

ニューラルネットワークと人間の情報処理の類似性

－数独の求解を例として－

日大生産工(院) ○大谷 哲広
日大生産工 松田 聖

1 まえがき

ホップフィールドネットワークを用いた組合せ最適化問題の求解が多く試されている。同様に、多くのニューラルモデルは連想記憶、音声認識、自然言語理解、視覚情報処理、推定などの人間の知的活動を提案してきた[1]、[2]。筆者らは、人間のすべての知的活動は最適化過程であるとの仮説に基づき、すでに代表的な人間の知的活動である意思決定のホップフィールドモデルを提案している[3]、[4]。

一方、ホップフィールドネットワークを用いたパズル解法の試みとして、Nクイーン問題や、4色問題などの例がある。パズルを解くことは知的な活動であり、最適化過程として見ることもできる。他方で、ホップフィールドネットワークの欠点として、ローカルミニマムに陥りやすいこと、ホップフィールドネットワークを用いた組合せ最適化問題を解く際に、問題の大きさと最適解発見困難の間に密接な関係があることがよく知られている。そこで、人間とホップフィールドネットワークの問題解法における類似点や相違点についての観点から、人間にとって簡単に解くことができる問題がニューラルネットワークにとって簡単なのか。または、その逆なのかということについて、研究していく。

2 ニューラルネットワーク

人間の脳にはニューロンと呼ばれる140億個の神経細胞が存在する。その神経細胞は互いに結合し情報を交換し合い、人間の記憶や判断などの精神活動を行っている。このような脳の情報処理機構を、神経細胞の働きを模擬したニューロンに多数配置することによってコンピュータ上で行おうとするのがニュー

3 ホップフィールドネットワーク

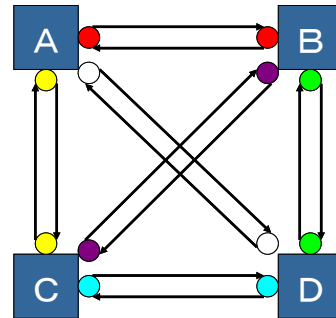


図.1

ラルネットワークである。ホップフィールドネットワークは相互結合型ネットワークの代表的なものである。1980年代の前半にJohn Hopfieldによって提案された。このネットワークの特徴はニューロン間に相互的な対称作用があることと、ネットワークにエネルギー関数を導入しているということである。ニューロン間に相互的な対称作用があるというのは、図.1を参照すると、ニューロンAからBにおいて、AからB、BからAへの結合荷重（図.1で結合荷重を円で表現している）が等しいということである。この対称作用がすべてのニューロンにおいて非同期的に行われている。エネルギー関数は、現在のネットワークの状態を知る方法として用いられ、John Hopfieldらによって、このネットワークの状態は、必ずエネルギー関数が減少する方向に変化することが示されている。エネルギー関数がどのような形状になるかは結合荷重分布など、ニューラルネットワークの構造によって変わるが、一般に多数の極小点を持つ多安定関数になる。

4 数独の解法

例題としてProblem Aを図.2に示す。数独は

Problem A-1

	9			1		2	7
		7	4		6		1
1		5	7	2		6	4
			1		8	3	4
4	8						
			6		5	8	2
7		9	2	8		1	6
		8	5		1		9
	1			3			8

図. 2

6	9	4	8	1	3	5	2	7
8	2	7	4	5	6	9	3	1
1	3	5	7	2	9	6	4	8
5	6	2	1	7	8	3	9	4
4	8	1	3	9	2	7	5	6
9	7	3	6	4	5	8	1	2
7	5	9	2	8	4	1	6	3
3	4	8	5	6	1	2	7	9
2	1	6	9	3	7	4	8	5

図. 3

9×9のマスを構成され、あらかじめいくつかのマスには数字が入っている。数独の目的は、行、列、太線で囲まれた3×3の範囲のすべてのマスに1～9の数字をそれぞれ一つずつ入れて、すべてのマスに数字を入れることである。数独は大抵簡単に解くことができるが、中にはとても難しい問題もある。

Problem Aは以下のように解くことができる。

始めに、例えば3番目の行には3つの空白があり、それにはこの行で、まだ出現していない3、8、9が入る。8と9のどちらもすでに第2列にあるため、このマスには8と9は入らない。よって、このマスには3が入る。同様に、6列には8があるため、数字2と6の間には8は入らず9が入り、最も右のマスに8が入る。こうして、Problem A-2の図. 4が得られる。次に、第6列に4つの空白のマスがあり、この列には2、3、4、7が入る。数字2はすでに第6列を含む上、下二つの3×3の範囲に存在しているため、第6列の5行目のマスに入る。数字3は一番下の3×3の範囲に存在しているため、第6列の1行目のマスに入る。数字7は7行目に存在するので、第6列の9行目のマスに入る。最後に4は余ったマスに入る。こうして、Problem A-2の図. 5が得られる。このような、同じ行、列や3×3

Problem A-2

	9			1			2	7
		7	4		6			1
1	3	5	7	2	9	6	4	8
			1		8	3		4
4	8							
			6		5	8		2
7		9	2	8		1	6	
		8	5		1			9
	1			3			8	5

図. 4

	9			1	3		2	7
		7	4		6			1
1	3	5	7	2	9	6	4	8
			1		8	3		4
4	8				2			
			6		5	8		2
7		9	2	8	4	1	6	
		8	5		1			9
	1			3	7		8	5

図. 5

の範囲にある数字から判断して空白のマスに数字を入れる手法を『scanning』と呼ぶ。

『scanning』を使うことで、Problem A-1の図. 2を解き、図. 3を得ることができる。

通常、我々は『scanning』を使って容易に数独を解くことができる。しかしながら、時に、困難な問題に出会うことがある。そのような、難しい問題は『scanning』だけでは解くことができなく、いわゆる『forcing chain』を必要とする。例[5]として、Problem Bに図. 6を示す。

Problem B[5]

x	2	3	4	5	6	7	8	9
7	8	5	2	1	9	4	6	3
6	4	9	7	8	3	5	1	2
2	3	7	6	9	8	1	5	4
8	1	4	5	7	2	9	3	6
9	5	6	3	4	1		f	e
4	7	a		3	5	6	d	b
5	6	8		2	4	3	e	
3	9			6	7		4	5

図. 6

『scanning』を使うと左上のx=1を見つけるこ

とができるが、とはいえ、aの値もfの値も『scanning』を使っても見つけることはできない。しかし、aの値は1か2のいずれかであり、下のようになずれの場合でも、fの値は2になる。

$$a=1 \rightarrow b=8 \rightarrow c=7 \rightarrow f=2$$

$$a=2 \rightarrow d=9 \rightarrow e=7 \rightarrow f=2$$

このような解法を、『forcing chain』という。『forcing chain』とは、推測を組合せて解を得る手法であり、解を求めたいマスとは違う行、列や3×3の範囲からの情報を基に、いくつかの数字を仮定して解を導きだすものである[5]。『forcing chain』を必要とする問題を人間は一般に難しいと感じる。

5 ホップフィールドによる数独

5. 1 ホップフィールドによる数独の解法

この節でホップフィールドによる数独の解法を試す。始めに、ホップフィールドの定義を示し、ホップフィールドによる数独の定式化を与える。

5. 2 ホップフィールドネットワーク

ホップフィールドネットワークのダイナミクスは次のように定義される[3]。

$$\frac{du_i}{dt} = \sum_j w_{ij} v_j + h_i (w_{ij} = w_{ji}) \quad (1)$$

$$v_i = \frac{1}{2} \left(1 + \tanh \left(\frac{u_i}{T} \right) \right), \quad (2)$$

$v_i \in [0,1]$, u_i , h_i はニューロンiのそれぞれ、出力値、内部値、バイアスを表し、 w_{ij} は、シナプスの重み(iとjは通例 $w_{ij} = w_{ji}$, $w_{ij} \neq 0$)、そしてTは定数を表す。ネットワークのエネルギーは(3)式により与えられる。

$$E = -\frac{1}{2} \sum_i \sum_j w_{ij} v_i v_j - \sum_i h_i v_i \quad (3)$$

次によく知られている(4)、(5)式を適用する。

$$\frac{du_i}{dt} = -\frac{\partial E}{\partial v_i} \quad (4)$$

$$T \frac{du_i}{dt} = -2v_i(1 - v_i) \frac{\partial E}{\partial v_i} \quad (5)$$

この表現はニューロンの変化値エネルギーEを減少させ、最終的には漸近の安定平衡

$dv_i/dt = 0$ に収束する。ここでの0や1またはいくつかの値は $\partial E / \partial v_i = 0$ の条件を満たす。

なお、N次元空間 $[0,1]^N$ はネットワーク超立方体、または単に、超立方体と呼ばれ、 $v \in \{0,1\}^N$ は超立方体の頂点と呼ばれる(Nはネットワークのニューロン数を表す)。ネットワークの状態は超立方体のポイントによって表現できる。

5. 3 ホップフィールドによる数独

ホップフィールドによる数独の定式化を与

える。エネルギーが最小解に収束することで、数独を解法することが期待できる。数独は以下のように表現できる。

$$E = \frac{1}{2} \left[\sum_{y=1}^9 \sum_{z=1}^9 \left(\sum_{x=1}^9 v_{xyz} - 1 \right)^2 + \sum_{x=1}^9 \sum_{z=1}^9 \left(\sum_{y=1}^9 v_{xyz} - 1 \right)^2 + \sum_{x=1}^9 \sum_{y=1}^9 \left(\sum_{z=1}^9 v_{xyz} - 1 \right)^2 + \sum_{z=1}^9 \left(\sum_{y=1}^3 \sum_{x=1}^3 v_{xyz} - 1 \right)^2 + \sum_{z=1}^9 \left(\sum_{y=1}^3 \sum_{x=4}^6 v_{xyz} - 1 \right)^2 + \sum_{z=1}^9 \left(\sum_{y=1}^3 \sum_{x=7}^9 v_{xyz} - 1 \right)^2 + \sum_{z=1}^9 \left(\sum_{y=4}^6 \sum_{x=1}^3 v_{xyz} - 1 \right)^2 + \sum_{z=1}^9 \left(\sum_{y=4}^6 \sum_{x=4}^6 v_{xyz} - 1 \right)^2 + \sum_{z=1}^9 \left(\sum_{y=4}^6 \sum_{x=7}^9 v_{xyz} - 1 \right)^2 + \sum_{z=1}^9 \left(\sum_{y=7}^9 \sum_{x=1}^3 v_{xyz} - 1 \right)^2 + \sum_{z=1}^9 \left(\sum_{y=7}^9 \sum_{x=4}^6 v_{xyz} - 1 \right)^2 + \sum_{z=1}^9 \left(\sum_{y=7}^9 \sum_{x=7}^9 v_{xyz} - 1 \right)^2 \right] \quad (6)$$

$$= \frac{1}{2} \left[\sum_{y=1}^9 \sum_{z=1}^9 \left(\sum_{x=1}^9 v_{xyz} - 1 \right)^2 + \sum_{x=1}^9 \sum_{z=1}^9 \left(\sum_{y=1}^9 v_{xyz} - 1 \right)^2 + \sum_{x=1}^9 \sum_{y=1}^9 \left(\sum_{z=1}^9 v_{xyz} - 1 \right)^2 + \sum_{x=1}^9 \sum_{y=1}^9 \left(\sum_{z=1}^9 v_{xyz} - 1 \right)^2 + \sum_{i=0}^2 \sum_{j=0}^2 \left\{ \sum_{z=1}^9 \left(\sum_{x=3i+1}^{3(i+1)} \sum_{y=3j+1}^{3(j+1)} v_{xyz} - 1 \right)^2 \right\} \right] \quad (7)$$

なお $v_{xyz} = 1$ は、数独のマス(x,y)にz(1~9)の数字が入ることを意味する。(6)や、(7)式の初めの3式はそれぞれの行や、列にz(1~9)の数字が入ることを意味し、(6)の4~12番目の式や、(7)の4番目の式は太枠の3×3の範囲の

各マスに $z(1\sim 9)$ の数字が入ることを意味する。ネットワークの状態が超立方体の頂点 $\mathbf{v} \in \{0,1\}^N$ の時に $E(\mathbf{v}) = 0$ となり数独の解が得られる。 $729(=9^3)$ のニューロンのネットワークダイナミクスは以下のように与えられる。

$$\begin{aligned} \frac{du_{abc}}{dt} &= -\frac{\partial E}{\partial v_{abc}} \\ &= -\left(\sum_{x=1}^9 v_{xbc} - 1\right) - \left(\sum_{y=1}^9 v_{ayc} - 1\right) \\ &\quad - \left(\sum_{z=1}^9 v_{abz} - 1\right) \\ &\quad - \left(\sum_{x=3(i_0+1)}^9 \sum_{y=3(j_0+1)}^9 -1\right) \\ &= -\sum_{x=1}^9 v_{xbc} - \sum_{y=1}^9 v_{ayc} - \sum_{z=1}^9 v_{abz} \\ &\quad - \sum_{x=3(i_0+1)}^9 \sum_{y=3(j_0+1)}^9 v_{xyc} + 4, \end{aligned} \quad (8)$$

$$v_{abc} = \frac{1}{2} \left(1 + \tanh \left(\frac{u_{abc}}{T} \right) \right), \quad (9)$$

なお、ここで、 $3i_0 + 1 \leq a \leq 3(i_0 + 1)$ かつ、 $3j_0 + 1 \leq b \leq 3(j_0 + 1)$ である。そして前もって与えられたニューロン値は計算中に変化させない。

6 シミュレーション

この節では、(8)や(9)で与えられた式によってホップフィールドネットワークが数独を解くことができるのかをシミュレーションによって調べる。筆者らはProblem A、Bの他に[6]に示された初級、中級、上級、番外編の問題を10問ずつのシミュレーションを行った。それらの問題を各ニューロンの初期値を0.5近くのランダム値に設定して10回ずつシミュレーションを行った。各々のニューロン値の更新は、(1)式の微分方程式よりも次の差分方程式を用いて非同期的に行った。

$$\Delta u_i = \left(\sum_j w_{ij} v_j + h_i \right) \Delta t \quad (10)$$

そのため、(9)式の代わりに(11)、(12)式を使用した。

$$\begin{aligned} \Delta u_{abc} &= -\frac{\partial E}{\partial v_{abc}} \Delta t \\ &= \left(-\sum_{x=1}^9 v_{xbc} - \sum_{y=1}^9 v_{ayc} - \sum_{z=1}^9 v_{abz} \right. \\ &\quad \left. - \sum_{x=3(i_0+1)}^9 \sum_{y=3(j_0+1)}^9 v_{xyc} + 4 \right) \Delta t \end{aligned} \quad (11)$$

$$v_{abc} = \frac{1}{2} \left(1 + \tanh \left(\frac{u_{abc}}{T} \right) \right), \quad (12)$$

なお、 $3i_0 + 1 \leq a \leq 3(i_0 + 1)$ かつ、 $3j_0 + 1 \leq b \leq 3(j_0 + 1)$ である。 Δt や T の値は次のように設定した。

$$\Delta t = 0.1$$

$$T = 0.5$$

シミュレーション結果は初級、中級、上級、番外編の問題全てで、数独の解を得ることができた。

7 まとめ

筆者らは、ホップフィールドネットワークによる数独の解法を試みた。今回の実験により、人間が難しく感じる数独をニューラルネットワークで解法可能という興味深い結果が得られた。

このことから、ホップフィールドネットワークも人間と同様に、scanningだけでなく forcing chainを必要とするような人間にとって難しい問題も解いていると考えられる。一方、8クイーン問題のような人間にとって簡単に解くことができる問題がニューラルネットワークで簡単に解ける訳ではないという事実もある。今後はこれらの相違関係を明らかにし、なぜ数独が解法可能なのかという研究を進めていきたい。

「参考文献」

- [1] K. Fukushima, "Neural network and information processing", Asakurapub, 1989, in Japanese.
- [2] M. Hassouned., "Associative Neural Memories" Oxford University Press 1992.
- [3] S. Matsuda, "Theoretical characterizations of possibilities and impossibilities of Hopfield neural networks in solving combinatorial optimization problems (Extended Abstract)", icnn' 94, vol. pp. 4563_4566, 1994.
- [4] S. Matsuda, "A neural network model of the decision-making process based on AHP" iicnn2005, 2005
- [5] Wolfram MathWorld, Sudoku, <http://mathworld.wolfram.com/Sudoku.html>
- [6] <http://www.ri-center.tsukuba.ac.jp/~kurata/nplc/nplmon/B020705.HTM>, in Japanese.