

## リカレントネットにおける感度調整学習時の感度とリアプノフ指数の関係

大分大学 ○江島拓哉 徳丸侑輝 柴田克成

**The relation between Sensitivity and Lyapunov exponent when Sensitivity Adjustment Learning is applied to Recurrent Neural Networks**

Takuya Ejima and Yuuki Tokumaru and Katsunari Shibata, Oita University

Abstract: As an essential technique towards the emergence of thinking, we have studied new reinforcement learning using a chaotic neural network. The new reinforcement learning works well when the network status is around the "edge of chaos" which is the border between chaos and non-chaos and so the control of chaoticity is required. We have proposed a local learning method named "sensitivity adjustment learning". The sensitivity is a local index defined in each neuron as the magnitude of the output gradient with respect to the input. In this learning, the weights are updated to increase the index. In this paper, we observed the relation between sensitivities and Lyapunov exponent, which shows the network chaoticity, for several cases of a flat RNN when the learning was applied. The results show that "log sensitivity", which is the natural logarithm of the average sensitivity over all neurons, corresponds reasonably well with the Lyapunov exponent until dynamics get chaotic regardless of number of neurons or connection rate and that suggests the learning can adjust the network dynamics around the "edge of chaos".

**1 序論**

現在、人工知能が注目され、多くの分野への応用が期待されている。それは Neural Network(以下 NN) に深層学習を行うことで、認識や自然言語処理等において他の手法に比べ人間に近い柔軟な機能の獲得ができるからである [1][2]。しかし、未だ人工知能はセンサからモータまでのトータルの機能としての柔軟性や適応性では人間の脳機能に敵わない。本研究室では以前より、より高次の機能の創発を目指すに当たり機能モジュール分割に問題があると考え、入力から出力までを NN で構成し、行動に対する報酬と罰を与える強化学習を行うことで人間の手を加えず自律的に機能を獲得させる「End-to-End 強化学習」を提唱してきた [3]。この手法によりこれまでネットワーク内部に認識、記憶、予測等の機能が他と調和する形で創発することが確認されており、近年では DeepMind による、ゲームにおける戦略獲得でもこの手法により非常に優れた結果が示されている [4][5]。

本研究室では NN を用いた End-to-End 強化学習による人間の持つ諸機能の創発を目指しており、現在、典型的な高次機能の一つである「思考」の創発に向けて基礎研究を行っている。「思考」には「外部からの影響を受けなくても合理性を持って状態が次々と変化していく性質」が必要と本研究室では考えているが、多段階の「状態遷移」を零からネットワークに学習させることは非常に困難である [6][7]。そのため、本研究室ではカオスダイナミクスによる不規則な状態遷移を強化学習における「探索」として利用し、これを学習によって合理的なものにすることによって「思考」のようなものを創発できるという

仮説を立てている。

しかし、Recurrent Neural Network(以下 RNN) の持つカオス性が強すぎると強化学習がうまくいかず、弱すぎると自律的な状態遷移を行うことができず探索できなくなるという問題がある。つまり、合理的な状態遷移の学習のためには、単に強化学習を行うだけでなく学習によって RNN 内部にカオスダイナミクスを生成し、学習もでき、かつ固定点や閉軌道でない(時間的に一定の状態に留まらない)状態遷移も続けることができる適切なカオス性の強さに調節することが必要であり、この適切なカオス性の強さは、収束と発散の境目である"Edge of Chaos"付近であると考えている [8][9]。また、このようにすることで、強化学習に用いる情報が収束発散どちらもせずネットワークにそのまま残りやすくなるという点でも有効であると考えている。

先行研究において、ネットワークを構成するローカルな個々のニューロンにおける指標である「感度」を、山登り法(最急上昇法)を用いて重み値を更新し大きくしていく感度調整学習により、RNN 内部にカオスを発生させられることが確認できている [10]。

本研究では『感度調整学習による個々のニューロンの感度の調整で、ネットワーク全体のカオス性の調整が行えるか?』の調査を主目的とする。そのため、層構造を持たないフラットな RNN に感度調整学習を適用し、ローカルな指標である感度とネットワークのグローバルなダイナミクスにおけるカオス性の指標である最大リアプノフ指数との関係を観察し、ニューロン数や結合割合によらずに Edge of Chaos の状態を実現できる可能性を調べた。

## 2 研究方法

本研究では、静的なニューロンを用いて層構造を持たないフラットなネットワークを構成した。外部入力やバイアスはなく、時刻  $t$  には前時刻  $t-1$  の出力が入力として入るフィードバックの構造を持つことでダイナミクスを持つ。また、活性化関数として  $\tanh$  を用いた。時刻  $t$  における  $N$  個中の  $j$  番目のニューロン出力  $o_{j,t}$  は前時刻の出力  $\mathbf{o}_{j,t-1} = [o_{1,t-1}, \dots, o_{j,t-1}, \dots, o_{N,t-1}]$  と重み値ベクトル  $\mathbf{w}_j$  を用いて以下のように示される。

$$u_{j,t} = \mathbf{w}_j \cdot \mathbf{o}_{t-1} \quad (1)$$

$$o_{j,t} = f(u_{j,t}) = \tanh(u_{j,t}) \quad (2)$$

ここで、 $u_{j,t}$  はニューロンの内部状態である。

ニューロンの現在の状態の近傍における収束発散を「感度」としており、これを学習における各ニューロンの評価関数  $V$  とした。

$$\begin{aligned} V_j &= |\nabla_{\mathbf{o}_{t-1}} o_{j,t}| = f'(u_{j,t}) \sqrt{\sum_i w_{j,i}^2} \\ &= f'(u_{j,t}) |\mathbf{w}_j| \end{aligned} \quad (3)$$

山登り法（最急上昇法）を用いて重み値の更新を行うことにより感度を調整する学習を感度調整学習としている。重み値の更新量を  $\Delta w$  とした時の更新量を式に示す。

$$\Delta \mathbf{w}_j = \eta \nabla_{\mathbf{w}_j} V_j \quad (4)$$

ここから、 $\tanh$  を用いて重み値の更新量  $\Delta \mathbf{w}_j$  は

$$\Delta \mathbf{w}_j = \eta (1 - o_j^2) \left( \frac{\mathbf{w}_j}{|\mathbf{w}_j|} - 2o_j |\mathbf{w}_j| \mathbf{o}_{t-1} \right) \quad (5)$$

と表される。

感度はニューロンのローカルな初期値鋭敏性を表しており、ネットワーク全体のグローバルな初期値鋭敏性を見るために以下のように求めた最大リアプノフ指数を用いる。調査時の各ニューロンの内部状態に、摂動として大きさが 0.001 の乱数を入れた状態と入れない状態の 2 状態を用意し、初期の 2 状態ユークリッド距離  $d_t$  を求める。T ステップ分フィードバックで入出力を循環させ、時刻毎に 2 状態ユークリッド距離  $d_{t+1}$  を求め以下の式で最大リアプノフ指数  $\lambda$  を算出する。

$$\lambda = \frac{1}{T} \sum_{t=1}^T \ln \frac{d_{t+1}}{d_t} \quad (6)$$

本研究では、 $T=15$  とした。ローカルな収束発散の指標である感度を大きくしていくことによって、グローバルな収束発散の指標である最大リアプノフ指数も大きくなっていくことに期待する。

## 3 シミュレーション

### 3.1 ネットワーク

本研究で用いた NN は、いずれの構造においてもニューロンの出力が次の時刻の入力として入ってくるリカレント構造を持つ。各ニューロンの結合部には、初期重み値を微小な一様乱数で与えており、この重み値を感度調整学習で更新することによってカオスダイナミクスを生成する。その他、本学習で使用した各パラメータを Table 1 に示す。内部ダイナミクスの変化のみを見るためにリカレント部以外の入力層や出力層は用いておらず、ニューロン数と結合割合を変えながら検証を行った。また、入力を加えない場合ネットワークの出力が全て 0 に収束し、重み値の更新を行っても変化が起きないため、100 ステップに 1 回の間隔で 100,000 ステップまでの間、トリガーとしての微小な乱数を内部状態に付与した。

Table 1: Parameters

名前	設定値
ニューロン数	100 or 50
結合割合 (%)	100 or 50
試行回数 (step)	200,000
重み値の学習率 $\eta$	0.00001
初期重み値の値域	[-0.1, 0.1]
トリガー (乱数) の値域	[-0.1, 0.1]

最大リアプノフ指数はネットワークに 1 つであるが、感度は個々のニューロンの入出力関係を表すため、ニューロンの個数分存在する。そこで全ニューロンの感度の平均を用いて観察を行った。さらに、式 (6) のように、リアプノフ指数は算出時に対数を取るため、比較の際には感度もリアプノフ指数に揃え、“全ニューロンの感度の平均値の自然対数をとったもの”を用いて比較した。本研究ではこれを“対数感度”とした。また、最大リアプノフ指数は 1,000 ステップに 1 回測定を行い対数感度と比較した。

### 3.2 ニューロン数 100 個・全結合の場合

まず、ニューロン数 100 個で全結合の場合の結果を示す。次の Fig. 1 にそれぞれ出力絶対値の最大値と平均値、Fig. 2 にあるニューロン 1 個に結合する重み値を 10 個取り出したもの、Fig. 3 に評価関数である感度の平均値、Fig. 4 に最大リアプノフ指数の学習による変化を示す。

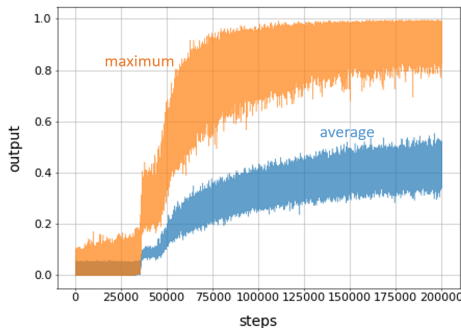


Fig. 1: Change of mean and maximum absolute outputs during learning

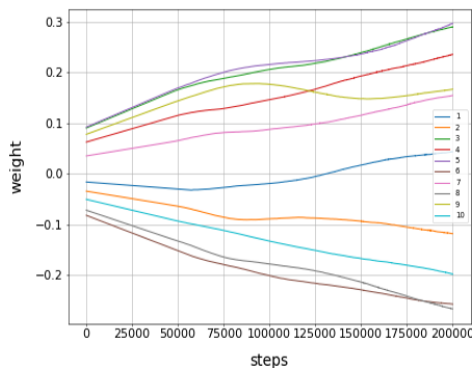


Fig. 2: Change of connection weights (sample 10 weights in one neuron)

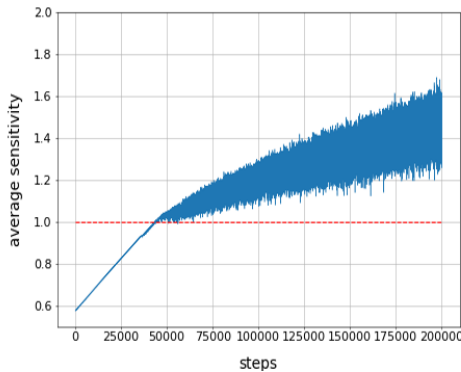


Fig. 3: Change of average sensitivity during learning

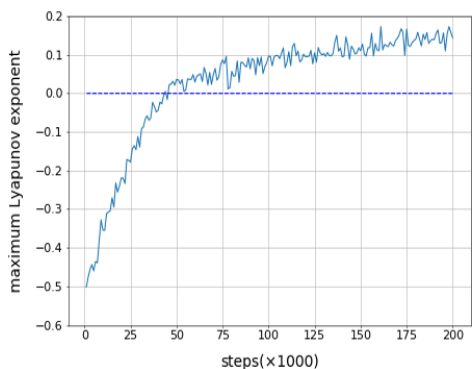


Fig. 4: Change of Lyapunov exponent during learning

Fig. 1 を見ると出力の最大値、平均値共に学習初期にはトリガーを加えてもニューロンの出力は収束しており、最大値は 30,000 ステップから 40,000 ステップ付近で急激に出力の最大値は増加し、平均値は 40,000 ステップから 200,000 ステップにかけて弧を描くように増加していく。感度調整学習では重み値を式 (5) により更新して感度を大きくしていく。この過程において重み値がどのように変化しているのかを 1 ニューロンから 1 つの重み値を取り出した例が Fig. 2 である。学習初期は一様乱数によって与えられた -0.1 から 0.1 までの重み値なのでその範囲に分布しているが、学習が進むに従ってそれぞれに分かれ大きくなっている。Fig. 3 の感度の平均値は出力が 0 の近くを離れ、高下するようになると式 (3) の  $f'$  の値が変化するため、それに伴った変化をするものの、全的に感度は増加している。Fig. 4 の最大リアプノフ指数も学習が進むに従って大きくなっている。

Fig. 5 に対数感度と最大リアプノフ指数の関係を示す。

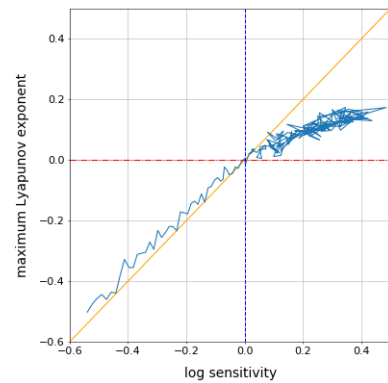
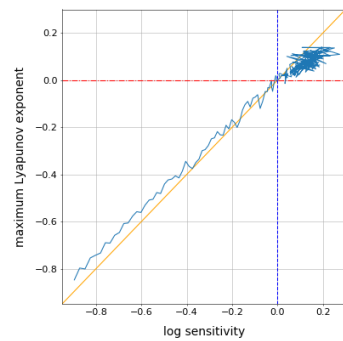


Fig. 5: Relation between the maximum Lyapunov exponent and log-sensitivity during learning (number of neurons:100, connection rate:100 %)

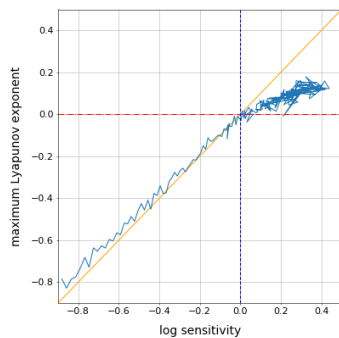
最大リアプノフ指数が 0 を超えるまでの非カオス領域を見ると、対数感度と最大リアプノフ指数は完全ではないものの数値的な一致が見られる。Edge of Chaos 付近においても最大リアプノフ指数と対数感度の一致が見られ、感度を用いた収束発散の境界と最大リアプノフが示す収束発散の境界が概ね一致することを示している。ただし、最大リアプノフ指数が 0 を超え、カオティックな領域に入ると対数感度と最大リアプノフ指数の対応は見られなくなった。

### 3.3 ニューロン数・結合割合を変えた場合

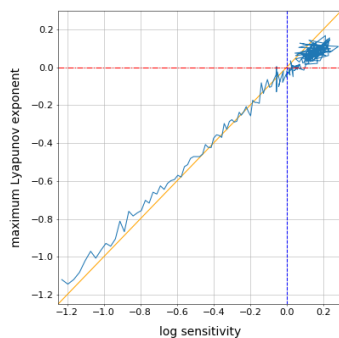
Fig. 6(a)~(c) にそれぞれ結合割合を 50 % に減らした場合、ニューロン数を半分の 50 個に減らした場合、その両方ともに減らした場合の対数感度と最大リアプノフ指数の関係を示す。結合割合やニューロン数を変化させたいずれの場合においても、非カオス領域である最大リアプノフ指数 0 以下では、対数感度と最大リアプノフ指数の関係は概ねの数値的な一致が見られた。



(a) number of neurons:100,  
connection rate:50 %



(b) number of neurons:50,  
connection rate:100 %



(c) number of neurons:50,  
connection rate:50 %

Fig. 6: Relation between the maximum Lyapunov exponent and log-sensitivity during learning (when the network architecture is varied)

## 4 結論

本研究では、感度というローカルなダイナミクスを最急上昇法を用いて増加させる感度調整学習を、単層の Recurrent Neural Network(以下 RNN) に対してニューロン数や結合割合を変更しつつ行い、グローバルなダイナミクスである最大リアプノフ指数がどのように変化するか検証した。そして、ニューロン数や結合割合を変化させたにもかかわらず、非カオス領域では対数感度と最大リアプノフ指数の数値的に概ねの一致が見られることが分かった。一方で、カオス領域ではその対応がなくなることも確認した。以上のことから、各ニューロンが入力ベクトルとその近傍との距離に対する出力が変化の最大値を一致させるように学習することによってニューラルネッ

ト全体のダイナミクスを Edge of Chaos の状態に持っていける可能性を示した。さらに、初期重み値を小さな乱数で設定し、感度調整学習を行い、感度が1になったところで学習を止めれば、ネットワークが Edge of Chaos の状態を容易に生成し、安定的に維持できる可能性を示唆している。

今後は、層構造の RNN の感度とリアプノフ指数がどのように対応するのか、確認していきたい。さらに、感度を見て学習を止めれば Edge of Chaos を始めとするカオス性の調整を行うことができるため、これと強化学習と組み合わせることがどの程度有効かも確認していきたい。

## 謝辞

本研究は JSPS 科研費 (15K00360, 20K11993) および 栢森情報科学技術振興財団研究助成金の補助を受けたものである。ここに謝意を表する。

## 参考文献

- [1] L. Jiao, F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, R. Qu, A survey of deep learning-based object detection, arXiv:1907.09408. (2019)
- [2] D. W. Otter, J. R. Medina, J. K. Kalita, A survey of the usages of deep learning in natural language processing, arXiv:1807.10854.(2019)
- [3] K. Shibata, Functions that Emerge through End-to-end Reinforcement Learning — The Direction for Artificial General Intelligence —, arXiv:1703.02239v2, RLDM 17 (2017)
- [4] V. Minh, K. Kavukcuoglu, D. Silver, et al. Human-level control through deep reinforcement learning, Nature, 518, pp. 529–533 (2015)
- [5] D. Silver, T. Hubert, et al., A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play, Science, Vol. 362, Issue 6419, pp. 1140–1144 (2018)
- [6] 田口優馬, 柴田克成: 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, SICE九州支部学術講演会, pp. 87-90 (2011)
- [7] 沢津橋由人, 柴田克成: リカレントネットを用いた強化学習における離散的かつ抽象的な状態表現の創発, 計測自動制御学会 システム・情報部門学術講演会 2012 講演論文集, 3B1-3.pdf, pp. 402-407 (2012)
- [8] 後藤祐樹, 柴田克成: カオスニューラルネットを用いた強化学習におけるカオス性の影響, 第36回計測自動制御学会九州支部学術講演会予稿集, pp. 97-100 (2017)
- [9] T. Matsuki and K. Shibata Adaptive balancing of exploration and exploitation around the edge of chaos in internal-chaos-based learning, Neural Networks, Vol. 132, pp. 19-29 (2020) (to appear)
- [10] 徳丸侑輝, 柴田克成: リカレントネットにおける感度調整学習でのカオスダイナミクスの生成と維持, 第38回計測自動制御学会九州支部学術講演会予稿集, pp. 75-78 (2019)