

リカレントネットによる状態遷移タスクの教師あり学習におけるカオス性の影響

大分大学 ○大石将人 松木俊貴 柴田克成

Effect of Chaos on Supervised Learning of State Transition Tasks Using Recurrent Neural Networks

○ Masato Ooishi, Toshitaka Matsuki, Katsunari Shibata (Oita University)

Abstract: We think that the emergence of rational multi-step state transition is critical to realize higher functions like 'thinking'. However, such learning was difficult for recurrent neural networks(RNNs) with only positive self-feedback that is good for memory-required tasks. On the other hand, in chaotic neural networks, state transitions autonomously though the dynamics is irregular. Then we think chaotic dynamics make the learning of multi-step state transitions easy. In this paper, we applied RNNs with random feedback connection weights to the learning of a multi-step state transition task varying the scale of the weight matrix. We show that learning performance is better when the RNN is at edge of chaos has or a weak chaoticity, while the performance is bad for the RNN with only positive self-feedback weight even though that makes the error propagation better.

1 序論

近年、人工知能が注目を集める中で、今後はさらに様々な状況で活動可能な、より人間に近い高次機能を持った人工知能が求められていくと予想される。そのような人工知能の実現には、人間が行う「思考」のようにエージェント自身が置かれている状況などを判断し、物事を順序立てて合理的に行動することが重要である。われわれは以前からリカレントニューラルネット (RNN) と強化学習を組み合わせることで記憶や予測といった機能の創発を簡単なタスクで確認 [1][2] しており、同様の枠組みで「思考」のような機能の形成を目指している。

頭の中で考えを巡らせる「思考」とは、脳内における合理的な多段階の状態遷移であるとわれわれは考えた。そしてそれを実現するためには、多段階の状態遷移を行うフロー型のアトラクタの形成が必要である [3]。しかし、フィードバック (FB) 結合重み値の小さい低活性な RNN では、時間経過とともに内部状態が減衰するダイナミクスとなってしまう、アトラクタの形成は容易ではない。

また、固定点アトラクタの形成を容易にし、かつ Back Propagation Through Time(BPTT) 法による誤差伝播の消失、発散を防ぐために、ニューロン自身への結合であるセルフフィードバック (selfFB) 結合の重み値を適度な正の値に設定することで、記憶が必要なタスクの学習が容易になる [1][2]。しかし、これでは状態遷移を伴うフロー型のアトラクタの形成は難しく、従来、学習できる状態遷移の数は少なかった [4]。

そこでわれわれはこのような RNN を学習させて状態遷移を内部に形成していくのではなく、元々不規則な状態遷移を行うカオスダイナミクスを RNN に導入したカオスニューラルネット (ChNN) を学習させることで、合

理的な状態遷移を作り出すことが容易になるのではないかと考えた。ChNN 内部のランダムに近い自律的な状態遷移を行うカオスダイナミクスが、学習によって固定点でないフロー型のアトラクタを形成することで、「思考」創発の手がかりが得られることを期待する。

本研究では ChNN を導入することにより、多段階の状態遷移を伴うフロー型のアトラクタが、学習によって形成が容易になるかを確かめるために、入力によって多段階に状態を遷移した後の状態を出力するタスクの教師あり学習を行い、カオス性の影響を調査する。

2 RNN による多段階状態遷移の教師あり学習

Fig. 1 に本研究における状態遷移タスクの概要図を示す。本研究で扱う多段階の状態遷移とは、Fig. 1 に示すようにエージェントが現在いる状態とその状態における入力 (= 行動) によって正しい出力 (= 状態) が異なるというものである。エージェントがこの状態遷移を行うには、エージェントが RNN の内部状態によって現在いる状態を把握し、入力による状態遷移を学習によって形成することが必要である。

RNN のある時刻 t における中間層と出力層の内部状態 \mathbf{u}_t^{hid} , \mathbf{u}_t^{out} および各層の出力 \mathbf{o}_t の計算式を示す。

$$\mathbf{u}_t^{hid} = \mathbf{W}^{FW} \mathbf{o}_t^{in} + \mathbf{W}^{FB} \mathbf{o}_{t-1}^{hid} \quad (1)$$

$$\mathbf{u}_t^{out} = \mathbf{W}^{out} \mathbf{o}_t^{hid} + \boldsymbol{\theta}_{out} \quad (2)$$

$$\mathbf{o}_t = \tanh(\mathbf{u}_t) \quad (3)$$

\mathbf{W}^{out} , \mathbf{W}^{FW} はそれぞれ出力、中間層の下層からの入力に対する重み値、 \mathbf{W}^{FB} は FB 結合の重み値を表す。 $\boldsymbol{\theta}_{out}$

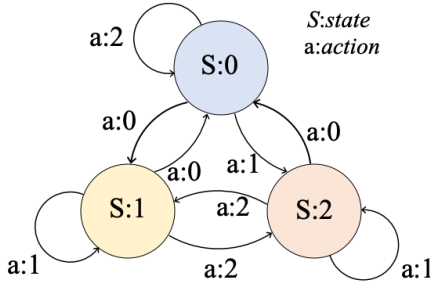


Fig. 1: State transition in the learning task.

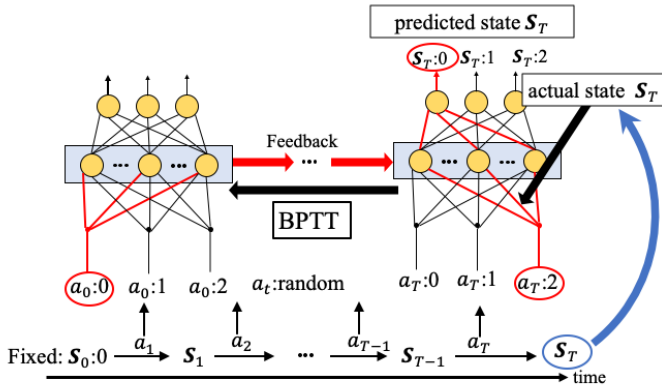


Fig. 2: Learning task using a recurrent neural network.

は出力層バイアスである。中間層の内部状態を計算する式 (1) においてランダムに決定する FB 結合 \mathbf{W}^{FB} のスケールを大きくすることで ChNN とすることができる。RNN の各種パラメータの更新は最終時刻 T に状態 S_{T+1} を教師信号として与え、時刻を遡って重み値を更新する BPTT 法を用いて教師あり学習を行う。Fig. 2 に RNN の入力と出力、教師信号のイメージを示す。Fig. 1 のように、各試行で時刻 $t = 0$ ではエージェントがいる状態 S_t を 0 に固定し、以後エージェントには行動 a_t を入力として与える。一方でエージェントがいる状態 S_t についての入力は与えない。これを毎時刻行い、各試行の最終時刻 T にのみ教師信号としてエージェントがいる実際の状態 S_T を与え、BPTT 法により学習を行う。学習できればエージェントは与えられる行動 a_t のみから自身が最終時刻にいる状態 S_T を出力できるようになる。そして RNN の FB 初期重み値スケールを変化させ、RNN の学習における勾配消失、発散の影響を考慮しながらカオス性による学習性能の比較を行う。

3 シミュレーション

Table 1 にシミュレーションで用いたパラメータを示す。出力層重み値 \mathbf{W}^{out} 、入力から中間層への重み値 \mathbf{W}^{FW} および出力層バイアス θ_{out} は Table 1 に示した値域の一

様乱数によって決定する。FB 結合 \mathbf{W}^{FB} はスペクトル半径が 1 となるランダムな正方行列と任意の FB 重み値スケール λ の積によって決定する。シミュレーションでは FB 結合重み値スケール λ を変えることで RNN のカオス性を変化させ、学習性能を比較する。

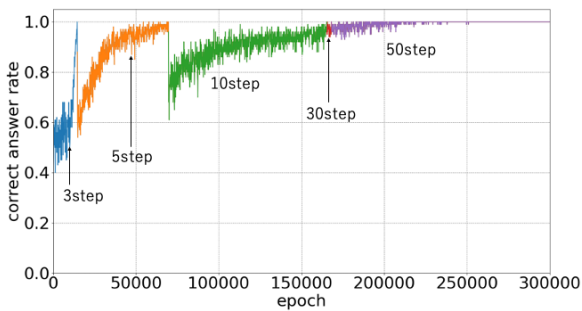
一度の学習は 300000 試行行われる。先述のように RNN による多段階の状態遷移の学習は難しいため、まず状態遷移の回数を 3 step で学習を行う。ある状態遷移数 T で 200 回連続で正答すると、その状態遷移数での学習が完了する。3 step の学習が完了すると 5 step, 10 step, 30 step, 50 step と、学習が完了する度に状態遷移の回数を増やしていく。

Fig. 3 に FB 結合重み値スケール $\lambda = 1.2$ で学習を行った場合の学習による正答率の変化、学習後の RNN の出力を示す。Fig. 3(a) より、学習が進むとともに正答率が上昇しており、遷移回数が増えて一度落ち込んでも再び上昇していることがわかる。Fig. 3(b) では学習を行っていない時刻 $t = 0$ 付近をのぞいて、ほとんどの時刻 t でエージェントが入力された行動に対して正しい状態 S_{t+1} を出力できていることがわかる。

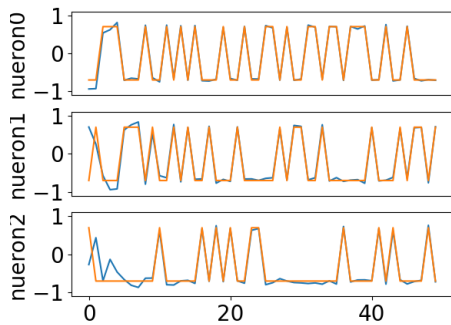
これらを踏まえて、FB 結合重み値スケール λ を変化した場合の学習性能を見ていく。Fig. 4 に FB 結合重み値スケール λ ごとの学習後のテストを行った際の正答率を示す。なお、正答率算出の際、300000 試行のうち 30 step までの学習が完了せず、50 step の学習を行っていないパラメータについては、正答率を 0 として扱った。Fig. 4 に FB 結合重み値に対する状態遷移タスクの学習性能の比較を示す。Fig. 4 より、FB 結合重み値スケール 0.2~0.8 の比較的小さい部分では FB 結合重み値スケール λ が大きくなるほど正答率が上昇しており、FB 結合重み値スケール λ が 1.0~1.8 で最も正答率が高くなっていることがわかる。また、FB 結合重み値スケール λ が

Table 1: Parameters for Simulation

Name		
Number of hidden neurons		100
Initial weight	hidden-output layer	[-0.2,0.2]
	inputs-hidden layer	[-0.25,0.25]
	feedback scale	varied
Learning rate	outputs bias	[-0.01, 0.01]
	output layer	0.01
	hidden layer	0.005
	feedback	0.00005
	outputs bias	0.05
Traceback times in BPTT		all time
Epoch of Learning		300000



(a) Learning curve showing the rate of correct answers.



(b) Outputs of RNN(blue) and signals representing actual state after training(orange).

Fig. 3: Learning results for state transition tasks. (feedback connect weights scale:1.2)

2.0 より大きい部分では FB 結合重み値スケール λ が大きくなるほど正答率が低下しており、学習性能が低いことがわかる。

カオス性が学習性能に与える影響を検証するため、RNN の FB 結合重み値スケール λ に対するリヤプノフ指数を Fig. 5 に示す。リヤプノフ指数はカオス性の指標であり、RNN の内部のカオス性を観察するために入力を与えずに 1000 step のシミュレーションで測定した。Fig. 5 より、FB 結合重み値スケール λ が 0.2~1.0 の比較的小さい部分ではリヤプノフ指数が負になっており、カオス性がないことがわかる。FB 結合重み値スケール λ が 1.0~1.5 ではリヤプノフ指数がほぼ 0 となっているが、この部分では RNN 内部のダイナミクスは収束も発散もしない、情報を保持するのに適したダイナミクスとなっている。FB 結合重み値スケール λ が 1.0 より大きくなるとリヤプノフ指数が徐々に大きくなっており、カオス性が強くなっていることがわかる。Fig. 4 と Fig. 5 を比較するとリヤプノフ指数がほぼ 0 となる FB 結合重み値スケール λ で正答率がよくなっている。このことから、収束も発散もしない edge of chaos の適度なカオス性が多段階の状態遷移の学習に適しているのではないかと考えられる。また edge of chaos となるような FB 結合重み値では誤差信号

の発散、消失も起きにくいいため、このことも学習性能が高いことに影響していると考えられる。

また、selfFB を設定することで BPTT 法における誤差の発散、収束が抑えられ、selfFB を設定しない場合と同様に学習が可能なのかを検証する。そのために selfFB を正の定数に設定して同様のタスクを行う。Fig. 6 に selfFB1.0 の学習による正答率の変化を、Fig. 7 に selfFB の値に対する状態遷移タスクの学習性能の比較を示す。

Fig. 6 より徐々に正答率は上がっており、学習は可能であることがわかる。しかし 10 step の学習が完了しておらず、最終的な 50 step の学習を開始できていないことがわかる。Fig. 7 については、ほとんどの場合で Fig. 6 と同様に 50 step の学習が開始できていないため、正答率が 0 となっており、Fig. 4 と比べても全体的に学習性能がかなり悪いことがわかる。このことから、selfFB を設定して誤差の伝播をよくすれば学習ができるわけではないことがわかる。selfFB を大きくしても RNN 内にカオスダイナミクスによる自発的な状態遷移は生じないため、カオス性の有無が Fig. 4 と Fig. 7 の学習性能に差が生じた原因ではないかと考えられる。このことから、状態遷移タスクでは誤差が消失、発散せず、edge of chaos や弱いカオスダイナミクスを生じる FB 初期重み値が適していると考えられる。

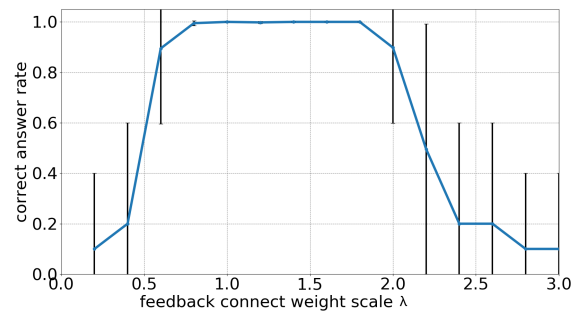


Fig. 4: Comparison of learning performance in the state transition task varying the feedback weight scale.

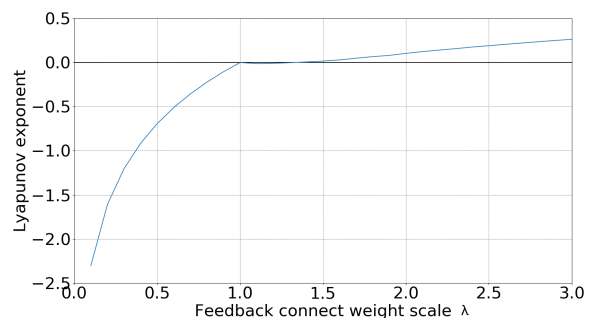


Fig. 5: Change of Lyapunov exponent according to the feedback weight scale.

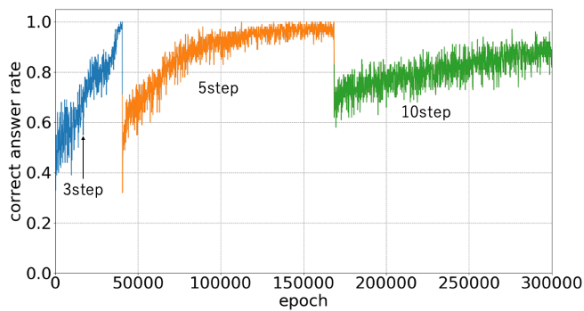


Fig. 6: Learning curve showing the rate of correct answers by learning. (Positive selfFB weight:1.0)

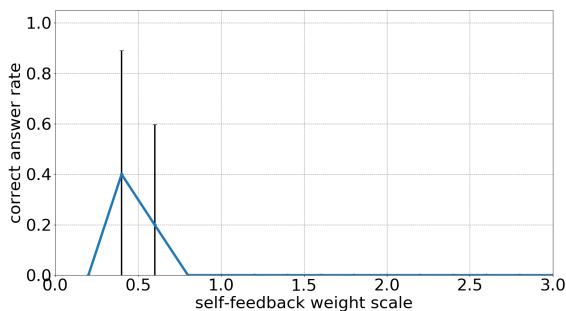
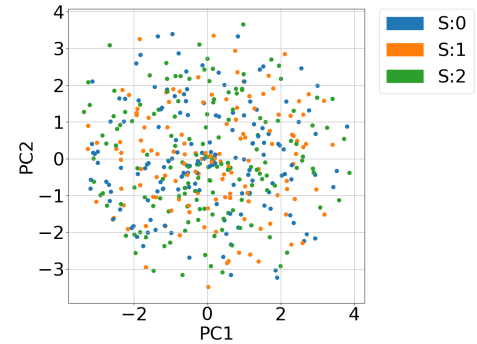


Fig. 7: Performance comparison of state transition learning with only positive self-feedback weights

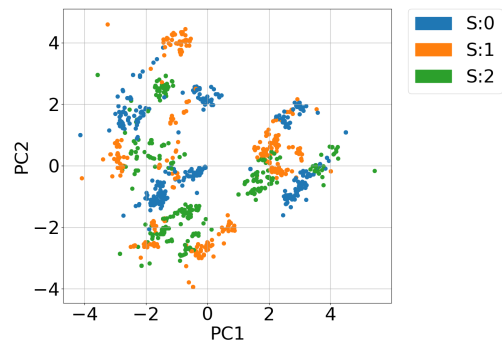
また状態遷移タスクを学習することで、ChNN が元々持つ不規則なカオスダイナミクスが学習によってどのように状態遷移するようになったのかを観察する。学習による RNN 内部の状態遷移の変化を見るために、Fig. 8 に FB 結合重み値スケール 1.2 の場合の RNN の中間層出力の主成分分析の結果を示す。学習前の内部状態である Fig. 8(a) をみると、3 状態が全体的に分布しており、ランダムに近い内部状態表現であることがわかる。一方、学習後の内部状態である Fig. 8(b) をみると 3 状態それぞれがある程度固まって分布しており、多段階の状態遷移の学習を行う過程で同一状態を近い表現で表すような内部状態表現が形成されていったと考えられる。

4 結論

本研究ではカオスによる不規則な状態遷移を持つ RNN が、不活性なものや固定点収束ダイナミクスの形成に有利な selfFB を正にする場合と比較して、状態遷移の学習に有利であることが示唆される結果が得られた。今後は入力がなくとも状態遷移を行うフロー型アトラクタの形成の確認、強化学習による多段階状態遷移の獲得などが課題としてあげられる。



(a) before learning



(b) after learning

Fig. 8: Difference of state representation between before and after learning (weight scale:1.2) PCA is applied to the hidden layer outputs

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。ここに謝意を表す。

参考文献

- [1] K. Shibata, H. Utsunomiya: Discovery of Pattern Meaning from Delayed Rewards by Reinforcement Learning with a Recurrent Neural Network, Proc. of IJCNN 2011, pp.1445-1452, N-0311. (2011)
- [2] K. Shibata, K. Goto : Emergence of Flexible Prediction-Based Discrete Decision Making and Continuous Motion Generation through Actor-Q-Learning, Proc. of ICDL-Epirob.ID 15 (2013)
- [3] 柴田克成: 「End-to-End 強化学習による知能創発と「思考」創発へ向けた新しい強化学習」第 62 回システム制御情報学会研究発表講演会講演論文集 (2018)
- [4] Y. Sawatsubashi, M. F. Samusudin, K. Shibata: Emergence of Discrete and Abstract State Representation through Reinforcement Learning in a Continuous Input Task, Proc. of RiTA 2012, M1C-2.pdf, pp.13-22 (2012)