

# リカレントネットによる内部状態遷移を要する 問題学習時の初期重み値の影響

大分大学 田口優馬 柴田克成

Effect of initial weight value when a problem requiring internal state  
transition is learned by a recurrent neural network

Yûma Taguchi, Katsunari Shibata, Oita University

**Abstract** In this paper, to solve the problem that it is difficult for a recurrent neural network to learn a task requiring multi-step discrete state transition, the initial weight for the feedback connections were investigated in the supervised learning of a counter task. It was observed after learning that multi-step state transition was mainly represented by the chain of excitation-suppression-excitation-suppression among hidden neurons. To make the learning easy, it is suggested that the weight for self-feedback connections is moderate and that for the neighbour connections is negative.

## 1 まえがき

現在のロボットは、自転車に乗ることが出来るなど巧みな行動を行うことができる。しかし実世界において様々な状況に柔軟に対応して適切な行動ができるロボットは見当たらない。このような能力は人間がプログラムで与えることは困難であり、ロボットが自律的に学習し、目的に応じた状況判断能力を自ら身につけて行くことが重要であると考えられる。

このような観点から当研究室では強化学習の自律学習能力、ニューラルネットの抽象的な内部表現獲得能力に注目し、両者を組み合わせた学習の研究を行ってきた。また記憶を扱うことができるリカレントネットを用いることで膨大なセンサ信号の中から必要な情報を抽出して記憶し、行動に反映させることも学習によって獲得できることを確認した<sup>1)</sup>。しかし、例えばロボットがスイッチを押してからゴールへと向かう行動の学習はできて、2個のスイッチを押した後にゴールにへ向かうというような多段階の状態遷移が必要な問題になると急激に学習が困難となった。しかし人間はセンサ情報が時々刻々と変化する実環境の中で柔軟に状況判断を行い、多段階の離散的な状態遷移に基づき行動していると考えられる。例えば車に乗る際、「鍵を開ける」「ドアを開ける」「座る」「エンジンをかける」というように人間は無意識のうちに目的に応じて状態を抽象化・分節化することで、効率的に適切な行動を学習しているように見える。

多段階の状態遷移の学習のためには、状態が遷移すべきでない場合はそのまま保持し、遷移すべき時には状態をガラッと変化させるというような一見相反することを

両立させなければならないという難しさがある。リカレントネットを用いて多段階の状態遷移が必要な簡単なタスクであるカウンタータスクを教師あり学習させても学習が容易でないことを考えると、リカレントネットの学習を改善する必要があるのではないかと考えられる。

これまで当研究室ではリカレントネットを用いて学習させた場合、記憶した情報を保持させ、かつ学習を安定させるという観点から適切と考えられる初期重み値を設定して学習させてきた。しかしこれでは1つの状態を安定して記憶させるためには良いが、異なる状態へと抜け出すことに対しては妨げになるために多段階の状態遷移を表現することが困難になる可能性が考えられる。

そこで本研究ではリカレントネットを用いて学習を行う際に、状態の保持と遷移の両立をさせるために、フィードバック部の初期重み値どのように与えることが良いかを模索する。

## 2 リカレントネットにおける初期重み値の重要性

過去の入力を現在の出力に反映させる必要があるタスクをリカレントネットを用いて BPTT ( Back Propagation Through Time ) で学習させる場合、フィードバック部の重み値をうまく設定しないと過去の入力が現在まで保持されず、さらに BPTT で誤差を過去に逆伝播させる際にも時間を遡ることで伝播誤差が小さくなったり、逆に発散してしまって学習されないことが考えられる。このことは文献<sup>2)</sup>でも指摘されており、彼らはそれを避けるための Long Short-Term Memory という特殊な構造のニュー

ラルネットを提案している。一方、本研究室では以下の根拠を元に通常のリカレントネット、セルフフィードバックは 4.0、それ以外は 0.0 という値を用いてきた。

まず、簡単のため Fig.1 のように外部入力がなく、セルフフィードバック結合のみを有する単一ニューロンの離散時間モデルを考える。ニューロンの内部状態を  $u$ 、ニューロンの出力を  $x$ 、重み値を  $w$ 、バイアスを  $\theta$ 、ニューロンの出力関数であるシグモイド関数を  $f$  とする。次の時刻  $t+1$  でのニューロンの出力  $x_{t+1}$  は現時刻での出力  $x_t$  を用いて次式で表される。

$$x_{t+1} = f(u_{t+1}) = f(wx_t + \theta) \quad (1)$$

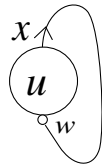


Fig. 1: One neuron only with a self-feedback connection.

Fig.2 にこの場合のリターンマップを示す。このニューロンが記憶を形成する場合、Fig.2 のように  $x_{t+1} = x_t$  の直線と 3 つの点で交わって 2 つの安定平衡点を形成し、双安定状態になる必要がある。そのためには少なくともシグモイド関数の傾きの最大値が  $x_{t+1} = x_t$  の傾きである 1 より大きくなる必要があるため、

$$\frac{dx_{t+1}}{dx_t} = wf'(wx_t + \theta) > 1 \quad (2)$$

を満たす必要がある。出力幅が 1 のシグモイド関数の傾きの最大値は 0.25 であるので、初期重み値として 4.0 を設定すると双安定状態となるかならないかの境目の状態となり、必要に応じて容易に記憶を形成できると考えられる。

一方、学習の観点から考えると、BPTT の学習におい

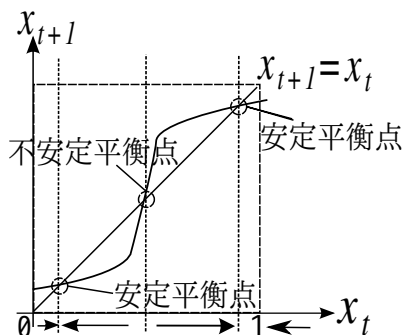


Fig. 2: Return map of one neuron only with a self-feedback connection.

て時間をさかのぼって誤差を伝播する際、伝播誤差  $\delta$  は

$$\delta_{t-1} = wf'(u_{t-1})\delta_t \quad (3)$$

と計算される。シグモイド関数の導関数  $f'$  の最大値は 0.25 であることから初期重み値が 4.0 であると積が 1.0 となり  $\delta$  が発散せず、また急激に減衰することなく効率的に過去への誤差伝播を行うことができる。他のフィードバックの結合については、ここまでの議論が成り立たなくなる可能性をなくすため、すべて 0 を初期値にして学習を行った。しかし、多段階の状態遷移の学習が困難なことから考えると、この初期重みの設定のままでは 1 つの状態を保持することは容易であるが、中間層同士の結合がないために多段階の状態遷移を学習を通して表現できるようになることは困難であると考えられる。

そこで、本研究では状態の保持と遷移の両立を促すために、従来初期重みの段階では重み値を 0 としていた中間層ニューロン同士をどのように結合させれば良いかを探った。

### 3 カウンタタスクの教師あり学習

学習には、リカレントネットとして中間層ニューロンの出力を次の入力として付加する Elman 型のリカレントネットを使用し、BPTT 法に基づいて時間をさかのぼり教師あり学習を行う。

状態遷移が必要なタスクとして数をかぞえるカウンタタスクを用いる。まず出力層ニューロンを Fig.3 のように 6 個配置し、その時カウントしている数に対応するニューロンに教師信号 0.9 を与え、それ以外のニューロンには教師信号 0.1 を与える。入力層ニューロンに 1 が入力された場合、出力層ニューロンの発火状態が隣のニューロンに順次遷移する。入力層ニューロンに 0 が入力された場合、遷移させずにその時点での状態を保持するよう前のステップと同じ教師信号を与える。そして毎ステップ BPTT で、カウントを開始した時刻まで遡って教師あり学習をさせる。入力は 0 が 3/4、1 が 1/4 の確率で入るように設定し、状態を保持する時間をある程度確保するようにした。6 つ目の出力ニューロンが発火している時に 1 が入力された場合、その試行を終了し、すべての中間層ニューロンの値を 0 と 1 の中間の値である 0.5 にリセットし再び 1 からカウントを始める。

ネットワークの構造は 3 層の Elman 型リカレントネット、中間層ニューロンの数は 10 個とし、ニューロンの出力は 0~1 の値域のシグモイド関数を用いた。また、各中間層ニューロンでは外部入力がない場合に 0.5 の出力

を保持するようにバイアスの初期値を設定した。例えば、セルフフィードバックの初期重み値が 4.0 で他のフィードバックが 0.0 の場合、バイアスの初期値を 2.0 とした。学習係数は 0.4、フィードバック部の学習係数は 0.02、バイアスは 0.04 とした。また、学習の成功条件は各出力ニューロンの出力と教師信号の誤差の絶対値が 0.1 以下の状態が 1000 回連続で続く場合とし、40 万ステップ学習をさせても成功条件を満たさない場合は失敗とし学習を打ちきった。

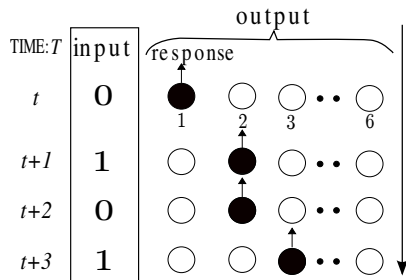


Fig. 3: Counter task as a multi-step state transition task.

## 4 学習結果

### 4.1 セルフフィードバックを変更した場合

初めに、セルフフィードバック部の初期重み値が大きいと状態遷移の学習を阻害してしまうかどうかを確認するためにセルフフィードバック部の初期重み値を変更し、比較した結果を Table1 に示す。セルフフィードバック以外の初期重み値はすべて -1 ~ 1 の乱数とし、セルフフィードバック部の初期重みを 0 ~ 4 の 5 パターンおよび、他のフィードバックと同様に -1 ~ 1 の範囲の乱数とした計 6 パターンについて、乱数系列を 300 パターン変えて学習性能を観察した。

Table 1: Learning performance according to the initial weight for self-feedback connections.

initial weight for self-feedback	success ratio(%)	average steps	standard deviation
0	62	291646	63062
1	76	267725	65009
2	80	255460	70682
3	80	226122	69893
4	69	288141	75794
-1 ~ 1	42	306916	56665

Table1 にそれぞれの場合の学習成功率と学習成功までの平均ステップ数を示す。この表より予想していたよう

な大きな差は見られなかったものの、セルフフィードバック部の初期重み値が 2 および 3 の場合がもっとも成功率が高い結果となり、学習成功までのステップ数も 3 の場合が最も少なくなった。これらの結果よりセルフフィードバック部の初期重み値は状態遷移の観点から 4 では大き過ぎ、また記憶の保持の観点から 0 では小さ過ぎると推測される。

### 4.2 中間層ニューロン同士の結合について

次に、状態遷移の学習を加速するために、中間層ニューロン同士の初期重み値をどのようにするのが良いかを調べた。特に、遷移を加速すると期待される初期重み値について、その効果を見る。ここでは、状態が次々と遷移するためには、中間層ニューロンを鎖状に、つまり隣同士に初期重み値を持たせ、その他を 0 とすることを考える。本来は、複数の状態に分岐することも考慮する必要があるが、ここでは、扱っているタスクがカウンタータスクであることもあり、簡単のため状態分岐については考慮しない。そして、セルフフィードバックの初期値を 3 に固定し、Table1 でも示したが、セルフフィードバック以外をすべて -1 から 1 の乱数とした場合を基準とし、セルフフィードバック以外の重み値をすべて 0 にした場合、次のニューロンに対して 0、前のニューロンに対して -1 で結合した場合、次のニューロンに対して +1、前のニューロンに対して -1 で結合した場合、隣のニューロンにはすべて -1 で結合した場合、さらに、他のニューロンに抑制の結合を持たせた場合に学習がうまくいったことから、セルフフィードバック以外の他のニューロンすべてに対し、フラットに -0.3 の負の重み値とした場合の 6 通りについてその結果を Table2 に示す。なお、中間層ニューロンの両端、今回の場合 10 番ニューロンと 1 番ニューロンは隣り合っているものとみなし、結合させている。

当初は、中間層ニューロンも出力ニューロンのように順次発火することを想定していたため、次のニューロンには正の結合、前のニューロンには負の結合を行うことが良いと予想していたが、実際には、隣同士ともに負の結合をした場合が最も学習成功率が高いという結果となった。さらに、この場合の、学習後の中間層ニューロンの各出力の時間変化の様子を Fig.4 に示す。この図より、中間層ニューロンの発火が徐々に遷移していくのではなく、カウントするたびに、たとえば、カウントが 4 に移ると 4 番ニューロンの値が大きくなり、カウントが 5 に移ると 3 番ニューロンの値が減り、カウント 6 では 2 番ニューロンの値が増大するという具合に、ニューロンごとに値

が大きくなるものと小さくなるものに分けられ、カウント3から4へ遷移する場合を除き、互いの抑制結合を利用して順次状態が遷移していていることがわかる。

Table 2: Learning performance according to the initial weight for neighbor connections.

initial weight of self-feedback 3					
to previous	to next	otherwise	success ratio(%)	average steps	standard deviation
0	0	0	90	189202	47968
0	-1	0	86	234789	68103
-1	1	0	64	259518	69030
-1	-1	0	95	212227	60190
-0.3	-0.3	-0.3	82	250135	50747
-1 ~ 1	-1 ~ 1	-1 ~ 1	80	226122	69893

セルフフィードバック以外をランダムに結合した場合についても同様に、あるニューロンが発火すると次に別のニューロンの発火が抑制され、それによりまた別のニューロンが発火するという状態遷移が中間層内に見られた。このような状態遷移を表現するためには、影響を及ぼすニューロン同士が負の結合をすると良いと考えられることから、隣同士のニューロンのみを負の初期重み値で結合することで、そのような中間層ニューロン間の状態遷移の学習を加速したと考えられる。また実際に Fig.4 を見ると発火・抑制・発火の伝播が隣同士のニューロンであることが多いことから裏付けられる。

## 5 まとめ

リカレントネットを用いた多段階の離散的な状態遷移を要する問題の学習において、学習を簡単にするための初期重み値を模索した。セルフフィードバック部の初期重み値については、値の保持という観点から設定した4.0という値よりも小さい2.0や3.0という値を設定した場合が学習の成功率が高く、記憶の保持と状態遷移を両立するためには大き過ぎず、小さ過ぎない値が適していると推測される。

一方、学習後の中間層ではニューロン間で発火・抑制・発火・抑制の伝播が交互に現れて状態遷移していく様子が観察された。そしてそれを加速すると考えられる隣同士のニューロンを-1の初期重み値で結合した場合の学習成功率が最も良かった。

また、リセットしないカウンタータスクや振動子の学習では、個々の中間層ニューロンにおいて出力値の増減

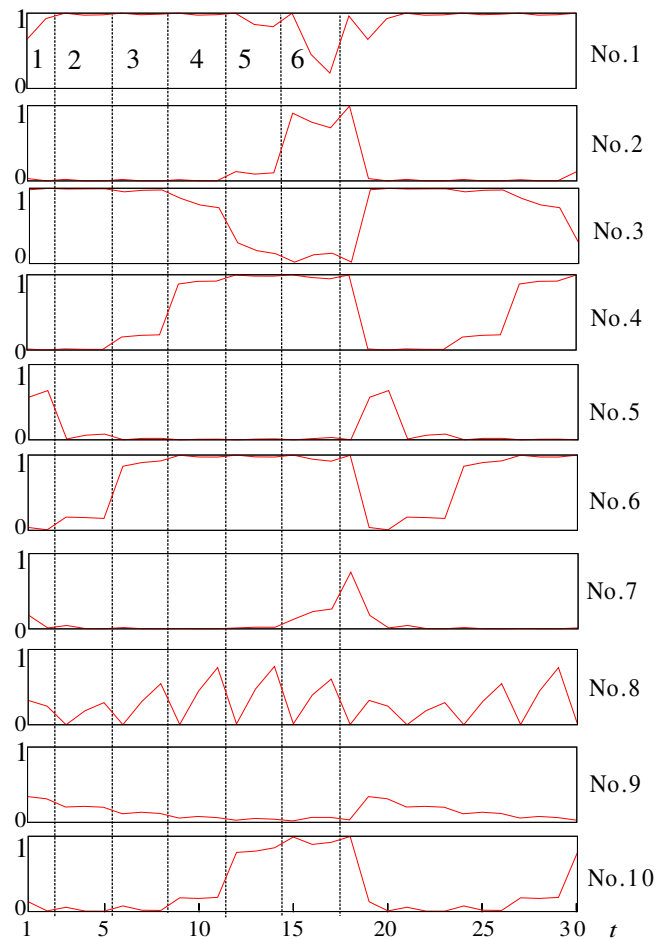


Fig. 4: The change of each hidden neuron's output. Each number in the figure indicates the counted number at the time step.

の両方を必要とするのでさらに学習が難しくなると推測される。これらのタスクや状態分岐のあるタスクの学習は今後の課題である。

## 謝辞

本研究は、日本学術振興会科技学術研究費補助金基盤研究(B)#19300070 および(C)#23500245の補助を受けた。

## 参考文献

- 1) H.Utsunomiya, K.Shibata:Advances in Neuro Information Processing, Lecture Notes in Computer Science, Vol. 5507, pp.970-978, 2009
- 2) S.Hochreiter, J.Schmidhuber:Long Short-Term Memory. Neural Computation,Vol.8, No.9, pp1735-1780, 1997.