解　説

様々な角度から見たニューラルネットワークの将来像

5. 神経回路網理論と統計力学

椎　野　正　寿†

1. はじめに

ニューラルネットワークの研究の目標は，われわれヒトの脳をはじめ生物の神経回路網で実際に行われている分散並列型の情報処理のメカニズムを探ることであるとともに，その過程においてえられる様々な知見をヒントにして，ニューロコンピュータとも呼ぶべき新しいタイプの情報処理様式にもとづいた機械を人工的に開発することであろう．そうしたことが満足ゆくレベルで実現した暁には，世界観・自然観だけにとどまらず生活様式などのすべての点で，人類がこれまで予想すらしえなかったような，状況の質的変化の起こることが必至となるに違いない．

しかし，そのような夢のニューロコンピュータが獲得することになろう，実際の脳などの神秘的ともみえる働きのおおもとはといえば，単に一言でニューラルネットワークという自由度の大きい非線形力学系の示す挙動にあるはずである，とするのは言い過ぎであろうか．ずいぶんと単純に割り切ってみる見方だと思われる節もないではないが，脳の複雑な働きをわれわれ自身の脳で，ある程度満足のいく理解をしようと思えばそのような立場に立ってみるのも一つのゆきかたではなかろうかと少なくとも筆者は考えるのである．ニューラルネットワークは，興奮と非興奮が閾値により決められるかたちの非線形素子であるニューロンより成り立っており，それらが多数お互いにシナプス結合により相互作用し合う多自由度力学系を構成しているもので，外界からの刺激に応じて生ずる電気的活動の時空間パターンが連想記憶，学習，思考，自己組織化などといった柔軟性に富んだ情報処理の基礎になっていると考えられるからである．

2. なぜ統計力学なのか？

一般に多自由度力学系の性質を見極めるには，自由度の大きいことを活用して統計的に眺めることにより，逆に扱うべき自由度の数を減らすという常套手段を物理学では用いている．現在のハイテク技術を支える半導体物理学や磁性物理学にしても，アボガドロ数程度の多数の電子やスピンを統計的に扱うことにより巨視的現象を記述するという統計力学に立脚しているものである．したがってニューラルネットワークに対しても，まず自由度の大きさの点で統計力学的な考え方に頼らざるを得ないことになってきそうだと考えるのは自然なことであろう．

物理系の平衡状態を扱う統計力学には，系全体のエネルギとともに温度の概念が入っており，熱的なゆらぎに対応したギブス分布が中心的な役割を演ずるものである．そのため，絶対温度 0K のようなきわめて低温ではエネルギが，また一般に有限温度では自由エネルギが最小化される状態が実現されることになっている．

ニューラルネットの連想記憶モデルをはじめ，パーセプトロンやバックプロパゲーションなどの学習モデル，また TSP などの探索問題などにおいても，リアプノフ関数ないしは，評価あるいはコスト関数としてのエネルギ関数を考えることができ，これらを最小化する形で問題設定のできるものがある．そのような場合，ニューロン数やあるいはシナプス結合の数といった自由度を大きくした極限での系の性質を議論したい問題については，統計力学的な考え方がおおいに活用できるところとなるもので，近年この方面からの精力的な研究が行われてきている[1)~4)]．ニューラルネ

† Neural Network Theory and Statistical Mechanics by Masatoshi SHIINO (Dept. of Applied Physics, Faculty of Science, Tokyo Institute of Technology).
†† 東京工業大学理学部応用物理学科

ットの理論では，記憶パターンないしは入力パターンについてランダムな取扱いがなされることが多いが，その場合，レプリカ理論と呼ばれる，ランダム系の統計力学において用いられる処方せんが，これまた一役買うことにもなっている．さらに，上にあげた最小化問題に帰着されるニューラルネットにおいて，エネルギ関数のローカルミニマムを回避する目的や学習過程において汎化能力を高めるために，系に熱ゆらぎに対応したノイズを加えてモデルの確率過程化をはかるような場合には，統計力学に現れる温度の概念が自然と導入されることになる．

3. 連想記憶モデルの場合

筆者が現在まで主に関わってきた分野である，ニューラルネットの連想記憶モデルについて，統計力学的なアプローチが実際どのように展開されてきたのかについてみてみよう．Hopfield[5]が Hebb の学習則をシナプス結合に取り入れた連想記憶モデルに対し，エネルギを持ち込むことにより記憶想起の問題に物理的な解釈の光を当てて以来，統計物理の研究者が，それまで情報科学あるいは工学の分野の問題であるとされてきたものに殺到するかたちで研究をはじめだしたことは記憶に新しい．いくつかの注目に値する研究結果がだされてきた中で Amit たちの AGS 理論[6]がある．

Amit たちは，1，-1 の2値の状態をとるイジングスピンようのニューロンからなるニューラルネットが，確率的な状態更新を行うとするいわゆるイジングスピンネット（ISN）モデルに基づき Hopfield モデルの解析を行った．ランダムな記憶パターンを仮定してレプリカ理論を用いることにより，記憶パターン数の全ニューロン数に対する割合αの最大値すなわち臨界記憶容量を相転移理論の枠組みで明確に捕えてみせたのである．すなわち，温度 T をもつ ISN の平衡分布であるギブス分布に基づく自由エネルギの算出を，全ニューロン数 $N\to\infty$ という熱力学極限のもとで行った結果，まず，特定の記憶パターンとの一致度を示す相関（オーバラップとも呼んで m で表すと $m=1$ が完全な記憶想起に相当する）がほぼ1である記憶想起状態なるものが，熱力学的安定ないしは準安定状態として存在すること．さらにαを0より増加させたとき，記憶パターンの干渉によるノイズ増大のためその記憶想起状態は突然不連続的に（相転移理論では1次の相転移と呼び，よく知られた例では固体・液体，液体・気体の相転移がある）α のある値 α_c で消失することになり，その α_c が臨界記憶容量となること，などが分かったのである．ちなみに，温度0の場合には $\alpha_0\sim 0.14$ と得られている．

ここで重要なことは次の2点である．まず，記憶想起状態であっても m は正確に1となってなく完全な記憶想起が行われるわけではないことである．これは単に有限温度のためというだけではなく温度0の場合でもそうであって，統計力学的議論の末はじめて分かるといってよい，パターンの干渉効果が m に反映されたことによるものなのである．第2は，臨界記憶容量が相転移点として捕えられたことである．相転移は非線形力学系のアトラクタについての分岐現象そのものであるため，今議論している連続記憶ニューラルネットはアトラクタニューラルネットの一つの型（固定点型アトラクタをもつ）として位置づけられるものであることが明瞭になる．すなわち，一つの記憶パターンが想起されるということは力学系（この場合，確率的な力学系）の対応したアトラクタに系の状態が落ち込んでゆくということで理解される．

4. アナログニューラルネットへの一般化と SCSNA

レプリカ理論に基づく統計力学の方法は，シナプス相互作用の対称性が仮定されている場合には，系を構成するニューロンがイジングスピン型ニューロンでなくてもある程度広く適用できる．入出力関係を表す伝達関数で特徴づけられるアナログニューロンは，神経生理学的に対応がとりやすいということで工学的にも多用されるモデルニューロンであるが，シグモイド型のような単調な伝達関数をもつものでは，臨界記憶容量がやはりレプリカの方法によりきちんと求められるものとなる[7),8)]．定性的にはアナログゲイン β と ISN の場合の温度の逆数 $(1/T)$ との間には非常に良い対応がとれるという結果になっており，それ自体ある程度直感的にも納得のいくものである．ところが，伝達関数が非単調なものであると仮定すると，シナプス結合が対称であっても，もはやリア

プノフ関数が存在せず，これまでのような統計力学理論が使えなくなってきてしまう．ギブス平衡分布に基づいて自由エネルギを求める形式の議論が不可能になるそのような場合にも，広い意味での統計力学的考え方を用いた方法によれば対処できることになるのである．

筆者らは，従来のシグナルノイズ解析の精神に沿ってはいるが少し違ったやり方で，すなわち各ニューロンのシナプス重率付き入力の総和である局所場をシグナル，ノイズ，出力比例項の三つの部分にセルフコンシステントに分離して，記憶想起状態を記述する基本方程式のセットを導出することで，任意の形の伝達関数をもつアナログネットの臨界記憶容量の解析が可能になることを示した．SCSNA (Self-Consistent Signal-to-Noise Analysis；自己無撞着シグナルノイズアナリシス）と呼ぶその新しい方法は[9),10)]，シグモイド型のような単調な伝達関数に対しては，これまでのレプリカ法に基づく統計力学理論の結果をもちろん正確に再現するものである．ところが，問題の非単調伝達関数に適用すると，臨界記憶容量に従来の3倍弱程度までの増加がみられる[10)~13)]ということとともに，定性的にまったく新たな現象が見い出されることになったのである．

それは，通常の Hebb 学習則を取り入れたシナプス結合に基づく連想記憶モデルであっても，α のある領域にわたり実質的なパターンオーバラップ[10),11)]が正確に 1 となって，誤差のない完全な記憶想起が保障される状態が存在しえるということで，これまでの常識を破ることが起こっているものであるといえなくもない．スーパリトリーバル状態[10),11)]と呼ぶこのような記憶想起状態は，伝達関数の適当な位置に急激な現象部分がある場合に現れ，それは，局所場におけるノイズがあるしかけにより消失してしまうためであるということができる．

蛇足ながら，SCSNA はレプリカ理論など他の統計力学的方法と同様，ニューロン数 $N\to\infty$ の仮定のもとに成立するものであるわけだが，その結果自体はわずか数百個程度のニューロンからなるネットワークについてもかなりよく当てはまることが数値シミュレーションから分かっており，Nの有限性に起因する実際上の問題は起こらないことになる．

連想記憶モデルにおいては臨界記憶容量についての議論だけでは不十分であり，エネルギ関数のローカルミニマムである偽記状態がどの程度の数存在するものであるのかも見極めておかねばならない．そのような準安定状態の個数を計算することも，ニューロン数 N がきわめて大きいという条件のもとで，鞍点法と呼ぶ統計力学理論の常套手段を用いて実行することができる[14)~16)]．その結果，それぞれアナログゲイン β と温度 $T=1/\beta$ で特徴づけられるアナログニューラルネット (ANN) と ISN との間の，連想記憶モデルとしての性能の比較なども議論できるようになっている[17)]．

これまで，シナプス結合を Hebb 則という一つの学習則で特定した場合についての連想記憶モデルの性質について言及してきたわけであるが，シナプス結合を他にもいろいろ変えたときの，記憶可能なパターン数の理論的上限いわゆる Gardner 容量がいくらになるかという解析についても，鞍点法を用いて相互作用空間上での体積を評価することにより実行可能となっている[18)~20)]．Gardner の仕事はこの方面における系統的な解析の先べんを付けたという意味も含めて大きな評価をもつものである．

5. 今後に向けて

統計力学のサイドからニューラルネットワークの数理的側面をながめたとき，現在までのところ，一般性のある理論的解析がある程度，成功を収めてきたといえそうなところは，アトラクタニューラルネットの一つのタイプである，力学系のフローが固定点型のアトラクタをもつような場合に限られているようである．そこでは，そうしたアトラクタの存在とともに，自由度の大きいことが中心的な役割を果たしており，平衡系の統計力学ないしはその関連の手法が使用可能になっている．いわば，統計力学を，形式はそのままにして，その適用対象を従来の物性物理学の分野から生物神経・情報処理に関するニューラルネット理論へ拡張するという形で，他分野へ一部分開放したということになるであろうか．

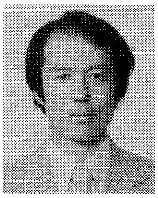
将来性に富んだニューラルネット理論のさらなる展開に向けて対応を取るべく，従来の平衡系統計力学の枠を越えて統計力学を構築することは，

物理学会外の分野への統計力学のより広範囲な開放を意味するとともに，それ自身をもより豊かな学問体系へと発展させることに通ずるものであると思われる．シナプス相互作用を非対称に取るなどして，生体の神経系に一歩近付けたモデルを構成すると，ニューラルネットにはエネルギの概念が適用できなくなってしまうわけであるが，そのことによりリミットサイクルやカオスといったアトラクタが生成される場合には，非平衡統計力学の枠組みで非平衡相転移などの議論を行うのも一つの方法となるであろう[21]．最近のいろいろと明らかにされつつある神経生理学実験に基づく知見からすると，ニューラルネット上に生ずる特に時間的な振動をともなう時空間パターンが，並列分散型の情報処理の鍵を握るものになるのではなかろうかとの感を強く抱かざるをえない．したがって，力学系の分岐現象をより鮮明に豊かにみせるそうした時間振動型のアトラクタに基づくアトラクタニューラルネットの数理構造を，情報処理機能との関連において系統的に解析する方向で，非平衡統計力学などの必要な道具の開発も含めて今後積極的な研究を推し進めていかなければならないように思われる．

参考文献

1) J. Phys. A.: Math. Gen. 22 (1989).
2) van Hemmen, D. and Schulten (Eds.): Models of Neural Networks (Springer-Verlag Berlin Heidelberg 1991).
3) Garrido, L. (Ed.): Statistical Mechanics of Neural Networks (Lecture Notes in Physics 368 Springer-Verlag Berlin Heidelberg 1990).
4) Watkin, T.L.H. and Rau, A.: Rev. Mod. Phys. 65, 499 (1993).
5) Hopfield, J. J.: Proc. Natl. Acad. Sci. USA 79, 2554 (1982); ibid. 81, 3088 (1984).
6) Amit, D. J., Gutfreund, H. and Sompolinsky, H.: Phys. Rev. Lett. 55, 1530 (1985); Ann. Phys. 173, 30 (1987).
7) Shiino, M. and Fukai, T.: J. Phys. A: Math. Gen. 23, L1009 (1990).
8) Kuhn, R., Bos, S. and van Hemmen, J. L.: Phys. Rev. A 43, 2084 (1991).
9) Shiino, M. and Fukai, T.: J. Phys. A: Math. Gen. 25, L375 (1992).
10) Shiino, M. and Fukai, T.: Phys. Rev. E 48, 867 (1993).
11) Shiino, M. and Fukai, T.: J. Phys. A: Math. Gen. 26, L831 (1993).
12) Morita, M.: Neural Networks 6, 115 (1993).
13) Yoshizawa, S., Morita, M. and Amari, S.: Neural Networks 6, 167 (1993).
14) Gardner, E.: J. Phys. A 19, L1047 (1986).
15) Fukai, T. and Shiino, M.: Phys. Rev. A 42, 7459 (1990).
16) Marcus, C. M., Waugh, F. R. and Westervelt, R. M.: Phys. Rev. A 41, 3355 (1990).
17) Fukai, T. and Shiino, M.: J. Phys. A: Math. Gen. 25, 2873 (1992).
18) Gardner, E.: Europhys. Lett. 4, 481 (1987); J. Phys. A 21, 257 (1988).
19) Gardner, E. and Derrida, B.: J. Phys. A 21, 271 (1988).
20) Shim, G., Kim, D. and Choi, M. Y.: J. Phys. A 26, 3741 (1993).
21) 合原一幸編著：ニューラルシステムにおけるカオス，東京電機大学出版 (1993) 第6章「ニューラルネットワークの統計力学とカオス」

（平成6年4月4日受付）



椎野　正寿

昭和20年生．昭和43年東京工業大学理工学部応用物理学科卒業．昭和 50 年同博士課程単位取得中退．理学博士．東京工業大学理工学部応用物理学科助手を経て現在，同助教授．専門は非線形・非平衡系統計物理学．応用統計力学の立場よりニューラルネットワークの研究に従事．日本物理学会会員．