

脳と知能の物理学

寺前順之介

京都大学情報学研究科 先端数理科学専攻

teramae@acs.i.kyoto-u.ac.jp

概要

我々の脳は約 1,000 億もの神経細胞からなる巨大なネットワークである。認知、予測、推定、判断、思考、運動制御、記憶、学習など、我々が持つ極めて高度な情報処理能力は、このネットワーク上で実現する集団現象にその実体を持つ。近年、神経系に対する実験的及び理論的技術が急速に進展したことにより、これまで長く未解明に残されてきた脳の動作原理に関して、多くの性質が具体的に明らかになり始めた。この講義では、特に興味深く、また極めて重要と思える脳の「自発揺らぎ」と呼ばれる現象を軸に、脳科学の最新の研究成果を紹介する。まず自発揺らぎが神経系において実現されるメカニズムを数理的に解説し、さらに自発揺らぎを利用することによって、現在機械学習などで利用されるものとは大きく異なる「脳の学習アルゴリズム」が得られることを示す。この学習アルゴリズムは、我々の脳（具体的には大脳皮質）に関する多く実験結果を統合的に説明することで、脳の情報処理とは一体何なのかという疑問に一つの答えを与えているように見える。それらを学習することを通じて、生命の「知能」の実体について議論していこう。

1 はじめに

1.1 脳とニューラルネットワーク

我々の脳は約 1000 億もの神経細胞（ニューロン）が、シナプスと呼ばれる部位で結合しあって構成される極めて大規模なネットワークである（図 1）。各神経細胞はこのネットワーク上で、「スパイク発火」と呼ばれる短時間のインパルス的な電気的信号をやり取りしており、そのスパイク発火の総体が、認知、予測、推定、判断、思考、運動制御などの我々が持つ高度な知能を実現する実体であると考えられている。

スパイク発火は、単に神経細胞間の情報伝達に使用されるだけでなく、脳のネットワーク構造自体にも影響を与えている。神経細胞間でのシナプス結合の有無や強度は、そ

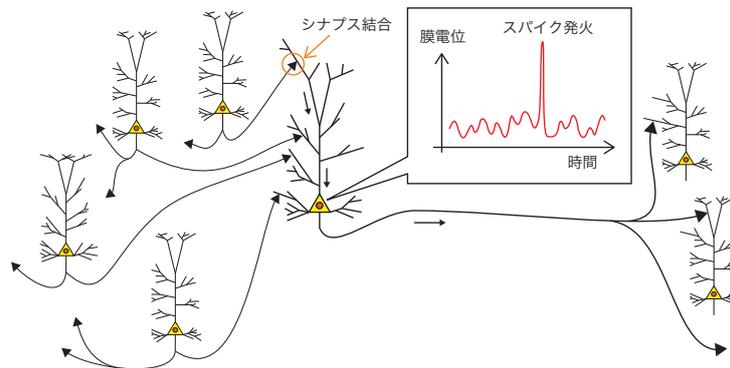


図1 我々の脳は、多数の神経細胞がシナプスと呼ばれる部位を介して結合した大規模ネットワークである。各神経細胞の膜電位がスパイク発火と呼ばれる一過的電位上昇を示すことで情報を伝達すると同時に、スパイク発火によってネットワーク構造が変化することで記憶や学習が実現する。

のシナプスを介して伝播するスパイク発火に応じて徐々に変化する「シナプス可塑性」と呼ばれる性質を有している。このシナプス可塑性によって、ネットワーク自身の構造が、環境や経験を反映して変化していくことが、我々の記憶形成や学習能力の実体であると考えられている。

ニューロンやシナプスが示すこれらの性質を大胆に簡略化してモデル化したものが、いわゆるニューラルネットワークと呼ばれるネットワークである。ニューラルネットワークには様々なバリエーションが存在するが、いずれも神経細胞に当たる多数のユニットを素子として、それらをシナプスに当たる重みを持つ結合によって相互に結合することで、パターン認識や時系列予測など、既存のコンピュータには実行が難しい問題の解決を目指すものである。近年著しい性能を達成して高い注目を集めているディープラーニング（深層学習）も、このようなニューラルネットワークの一つである。

ところが、最新のディープラーニングを含む人工ニューラルネットワークが基盤としている脳の知見は、必ずしも我々人類が手に入れている脳の最新の知見ではない。例えばディープラーニングで頻繁に用いられる畳み込みニューラルネットワークというネットワーク構造は、1960年頃に実験的に報告された「受容野」という脳の視覚野の神経細胞が持つ性質 [1] に基づいて、1980年代に最初に提案され [2]、その後2010年代に著しい性能向上を見せて一気に普及したものである [3]。つまり畳み込みニューラルネットワークが基盤にしている脳の知見は、実は60年以上前に得られた結果なのである。一方で、畳み込みニューラルネットワークが一気に普及する直前の2000年代ごろから、神経系の実験技術が急速に著しい進展を見せ始め、それまで実験的には測定が困難だった脳内での実

際の神経ネットワークの構造や、多数の神経細胞のスパイク発火のリアルタイムでの動態などが明らかになってきた。それら近年の実験技術の急発展によって得られた脳の実験結果のほとんどは、少なこともまだ、最新の畳み込みニューラルネットワークにも取り入れられていないのである。

特に興味深いことは、これら実験技術の急速な向上によって明らかにされた脳の知見は、多くの点で最新のディープラーニングを含む人工ニューラルネットワークの動作様式とは整合していないことである。脳で発見され始めた最新の知見のうち、どれが神経情報処理に本質的に重要でどれが副次的な結果なのかは、まだ研究途上であり明らかになっていない点も多い。しかしそれでも、最新の実験結果の多くは、ディープラーニングを含む現在の人工ニューラルネットワークの動作原理とは大きく異なっており、ディープラーニングなどとは異なる何らかの全く新しいニューラルネットワークの動作原理の存在を示唆しているように見えるのだ。

著しい成果を更新し続けている最新のディープラーニングであっても解決すべき問題は山積みである。学習に必要なデータサイズや計算量の多さ、新規環境への適応性、言語や動画など時間構造が本質的な課題の学習、記憶の形成や適切な利用、目的の設定や自発性などは特に重要な課題だろう。性能が単一の指標で簡単に測れるような画像処理やゲームなど、ルールと目的が外部から明示的に与えられる一部の問題以外では、我々人間を含む動物の脳の能力にはまだまだ全く及んでいないのが現状である。

おそらく、現在の人工ニューラルネットワークが持つこの限界や問題点を超えるには、生物が持つ実際の脳の知見を理解し、応用していくことが必要なはずだ。現在のディープラーニングの著しい発展が、視覚野の神経細胞ネットワークの性質を取り入れることで生まれたように、最新の人工ニューラルネットワークにも取り入れられていない実際の脳の特性を知り、原理を解明して利用できれば、現在の人工ニューラルネットワークを超える本当に脳型な新たな人工知能を作り出すことも可能になると期待できる。それは既存の人工知能が抱える多くの問題を解決すると期待されるとともに、脳の動作原理の解明によって、生命知能と人工知能の枠を超え、「知能とは何か」という大きな疑問の解明も可能にしていくのだろう。

1.2 脳の自発揺らぎ

この講義では、実際の脳が示す特性のうち特に重要だと考えられる「持続的自発発火活動（自発揺らぎ）」と呼ばれる脳（具体的には大脳皮質）の性質に着目する [4, 5, 6, 7]。

大脳皮質は我々の脳の表面をシート状に覆う脳の部位であり、「高次機能」と呼ばれる、

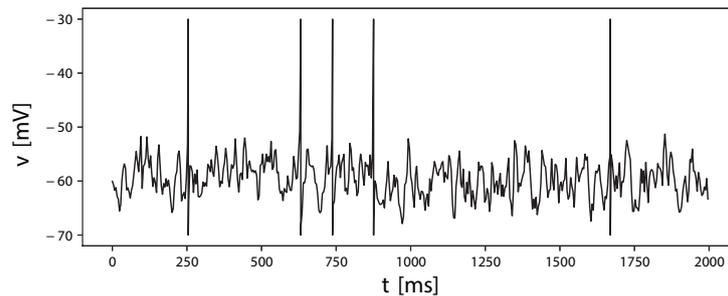


図 2 脳内での持続的自発発火活動の概念図。膜電位が大きく変動し、1 ヘルツ程度のランダムなタイミングでのスパイク発火が持続する。

認知、予測、推定、記憶、計画立案など特に高度な機能の実現を担っていると考えられている。ところが大脳皮質の神経細胞が示すスパイク発火を大規模に測定すると、それらが互いに独立に、しかもほぼランダムなタイミングで発火し続けていることが分かる。この自発的に持続するランダムなスパイク発火が「持続的自発発火活動」であり、そのランダム性から「脳の自発揺らぎ」とも呼ばれる神経活動である (図 2)。

自発揺らぎで生成されるスパイク発火の頻度 (発火率) は、各細胞あたり 1 秒間に 1 回程度、つまり 1 ヘルツ程度である。神経細胞の持つ典型的な時間スケールがミリ秒程度であるため、1 ヘルツという発火率を持つ自発揺らぎは神経細胞の活動としては非常に遅い現象であることがわかる。

不思議なのは、この低頻度スパイク発火活動が、私達ヒトや動物が外部から明示的な感覚刺激を受けず、また運動などを生成していないときでも自発的に持続している点である。たとえ低頻度でも、自発揺らぎは脳内で常時持続しているため、全体としては膨大なエネルギーを消費しており、脳で消費するエネルギーの実に 8 割もが、このランダムなスパイク発火に利用されていると言われている [8]。ヒトの消費エネルギーの 4 分の 1 が脳で消費されると言われるため、私達が摂取する総カロリーの 5 分の 1 が、この単にランダムなノイズに見える自発揺らぎに利用されている割合だ。

脳も生物であり神経細胞も小さな細胞なので、例えば熱揺らぎなどの影響で、多少のランダムな揺らぎが存在するのは当然と考えるかもしれない。しかしそれは違う。実はネットワークから単離した単一の神経細胞は極めて信頼性の高い素子として動作し [9]、実際の大脳皮質で観測されるような揺らぎや不規則性は示さないことがわかっている。つまり脳内で観測される自発揺らぎは不可避な現象ではなく、むしろ神経ネットワークの何らかの効果によって出現していると考えられるのだ。

さらに実は、神経ネットワークが自発揺らぎを生成し維持することが、極めて難しいこ

ともわかっている [10]。簡単な仮定に基づいて計算すると、神経ネットワークが持つ多数の正のフィードバック構造によって、低頻度の自発揺らぎが存在すると、スパイク発火の過度な連鎖が発生し、ネットワーク全体が同期的スパイク発火を繰り返す発作的な活動状態に遷移することが分かっている。

つまり私達の脳は、生体の全エネルギーの約 20 % ものエネルギーを消費して、発作的な状態に陥る極めて危険性の高いランダムな低頻度活動を維持しているといえる。なぜ？そしてどのように、私達の脳は、このランダムなスパイク発火活動を維持しているのだろうか？脳のネットワークが低頻度のランダム発火を安定して維持するメカニズムは一体何で、それにはどんな機能的意義があるのだろうか？

この講義を通じて、これら 2 つの問の両方に答えをだしたいと思う。まず第 2 章で自発揺らぎが安定して維持されるメカニズムを解明し [11]、さらに第 3 章で、自発揺らぎこそが脳の情報処理と学習の本質であると考えられることを示そう。

2 自発揺らぎと脳のネットワーク構造

2.1 神経ネットワークの微分方程式

単一神経細胞の電気的状態は、神経細胞の細胞膜内外での電位差（膜電位） v によって記述される。膜電位は、通常、細胞外を基準にして細胞内がマイナス 70 ミリボルト程度に維持されている。この値は、他の神経細胞からの入力によって上下し、それがマイナス 50 ミリボルト程度の発火閾値と呼ばれる値まで上昇すると、活動電位、あるいはスパイク発火と呼ばれる膜電位の急激な一過的上昇が発生する。このスパイク発火が神経細胞にとっての情報表現である（図 1, 2）。

閾値以下であれば、膜電位は、微分方程式

$$C \frac{dv}{dt} = -g_L (v - V_L) + I \quad (1)$$

に従う。ここで C は細胞膜の電気容量、右辺第一項 $g_L (v - V_L)$ は、膜に存在するイオンチャンネルを通じて定常的に流れるイオン電流（漏れ電流、リーク電流）を表しており、 g_L がリークコンダクタンス、 V_L がリーク電流の反転電位と呼ばれる。 I は外部からの入力電流を表す。式 (1) は線形の微分方程式であるため、この方程式だけでは、膜電位 v が、時定数 C/g_L で $V_L + I/g_L$ に緩和するだけで、スパイク発火は表現されない。そこで、膜電位が発火閾値 V_{thr} に達することをスパイク発火とみなし、発火後の膜電位を強

制的にリセットするメカニズム

$$v(t) = V_{\text{thr}} \rightarrow v(t + dt) = V_{\text{reset}} \quad (2)$$

を導入する。式 (1) と (2) をあわせたものを神経細胞の「積分発火モデル」と呼ぶ。この簡単なモデルで神経細胞の多くの電氣的性質がうまく説明できることが知られている。

入力電流 I が定数であれば、十分大きい I で、積分発火モデルは周期的なスパイク発火を生成するが、実際の脳内では、神経細胞が受ける入力は定数ではなく、シナプス結合を介した他の神経細胞のスパイク発火によって決定されるはずである。そこでこの効果を取り入れることで、神経ネットワーク内での神経細胞の膜電位が従う方程式を導出しよう。

神経細胞には、自身がスパイク発火を起こすことによって、他の神経細胞の膜電位を増加させる興奮性神経細胞 (Excitatory neuron) と、抑制させる抑制性神経細胞 (inhibitory neuron) の二種類が存在する。このことに注意して、 $i = 1, \dots, N_E$ を興奮性神経細胞、 $i = N_E + 1, \dots, N_E + N_I$ を抑制性神経細胞とする $N_E + N_I$ 個の神経細胞からなるネットワークを考えよう。詳しくは講義で説明するが、興奮性神経細胞と抑制性神経細胞は、それぞれ受け手側の神経細胞の興奮性コンダクタンスと抑制性コンダクタンスを増加させることで、間接的に受け手側神経細胞の膜電位を上下させる。 i 番目の神経細胞の膜電位を v_i 、興奮性コンダクタンスを g_{Ei} 、抑制性コンダクタンスを g_{Ii} とすれば、これらの変数が従う連立微分方程式

$$C \frac{dv_i}{dt} = -g_L (v_i - V_L) - g_{Ei} (v_i - V_E) - g_{Ii} (v_i - V_I) \quad (3)$$

$$\frac{dg_{Ei}}{dt} = -\frac{1}{\tau_g} g_{Ei} + \sum_{j=1}^{N_E} G_{ij} \sum_s \delta(t - t_{js}^{\text{spike}} - d_{ij}) \quad (4)$$

$$\frac{dg_{Ii}}{dt} = -\frac{1}{\tau_g} g_{Ii} + \sum_{j=N_E+1}^{N_E+N_I} G_{ij} \sum_s \delta(t - t_{js}^{\text{spike}} - d_{ij}) \quad (5)$$

$$v_i(t_{is}^{\text{spike}}) = V_{\text{thr}} \rightarrow v_i(t_{is}^{\text{spike}} + dt) = V_{\text{reset}} \quad (6)$$

が得られる。ここで V_E 、 V_I は、それぞれ興奮性チャンネルと抑制性チャンネルの反転電位であり、 τ_g は興奮性、抑制性コンダクタンスの緩和時定数である。 t_{js}^{spike} は j 番目の神経細胞の s 番目のスパイク発火時刻、 $\delta(t)$ はディラックのデルタ関数であり、式 (4)、(5) は、それぞれ、 j 番目の神経細胞が興奮性か抑制性かに応じて、その神経細胞のスパイク発火時刻から d_{ij} 遅れた時間に、 i 番目の神経細胞の興奮性あるいは抑制性コンダクタンスが G_{ij} 増加することを意味している。つまり、 G_{ij} が j 番目の神経細胞から i 番目の神経細胞へのシナプス結合強度にあたり、 d_{ij} がそこでのシナプス遅延時間を表す。

2.2 平均場近似

ネットワーク全体のシナプス結合 G_{ij} やシナプス遅延 d_{ij} が与えられれば、原理的には、連立微分方程式 (3) ~ (6) を解くことで、神経ネットワーク上でのスパイク発火ダイナミクスの予言が可能になる。では、大脳皮質の最も基本的な活動である自発揺らぎも、この連立微分方程式で説明できるだろうか。

特に重要なのは自発揺らぎの約 1 ヘルツ程度という低い発火率である。この遅い時間スケールが、ネットワークの効果として出現し、自発的な活動として維持されるかどうかを検討しよう。具体的には、自発ゆらぎのような不規則でかつ定常的なスパイク発火がネットワークの全体的な活動として持続していると仮定してその発火率を求めてみよう。

そのために、多体問題である式 (3) ~ (6) を一体問題で近似する平均場近似を利用する [10]。ネットワーク内の興奮性神経細胞と抑制性神経細胞の平均発火率を、それぞれ r_E 、 r_I とし、例えば、ネットワーク内の興奮性神経細胞間の平均結合強度を G_{EE} 、抑制性神経細胞から興奮性神経細胞への平均結合強度を G_{EI} などと記述する。また興奮性神経細胞に入力する抑制性神経細胞数を K_{EI} などと記述する。各神経細胞のスパイク発火が独立で不規則であることから、各神経細胞のスパイク発火時刻は独立なポアソン過程として近似できる。独立なポアソン過程の重ね合わせもまたポアソン過程に従うため、平均場近似の結果、ネットワーク内のある一つの神経細胞の膜電位が従う方程式として

$$C \frac{dv}{dt} = -g_L(v - V_L) - g_E(v - V_E) - g_I(v - V_I) \quad (7)$$

$$\frac{dg_E}{dt} = -\frac{1}{\tau_g} g_E + G_{XE} \sum_s \delta(t - t_s^{XE}) \quad (8)$$

$$\frac{dg_I}{dt} = -\frac{1}{\tau_g} g_I + G_{XI} \sum_s \delta(t - t_s^{XI}) \quad (9)$$

$$(10)$$

を得る。ここで X は着目する神経細胞が興奮性か抑制性かに応じて $X = E, I$ とし、 t_s^{XE}, t_s^{XI} はそれぞれ到着率 $r_E K_{XE}$ と $r_I K_{XI}$ のポアソン過程である。

各神経細胞への入力神経細胞数 K_{XY} が十分大きく、かつ平均結合強度 G_{XY} が十分小さければ、式 (8) (9) の第 2 項は、平均 $r_Y G_{XY} K_{XY}$ 、分散 $\sigma_{XY}^2 = r_Y G_{XY}^2 K_{XY}$ の白色ガウスノイズで近似できるため、 g_Y ($Y = E, I$) の平均は $g_{Y0} = \tau_g r_Y G_{XY} K_{XY}$ とな

り、 $g_Y = g_{Y0} + \eta_Y$ とすれば、確率微分方程式

$$C \frac{dv}{dt} = -g_L(v - V_L) - (g_{E0} + \eta_E)(v - V_E) - (g_{I0} + \eta_I)(v - V_I) \quad (11)$$

$$\frac{d\eta_E}{dt} = -\frac{\eta_E}{\tau_g} + \sigma_{XE}\xi_E \quad (12)$$

$$\frac{d\eta_I}{dt} = -\frac{\eta_I}{\tau_g} + \sigma_{XI}\xi_I \quad (13)$$

$$(14)$$

を得る。ここで ξ_E と ξ_I は、互いに独立な平均 0 分散 1 の標準白色ガウスノイズである。

計算をさらに進めるために、第 1 式に含まれる非線形項 $\eta_X(v - V_X)$ の v を v の平均値で置き換える実効時定数近似 (effective-time-constant approximation) と呼ばれる近似を利用しよう [12]。 v の平均値は、上式でノイズ強度 σ_{XY} を 0 とすることで $V_0 = (g_L V_L + g_{E0} V_E + g_{I0} V_I) / (g_L + g_{E0} + g_{I0})$ と得られるため、 $\eta_E(V_0 - V_E) + \eta_I(V_0 - V_I)$ を新たに η とすることで、2 変数の確率微分方程式

$$\frac{dv}{dt} = -\frac{v - V_0}{\tau_e} + \eta \quad (15)$$

$$\frac{d\eta}{dt} = -\frac{\eta}{\tau_g} + \sigma\xi \quad (16)$$

を得る。ここで、 $\tau_e = C / (g_L + g_{E0} + g_{I0})$ であり、 $\sigma^2 = \sigma_{XE}^2 (V_0 - V_E)^2 + \sigma_{XI}^2 (V_0 - V_I)^2$ とした。式 (15) (16) では明示的に記載していないが、 V_0, τ_e, σ の各パラメータの値は、着目している神経細胞が興奮性か抑制性か、つまり X の値、により異なる値を取る。

神経細胞の発火率は、膜電位 v が発火閾値 V_{thr} を単位時間当たりには通過する割合、つまり v_X 方向の定常確率流にあたるため、スパイク発火を表す膜電位の閾値とリセットに注意すれば、確率微分方程式 (15) (16) に従う膜電位の定常確率密度分布と発火率が、近似的にはあるが

$$P_0(v) = C \exp\left(-\frac{(v - V_0)^2}{\tau_g^2 \tau_e \sigma^2}\right) \begin{cases} \int_{s_1}^{s_2} \exp(s^2) ds, & v < V_{reset} \\ \int_{s(v)}^{s_2} \exp(s^2) ds, & V_{reset} < v < V_{thr} \end{cases} \quad (17)$$

と

$$r_X^{out} = J_X(V_{thr}) = \left(\tau_{ref} + \tau_e \sqrt{\pi} \int_{s_1}^{s_2} \exp(s^2) (1 + \operatorname{erf}(s)) ds \right)^{-1} \quad (18)$$

の形で与えられる [13]。ただし、神経細胞はスパイク発火直後には続けて発火できない不応期 (refractory period) という性質を持つため、その大きさを τ_{ref} として取り入れた。

平均場近似の仮定によれば、この発火率は、入力発火率である r_E あるいは r_I に一致するはずである。よって定常状態の発火率に関する自己無撞着方程式

$$r_X^{\text{out}}(r_E, r_I) = r_X \quad (19)$$

を得る。この解を数値的に求めることで、自発揺らぎのように非同期かつ不規則にスパイク発火活動が持続する活動の発火率を求めることができる。

ところが、生物学的に知られた値を代入して、この自己無撞着方程式の解を求めても、1Hz 程度という自発ゆらぎの発火率は全く再現できない。解として現れるのは、発火率 0 というスパイク発火が持続しない状態か、あるいは持続するのであれば、数十から数百 Hz という実際の脳の状態とは桁違いに大きな発火率の場合だけである。つまり我々の脳の 8 割ものエネルギーを消費する最も基礎的なスパイク発火状態は説明できないのだ。

2.3 神経ネットワークの不均一性と対数正規分布

この謎を解く鍵は、結合行列 g_{ij} の分布にある。上述したように、結合行列の値は平均的には極めて小さいため、各神経細胞の入力はガウス近似可能だと考えられてきた。ところが、たとえ平均は小さくとも、全ての結合が小さいわけではない。実はシナプス結合には、極めてまれに、平均的結合強度の 10 倍以上もの効果を与える非常に強い結合が存在するのである [14, 15]。

神経細胞間の結合強度は、具体的にはシナプスの前側の神経細胞 (前シナプス神経細胞) がスパイク発火を起こしたときに、シナプスの後側の神経細胞 (高シナプス神経細胞) の膜電位がどれだけ上昇するかで表されるため電位の単位で測定され、その値は興奮性後シナプス電位 (EPSP, Excitatory PostSynaptic Potential) と呼ばれる。Song らは、数千の興奮性神経細胞のペアに対して、EPSP の値を網羅的に測定することに成功し、確かにほとんどの EPSP は 1 ミリボルト以下と極めて小さいこと、しかし EPSP 全体の分布は、10mV 程度まで右に長い裾をひいたいわゆるロングテール分布になっており、具体的には対数正規分布 (図 3)

$$P(x) = \frac{1}{\sqrt{2\pi\sigma x}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) \quad (20)$$

で良く近似できることを発見した [14]。

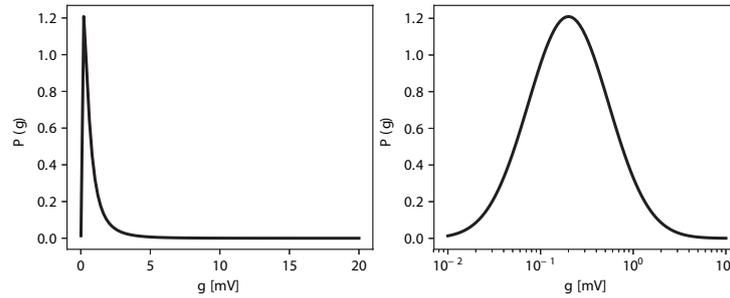


図 3 対数正規分布。左図は実軸、右図は横軸を対数軸で表示したもの。大多数の EPSP は 1 ミリボルト以下だが、稀に 10 ミリボルト程度に達する非常に大きな EPSP も存在する。

EPSP が裾の重い対数正規分布に従うという事実は、脳のネットワークが大多数の弱い EPSP と、少数だが無視できない極めて強い EPSP が共存する不均一性の強いネットワークであることを意味する。この著しい不均一性の存在によって、低頻度の自発揺らぎが脳内で安定して実現することを見ていこう。そのために、ネットワークが弱い EPSP だけから構成されると仮定して導出した発火率の式 (18) に、少数の極めて強い EPSP の効果を取り込むことを考える [11]。

小さい EPSP しか存在しない場合では、入力によってスパイク発火が誘発されるためには、受け手側の神経細胞の膜電位が、発火閾値近傍に存在している必要がある。これが、式 (18) が発火閾値近傍の確率密度分布を用いて記述される理由である。ところが大きい EPSP であれば、強い入力によって、膜電位は不連続に増加し、その値が発火閾値を超えれば良いために、膜電位が発火閾値から離れた神経細胞でも発火させることが可能になる。具体的には EPSP の大きさ g に対して、膜電位 v が $V_{\text{thr}} < v + g$ を満たしていれば良い。よって膜電位の定常確率密度分布 $P_0(v)$ を用いて、神経細胞が強度 g の強い EPSP を受けて発火する確率が

$$P_g = \int_{V_{\text{thr}} - g}^{V_{\text{thr}}} P_0(v) dv \quad (21)$$

と求められる。

このことから、弱い EPSP と少数の強い EPSP の両方を受ける神経細胞の出力発火率は、興奮性入力の発火率 r_E を用いて

$$r_E^{\text{out}} = J(V_{\text{thr}}) + r_E \sum_{g \in S} P_g \quad (22)$$

のように記述できる。ここで S は、ガウス分布で近似できない少数の極めて強い EPSP

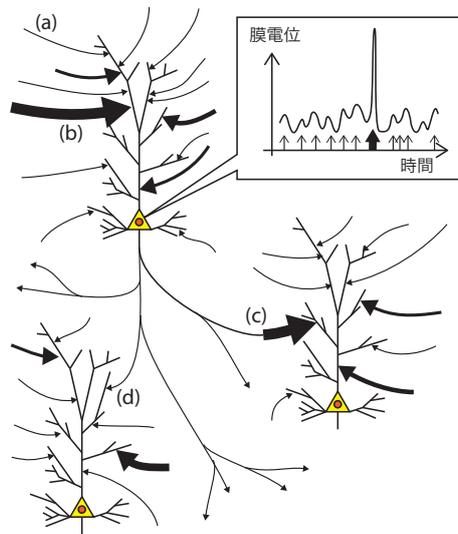


図4 自発揺らぎは多数の弱い EPSP と少数の極めて強い EPSP の共存により維持される。(a) 多数の弱い EPSP は膜電位を揺らぎを伴って上昇させ、(b) 強い EPSP によるスパイク発火を可能にする。(c) スパイク発火は再び他の神経細胞への強い入力を引き起こすと同時に、(d) 多数の弱い EPSP を介してネットワークのノイズ源にもなる。

の集合である。よって発火率の自己無撞着方程式は

$$J_E(r_E, r_I) + r_E \sum_{g \in S} P_g = r_E \quad (23)$$

$$J_I(r_E, r_I) = r_I \quad (24)$$

に変更される。

この自己無撞着方程式 (23) (24) は、式 (19) と異なり、生物学的に妥当な広いモデルパラメータの範囲で、1-2Hz 程度の実際の自発揺らぎに良く一致する発火率での解を持つことがわかる。しかも、現象論的な安定性解析を用いることで、その解が実際に安定解になっていることもわかる。

講義で詳しく説明するが、自発揺らぎの維持には大多数の弱い EPSP と少数の強い EPSP の両方が必要なことも分かる (図4)。少数の極めて強いシナプスは発火を高い信頼性で伝播させることで、自発的なスパイク発火を維持する働きを持つが、この強い EPSP だけではスパイク発火が同期してしまい、ネットワーク内の多くの神経細胞が同期的に発火する発作的状態に陥ってしまう。この発作的活動を抑制するためには、スパイク伝播が引き起こす同期を崩す機能が必要である。この機能を担うのが多数の弱い EPSP である。ある神経細胞がスパイク発火を起こすと、その神経細胞は強い EPSP を少数の神経細胞に送ると同時に、多数の弱い EPSP を多くの神経細胞に送る。受け手側の神経細胞

の立場から見ると、各神経細胞は少数の強い EPSP を受け取りながら、同時に多数の弱い EPSP を不規則なタイミングで継続的に受けることになる。このため多数の弱い EPSP は、各神経細胞の膜電位に対する内的なノイズの役割を果たし、その効果によって複数の神経細胞間での同期的なスパイク発火が抑制される。これによってスパイク発火の同期を抑制しながら、しかも持続的にスパイク発火を伝播させ続けることが可能になるのである。我々の脳は、この両極端な EPSP 分布によって発作的活動を抑えながら安定した揺らぎを維持しているのである。

3 自発揺らぎの機能

3.1 揺らぎが可能にする脳型情報処理と学習

前章では、ネットワーク自身が発作的状態に陥る危険をあえて犯してまでランダムで無規則なスパイク発火を維持しているように見えることを説明した。では、そもそもなぜ脳はこのようなランダムな活動を維持しているのだろうか？これまで、ランダムな神経活動が与える利点について幾つかの仮説が提案されてきたが、このランダムな活動が、脳の全エネルギーの8割を消費してまで維持される理由としては、どれも説得力が足りないのが現状であり、脳が積極的にランダム性を維持している理由は長く未解明に残されていた。

この章では、このランダムな活動こそが脳の情報処理の本質であり、脳の学習の本質であることを極めて強く示唆する最新の研究成果を紹介しよう。そのために重要になるのが、スパイク発火時系列に見られるランダムネスに加えて脳で普遍的に見られるもう一つのランダムネス、シナプス活動のランダムネスである。

第1章で簡単に述べたように、シナプス結合強度は、そのシナプスを介して伝播するスパイク発火によって適応的に変化することが知られている。ところが詳細に観測すると、この時間発展も強い揺らぎを伴って、非単調で確率的に変化することが実験的に観測されているのだ [16, 17, 18, 19, 20, 21, 22, 23]。

ニューロンのランダムなスパイク発火の時間スケールがミリ秒から長くても秒程度なのに対して、シナプスの確率的変化は一般には秒以上から分、あるいは数時間や数日といった長いスケールで起きる現象である。シナプス結合は我々の記憶や学習の実体だと考えられているため、シナプス強度が確率的に変化し続けているという事実は、スパイク発火のランダム性と同様に、かなり不思議な結果に思える。

その謎の解明を目指して、ニューロンが生成するスパイク時系列のランダム性や、シナプス結合強度の時間発展が見せるランダム性の機能的意義を提案する研究も幾つか行われ

てきた。しかし、これら2つのランダム性は独立に研究されてきており、両者がどのような関係にあるのか？なぜ2つのランダム性が必要なのか？2つのランダム性によってどのような脳の情報処理が可能になるのかなどは未解明に残されていた。

筆者はごく最近、これら2つのランダム性を統合することで、既存の人工ニューラルネットワークとは異なる原理で動作するニューラルネットワークのあらたなアルゴリズムが導出可能なことを発見した [24]。しかも、この新たなアルゴリズムは、脳の知見と整合しないディープラーニングなどの既存のニューラルネットワークとは異なり、実際の脳で実験的に観測される多くの知見と、きちんと整合しているのだ。

3.2 神経細胞とシナプスの二重の確率性

以下では、前章とは異なり、神経細胞を0か1の2値を取る確率変数として定式化しよう。0はその神経細胞が現在発火していないことを表し、1はスパイク発火が発生していることを表すとする。一般のニューラルネットワークと同様に、各神経細胞は、他の神経細胞からの入力（つまりスパイク発火）を、結合強度で重み付けして受け取るとし、スパイク発火はその入力強度（神経細胞の膜電位に当たる）に応じた確率でスパイクを発生するとする。よって、 i 番目の神経細胞から j 番目の神経細胞への結合強度を w_{ij} とすれば（前章とは添字の順序が逆になっているので注意して欲しい）、ネットワーク中の j 番目の神経細胞が発火する確率は

$$P(x_j = 1) = \sigma(v_j) = \sigma\left(\sum_{i=1}^N x_i w_{ij}\right) \quad (25)$$

で与えられる。ここで v_j は j 番目の神経細胞の膜電位であり、 N はネットワーク内の全神経細胞数である。ネットワークの構造は結合行列 w_{ij} で決まるが、フィードフォワード型や対称結合しか扱えない既存のニューラルネットワークとは異なり、ここではリカレント結合を含む任意の結合行列 w_{ij} を考えて良い。 $\sigma(x)$ は神経細胞の発火確率を決める活性化関数であり、以下ではシグモイド関数 $\sigma(x) = (1 + \exp(-x))^{-1}$ を用いる。

神経細胞間の結合を詳細に調べると、神経細胞間の結合は、一般には単一ではなく、複数のシナプスによって冗長的に表現されていることが知られている [25, 26, 27, 28, 29] (図5)。そこでこの特性をとりいれて、 w_{ij} で表される i 番目の神経細胞から j 番目の神経細胞への結合

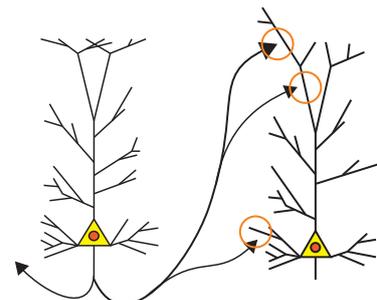


図5 神経細胞間の結合は複数のシナプス結合によって冗長的に表現されている。

も、複数のシナプスによって

$$w_{ij} = \sum_{k=1}^K a_{ijk} s_{ijk} \quad (26)$$

のように K 個のシナプスの重ね合わせとして表現されると仮定しよう。ここで s_{ijk} は、 i 番目の神経細胞から j 番目の神経細胞への k 番目のシナプス結合の状態を表すとし、ニューロンと同様に、0 か 1 の値を取る確率変数とする。 $s_{ijk} = 1$ はそのシナプス結合が実際に結合していることを示し、 $s_{ijk} = 0$ は結合していないことを示す。 a_{ijk} は該当するシナプスが結合したときに w_{ij} に寄与する大きさ（振幅）を表す定数とする。実際には、 a_{ijk} は適当な小さな値に固定しておけば良い。

人工ニューラルネットワークでは、学習によって結合重み w_{ij} を適切に調整することが必要になるが、今回の我々の設定では、その調整は各 s_{ijk} が 1 あるいは 0 になる確率を適切に調整することで実現されると考える。 a_{ijk} の値は、事前に設定した値に固定されており、学習の過程では変化しないことに注意して欲しい。

3.3 事後分布からのサンプリング

神経細胞とシナプスの両方が、多大なコストを払ってまで確率的な挙動を継続しているという脳の著しい特性に基づいて、その確率的挙動こそが脳の学習や情報処理の実体であると考えてみよう。中心となるアイデアは、神経細胞とシナプスの活動を、それらの状態が外界からの入出力に整合するように自身の状態を確率的にサンプリングすること、すなわち外界の状態が与えられた条件下での事後分布からのサンプリングであると考えてことだ。

外界からの入力、ネットワーク内の一部のニューロンの状態を与えられたデータの値に固定することで与えられるとする。ここで言う入力には、例えば教師あり学習の教師データのような、ネットワークが出力すべき目標出力も含まれる。これら外界からの入力を受けるニューロン群を可視ニューロン (Visible neurons) とし、 V で表し、外界から直接入力を受けないニューロンを隠れニューロン

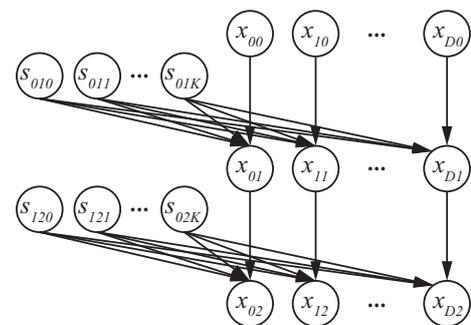


図6 神経細胞の発火状態 x_{di} とシナプス結合状態 s_{ijk} の依存関係。シナプス結合は与えられたデータセット内の全データに対して共有される。

(Hidden neurons) として H で表す。一般に隠れ層の神経細胞の状態 x_i は可視層に与えられた個々の入力データに依存して変化することに注意しよう。一方シナプスの状態は多数の入力データ、つまりデータセット、を整合的に説明可能なように決まるはずなので、その状態はデータセット内のデータ全体に共有されることに注意して欲しい。つまり各データをインデックス d で区別すれば、神経細胞の状態はデータ依存性を含めて x_{di} と書けるが、シナプスの状態 s_{ijk} はデータに対するインデックス d を持たない。図 6 は、最も簡単な 3 ニューロンからなるフィードフォワードネットワークの場合について、変数の依存性をグラフィカルモデルとして記述したものである。

この設定のもとで、入力データが可視層に与えられた条件下での神経細胞およびシナプスの状態の事後分布は、

$$P(\{x_{di}\}_{i \in H}, \{s_{ijk}\} | \{x_{di}\}_{i \in V}) \quad (27)$$

で与えられる。しかしこの分布は一般には極めて高次元の確率分布であり、変数 $\{x_{di}\}$ と $\{s_{ijk}\}$ も互いに依存しあうため、式 (27) に従うサンプルの直接的な発生は困難である。

3.4 ギブスサンプリング

この問題を解決するために、脳がこの高次元のサンプリングをギブスサンプリングで実装していると仮定しよう。ギブスサンプリングは、同時サンプリングが困難な、一般には多変数の同時分布からのサンプリングが、多数の一変数サンプリングの繰り返しに置き換えられることを保証する手法である [30]。ギブスサンプリングは、全変数の同時サンプリングの代わりに、他の全ての変数が与えられたとする条件付き確率分布からのサンプリングを、全ての変数について何度も繰り返せば、生成された変数の集合が、全変数の同時分布からのサンプリングに一致することを保証する。この時、各変数のサンプリング順序は固定されたものでも、ランダムな順番でも構わず、サンプリング頻度が変数ごとに異なっても良い。つまり、神経細胞の状態とシナプスの状態に当たる各変数を、非同期かつ局所的に

$$P(x_{di} | \{x_{dj}\}_{dj \neq di}, \{s_{ijk}\}) \quad (28)$$

$$P(s_{ijk} | \{x_{dj}\}, \{s_{lmn}\}_{lmn \neq ijk}) \quad (29)$$

からサンプリングすることを繰り返せば良い。ここから、神経細胞とシナプスの状態のサンプリング則が、ネットワークの大域的状態を考慮に入れる必要のない局所的かつ非同期的な状態更新則として導出される。

更新則を具体的に求めるために、まず神経細胞の状態 x_{di} について、他のすべての変数が与えられた条件下での対数尤度比を計算しよう。ベイズの公式を用いることで、

$$\begin{aligned}
 \log \frac{P(x_{di} = 1 | \dots)}{1 - P(x_{di} = 1 | \dots)} &= \log \frac{P(x_{di} = 1 | \dots)}{P(x_{di} = 0 | \dots)} \\
 &= \log \frac{P(x_{di} = 1)}{P(x_{di} = 0)} \prod_j \frac{P(x_{dj} | x_{di} = 1, \dots)}{P(x_{dj} | x_{di} = 0, \dots)} \\
 &= \frac{\sigma(v_{di})}{1 - \sigma(v_{di})} \\
 &+ \sum_j \begin{cases} \log \sigma(v_{dj, -i} + w_{ij}) - \log \sigma(v_{dj, -i}), & x_{dj} = 1 \\ \log(1 - \sigma(v_{dj, -i} + w_{ij})) - \log(1 - \sigma(v_{dj, -i})), & x_{dj} = 0 \end{cases} \\
 &= v_{dj} + \sum_j \begin{cases} (1 - \sigma(v_{dj})) w_{ij}, & x_{dj} = 1 \\ -\sigma(v_{dj}) w_{ij}, & x_{dj} = 0 \end{cases} \\
 &= v_{dj} + \sum_j (x_{dj} - \sigma(v_{dj})) w_{ij} \tag{30}
 \end{aligned}$$

を得る。ただし簡単のために、条件付き確率 $P(x_{di} | \{x_{dj}\}_{dj \neq di}, \{s_{ijk}\})$ の条件部分を $P(x_{di} | \dots)$ のように略記した。ここで $v_{dj, -i} = \sum_{k \neq i} x_k w_{kj}$ であり、4行目を得るために $v_{dj, -i} \gg w_{ij}$ を仮定して和の中を w_{ij} の一次式で近似した。上式を $P(x_{di} = 1 | \dots)$ について整理すれば神経細胞のスパイク発火状態の更新則として

$$P(x_{di} = 1 | \dots) = \sigma(v_{di} + b_{di}) \tag{31}$$

$$b_{di} = \sum_j (x_{dj} - \sigma(v_{dj})) w_{ij} \tag{32}$$

を得る。

シナプス結合変数 s_{ijk} についても、同様の計算を行うことで

$$\begin{aligned}
 \log \frac{P(s_{ijk} = 1|\cdots)}{1 - P(s_{ijk} = 1|\cdots)} &= \log \frac{P(s_{ijk} = 1|\cdots)}{P(s_{ijk} = 0|\cdots)} \\
 &= \log \frac{P(s_{ijk} = 1)}{P(s_{ijk} = 0)} \prod_d \frac{P(x_{dj}|s_{ijk} = 1, \cdots)}{P(x_{dj}|s_{ijk} = 0, \cdots)} \\
 &= q_{0,ijk} \\
 &+ \sum_d \begin{cases} \log \sigma(v_{dj,-ik} + x_{di}a_{ijk}) - \log \sigma(v_{dj,-ik}), & x_{dj} = 1 \\ \log(1 - \sigma(v_{dj,-ik} + x_{di}a_{ijk})) - \log(1 - \sigma(v_{dj,-ik})), & x_{dj} = 0 \end{cases} \\
 &= q_{0,ijk} + \sum_d \begin{cases} (1 - \sigma(v_{dj})) x_{di}a_{ijk}, & x_{dj} = 1 \\ -\sigma(v_{dj}) x_{di}a_{ijk}, & x_{dj} = 0 \end{cases} \\
 &= q_{0,ijk} + a_{ijk} \sum_d x_{di} (x_{dj} - \sigma(v_{dj})) \tag{33}
 \end{aligned}$$

を得る。ここでも、 x_{dj} の計算と同様に、条件付き確率の条件部分を $P(s_{ijk}|\cdots)$ のように略記した。ここで $v_{dj,-ik} = \sum_l x_l w_{lj} - x_i s_{ijk} a_{ijk}$ であり、4行目を得るために $v_{dj,-ik} \gg a_{ijk}$ を仮定して3行目の和の各項を a_{ijk} の1次式で近似した。また s_{ijk} の事前分布 $P(s_{ijk})$ の対数尤度比を $q_{0,ijk}$ とした。シナプスの結合状態について特に事前の偏りを導入する必要がなければ、単に $q_{0,ijk} = 0$ とすれば良い。実際以降の結果は、全て $q_{0,ijk} = 0$ として得られたものである。上式を $P(s_{ijk} = 1|\cdots)$ について整理すれば神経細胞のスパイク発火状態の更新則として

$$P(s_{ijk} = 1|\cdots) = \sigma(q_{0,ijk} + a_{ijk}q_{ij}) \tag{34}$$

$$q_{ij} = \sum_d x_{di} (x_{dj} - \sigma(v_{dj})) \tag{35}$$

を得る。

式(35)では和が与えられたデータセット内の全てのデータ d にわたって取られている点に注意して欲しい。これに対して、値がデータごとに決まる神経細胞の状態更新則(32)では、和は同じデータ d を受けるときの他の神経細胞 j にわたって計算される。つまり、神経細胞の状態は、現時刻で与えられたデータごとに更新されるのに対して、シナプス結合はデータセット内の全データを与えられた結果として更新される。これは、データセット内のデータが逐次的に与えられる我々の脳のような状況では、各データの効果を積算する分、シナプスの更新則が神経細胞のスパイク発火の更新則よりも遅いことを意味しており、これによってシナプスと神経細胞のスパイク発火の時間スケールの大きな違いが自然に説明される。

シナプス s_{ijk} は、 i 番目の神経細胞から j 番目の神経細胞への結合を表していたため、シナプス更新則 (35) は、 i 番目の神経細胞（前シナプス神経細胞）が発火し、 j 番目の神経細胞（後シナプス神経細胞）が発火すればその結合が強化され、逆に前シナプス神経細胞が発火したのに後シナプス神経細胞が発火しなければ結合が弱化されることを意味している。この性質は、実験的に観測されているシナプスのスパイク時間依存可塑性（Spike-Timing Dependent Plasticity, STDP）と呼ばれる性質と極めて良く一致している。

式 (32) や (35) では、他の神経細胞のスパイク発火状態の影響が、神経細胞の発火特性 b_{di} やシナプスの時間変化 q_{ij} に即座に影響を与えている。しかし実際の神経系では、スパイク発火の影響は、何らかの有限の時間スケールでの積算として決まるほうが自然だろう。これは、式 (32) や (35) を

$$b_{di} \leftarrow (1 - r_b) b_{di} + r_b \sum_j (x_{dj} - \sigma(v_{dj})) w_{ij} \quad (36)$$

$$q_{ij} \leftarrow (1 - r_q) q_{ij} + r_q \sum_d x_{di} (x_{dj} - \sigma(v_{dj})) \quad (37)$$

と置き換えることで実現できる。ここで r_b, r_q は、 $0 < r_b, r_q \leq 1$ を満たすパラメータであり、それぞれ b_{di} と q_{ij} の時間発展の時間スケールを調整する。もとの即時更新則は $r_b = r_q = 1$ にあたる。以下で示す数値計算の結果などから、この有限時間スケールの導入によって、学習が安定化することが分かっている。

式 (31) (34) (36) (37) を用いて、機械学習で標準的に利用される手書き文字認識の教師あり学習をリカレントニューラルネットワークに学習させた結果が図 7 である。学習の進展とともに認識精度が速やかに上昇し、学習に用いた訓練データセットと、用いなかったテストデータセットの両方に対して高い認識精度を示すようになる事がわかる。既存の学習手法のように直接的な誤差関数の最適化を行っていないために過学習が生じていないことにも注意して欲しい。

特に重要なのは、認識精度はほぼ一定の高い値を保ち続けているにもかかわらず、ネットワークの結合強度も、またネットワーク内でスパイク発火を続けるニューロンの状態も、常にほぼランダムに活動を続けている点である。ネットワークが再現性の高い出力を生成していれば、その内部構造も安定しており再現性が高いと素朴には考えられるかも知れないが、実はそうではない。むしろ内部変数が確率的に変動しているからこそ、ネットワークは要求される適切な入出力関係を生成することが可能になるのだ。神経系の実験によって観測されたスパイク発火やシナプス結合の絶え間ないランダムな変動は、この事後分布からのギブスサンプリングを実現するダイナミクスとして整合的に理解される。

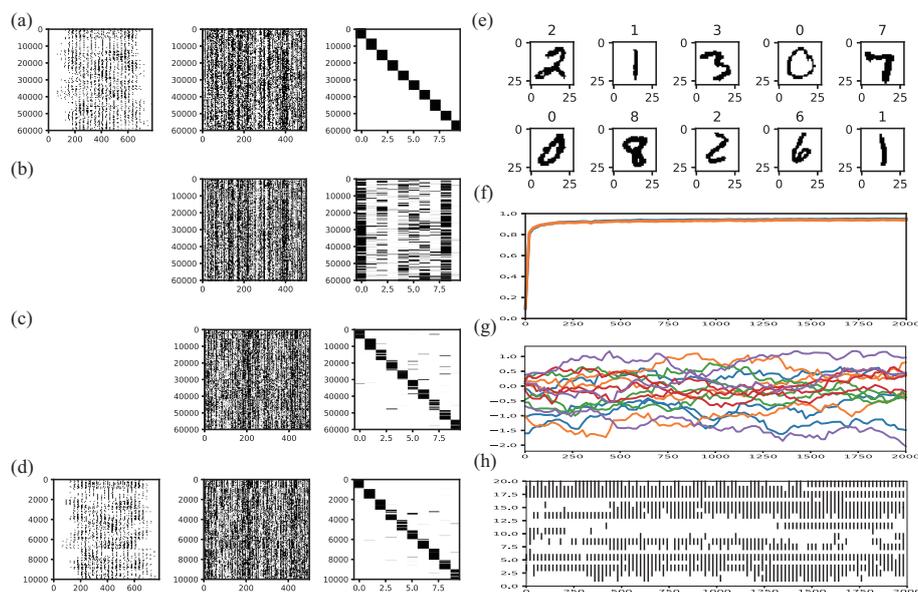


図7 手書き文字認識の学習。(a) 左：入力データ、中：隠れ層の神経細胞の発火状態、右：教師データ、(b) 左：学習前の隠れ層の状態、右：学習前の出力、(c) 左：学習後の隠れ層の状態、右：学習後の出力、(d) 左：テストデータ、中：学習後のテスト時の隠れ層の状態、右：学習後のテスト時の出力、(e) 手書き文字の例、(f) 学習精度の時間発展、(g) 結合強度の時間発展、(h) 隠れ層の神経細胞のスパイク発火の時間発展。

さらに、このようにして確率的ダイナミクスによって学習されたネットワーク中では、それぞれのニューロンが、実際の神経細胞が示すような受容野構造を自然に獲得していること、あるいは学習後のネットワーク構造の統計的な性質が実際の大脳皮質の局所ネットワークが示す統計的特性を極めて良く説明できることなども明らかになっている [24]。

3.5 まとめ

脳内の神経細胞やシナプスが多大なコストを支払ってまでランダムな活動を維持し続けるのは、決して無駄でもなければ不可避でもない。本稿の結果は、このランダムな活動こそが我々の学習の実体であり、脳の情報処理の実体であることを強く示唆している。

通常の機械学習の文脈では、学習は、ネットワークの結合強度などシステムの入出力関係を決定づけるパラメータの最適化として定式化されてきた。しかしこれまで見てきたように、脳型の学習は最適化ではない。最適化を用いなくてもニューラルネットワークに所望の機能を学習させることは可能なのだ。鍵は揺らぎを用いた事後分布からのサンプリングである。我々の脳は、与えられた入力や環境に最もよく合致する内部状態を、自分自身

でサンプリングし生成し続けるシステムとして記述できるといえる。学習や計算を、最適化ではなく、内的で確率的な情報の生成過程として理解することは、脳に限らず、広い意味での生命全般の知能や情報処理の解明にとっても、今後特に重要な視点になるのではないだろうか。

参考文献

- [1] Hubel, D. H., Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, 148(3), 574-591.
- [2] Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, 36(4), 193-202.
- [3] LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436.
- [4] Softky, W. R., Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *Journal of Neuroscience*, 13(1), 334-350.
- [5] Hromádka, T., DeWeese, M. R., Zador, A. M. (2008). Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS biology*, 6(1), e16.
- [6] Sakata, S., Harris, K. D. (2009). Laminar structure of spontaneous and sensory-evoked population activity in auditory cortex. *Neuron*, 64(3), 404-418.
- [7] Destexhe, A., Rudolph, M., Par, D. (2003). The high-conductance state of neocortical neurons in vivo. *Nature reviews neuroscience*, 4(9), 739.
- [8] Raichle, M. E. (2006). The brain's dark energy. *Science*, 314(5803), 1249-1250.
- [9] Mainen, Z. F., Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268(5216), 1503-1506.
- [10] Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3), 183-208.
- [11] Teramae, J. N., Tsubo, Y., Fukai, T. (2012). Optimal spike-based communication in excitable networks with strong-sparse and weak-dense links. *Scientific reports*, 2, 485.
- [12] Burkitt, A. N. (2006). A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. *Biological cybernetics*, 95(1), 1-19.
- [13] Fourcaud, N., Brunel, N. (2002). Dynamics of the firing probability of noisy

- integrate-and-fire neurons. *Neural computation*, 14(9), 2057-2110.
- [14] Song, S., Sjstrm, P. J., Reigl, M., Nelson, S., Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS biology*, 3(3), e68.
- [15] Lefort, S., Tomm, C., Sarria, J. C. F., Petersen, C. C. (2009). The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. *Neuron*, 61(2), 301-316.
- [16] Holtmaat, A., Svoboda, K. (2009). Experience-dependent structural synaptic plasticity in the mammalian brain. *Nature Reviews Neuroscience*, 10(9), 647.
- [17] Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J., Kasai, H. (2008). Principles of long-term dynamics of dendritic spines. *Journal of Neuroscience*, 28(50), 13592-13608.
- [18] Stettler, D. D., Yamahachi, H., Li, W., Denk, W., Gilbert, C. D. (2006). Axons and synaptic boutons are highly dynamic in adult visual cortex. *Neuron*, 49(6), 877-887.
- [19] Rokni, U., Richardson, A. G., Bizzi, E., Seung, H. S. (2007). Motor learning with unstable neural representations. *Neuron*, 54(4), 653-666.
- [20] Statman, A., Kaufman, M., Minerbi, A., Ziv, N. E., Brenner, N. (2014). Synaptic size dynamics as an effectively stochastic process. *PLoS computational biology*, 10(10), e1003846.
- [21] Loewenstein, Y., Yanover, U., Rumpel, S. (2015). Predicting the dynamics of network connectivity in the neocortex. *Journal of Neuroscience*, 35(36), 12535-12544.
- [22] Mongillo, G., Rumpel, S., Loewenstein, Y. (2017). Intrinsic volatility of synaptic connections—a challenge to the synaptic trace theory of memory. *Current opinion in neurobiology*, 46, 7-13.
- [23] Ziv, N. E., Brenner, N. (2018). Synaptic tenacity or lack thereof: spontaneous remodeling of synapses. *Trends in neurosciences*, 41(2), 89-99.
- [24] Teramae, J. (2019) Dual stochasticity in the cortex as a biologically plausible learning with the most efficient coding, bioRxiv 811646. doi: <https://doi.org/10.1101/811646>
- [25] Markram, H., Lbke, J., Frotscher, M., Roth, A., Sakmann, B. (1997). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones

- in the developing rat neocortex. *The Journal of physiology*, 500(2), 409-440.
- [26] Feldmeyer, D., Egger, V., Lübke, J., Sakmann, B. (1999). Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single ‘barrel’ of developing rat somatosensory cortex. *The Journal of physiology*, 521(1), 169-190.
- [27] Lee, W. C. A., Bonin, V., Reed, M., Graham, B. J., Hood, G., Glattfelder, K., Reid, R. C. (2016). Anatomy and function of an excitatory network in the visual cortex. *Nature*, 532(7599), 370.
- [28] Schmidt, H., Gour, A., Straehle, J., Boergens, K. M., Brecht, M., Helmstaedter, M. (2017). Axonal synapse sorting in medial entorhinal cortex. *Nature*, 549(7673), 469.
- [29] Hiratani, N., Fukai, T. (2018). Redundancy in synaptic connections enables neurons to learn optimally. *Proceedings of the National Academy of Sciences*, 115(29), E6871-E6879.
- [30] Geman, S., Geman, D. (1987). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. In *Readings in computer vision* (pp. 564-584). Morgan Kaufmann.