデルは身体の状態を忠実に反映していなくてはならない。短期的には外部の負荷

や疲労，長期的には怪我や成長といった要因により身体は変化するため，内部逆モデルもその変化に対応して学習する必要がある．このような学習はロボティクスでは適応学習と呼ばれる．逆モデルにおいて特に問題となるのは，与えられる教師信号が逆モデルの出力である運動指令空間ではなく感覚系で処理されるタスク空間で表現されているということである．

フィードバック誤差学習は，フィードバック運動指令で教師信号を近似しフィードフォワード制御器である内部逆モデルを修正することを提案する（図 1・4）¹⁶⁾．内部逆モデルが身体を忠実に記述していない場合，出力されるフィードフォワード運動指令は正しくないため，運動が理想軌道から外れ，それに応じて筋肉の伸張反射に基づくフィードバック制御の補正が加わる．このフィードバック信号は逆モデルの出力がどれくらい良かったか（もしくは悪かったか）の評価に使えるので，これを用いて逆モデルを学習させればよい．このフィードバック誤差信号は小脳皮質プルキンエ細胞の登上線維入力で表現されていると考えられている¹⁷⁾．

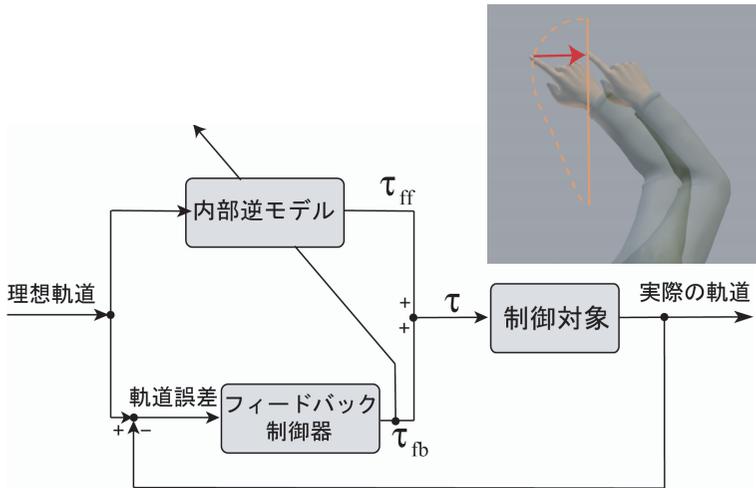


図 1・4 フィードバック誤差学習の概念図．運動指令 τ は，内部逆モデルからの出力であるフィードフォワード成分 (τ_{ff}) と軌道誤差から計算されるフィードバック成分 (τ_{fb}) から成る．内部逆モデルの学習前にはフィードフォワードの制御がうまくいかず，実際の手先軌道（右上オレンジ破線）は理想軌道（オレンジ実線）からずれてしまい，フィードバックによる制御（赤の実線）が働く．このフィードバック成分は，内部逆モデルの制御がどれくらいずれていたかを運動指令の空間で示しており，逆モデルの教師信号として用いることができる．

(4) モザイクモデル

身体ダイナミクスは高次の非線形性をもち，かつ同じ身体をその時々必要に応じて使い分ける必要がある．複雑でしかも切り替えの必要な制御問題を，汎用的だが複雑な単一のネットワークで処理するより，単純な局所モジュールを多数用意して状況に合わせて使い分ける方が学習においても処理速度においても有利であろう．このような思想に基づき，モザイク

モデルでは予測器である内部順モデルと制御器である内部逆モデルの対を複数備えたものから構成される¹⁸⁾。

与えられた状況でどの対を使うかというのは、それぞれの順モデルがどれくらいよい予測をしたか（責任信号）によって決まる。個々の対は単純な線形の予測器・制御器であっても、それらを状況ごとに使い分けることで非線形ダイナミクスや状況依存の制御問題に対応し得る。またどれくらい学習するかは責任信号の大きさによって決まり、個々の内部モデルが線形でかつ独立であるがゆえ学習も単純で干渉も起きない。つまり、モザイクモデルは複数の単純な対を柔軟に組み合わせることで、非線形制御と学習を可能にしているのである。

1-2-5 結 語

本稿では工学理論がどのように脳機能を理解する枠組みを与えるか、また脳に代表される生物の情報処理がどのように新たな工学理論を促してきたかについて解説した。1960年代にパーセプトロンの研究者は、「単純なニューラルネットにより脳の原理は理解されており、コンピュータが十分速くなればヒトの脳機能を再現できる」と楽観的に考えていたそうである。現実にはコンピュータが何万倍も速くなった現在ですら、パターン認識・身体制御といった脳がたやすくこなせる課題においても、コンピュータの能力はヒトに比べるべくもない。これは脳の原理が未だに完全には理解されていないということであり、その理解には更なる脳科学と工学理論の相互作用が必要と思われる。

参考文献

- 1) D. Marr, "Vision," Freeman Publishers, 1982.
- 2) 川人光男, "脳の計算理論," 産業図書, 1995.
- 3) S. Thrun et al., "Probabilistic Robotics," The MIT Press, 2005.
- 4) Rosenblatt, "Principles of Neural Dynamics," Spartan, 1962.
- 5) T. J. Sejnowski and C. R. Rosenberg, Complex Systems, **1**, 145-168, 1987.
- 6) N. K. Logothetis and J. Pauls, Cerebral Cortex, **4**, 401-414, 1995
- 7) R. A. Jacobs et al., Neural Computation, **3**, 79-87, 1991.
- 8) C. E. Rasmussen and C. Williams, "Gaussian Process for Machine Learning," The MIT Press, 2006.
- 9) T. P. Jung et al., Proc IEEE, **89**, 1107-1122, 2000.
- 10) B. A. Olshausen and D. J. Field, Nature, **381**, 607-609, 1996.
- 11) Y. Weiss et al., Nature Neuroscience, **5**, 598-604, 2002.
- 12) M. Kawato et al., Network, **4**, 415-422, 1993.
- 13) R. P. N. Rao and D. H. Ballard, Nature Neuroscience, **2**, 79-87, 1999.
- 14) M. I. Jordan, D. M. Wolpert, In "The New Cognitive Neuroscience," (ed.) M. Gazzaniga, The MIT Press, 1999.
- 15) M. Kawato, Current Opinion in Neurobiology, **9**, 718-727, 1999.
- 16) M. Kawato et al., Biological Cybernetics, **57**, 169-185, 1987.
- 17) Y. Kobayashi et al., J Neurophysiol, **80**, 832-428, 1998.
- 18) D. M. Wolpert and M. Kawato, Neural Networks, **11**, 1317-1329, 1998.

脳と神経モデル - 1 章

1-3 脳と複雑系モデル

(執筆者：津田一郎)[2008年5月受領]

1-3-1 要約

脳はニューロンの集合体であり、変化し得る結合強度をもつシナプスによってネットワークを構成している。しかし、結合は機能的であり、それゆえに機能単位はそれによって構成されるシステムの機能によって自己無撞着的に規定される。すなわち、脳というシステムを構成する機能単位はレディーメイドではなく、オーダーメイドである。こういったシステムの構造は複雑系の主たる特徴である。本節では複雑系科学において発展してきた数理概念や数理モデルを用いて脳のダイナミクスの意味を探ってみたい。

1-3-2 はじめに

近年、脳神経生理学の幅広い分野において脳のダイナミクスに注目が集まっている。神経回路網理論のダイナミクスをはじめ脳の理論的研究においては早くからダイナミクスが注目され、長い研究の歴史がある。神経回路網を力学系として扱うことで、その動力学を論じるといふ研究動向は1960年代後半から1980年代にかけて、典型的には甘利俊一¹⁾、Kohonen²⁾、Malsburg³⁾らによって連想記憶、自己組織化マップ、Hebbの動的集成体仮説などを理論化していく過程で起こった。1980年代初頭にはHopfield⁴⁾による連想記憶のスピングラスモデルが提案され、そのアトラクターダイナミクスとともに応用が注目を集めた。更に、1980年代から90年代にかけて、川入光男らによる運動制御の動的理論がMarr-Albus-伊藤正男と続く小脳関連運動制御に関する理論と実験事実を基盤にして発展した⁵⁾。

筆者も1980年代半ばから脳研究に参入し、脳のダイナミクスに焦点を当て、脳活動を動力学的な見方で研究することの重要性を一貫して強調してきた。特に、その中で力学系のカオスの情報力学に関する理論を構築し、脳活動にカオスが現れる可能性、及びカオスが脳の動的な情報表現に重要な役割を果たすであろうという予想を行った。脳のダイナミックな情報処理様式をカオス力学系の観点から研究するという意味で、これを標語的に“カオスの脳観”と呼んだ⁶⁾。

1980年代後半には、Freeman⁷⁾⁸⁾がウサギやラットの嗅球においてカオス的な振動現象を発見し、カオスのもつ軌道不安定性が臭い知覚におけるパターン分離能力の大きさを保証しているという説を出して注目を集めた。Freeman, Kay, Kozmaらは、臭い情報処理に係る嗅球、嗅皮質、内嗅野、海馬の同時測定によって、これらの部位にカオス的活動を観測した。動物がモチベーションをもち臭いをかぎ、知覚し、判断し、行動決定し、実際に行動を起こす、という一連の機能発現の過程において、これら脳部位に局所脳波、特に γ と β 帯域の活動が表れ、その波がカオスの海の中で伝播していく様子が明らかになった。

カオスの中を情報が伝播するという事実は、カオスが状態の混合を引き起こす性質をもっていることから驚くべきことのように思われるが、これはまさに私たちが1980年代の半ばに構築したカオスの情報理論から十分に説明できることである。これは1-3-5項で説明しよう。神経細胞がカオス的な信号を生み出すことは1980年代初頭に主に日本では林初男と石塚智²⁾、松本元と合原一幸³⁾の二つのグループにおいて発見された。この発見をもとに、合

原は南雲・佐藤のニューロンモデルをその基本的な分岐構造を保存したまま可観測力オスモ出現できるように拡張し、カオスニューロンモデル¹¹⁾を構築するとともに、カオスニューラルネットワークを構成し、より多様な神経系のダイナミクスを研究した。

また、やはり1980年代の半ばから、清水博、山口陽子、矢野雅文と筆者が中心になって非線形振動子の引き込みを使った記憶や知覚情報処理のモデルを提案した¹²⁾。これは、脳活動には非線形振動子で表現されるような振動状態が現れ、それらの間に振動数の引き込みや位相の同期が起こることで知覚が成立するであろうという作業仮説に基づいて提案されたものである。この理論の少し後になって、Freemanの発見に刺激を受けたSingerやGrayら¹³⁾はネコの視覚野において、スパイク同期に基づくと思われる振動状態の同期現象を見つけ、自律的な特徴抽出の機構との関係を論じた。

脳に対する動的な見方の一つの例として受容野が動的に変化するという塚田稔らの先駆的な研究¹⁴⁾がある。これは1990年代に入ってDinse¹⁵⁾によって実験的に確認された。これに関連して、塚田は時系列情報を神経細胞が如何に処理しているかを考え、マルコフ性に焦点を当てた統計的な解析を行い、海馬CA1における時空間学習則を提案した¹⁶⁾。これは後に筆者との共同研究につながる。

脳のダイナミクスに焦点を当てた過去の研究はまだ多くあるが、それらすべてを紹介することが本稿の目的ではない。本稿では、次の1-3-3項において、こういった脳のダイナミクスの基盤の一つとなる非線形・非平衡系の研究と力学系の関係について簡単に解説する。1-3-4項において、非線形非平衡系の研究の一つの源流をおく複雑系科学の成り立ちと特徴を述べる。1-3-5項ではカオスネットワークの情報理論のエッセンスを簡単に述べる。1-3-6項において、最近脳科学で話題になっているアトラクターダイナミクスと遷移ダイナミクスを非線形力学系の観点から解説し、記憶や知覚との関係を述べる。1-3-7項は、塚田グループとこの10年間行ってきた共同研究についての概説と将来展望にあてる。

1-3-3 非線形・非平衡系と力学系

物理学は平衡系や閉鎖系の現象に関しては成功を収めた。システムにエネルギーや物質の流れが常に注がれるとき、システムはもはや熱的にも力学的にも平衡状態ではなく、平衡から遠く離れた非平衡状態に置かれる。このとき、システムに起こる構造化は平衡構造に勝るとも劣らないほど多様である。システムを制御するパラメータを一方向に変化させたとき、システム内で生成される秩序構造は臨界点を越えると不安定になり、新しい秩序構造に取って代わられる。パラメータの変化とともに、このような転移が限りなく繰り返される。

このように、平衡相転移が通常はパラメータの変化に対して無秩序から秩序への変化に留まるのに対して、非平衡系で起こる秩序状態の転移は多くの場合繰り返し起こる。この転移をPoincaréにちなんで分岐と呼ぶ。非平衡系では分岐の多様性が構造の多様性の主たる要因である。こういった非平衡系の分岐構造は多くの具体的なシステムで詳しく研究されてきた。流体系、光系、化学反応系はもとより、多くの生命現象においても研究が進んだ。心臓の拍動、糖代謝系のダイナミクス、指尖脈波、神経細胞のパルス発火やバースト発火、更には種々の脳波のダイナミクスなどが詳しく研究されてきた。

神経の活動状態は閉鎖系の平衡状態ではなく、開放系の非平衡状態である。非平衡系の秩序状態は巨視的物理量の平均的なダイナミクスに支配されると考えられるので、数学的には

微分方程式や写像（写像の繰り返しによる差分方程式）がよい記述を与える．特に、非平衡系の秩序形成は系に内在する非線形性によって引き起こされるので、これら微分方程式や差分方程式の一価解析的な解を陽に求めることは一般にはできない．そこで、こういった方程式の定性的な解法が必要になる．これは 100 年以上前に Poincarè によって天体力学の 3 体問題の解法のために創始された方法である．今日では、まさに Poincarè が発見したカオス的な解の構造までも明らかにする理論として洗練された理論体系が出来上がっている．力学系という数学の分野である．

1-3-4 複雑系科学の成り立ちと複雑性

複雑系、あるいは複雑系科学に関しては多くの場所で論じてきた¹⁷⁾¹⁸⁾¹⁹⁾．詳しくはそちらを参考にしていただくとして、本項ではそれらのエッセンスだけを述べる．上記参考文献で与えられた見解は、生命こそが複雑系だという観点である．しかし、例えば、認知科学者で経済学者の Herbert Simon は複雑系を人間がつくった人工システムすべてだと考えている．こういった違いはあるが、Simon が示した今日の複雑系研究につながる研究史は大変参考になる．ここでは Simon の著作²⁰⁾を参考にして筆者の考察を加えながら、まずは複雑系研究の歴史を概観する．

サイモンによれば複雑系研究は三期から成る．第一期は第一次大戦後にヨーロッパで起こった．そこでは、ゲシュタルト心理学が代表的な複雑系研究としてあげられている．物の認識において、系を構成する部分に還元して理解するのではなく、全体を全体として形という観点から理解しようという心理学の研究分野が起こったのである．日本では、明治時代に寺田寅彦が身近な物理現象に根本的な複雑さが現れることを看破し、いわゆる寺田物理学を築いた．現代日本には、この流れを汲む「形の物理学」²¹⁾や「パターン形成」²²⁾の研究がある．更には、現代の生命科学につながる研究として、Henri Bergson の創造的進化の研究があげられる．Bergson は生命をある種の“心的”状態の連続性として捉えた．断絶できないもの、時間の連続体として生命を捉えた．このように、第一期の研究対象である複雑性の特徴は全体性、連続性にあると言えるだろう．

第二期は第二次大戦後にアメリカ合衆国で起こった．サイバネティックスの運動が典型的である．サイバネティックスは諸科学をつなぐ科学であり、現代の複雑系科学につながっていくものである．Nobert Wiener は複雑で予測が難しい時系列の解析を可能にする数学理論を構築することで、予測理論、フィードバック制御理論の基礎を築いた．サイバネティックスの成果の一つとして神経系の数学モデルがあげられる．McCulloch-Pitts の形式ニューロンである．この形式ニューロンで適当なネットワークを構成すれば、万能チューリングマシンと等価であることが証明される．万能チューリングマシンはアルゴリズムでかけるものならずべて計算できるマシンである．今のデジタル計算機は万能チューリングマシンの一種である．すなわち、単純な形式ニューロンの集まりでさえ、既にデジタル計算機の能力があることが保証されているのである．このように、第二期の研究対象である複雑性の特徴は予測と普遍性にあると言えるだろう．

第三期は 1980 年以降に起こった．非線形・非平衡系の研究に代表される研究動向である．Ilya Prigogine は研究の指針をベルグソンの創造的進化においたことは興味深い事実である．Prigogine は非線形・非平衡系の熱力学を拡張して散逸構造論²³⁾を打ち立てた．Herman Haken

はサイバネティックスの流れを汲むシナジェティックス²⁴⁾という学問を提唱した。これは諸科学をその共通普遍項としての相転移という概念でつなげてみようという試みである。既に述べたように、非線形・非平衡系の相転移的な状態変化は数学的には分岐現象と呼ばれている。

カオスは非平衡状態の分岐現象の一つであり、系の巨視的な性質を表している。一般に、系の巨視的性質は平均化方程式によって記述される。平均化方程式は確率的要素を含まない決定論的な方程式である。この決定論的な方程式の解の中には非周期的で予測不可能な振る舞いをする解が存在することがある。これがカオスと呼ばれている予測不可能な現象を表す。また、カオス軌道を形成するアトラクタはフラクタル構造を含む。いくつかの典型的なカオス力学系においては、カオスとフラクタルの双対関係が知られている。このように第三期においては、確率現象も決定論的な力学から生成され得るという点が強調された。したがって、第三期の研究対象である複雑性の特徴は決定論的な不確実性にあると言えるだろう。

さて、系を各部分に分けて考えることができる系を単純系ということにしよう。つまり、系から部分を分離してもなおかつその部分は系に組み込まれていたときの機能を発揮できるような部分から成り立つ系を単純系だと言うことにする。要は、機能をもつ部分の単なる集合と等価であるような系である。仮に部分系同志が相互作用したり、結合したりしていても、部分系の機能は不変である。単純系をこのように考えると、複雑系はまさに単純系ではない。

具体例をあげよう。単純系の典型的な例は、工業用ロボットである。例えば、物体の位置を認識して物を一定の場所に集める工業用ロボットを考えよう。ロボットは個別特異的な機能をもつ部分から成る。頭には視覚情報を処理する目が部分系として取り付けられ、物をもったり移動したりする手と足の部分系がある。各部分系を構成する物質のより小さなレベルではもはやそれは目でも手でも足でもないで、系の機能は機能的な部分系（目、足、手）に還元され、それらを総合することで理解される。また、そのようにしてロボット全体は作られるのである（脳の機能を研究するための脳型ロボットはこのように単純ではないことは既に承知の事実である）。

一方、複雑系の典型例は、脳の機能分化や胚発生の過程での細胞分化に典型的に現れる。これらは最初に部分系である機能モジュールができてその後にそれらが合成されて系全体の機能が作られるわけではなく、系全体が機能的になるように発展する過程の中で各機能モジュールが部分系として生成される。しかも、この過程を決めている法則は長い生物進化の中で決まってきたことである。どのようにして、部分系が作られるのか、その機構を解明するのが複雑系科学の大きな使命である。

1-3-5 カオスネットワークの情報理論

筆者は松本健司, John S. Nicolis とともにカオスの情報理論²⁵⁾²⁶⁾²⁷⁾²⁸⁾をつくってきた。特に、カオスに外部から雑音を印加すると秩序状態が出現するノイズインデューズドオーダー²⁹⁾の発見が切っ掛けで、カオスを情報チャンネルとみなしてその容量と情報の流れの間の関係を突き止める必要性に迫られた。カオスの情報生成の理論は大野克嗣の先駆的な仕事³⁰⁾の後、R. Shaw³¹⁾やJ. Crutchfield, 松本健司, 金子邦彦と筆者によって発展された¹⁸⁾。筆者とNicolisはカオスの情報容量をリアプノフ指数で算出する方法を考え、これを脳の短期記憶容量と関係づけた。ここでは、カオスやカオスネットワークの中の情報の伝播の発見に関する基礎理論を簡単に解説する。

カオスは軌道不安定性をもっているのです、わずかの誤差が指数関数的に拡大する。この事実が、多くの人にカオスの中では情報は単に減衰するだけであるのでカオスは情報処理装置としては機能しない、という命題を信じさせた。しかし、我々はすべてのカオスがそうではなく、非一様なマルコフ分割をもつようなカオス力学系ではむしろカオスは情報をダイナミックに保持し、それと結合した他のカオス力学系に適切に情報を伝播させる機能をもつことを発見した。カオスの情報理論を構築するに当たって、我々はカルバック情報量を基礎においた。これによって、カオスの不変測度に相対的に時々刻々の情報量が計算できる。更に、情報流という新しい概念をカルバック情報量の時間変化（写像の場合は一回の写像の前後の差であり、ベクトル場の場合は時間微分）で定義した。これにより、平均情報流と情報流のゆらぎが定義できる。カオスの場合は、初期分布がもっていた情報は平均として失われる。例えば、これをビット空間で見れば、下の桁にあった情報は上の桁に流れたり、下の桁に戻ったりしながら、平均として上の桁に流れていき、最も上の桁で失われる。このように情報流にはゆらぎがあるが、どんなカオス系も平均としては下の桁から上の桁に情報が流れ、最上桁で失われる。興味深いことに、情報流のゆらぎが小さいときは、情報流はリアプノフ指数に一致することを証明できる。これは多くの人が描いていた平均的なカオスの描像に一致する。

それでは、情報流のゆらぎが大きいとどうなるか。このときはもはやカルバック情報量ではカオスの情報構造の詳細は見えない。それを見るために、時間付き相互情報量を導入した。これを単体のカオス系に適用すると、相空間（状態空間）の分割のあるセルと別のセルが共通にもつ情報量が計算できる。時間変化を考慮するとどのセルからどのセルに情報が流れたかを概略推定できる。また、状態を二進展開し、ビット空間で相互情報量を考えるとあるビットから別のビットへどれだけの情報が流れたかを計算できるのである。また、時間付き相互情報量をカオス結合系（カオスネットワーク）に適用すれば、どのカオスからどのカオスへどれだけの情報が流れたかを計算できる。時間付き相互情報量は大変便利な量である。なぜなら、それは情報流のゆらぎが小さい場合を含んでいるからである。すなわち、これ一つでどんなゆらぎの場合も情報流を計算できる。時間付き相互情報量の時刻0での（減衰の）傾きが最大リアプノフ指数を与える。また、情報流のゆらぎが小さい場合、時間付き相互情報量は時間に対してほぼ線形で減衰する（減衰しきった最後はむしろ指数的である）。これは、毎時間同じ量だけ情報が失われることを意味する。すなわち、ビット空間で考えると情報は下の桁から上の桁にトコロテン式に流れるのである。それに対して情報流のゆらぎが大きい場合、時間付き相互情報量は時間に対して指数減衰をする。これは毎時間同じ割合で情報が減衰することを意味する。こういうことが起こるのは情報流のゆらぎが大きいからである。

情報流の大きなゆらぎは、様々な桁にあった情報を混合する。それに加えて平均として情報流が下の桁から上の桁に向かい失われるので、毎時間同じ割合で情報が失われるのである。この情報混合の概念は重要である。なぜなら、情報混合が起こると、どのビットを見てもほとんどすべてのビットの情報が入っているからである。すなわち、多くのビットの情報が仮に壊れたとしても少なくとも一つのビットが生き残っていれば、もとの情報を再構成可能であるからである。この性質が効果を発揮するのは、カオスを結合した場合である。入力情報をこの結合系に与えてそれがカオス結合系に保たれるかどうかはまさに個々のカオス系が情報混合をもつかどうかに依存する。実際、要素カオス系が情報混合をもつならば、入力情報はこの結合系の中にダイナミックに保たれる。それは入力された情報がカオスに入り減衰し

てしまう前に情報混合によってほとんどのビットにばら撒かれて結合した別のカオス系に伝えられるからである。このようにして入力情報は減衰なくカオス結合系を伝播する。最近分かったことは、例えばカオス結合系がフィードバック回路をもち複雑なネットワークを構成していても、この性質は保たれるということである。

1-3-6 脳におけるアトラクターダイナミクスと遷移ダイナミクス

最近、O'Keefe のグループはラット海馬の CA1, CA3 における場所ニューロンのダイナミックな振る舞いを観測し、アトラクターダイナミクスが働いていると主張した³²⁾。これと関連して、Moser のグループと McNaughton のグループが共同で類似の実験を行い³³⁾、アトラクターダイナミクスと解釈される場所ニューロンの活動の他に、力学系の他のタイプのダイナミクスあるいは確率過程的なダイナミクスと解釈される^{34, 35)}ものも見つかった。実空間での場所を連続的に表現するためには、場所ニューロンの相空間でのダイナミクスはアトラクターが連続的につながったような連続アトラクター³⁵⁾であるか、もしくは連続的につながった非アトラクター³⁶⁾かであろうと推論されている。

他方で、アトラクターダイナミクスが不安定化し、幾何的アトラクターによる記憶の表現の描像は崩れ、ミルナーアトラクターによる記憶痕跡の表現が現実性を帯びてくる。更に、ミルナーアトラクターが不安定化することで遷移ダイナミクスが働くようになる。この遷移ダイナミクスがエピソードの脳内再構築の機構を与えるとする筆者らの提案がある³⁷⁾。実際、私たちの非平衡神経回路モデルにおいては³⁸⁾、強度の強い入力に対しては通常の単発の連想が起こり、そうでない場合には連続連想が起こることが分かっている。更に、エピソード記憶に対して場所の記憶は手がかりを与えるが、場所ニューロンがこのような遷移ダイナミクスに従わねばならないという先見の理由はない。遷移ダイナミクスは場所ニューロンではない別のニューロンの活動である可能性もあり、それがエピソードの脳内再構築であるかもしれない。

私たちは高次元力学系の典型的な遷移ダイナミクスとしてカオスの遍歴という概念を提唱した³⁹⁾⁴⁰⁾⁴¹⁾。これは秩序状態間をカオス的に遷移するもの一般をさす。数学的な機構がいくつか提案されている。カオス的遍歴的な遷移は、Freeman や Kay によってラットやウサギでの臭い情報処理と関係して観測されている。そこでは、カオス状態の中を γ 波や δ 波に乗って情報が伝播する様子も示された。前小節で述べたことが動物の情報処理で観測されたのである。

最近、池谷ら⁴²⁾は CA3 の培養環境において異なる神経活動状態の間の遷移を発見した。観測された状態は、ランダムで非同期的な神経発火状態、アップ - ダウン状態、 θ リズムの周波数での同期状態、部分同期状態、全同期状態などである。ムスカリン受容体のアゴニストであるカルバコールをバスで与えたときにこれらの状態間の動的な遷移が見られた。遷移はしばしば非周期的に見えるが、確率的でもある。逆に、ムスカリン受容体のアンタゴニストであるアトロピンを投与するとこれらの自発的な遷移は消失し、単一の状態に留まる確率が圧倒的に増加した。この現象の背後に大自由度力学系の遷移機構が関与しているであろう。この遷移がカオス的遍歴であるかどうかは分かっていない。ヘテロクリニックサイクルの可能性は高いかもしれないが、これが可観測であるためには系に対称性がなくてはならない。どのような対称性があるのか、またそうでないならば、カオス的遍歴との関係はどうか、な

ど興味は尽きない。

1-3-7 複雑系からの将来展望

筆者はおよそ 10 年前から玉川大学の塚田 稔のグループと二つの共同研究を行ってきた。一つは、海馬 CA1 におけるカントールコーディング仮説の実証であり、他の一つは動物、特に高等哺乳動物の推論能力に関する研究である。紙数の都合で詳しい説明はここではできないが、概略を述べておこう。

筆者は 1995 年から力学系におけるカオスとカントール集合の双対関係の研究を開始した。O.Rössler との共同研究が切っ掛けである。簡単なニューロンモデルを使ってカオスニューロンが生成した時系列情報を別の安定なニューロンの膜電位に如何にコードするかを研究し、膜電位に現れるカントール集合がまさに時系列情報を正確にコードする様子を調べた。これを海馬 CA3・CA1 の相互作用系に適用するために簡単な数学モデルでこれを試した結果、モデル CA1 にカントール集合が現れ、それが CA3 の出力である時系列の情報をコードしていることを突き止めた。そこで、塚田グループと共同実験を行うことにした。この過程を通じて実験とモデル計算の結果を突き合わせて議論できるように、モデルも神経生理学的に意味のあるものに直して計算し直した。実験は福島によってラット海馬の CA1 で行われた。CA1 錐体細胞はその膜電位に入力時系列を履歴に従って階層的に埋め込んでいることが明らかになったのである⁴³⁾。

もう一つのテーマは動物の推論機構に関するものである。1997 年に推論の数学理論の骨子を作成し、塚田、斉藤と議論した結果、共同研究を行うことになった。その後、数学理論は一つのケーススタディーとして結実した。サルの推論実験はいろんな意味で難航した。坂上雅道が実験を主導し、Pan が実験を行った。タスクを決定するまでもに紆余曲折があった。彦坂興秀氏にも議論に加わってもらい様々なコメントをいただいた。結局は二重連想の課題になったが、坂上の様々な工夫で、刺激に依存して報酬予測を行うなどの行動が実現されることと対応するニューロンが前頭前野で見つかった⁴⁴⁾。推論実験のスタートがやっと切れたことになる。

最近では、人の創造活動における脳活動の研究も行えるようになってきた。私たちは、創造性と脳のダイナミクスの関係を是非知りたいと思っている。上の推論実験もまだ緒についたばかりである。神経細胞のダイナミクスと推論能力の関係がまだ明らかにされていない。それが分らなければ、真に推論機構を理解したことにはならないと思うのである。また、海馬におけるエピソード記憶形成機構の研究の別の重要性は海馬性の認知症の機構との関係である。これに対しても、単に場所やニューロンの特定だけではなく、その神経メカニズムが明らかにされる必要がある。このとき、神経ダイナミクスが如何に関わっているかは極めて重要な問題である。様々な神経症状に関係する分子や遺伝子の同定だけでは根本的な解決には到らないであろうことは明らかのように思われる。なぜなら、脳神経系は複雑なネットワーク系であり、因果関係は決して一種類の分子や遺伝子に還元されるものではないからである。そこでは、動的に変化する様々な物質、神経活動の関係性によって因果関係が定められているはずである。したがって、この動的に変化する関係性を観測できるような新しい技術の開発や新しい脳の理論的枠組みの構築が、脳の真の理解と認知症などの神経関連疾患の治療には不可欠になると思われるのである。

参考文献

- 1) S. Amari, "Neural theory of association and concept-formation," *Biol. Cybern.* **26**, 175-185, 1977.
- 2) T. Kohonen, *Associative Memory – A System Theoretical Approach*, Springer-Verlag, Berlin, Heidelberg, New York, 1978.
- 3)
- 4) J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. of Sci. U.S.A.*, **79**, 2254-2258, 1982.
- 5) 例えば, 川入光男, "脳の計算理論," 産業図書, 1996.
- 6) 津田一郎, "カオスの脳観," サイエンス社, 1990.
- 7) W. J. Freeman, *Societies of Brains – A Study in the Neuroscience of Love and Hate*, Lawrence Erlbaum Associates, Inc., Hillsdale, 1995.
- 8) W. J. Freeman, *How Brains Make up Their Minds*, Weidenfeld & Nicolson, London, 1999.
- 9) 例えば, 林 初男, "脳とカオス," 裳華房, 2001.
- 10) 例えば, K. Aihara, I. Tsuda, and H. Fujii, "Prologue–The pioneering work of the late Gen Matsumoto," *J. Integrative Neuroscience*, **3**, 111-113, 2004.
- 11) K. Aihara, T. Takabe and M. Toyoda, "Chaotic Neural Networks," *Phys. Lett. A*, **144**, 333-340, 1990
- 12) 例えば, Y. Yamaguchi and H. Shimizu, "Pattern recognition with figure-ground separation by generation of coherent oscillations," *Neural Networks*, **7**, 49-63, 1994.
- 13) 例えば, C. Gray, A. K. Engel, P. Koenig, and W. Singer, "Synchronization of oscillatory neuronal responses in cat striate cortex: Temporal properties," *Visual Neuroscience*, **8**, 337-347, 1992.
- 14) 水野 真, 寺沢充夫, 塚田 稔, 福島弘毅, "ネコの網膜神経回路の非線形解析," 電子通信学会, *MBE79-83*, 147-152, 1980.
- 15) H. R. Dinse, K. Krüger, and J. Best, "A temporal structure of cortical information processing," *Concepts in Neuroscience*, **1**, 199-238, 1990.
- 16) M. Tsukada, T. Aihara, H. Saito, and H. Kato, "Hippocampal LTP depends on spatial and temporal correlation of inputs," *Neural Networks*, **9**, 1357-1365, 1996.
- 17) "岩波数学思想事典," 岩波書店, 1998.
- 18) 金子邦彦, 津田一郎, "複雑系のカオスのシナリオ," 朝倉書店, 1996 年; K. Kaneko and Tsuda, *Complex Systems: Chaos and beyond*, Springer-verlag, 2001.
- 19) "岩波数学辞典 (第 4 版)," 岩波書店, 2007.
- 20) ハーバート A. サイモン, "システムの科学 (第 3 版)," パーソナルメディア, 2001.
- 21) 例えば, 小川— 泰 他編, "形の科学百科事典," 朝倉書店, 2004.
- 22) 例えば, 西浦廉政, "自己複製と自己崩壊のバターンダイナミクス, 岩波講座「物理の世界」物理と数理解 5, 岩波書店, 2003.
- 23) G. ニコリス, I. プリゴジーン, "散逸構造," 岩波書店, 1977.
- 24) H. ハーケン, "協同現象の数理解," 東海大学出版会, 1980.
- 25) K. Matsumoto and I. Tsuda, "Information theoretical approach to noisy dynamics," *J.Phys.A:Math.Gen.*, **18**, 3561-3566, 1985.
- 26) K. Matsumoto and I. Tsuda, "Extended information in one-dimensional maps," *Physica*, **26D**, 347-357, 1987.
- 27) K. Matsumoto and I. Tsuda, "Calculation of information flow rate from mutual information," *J. Phys. A: Math. Gen.*, **21**, 1405-1414, 1988.
- 28) J. S. Nocolis and I. Tsuda, "Chaotic dynamics of information processing: The magic number seven plus-minus two revisited," *Bulletin of Mathematical Biology*, **47**, 343-365, 1985.
- 29) K. Matsumoto and I. Tsuda, "Noise-induced order," *J. Stat. Phys.*, **31**, 87-106, 1983.
- 30) Y. Oono, "Kolmogorov-Sinai entropy as disorder parameter for chaos," *Prog. Theor. Phys.*, **60**, 1944-1947, 1978.

- 31) R. Shaw, "Strange attractors, chaotic behavior, and information flow," *Zeit. für Naturf.*, **36a**, 80-92, 1981.
- 32) T. J. Wills, C. Lever, F. Cacucci, N. Burgess, and J. O'Keefe, "Attractor dynamics in the hippocampal representation of the local environment," *Science*, **308**, 873-876, 2005.
- 33) J. K. Leutgeb, S. Leutgeb, A. Treves, R. Meyer, C. A. Barnes, B. L. McNaughton, M-B. Moser, and E. I. Moser, "Progressive transformation of hippocampal neuronal representations in "morphed" environments," *Neuron*, **48**, 345-358, 2005.
- 34) M. Tsodyks, "Attractor neural networks and spatial maps in hippocampus," *Neuron*, **48**, 168-169, 2005.
- 35) J. A. Goldberg, U. Rokni, and H. Sompolinsky, "Patterns of ongoing activity and the functional architecture of the primary visual cortex," *Neuron*, **42**, 489-500, 2004.
- 36) A. Treves and E. T. Rolls, "Computational analysis of the hippocampus in memory," *Hippocampus*, **4**, 374-391, 1994.
- 37) I. Tsuda, "Towards an interpretation of dynamic neural activity in terms of chaotic dynamical systems," *Behav. Brain Sci.*, **24**, 793-847, 2001.
- 38) I. Tsuda, "Dynamic link of memories – chaotic memory map in nonequilibrium neural networks," *Neural Netw.*, **5**, 313-326, 1992.
- 39) K. Ikeda, K. Otsuka, and K. Matsumoto, "Maxwell-Bloch turbulence," *Prog. Theor. Phys. Suppl.*, **99**, 295-324, 1989.
- 40) K. Kaneko, "Clustering, coding, switching, hierarchical ordering, and control in network of chaotic elements," *Physica D*, **41**, 137-172, 1990.
- 41) I. Tsuda, "Chaotic itinerancy as a dynamical basis of Hermeneutics of brain and mind," *World Futures*, **32**, 167-185, 1991.
- 42) T. Sasaki, N. Matsuki, and Y. Ikegaya, "Metastability of active CA3 networks," *J. Neurosci.*, **17**, 517-528, 2007.
- 43) Y. Fukushima, M. Tsukada, I. Tsuda, Y. Yamaguti, and S. Kuroda, "Spatial clustering property and its self-similarity in membrane potentials of hippocampal CA1 pyramidal neurons for a spatio-temporal input sequence," *Cogn. Neurodyn.*, **1**, 305-316, 2007.
- 44) X. Pan, K. Sawa, I. Tsuda, M. Tsukada, and *M. Sakagami, "Reward prediction based on stimulus categorization in primate lateral prefrontal cortex," *Nature Neuroscience*, in press, 2008.