

アナログ神経回路網 vs. ボルツマン機械

1E-2

深井 朋樹¹ 椎野 正寿²
 1 群馬女子短期大学 2 東京工業大学

神経回路網は並列分散処理モデルとして盛んに研究されてきたが、^{1,2)}一方では物理学におけるスピン系と離散値をとる形式ニューロンの系との類似性から、物理学的な方法を駆使した基礎的研究が進められている。³⁾しかしながら、処理ユニットが離散型であることは少なからず生理学的なモデルとしての現実性を損ない、また最適化問題等への応用の面から見ても想起過程が準安定状態に陥り易い等の欠点を伴う。準安定状態を避け、記憶された状態をうまく想起するために様々な工夫がされているが、それらは大きく二つの組に分類されるだろう。一つはボルツマン機械に代表される離散型処理ユニットに確率的状態更新規則を組み合わせた確率過程モデルであり、⁴⁾他の一つは情報処理ユニットの状態及び出力が連続値を取るようなアナログ神経回路網である。⁵⁾前者では確率的揺らぎが準安定状態から系の状態を抜け出させる役目を果たし、後者ではいわゆるエネルギー面の凹凸を滑かにすることで準安定状態が消去されるものと考えられる。その構成上の違いにも関わらず両者は似通った点を持つが(例えばどちらも”温度”に対応するパラメータを通して記述される)、この二つのモデルはどの程度性質を共有し、またどのような優劣の差があるのであろうか。本講演では、シナプス結合が対称な場合について準安定状態の密度分布^{6,7)}を理論的に求め、二つのモデルの性質、動作を論じ上述の疑問に一つの回答を与えたい。

N 個の処理ユニットを持つ神経回路網と、通常のヘップ則により p 個のランダムパターン $\{\xi_i^\mu = \pm 1\}$ ($\mu = 1 \dots p$) を埋め込んだシナプス結合 $J_{ij} = (1/N) \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu$ ($i \neq j$), $J_{ii} = 0$ を考える。ボルツマン機械においては形式ニューロンを $\{S_i = \pm 1\}$ ($i = 1 \dots N$) とすると、時刻 $t+1$ での神経 i の状態 $S_i(t+1)$ は以下の確率 P_+ で $+1$ となる。

$$P_+ = \frac{1}{1 + \exp(-\beta \sum_{j \neq i} J_{ij} S_j(t))} \quad (1)$$

ここでパラメータ β は統計力学との対応では温度の逆数と解釈される。一方アナログ神経回路網では神経ユニットの状態は以下の発展方程式に従う連続的変数 u_i で表現される。

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j \neq i} J_{ij} z_j, \quad z_i = f(u_i) = \tanh(\beta u_i). \quad (2)$$

ここで f は処理ユニットの入出力関数である。上式の β はやはり温度の逆数のような働きをするので、(1) 式のパラメータと特に区別せずに用いる。どちらのモデルも $\beta \rightarrow \infty$ では温度 0 の決定論的な離散型神経回路網に移行する。良く知られているようにアナログ神経回路網ではエネルギー関数を最小にするような状態発展が起こるが、⁵⁾ボルツマン機械においてもその平衡状態では、温度 0 の離散型神経回路網のエネルギー関数²⁾にエントロピーを考慮した自由エネルギーが最小化される。ボルツマン機械を最適化問題に適用する場合は一般に冷却過程(アニーリング)を同時に伴うので想起は非平衡過程であるが、冷却は極めてゆっくり行われるので⁴⁾平衡状態での自由エネルギーの準安定状態分布を調べることは十分意味がある。結局以下のようなエネルギー(自由エネルギー)関数の準安定状態の密度分布を求める問題に帰着する。

$$E = -\frac{1}{2} \sum_{i \neq j} J_{ij} z_i z_j + \frac{1}{2\beta} \sum_i [(1+z_i) \ln(1+z_i) + (1-z_i) \ln(1-z_i)] + \gamma \left(\frac{\alpha z_i}{1 - \beta(1 - \sum_i z_i^2/N)} - \alpha z_i \right). \quad (3)$$

但し $\alpha = p/N$ であり、ボルツマン機械に対しては $z_i \equiv \langle S_i \rangle$ とする。右辺第三項はスピン系の物理学では Onsager 反作用項と呼ばれ、⁸⁾ $\gamma = 0$ はアナログ神経回路網に、 $\gamma = 1$ は平衡状態でのボルツマン機械に対応する。エネルギー関数が与えられたので、準安定状態の密度分布は次の式で求まる。⁹⁾

$$n(q, \alpha, \beta, \gamma) \equiv \int_{-1}^1 \prod_i dz_i \prod_i \delta\left(\frac{\partial E}{\partial z_i}\right) \left| \det \frac{\partial^2 E}{\partial z_i \partial z_j} \right|^{-1} \exp[NG(q, \alpha, \beta, \gamma; \text{積分パラメータ})]. \quad (4)$$

ここで $q \equiv \sum_i z_i \xi_i^{(1)}/N$ は任意の記憶パターンと準安定状態とのパターンオーバーラップで相関の度合を示すパラメータである。すなわち $N \rightarrow \infty$ では、関数 $G(q, \alpha, \beta, \gamma)$ が正になるようなパラメータ領域にのみ準安定状態が分布し、負

Analog Neural Networks versus Boltzmann Machines

Tomoki FUKAI¹, Masatoshi Shiino²

1 Gumma Women's College, 2 Tokyo Institute of Technology

になるような場合には存在しない。上式(4)ではいくつかの積分パラメータを除きほとんどの計算が解析的に遂行され、右辺の様な指数関数的分布が得られるのであるが、 $N \rightarrow \infty$ では残りの積分も鞍部点法により容易に数値的に実行できる。

典型的な場合 ($\alpha = 0.05, \beta = 3.0$) について、計算結果は図1の様になる。アナログ神経回路網では $q \approx 1$ で埋め込みパターンとの相関が極めて大きい安定想起状態、それより相関が小さい状態(これらは不安定であることが別の解析から解る⁹⁾)及び $q \approx 0$ で相関を持たない準安定状態が存在する。各ピークの高さは 10^{-8} 程度である。第三番目の状態の密度が極めて小さい事はアナログ神経回路網の想起能力の優秀さを示している。一方ボルツマン機械ではこの様な準安定状態の極端な減少は見られず、密度分布は温度0の離散型神経回路網のそれ^{6,7)}に近い。つまりボルツマン機械では、アナログ神経回路網で存在した $q = 0$ 近傍の準安定状態と、中間に位置する不安定状態との間の $G < 0$ の谷間が消失している。ボルツマン機械で冷却を極めて緩やかにする必要があるので、有限温度においてもこのように多数の準安定状態が存在しているためであると考えられる。さらに低温側では、アナログ神経回路網においても $q = 0$ 近傍の分布が増大するのが観測されるが、この増大は緩やかである(図2)。これらの理論的結果から、記憶想起においてアナログ神経回路網は極めて優秀な能力を有するが、ボルツマン機械ではアナログ神経回路網程の能力は期待できないと結論される。

α が増化すると、アナログ神経回路網では $q \neq 0$ の二つの安定、不安定状態のピークが合流しある値でついに負となり想起状態が消失する。一方ボルツマン機械ではある値で想起状態に対応するピークが $q = 0$ 側の大きな分布に連続的に接続する。これは想起状態が準安定状態と区別できなくなることを意味する⁶⁾。そこでこれらの値でそれぞれのモデルの臨界記憶容量を定義すると、図3の曲線が得られる。先に述べたように温度0の極限は一致するから、アナログ神経回路網の場合も非常に低温 ($\beta \gtrsim 40$) ではボルツマン機械と同等の判定基準に頼らざるを得なくなり、記憶容量の曲線も内側に切れ込んでくるのだが ($\beta = \infty$ で $\alpha_c = 0.113^6)$ この図には明示していない。図3と各々のモデルの熱力学的相転移曲線として求まる記憶容量³⁾とを比較すると、アナログ神経回路網では(上述の低温領域を除いて)曲線は一致し、ボルツマン機械に関してはアナログ神経回路網との開きが縮まる⁹⁾。いずれにしろ、記憶容量に関してもアナログ神経回路網が優位であるという結論が得られる。

参 考 文 献

- 1) S.Amari, and K.Maginu, Neural Networks 1 63 (1988);
- 2) J.J.Hopfield, Proc. Natl. Acad. Sci. USA 79 2554 (1982).
- 3) D.J.Amit, H.Gutfreund and H.Sompolinsky, Phys. Rev. Lett. 55 1530 (1985).
- 4) S.Geman and D.Geman, IEEE Trans. PAMI 6 721 (1984).
- 5) J.J.Hopfield, Proc. Natl. Acad. Sci. USA 81 3088 (1984).
- 6) E.Gardner, J. Phys. A19 L1047 (1986).
- 7) T.Fukai, J. Phys. A23 249 (1990).
- 8) D.J.Thouless, P.W.Anderson, and R.G.Palmer, Philos. Mag. 35 593 (1977).
- 9) T.Fukai and M.Shiino, J. Phys. A23 L1009 (1990); Phys. Rev. A 掲載決定.

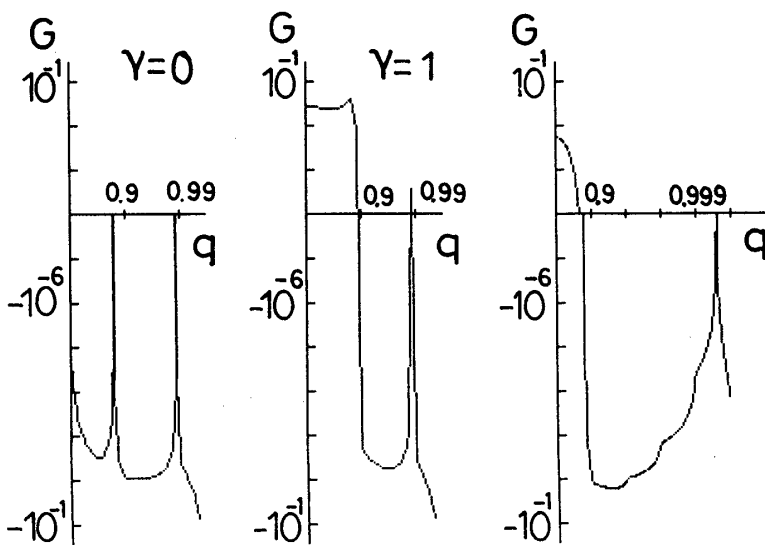


図1. $G (\alpha = 0.05, \beta = 3, \gamma)$

図2. $G (\alpha = 0.05, \beta = 10, \gamma = 0)$

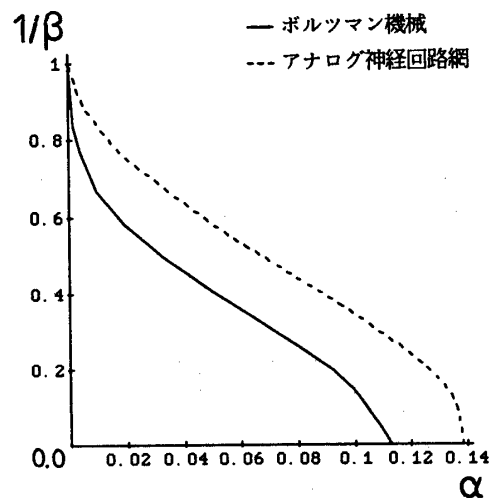


図3. 臨界記憶容量