

自分の脳がどうなっているかは誰でも気になる。どうして何年も前の光景を鮮やかに思い浮かべる事が出来るのだろうか。しかし頭を開いていくら詳しく中身を観察しても心の動きは見えてこない。高度な脳の機能と微視的な実験の間を結び付ける原理が必要である。うまくいくと、その原理を探求する過程でいままでのコンピュータとは違った方式の情報処理装置の作り方が見つかるかも知れない。物理をやってきた者にも何かこの分野で出来ることがあるだろうか。

### 1. なぜニューラルネットワークを？

Q: 西森さん、最近どんなことをやってるんですか？

N: ニューラルネットワークです。

Q: 何ですかそれ？

N: 簡単に言えば、脳の仕掛けを真似した系の理論的な解析です。日本語では神経回路網です。

Q: なんでもまた脳やら神経やらに手を付けたんですか？

N: 私自身は脳の高度な機能の解明を直接の目標にしてやっているのではなくて、いろんなモデルの数理的な構造を調べることが面白そうでした。要するに非線形の多体問題ですよ。もちろん長い目で見ればその成果が脳を理解するのに役立ったり、ニューロコンピュータみたいなもの設計原理になったりすることは願っていますし、それが研究費の獲得の名目にもなっていますが。

ニューラルネットワークは、基本になる単位（ニューロン：神経細胞）がたくさん集まって互いに信号をやり取りする系である。その相互作用の様子に応じてさまざまな情報処理が行われる。この分野は長い歴史を持っているが、物理出身の研究者が目立つようになったのはここ 10 年ほどである。そのきっかけの一つが、Hopfield によるニューラルネットワークとスピングラス（磁性体の一種）の理論との類似についての指摘であった。ニューロンがスピニンに対応し、ニューロン間の信号伝達効率がスピンの間の相互作用に相当するというのである。もちろんニューラルネットワークに手を染めた物理学者がみんなスピングラスの理論の応用を志していたわけではない。しかし、Hopfield の論文によって物理学者がこの分野に参入する障壁が大いに低くなったのも事実である。

Q: それで、物理出身の人たちはどういった仕事をし

たのですか？

N: さっき言ったみたいに、ニューラルネットワークの研究には二つの大きな目標というか方向があるんです。脳の機能を解明するために、特定の機能の発現に必要な要素を生理学や解剖学の成果から取ってきて、それを基にして定式化するやり方が一つ。もう一つは、実験的な事実は大切なヒントにはするけれど、例えばパターン認識をする装置を作るのに脳がやっていることをそっくりそのまま真似する必要はなくて、うまい設計思想を思い付けばそれでよいという方向。いずれにしても、新しい局面を開いてきた人達の中に物理の教育を受けてきた人は少なくありません。

### 2. キーワード

#### 2.1 ニューロン

脳での情報処理の単位はニューロン（神経細胞）である。ニューロンは本体から伸びた多くの突起（樹状突起）と、長い一本の繊維状の部分（軸索）が特徴的な細胞である（図1）。数多くの他のニューロンからの信号が樹状突起に入ってくる。これらの入力信号の和がしきい値を越すと、

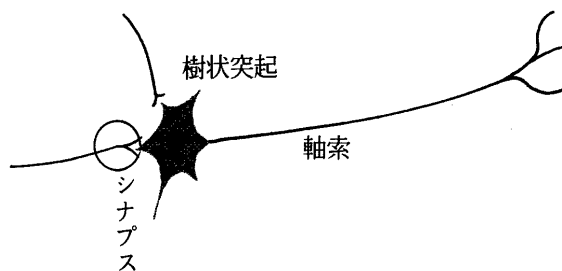


図1 ニューロンの模式図。信号は樹状突起からシナプスを通して入り、軸索から出ていく。

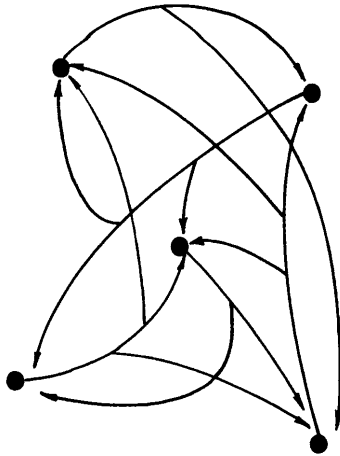


図2 相互結合型のニューラルネットワーク。各ニューロンは、結合を通して受け取った信号によって絶えず状態を更新している。

そのニューロンは軸索を通して次のニューロンへ信号を送るようになる。これを興奮あるいは発火と言う。まわりの人達の意見を聞いて自分の考えをひとつにまとめ、他人に伝えるというわけである。

## 2.2 シナプス

ニューロン間の接合部をシナプスと言う。シナプスは多くの場合、樹状突起にある。シナプスで受渡しされている信号は電気的なものである。たくさんの入力信号はニューロン内で単純に足し合わされるのではなく、それぞれの信号が符号・強さの重み付きで寄与する。例えば、ニューロンAからニューロンBへの入力はBを興奮させ、Bが自ら信号を出す方向に持っていき、別のニューロンCはBの興奮をおさえ、Bからの信号の送出手を抑制する傾向がある。前者の傾向を持つニューロン間の接合部を興奮性シナプス、後者を抑制性シナプスと言う。こうした競合するシナプスの存在によって多彩な情報処理が可能になっている。

## 2.3 相互結合型ネットワークと記憶

多くのニューロンが互いに結合して信号のやりとりをしているとしよう(図2)。ある瞬間には特定のニューロン群が興奮しており、次の瞬間には別のニューロン群が興奮している、という風に興奮のパターンが時々刻々変化している。たくさんの電球がチカチカと点滅を繰り返している様子を思い浮かべよう。何かを記憶しているというのは、単純に言い切ってしまうと、ある一定の点滅パターンがネットワークの安定な定常状態になっていることだと考えられている。各ニューロンは2.1, 2.2節で述べたやり方で、まわりからの信号を受けながら状態を更新して行く。そうした変化のあげくに点滅のパターンが定常状態に達したら、それに相当する記憶を思い出した(想起に成功した)ということになる。

信号 →

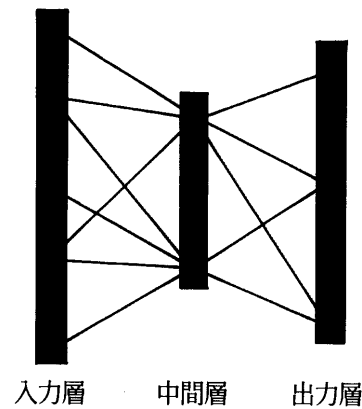


図3 フィードフォワード型のネットワーク。信号は入力層から出力層へ向かって一方に流れる。

## 2.4 フィードフォワード型ネットワークと学習

今度はニューロン群が何層かの層状に並んでいて、信号が入力層から出力層へ方向の流れを示すフィードフォワード型のネットワークを見てみよう(図3)。フィードフォワード型のネットワークの主な機能は、入力信号の種類に応じて出力層の異なるニューロンを興奮させること(入力信号の分類)である。例えば「あ」という文字が入力層のニューロン群の興奮パターンとして入力されたら、中間層での信号処理の後、出力層のある特定のニューロンが発火する。「い」という文字なら出力層内の別のニューロンの発火となる。こうした機能を実現するためのニューロン間のシナプス結合をどう与えるかが課題になってくる。これが学習の問題である。

## 3. 記憶していることを思い出す

相互結合型のネットワークでの記憶と想起について見てみよう。

### 3.1 自己連想記憶

ニューロン  $i$  が興奮している状態を  $S_i=1$  とし、興奮していないときを  $S_i=-1$  で表すことにする。ニューロン  $j$  から  $i$  へのシナプス結合の強さを  $J_{ij}$  とする。 $i$  が  $j$  から受ける入力は  $J_{ij}(S_j+1)/2$  であるから ( $S_j=\pm 1$  を入れてみよう)、入力の総和は  $\sum_j J_{ij}(S_j+1)/2$  となる。1/2の因子は積極的な役割を果たさないので省略しよう。上記の入力信号がしきい値  $\theta_i$  より大きいならば次の瞬間に  $S_i$  は1(興奮している)、そうでないなら-1(興奮していない)となる。式で書くと

$$S_i(t+\delta t) = \text{sign} [\sum_j J_{ij}(S_j(t)+1) - \theta_i] \quad (1)$$

である。  $\text{sign}[x]$  は  $x$  が正なら+1、負なら-1という関数。簡単のためしきい値  $\theta_i$  を  $\sum_j J_{ij}$  に等しく選ぶと、(1)式は

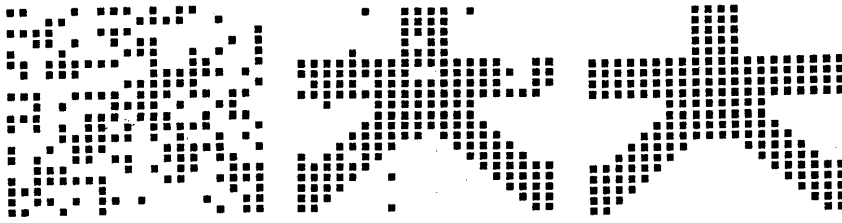


図4 20×20個のニューロンからなる相互結合型ネットワークに、一番右のパターンを含む6個のパターンを埋め込んで想起させる。興奮したニューロンを紫で表し、興奮していないニューロンの位置には何も書かない。初期状態(一番左)は、一番右のパターンの中でランダムに40%のニューロンを選んで状態を反転させたもの。6ステップ/ニューロンで想起に成功した、真ん中は2ステップ/ニューロン目での状態。

$$S_i(t+\delta t) = \text{sign}[\sum_j J_{ij} S_j(t)] \quad (2)$$

となる。これが系のダイナミクスを決める。(2)式によって、結合  $J_{ij}$  と初期条件に応じて各ニューロンの状態が時間発展をしていく様子が記述される。したがって、結合  $J_{ij}$  をどう選べば目的とする記憶(各ニューロンの特定の点滅状態)が定常状態として実現されるかが問題となる。

シナプス結合の選び方でよく使われるのが、一般化された Hebb 則と呼ばれる次の規則である。記憶させたいパターンが  $p$  個あり、 $\mu$  番目のパターンでの  $i$  番目のニューロンの状態が  $\xi_i^\mu = +1$  あるいは  $-1$  で表されているとしよう。このとき  $J_{ij}$  を次の式で決めるのが一般化された Hebb 則である。

$$J_{ij} = \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \quad (3)$$

(2)式に(3)式を入れるとうまくパターンの想起ができることが分かっている。図4に示した例では、最下段のパターンが他のいくつかのパターンとともに20×20個のニューラルネットワークに埋め込まれている(記憶させることをよくこう呼ぶ)。不完全なパターンから次第に完全なパターンを思い出す(想起する)ことに成功している。実際には、想起がうまくいくには、埋め込んだパターンの数  $p$  や初期条件、たくさんのパターン相互の間の類似性についての条件がこまごまとあることを注意しておく。

以上のような、自分自身に与えられたヒント(不完全なパターン)からの想起をするネットワークを自己連想記憶という。

### 3.2 Hopfield 模型

(2)式はよく見ると磁性体のイジング模型でのスピンの運動の規則と同じ形をしている。sign 関数の中身は、スピン  $i$  に他のスピンから掛かる有効磁場に他ならない。各スピンはまわりからの有効磁場の方を向くように sign 関数によって自分の向きを決めている。これは、エネルギー  $H = -\sum_{i,j} J_{ij} S_i S_j = -\sum_i S_i (\sum_j J_{ij} S_j)$  を出来るだけ低くするように運動する、絶対零度でのイジング模型のダイナミクスである。なぜなら、 $S_i$  と  $\sum_j J_{ij} S_j$  を同符号にすると  $H$  がより低い値を取るようになるからである。だから自己連想記憶の問題は、(3)式のような相互作用を持つイジング模型の統計力学になる。しかも(3)式の  $J_{ij}$  の符号は正にも負に

もなり得る。これは実は、スピンの向きがランダムに凍結したスピングラスと呼ばれる磁性体の基本的な特徴なのである。

Hopfield はこうしたことを指摘し、さらにある種の最適化問題(多変数関数を、拘束条件のもとで最小化あるいは最大化する問題)が相互結合型のニューラルネットワークで表現できることも示した。これが、たくさんの人たちがニューラルネットワークに取り組むきっかけのひとつになった。そこで、シナプス結合が(3)式で与えられる相互結合型ニューラルネットワークを Hopfield 模型という。実際には、Hopfield 以前にもこの模型は知られていて様々な解析がされていたことは忘れてはならない。

### 3.3 統計力学

さてイジング模型だということがわかると、統計力学によって相転移を調べてみようということになる。幸運なことに、Hopfield 模型では平均場理論が厳密な解を与える。(正確に言えば、そう信じる理由がある。)したがって、平衡状態に関しては理論的な解析がほぼ完全に実行できるのである。その結果が図5である。

縦軸はシナプスでのノイズレベル(信号伝達の不確かさ)を示すパラメータで、統計力学との類推から温度と呼ばれる。温度  $T$  が0でないと、ニューロンの状態更新は(2)式とは少し違って確率的な動きになる。すなわち、入力信号((2)式の sign 関数の引数)が正であっても、必ずしも  $S_i(t+\delta t)$  は  $+1$  にならないのである。(2)式は  $T=0$  の場合に相当する。図の横軸は1ニューロン当たりの埋め込まれたパターン数である。横軸  $p/N$  に直接対応する統計力学の量はない。あえて言えば、 $p$  は埋め込まれたパターン数だから、 $p/N$  の増加とともにパターン間の干渉による混乱(思い出しにくいこと)が増えるという意味で、ランダムさの度合と考えられるかもしれない。しかしながら、あまりこの類推は強調すべきではない。

さて、三つの異なる相が現れていることが見て取れる。図の上の方にある相(磁性体の常磁性相に相当)では、高いノイズレベルのために各ニューロンはランダムに興奮と静止を繰り返して連想記憶としての機能は果たさない。一方、左下の低ノイズ、少ないパターン数の領域では想起が可能である(想起相)。これらの二つの相の間にスピング

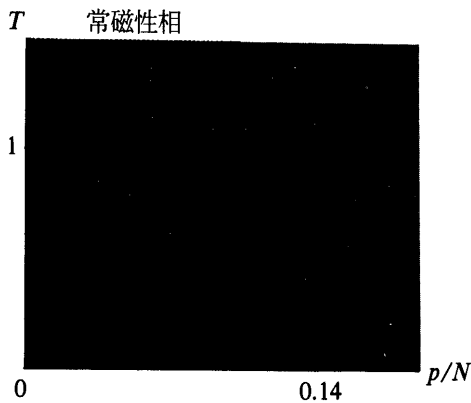


図5 Hopfield 模型の相図.  $T$ はシナプスでのノイズレベル,  $p/N$ は1ニューロン当たりの埋め込まれたパターン数. 想起は左下の想起相でのみ可能である.

ラス相がはさまれている. 図4のパラメータ ( $T=0, p/N=6/400=0.015$ ) は想起相内に位置しているので想起が出来たのである.

スピングラス相では, 常磁性相のようにニューロンの状態が刻一刻変わることはない. かと言って埋め込まれたパターンのどれかを想起しているわけでもない. 教えてもいないことに頭が凝り固まってしまったのである. 同じノイズレベル  $T$  の想起相と比べる (図を横に見て行く) とその原因が分かる. 記憶したパターン数 (横軸の値) がある限界値を越えるとスピングラスに相転移している. たくさん憶えすぎて全部忘れてしまったのである. しかも, この相転移はある種の1次転移である. 2次転移と違って, 前触れ (臨界ゆらぎ) もなく想起が突然不可能になるのである. 詰め込みの弊害はある日突然現れるということになる.

### 3.4 想起のダイナミクス

熱平衡状態の統計力学を使って解析するだけでは, ニューラルネットワークの本質を見失ってしまう. 本来, 式(2)のダイナミクスが基本であって, 時間とともに状態がどう変わるかが重要な関心事なのだ. 初期条件への依存の仕方 (どこまでヒントを悪くしても正解が思い出せるか) といった問題は, 時間依存性をあらわに含んだダイナミクスを直接調べないとはっきりした答えが得られない.

(2) 式をきちんと解いて  $\{S_i(t)\}$  の時間依存性を出せばすべて分かったことになるが, それは難しすぎて出来ない.  $i$  は1からニューロンの個数  $N$  までの値を取るが,  $N$  はひどく大きな数なのである (ヒトの脳だとすると  $10^{12}$  前後). そこで, もし数個のパラメータでネットワーク全体のマクロな性質の記述が出来れば, 第1近似としてはうまいだろう. まず考え付く重要なパラメータは, ネットワークの現在の状態が特定のパターンとどれだけ重なりを持っているかを表す  $m^\mu(t) = (1/N) \sum_i S_i(t) \xi_i^\mu$  である. さらに, 現在想起しようとしているパターン以外のパターンの影響がどれだけ入力信号を乱しているかを表す量  $\sigma(t)$  を導入するとよいことが明らかにされた. すなわち,  $m^\mu(t)$  と  $\sigma(t)$  の時

間発展を閉じた形で書き下せるのである. こうして, あるパターンを想起している途中のダイナミクスをコンパクトに記述できるようになった.

その結果, 想起相の中 (埋め込んだパターンがそのまわりのゆらぎに対して安定なパラメータ領域) であっても, 埋め込んだパターンから大きくずらせた初期条件から出発する時には, 初期条件とパターンとの類似性  $m^\mu(t=0)$  がある程度以上大きくないと実際には想起が出来ないことが示された. また想起の成功に導く初期条件の集合は, 状態空間の中で非常に入り組んだ部分空間を作っていることも明らかになった.

### 3.5 実際にはもっといろいろなことがある

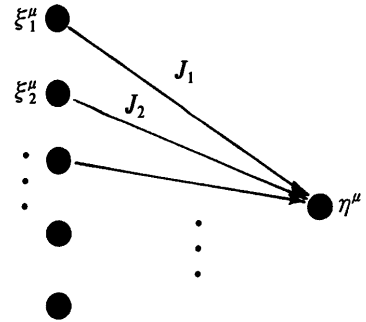
これまでは (3) 式のシナプス結合を持つネットワークについて説明してきた. しかし考えてみると, このネットワークは脳のモデルとしては相当非現実的だし, パターン認識装置としても制約が強すぎて役に立ちそうにもない. 例えば, (3) 式のシナプス結合は  $i$  と  $j$  の入れ換えに対して対称になっている. ニューロン  $i$  からニューロン  $j$  への信号伝達効率, 逆方向の信号伝達効率と全く同じであるということになる. こんなことは実際にはありえない. また図5で想起相が一番右に伸びている絶体零度 (横軸上) では, 埋め込まれたパターン数のニューロン数に対する比率は約 0.14 である. ニューロンの数の14%の数のパターンしか覚えられないというのは, いかにも情けない. こうした点をどう改善すれば良いかについては, たくさんの研究があって紹介しつくすことはとても出来ないが, ごくかいつまんでいうと次のようになる.

シナプス結合を Hebb 則 (3) から何らかのやり方でずらせて行く. 最初は 3.3 節で説明した固定パターンの想起状態が唯一の定常状態であるが, ずらせ方がある程度以上になるとリミットサイクルやカオスなどの非平衡な状態が現れ, これらが次第に支配的になる. リミットサイクルやカオスなどで表される非平衡の興奮パターンが脳の中での情報処理に積極的な役割を果たしている可能性も指摘されている. また, (3) 式とは違った方式でパターンを埋め込むことによって, 埋め込めるパターン数やパターン間の類似性 (相関) についての制約を大幅に緩められることも分かっている.

## 4. 新しいことを学習する

ニューラルネットワークに何かをやらせたいとき, どうやったらそれが可能になるかを教え込まないといけな

図6 一番単純なパーセプトロンの構造。( $\xi_1^1, \xi_2^1, \dots$ )には $\eta^1$ を、( $\xi_1^2, \xi_2^2, \dots$ )には $\eta^2$ を出力するように $J_1, J_2, \dots$ としきい値 $\theta$ を調整する。



3節では、(3)式のHebb則を使うと記憶と想起ができると述べた。もっと別のいろいろな機能を実現するにはどうすれば良いだろう。これが学習の問題である。

普通のコンピュータに何かをやらせたい時には、細部にわたるまで具体的な処理手続きを明示したプログラムを与えなければならない。ところが私たちの脳は全く別の仕方ですべての必要な機能を獲得している。学校で習うのは言語で表現された知識であって、頭を開けてどのニューロンとどのニューロンをつなぐかを実習したりするのではない。ブラックボックス(脳)の中では、環境からの入力(刺激)と生きていくのに必要な出力(行動)とをうまく対応させるようにシナプス結合が自然に形成・修正されている。

この節で扱うのは3節の自己結合型のネットワークではなく、信号の流れが一方向のフィードフォワード型のネットワークであることを注意しておこう。ダイナミクスは(2)式と同じであるが記憶容量を問題にするのではなく、シナプス結合の逐次的な修正によって、入力と出力の間に適切な関係を成立させたいのである。

#### 4.1 パーセプトロン

単純な学習課題として、入力された $\pm 1$ の組(パターン)を何らかの基準で2種類に分類する問題を考えてみよう。例えば、画面上に白黒のドットの集合として与えられたパターンが人の顔か否かを識別するのである(この例は実はあまり単純ではないのだが)。入力層に提示されたパターンが、可塑性を持った結合を通して出力ニューロンに送られるとする(図6)。これはパーセプトロンと呼ばれる入力識別ネットワークの中で一番簡単なものである。入力を $\{\xi_j^{\mu} = \pm 1\}$ 、望ましい分類結果を $\eta^{\mu}$ ( $=1$ あるいは $-1$ )としよう。 $\mu$ は入力パターンの番号、 $j$ は入力層内のニューロンの番号である(図6)。 $j$ 番目のニューロンと出力ニューロンとの結合 $J_j$ としきい値 $\theta$ をうまく調節して

$$\eta^{\mu} = \text{sign}[\sum_j J_j \xi_j^{\mu} - \theta] \quad (4)$$

を実現したい。顔の識別の例で言えば、 $\mu$ 番目のパターンが顔なら $\eta^{\mu} = 1$ 、違う時には $\eta^{\mu} = -1$ となるようにしたのである。それには次のようにすればよい。

ともかくも何らかの結合の値を初期値として与え、それに基づいて(4)式の右辺によって出力を計算し、その結果を $s^{\mu}$ とする。 $s^{\mu}$ は必ずしも $\eta^{\mu}$ とは一致しない。もし一致していたならば $J_j$ と $\theta$ は変更しない。違っていたら誤りを訂正する方向にパラメータを修正する。具体的には、 $s^{\mu} = -1$ ( $\eta^{\mu} = 1$ )なら(4)式の $\text{sign}$ 関数の引数を増加させ

$\eta^{\mu} = 1$ が出やすくなるよう結合を強め( $\Delta J_j = \epsilon \xi_j^{\mu}$ )、 $s^{\mu} = -1$ ( $\eta^{\mu} = -1$ )なら弱める( $\Delta J_j = -\epsilon \xi_j^{\mu}$ )。 $\epsilon$ は微小な正の定数。 $\theta$ も同じような方法で正しい方向に修正して行く。もしすべての $\mu$ について(4)式を満たすパラメータ $\{J_j\}$ と $\theta$ が存在するならば、この学習方式によって正しいパラメータのどれか一つに到達できることが知られている(パーセプトロンの収束定理)。以上のような、ネットワークの反応と正解を比べてネットワークの結合状態を望ましい方向に修正していく方式を、教師付き学習という。間違えたら、そばで見張っている教師に正しい方向に向かってぶたれるということである。

問題は正しいパラメータが存在するかどうかである。(4)式の意味を考えてみよう。 $N$ 次元空間( $x, y, z, \dots$ )の中に $J_1 x + J_2 y + J_3 z + \dots - \theta = 0$ なる超平面を定義する。(4)式が成立するなら、任意の入力( $x = \xi_1^{\mu}, y = \xi_2^{\mu}, z = \xi_3^{\mu}, \dots$ )がその超平面の上にあるか( $\eta^{\mu} = 1$ )、下にあるか( $\eta^{\mu} = -1$ )がきちんと決る。つまりそうなるように $\{J_j\}$ と $\theta$ を選べるということである。これを線形分離性と言う。世の中が一枚の平面で白か黒かにすっぱりと分かれてしまうかという、一般にはそうはいかない。線形分離可能な問題はごく限られていて、単純なパーセプトロンでは役に立たないことが多い。例えば二つの(1, 0)の組の排他的論理和XORを取るという簡単な課題でも、図7に示すように線形分離は不可能で、単純なパーセプトロンではどんなに学習を繰り返しても実現できない。何か対策はあるだろうか。入力層と出力の間に中間層をいくつかはさみ、入力信号の分離性を高めれば良い。そこで多層のフィードフォワードネットワークに目を向けよう。

#### 4.2 バックプロパゲーション

図8のような3層のフィードフォワードネットワークを取り上げる。入力層(層内でのニューロンの番号を $m$ で表す)、中間層(ニューロン番号 $k$ )、出力層(ニューロン番号 $i$ )のいずれの間のシナプス結合も学習によって変化させる。学習の目標は、シナプス結合の修正によって入力信号 $\{\xi_m^{\mu}\}$ に出力信号 $\{\eta_i^{\mu}\}$ を正しく対応させることである。学習の達成度の目安になるのは、望ましい出力 $\{\eta_i^{\mu}\}$ と実

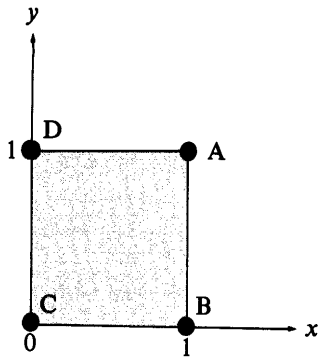


図7 排他的論理和 XOR は線形分離不可能な問題である。A, B, C, D の各点は (1, 1), (1, 0), (0, 0), (0, 1) に対応し、それらに XOR を作用させると A, C では 0 (青丸), B, D では 1 (赤丸) になる。一本の直線  $J_1x + J_2y - \theta = 0$  で青丸と赤丸を分けることは出来ない。

際の実出力  $\{s_i^l\}$  との差から作った 2 乗誤差  $E = \sum_i (s_i^l - \eta_i^l)^2$  であろう。これが小さいほど良い。そこで  $E$  を減らす方向にシナプス結合を変化させることになる。  $E$  は入力層と中間層の結合  $\{J_1 \equiv J_{km}\}$  および中間層と出力層の結合  $\{J_2 \equiv J_{ik}\}$  の陰関数である。これを減少させるには、  $E$  が一定の超曲面 (等高線の高次元版) に垂直な傾き (グラディエント) の方向  $-\partial E / \partial J_1$ ,  $-\partial E / \partial J_2$  に  $J_1$  と  $J_2$  を変えればよい。

このような考え方から、適当に与えた初期値から出発して、まず出力層とそれに隣接する中間層の間の結合を  $\Delta J_2 = -\epsilon \partial E / \partial J_2$  だけ変化させる。次に、中間層と入力層の結合を同様の公式  $\Delta J_1 = -\epsilon \partial E / \partial J_1$  に基づいて調整する。後者の微分を実行するに当たって、  $E$  は出力層でのニューロンの状態  $\{s_i^l\}$  のみにあらわに依存しているため、  $\partial E / \partial J_1$  を求めるには、入力層と中間層の結合  $J_1$  の変化が出力に及ぼす影響が分かってないといけない。これは当然ながら、中間層と出力層のシナプス結合  $J_2$  を経由した影響となる。それゆえ、まず中間層と出力層との結合を調整・決定してから次に入力層と中間層の結合の調整に入るという順序が重要である。言い換えれば、学習中は誤差の影響が出力層から中間層へ、それから入力層へと、学習後の信号伝達とは逆向きの順序で伝わる。これがこの学習法の名前バックプロパゲーション (誤差逆伝播法) の由来である。中間層が必ずしも一つでなくても要領は同じである。学習後の動作時に信号が伝わるのと逆向きに、学習による結合の変化を進行させる。

これですべてがうまくいくだろうか。うまくいくこともあるが保証の限りでない。単純なパーセプトロンの場合のような収束定理がないのである。線形分離性についての制約は取れたものの、今度は収束の保証がなくなった。また実際に収束する場合でも、層の数、ニューロンの数、微小量  $\epsilon$  などのパラメータの選び方は経験に頼る、いわばノウハウの部分が多い。それに、うまくいった例は発表される

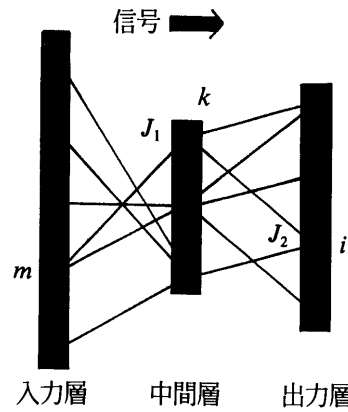


図8 3層のフィードフォワード型ネットワーク。信号は入力層から出力層へ流れるが、結合を修正する学習の最中には、修正情報は出力層から中間層、入力層へという向きに伝わる。

が失敗には口をつぐむのが普通だから、なかなか失敗例から学べない。ここらへんを体系化し、ノウハウではなくて学問にしようとする努力もずっと続けられている。

バックプロパゲーションがうまくいった例を見てみよう。有名なのは、文字で書かれた英語の文章を発音になおす NETalk というネットワークである。英語のつづりと発音の関係は例外に満ちており、明示的な形で規則を網羅してコンピュータに与えるのは容易なことではない。フィードフォワード型のニューラルネットワークでこれを実現するには、つづりとその発音の組を入力側と出力側にそれぞれ多数与え、バックプロパゲーションでシナプス結合を修正していく。規則は一般的明示的な形で与えられるのではなく、多数の例を通して暗示されるだけである。

結果は、学習に使った例については非常に良く発音が出来るようになり、しかも学習してないものに遭遇しても、かなりの確率で正しい発音をするようになった。例を通してのみ暗示された規則を、シナプス結合強度による表現という形で一般化することに成功したのである。これはなんとなく人間の学習に似ているではないか。バックプロパゲーションによる学習が脳で実現されていると考える生理学的な理由はない。しかしながら、多層ネットワークの入出力の間を可塑的な結合で結ぶことで線形分離不可能な課題を学習し、さらに一般化の能力をも持つシステムが出来る可能性を具体的に示した意義は大きい。

### 4.3 自己組織化

次に教師なしの学習の例を簡単に紹介しよう。教師なし学習というのは、出力結果と正しい答の比較をしたりせず、入力信号だけを使って外界の情報の特徴をネットワークの構造に反映させていくプロセスのことである。外部世界的表現を自分自身で作りに上げていくことから、自己組織化とも呼ばれる。

自己組織化にはたくさんの方式があるが、入力層に与えられた入力信号に一番強く反応するニューロンと入力層との結合を強化することによって、各入力信号に特異的に反応するニューロンを育てることが多い。その際、一番強く反応するニューロンとそのすぐまわりのニューロンについて結合を強化し、もう少し離れたものについては弱めるという操作をすれば、似通った入力に反応するニューロンが近い場所に並ぶトポロジカル・マッピングが実現する。

これは、脳の中での視覚や聴覚、触覚の処理の初期段階についての知見と類似している。例えば、網膜上の視細胞に光刺激が与えられたときに反応する大脳皮質上の視覚野のニューロンの配列は、網膜上で隣り合った刺激には視覚野上でも隣り合うニューロンの発火が対応するようになっている。こうした構造はおおまかには遺伝によって決定されているが、生後の環境の影響も重要な調整の役割を果たしている。理論的にもこうした調整が可能であることが、自己組織化のモデルの研究から明らかにされたのである。

## 5. 関連した話題

ニューラルネットワークの理論そのものではないが、関連した興味ある話題をいくつか紹介しよう。

### 5.1 夢と反学習

私たちはなぜ夢を見るのだろうか。心の奥底に潜んでいる願望が、抑圧の弱い睡眠中に表れてくるのだろうか。詳細なメカニズムが何であれ、脳の中で起こっていることから、ニューロンやシナプスといったレベルでの記述が究極的には出来てしかるべきだろう。これについて興味ある仮説がある。よけいな記憶の整理をするために夢があるというのである。

3節で、たくさん記憶しすぎるとスピングラス状態になってしまい、連想記憶として機能しなくなると述べた。(3)式によってパターンを次々と埋め込んでいく(シナプス結合を強める)とこういうことなる。実は、スピングラス状態に完全に移行してしまう前でも、埋め込んだ記憶以外に架空の定常状態がある程度生じている。埋め込んだ記憶の間の相関(類似性)が混乱を引き起こしているのである。この掃除をしておかないと、あとの学習に差し支える。夢はこの掃除に当たるというのである。(3)式で $p$ を増やして次々覚えていくのを学習と呼ぶなら、(3)式の和の中の1項にマイナスの符号を付けてシナプス結合を弱める反学習と呼ぶべきプロセスが夢に相当するというわけだ。昼間経験した事柄やそれらから派生したよけいな考え(架空

の定常状態)をその強さに応じて呼び出して、それに相当する分だけシナプス結合を弱める。そうすれば、特定の記憶や概念がやたらに大きな影響を持つことが避けられて、バランスの取れた心が保てる。斬新で魅力的な仮説だが、生理学的、心理学的な直接の裏付けがあるわけではない。

### 5.2 適応と淘汰

ヒトの能力は遺伝で決めるのだろうか、それとも環境が決定的な役割を担っているのだろうか。もちろん、どちらか一方だけということはありません。遺伝子は、ニューロンの配置やシナプスの形成をすべて制御しているのではなく大まかな作りを指定し、後は環境との相互作用に調整を任せているのである。その調整の方式として、ニューロンやシナプスを余分に作っておいて、環境からの刺激を強く受けたものだけが生き残るという戦略を取っている。ニューロンやシナプスのレベルで淘汰が起きているのである。もし適応と淘汰が、胚の形成後のニューロン・シナプスレベルでは起きずに遺伝子レベルのものだけだったら、数時間から数年のスケールで変化する環境への適切な対応は望めない。遺伝子の突然変異には数世代以上の時間が必要だからである。したがって、環境への適応を遺伝子の突然変異だけに頼っていたのでは、ヒトは地上から消滅してしまうかもしれない。シナプスの淘汰や可塑性を取り入れることによって、遺伝子自身への負担を最小限に押え、しかも速い周期で変化する環境に適応することが可能になっている。

さらに他の動物と違い、ヒトは学習の成果を他の個体に言語を通じて相当正確に伝達出来る。つまり学習するべき経験に直接遭遇することなしに、来るべき事態に備えられる。環境への適応度を向上させる手段を、遺伝子やニューロン、シナプスといった体の中の構造物だけでなく、体の外にも獲得したのである。しかも印刷や通信の発達によって、適応度向上のための情報が時間や空間を超越して共有されるようになった。ヒトの遺伝子の構造は数万年以上の間基本的な変化をしていないが、遺伝子の外での適応に関して、このように大きな変化が起きている。言語や文明の成立と発展は、したがって、遺伝子レベルでの進化に匹敵する重要な意義を持っていると言っても過言ではないだろう。

## 6. おわりに

ニューラルネットワークの研究は既存の分野を越えた交流によって大きな刺激を受けて進んで来た。筆者自身も研究の様々な局面で数理工学、生命科学、認知心理学などの

基礎知識の乏しさを実感する事がたびたびである。それと同時に、いろんな学問分野の、伝統的な物理とは違った面白さに惹かれることも少なくない。この解説を目にされた読者によって分野間の交流がさらに進むことを願っている。

最後に、草稿に有益なコメントを下された多くの方々に感謝する。

参考文献

本稿は初めてこの分野に接する人を対象としているので、全般的な解説

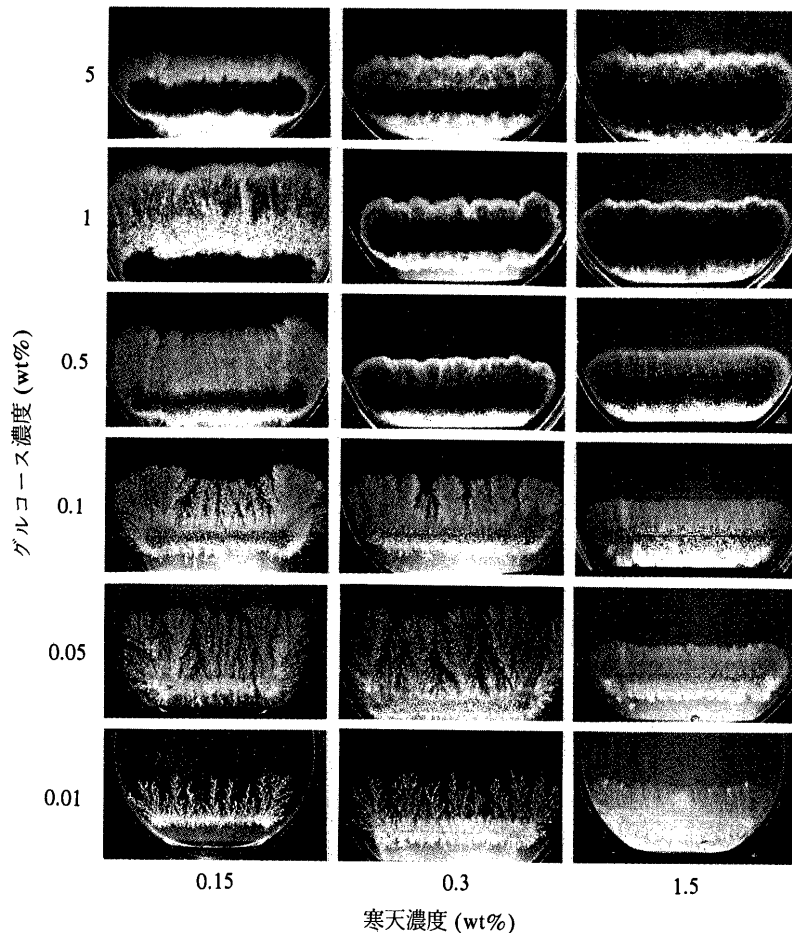
をした単行書に絞って文献を紹介する。原論文はそれぞれの本に引用されているものを参照してほしい。

- 1) P. Peretto: *An Introduction to the Modeling of Neural Networks* (Cambridge Univ. Press, 1992).
- 2) J. Hertz, A. Krogh and R.P. Palmer: *Introduction to the Theory of Neural Computation* (Addison-Wesley, 1991).
- 3) 中野 馨編: ニューロコンピュータの基礎 (コロナ社, 1990).
- 4) 甘利俊一: 神経回路網の数理 (産業図書, 1978).
- 5) M. A. アービブ著, 金子隆芳訳: ニューラルネットと脳理論 (サイエンス社, 1992).

# フォトサロン

## コウジカビ (*Aspergillus oryzae*) のコロニー (菌糸集落) の形態

種々のグルコース濃度 (即ち栄養分濃度) および寒天濃度 (即ち培地の硬度) の合成寒天培地 (Czapak agar medium) に, *A. oryzae* の胞子を直線上に接種して 24°C で培養した。各写真は、培地を入れた直径約 9 cm のシャーレの底の側から菌糸体に照明をあて、真上から撮影したものである。写真の下側の水平な線が長さ約 5.5 cm の接種線である。*A. oryzae* の菌糸体は複雑に分岐した細長い菌糸からなる。コロニーはこの菌糸が密集して形成されている。各コロニーの培養日数は以下の通り。左上から左下に向かって (寒天濃度 0.15%: 半液状培地) 順に, 10, 20, 20, 30, 30 日。中央上から下に向かい (寒天濃度 0.3%: 軟培地) 順に, 10, 10, 10, 15, 30, 30 日。右上から下へ (寒天濃度 1.5%: 標準硬度培地), 8, 8, 8, 8, 8, 10 日。栄養分濃度が低くなるにつれ菌糸密度は低くコロニーの拡大速度も低くなる。培地の硬度が低くなるとコロニー成長は緩慢になり、さらに栄養分濃度が低くなると、コロニーは樹枝状になるのが見られる。コロニーの中に黒く写っているのは成熟した分生子 (孢子) が豊富に形成されている部分である。*A. oryzae* の分生子は成熟するにつれて黄色から緑、茶色と色に変化していく。菌糸は空気中に立ち上がって伸びるものや、基質の中に潜行するものなどがあり、3次元的に伸長成長するが、コロニー全体としては培地の表面に沿って平面的に拡大していく。コロニーの底部や周縁部の培地からはグルコースが菌糸体に吸収され、標準硬度培地での 6 日間以内の培養実験でグルコースの濃度勾配ができることが知られている。



松浦 執 (東海大開発工), 宮島佐介 (中部大工)