

再帰結合神経回路モデルへのスパース構造導入による学習能力の向上

粟野 皓光[†] 尾形 哲也[†] 有江 浩明[‡] 谷 淳[‡] 高橋 徹[†] 奥乃 博[†]

[†] 京都大学大学院 情報学専攻 知能情報学専攻 [‡] 理化学研究所 脳科学総合研究センター

1. はじめに

本研究では再帰結合神経回路モデルにおけるスパース構造導入の効果について報告する。従来の神経回路モデルでは、ニューロン同士の結合に全結合を用いるのが一般的であり、スパース構造を導入した例は少なかった。一方、スパース構造の有効性は、連想記憶と関連して議論されている。Olshausen らがこれらの研究をまとめている [1]。これによると、初期の連想記憶モデルからスパース構造の有効性が議論されている [2]。これはパターン間の衝突を少なくすることにより記憶容量の向上に繋がるからである。近年では、再帰結合をもつネットワークに対して、その結合をスパースにした例が報告され始めている。例えば Echo State Network (ESN) ではリカレント結合部をスパース構造とすることで、保持しているダイナミクスに多様性を持たせカオス系列の予測学習に利用している [3]。

しかし ESN のようなスパース構造を持つ再帰結合神経回路において、リカレント部分の結合重みの学習は固定されているか Hebb 学習のように局所的に重みを更新する物であった [4]。我々は、ロボットの関節角度データ等、複雑な構造を持った時系列の予測学習に神経回路モデルを利用している。そのために、内部の結合重みも学習によって獲得する必要があると考えている。そこで、本研究では、誤差逆伝播法を用いて結合重みを学習するネットワークに対して、結合の一部を取り除くことにより、スパース構造を導入し、その影響を調査した。

2. 神経回路モデル

今回は、スパース構造を導入する対象として、Multiple Timescales Recurrent Neural Network (MTRNN) [5] を選択した。

2.1 MTRNN

MTRNN は Jordan 型 RNN の一種で、現状態 $IO(t)$ を入力すると次状態 $IO(t+1)$ を出力する予測器として動作する。コンテキスト層は C_F , C_S , PC_S に分けられる。それぞれのニューロングループは $IO(t)$ から C_S , PC_S と C_S , PC_S から $IO(t)$ を除いて、相互に結合している。

MTRNN を構成するニューロン群は異なる時定数を持っている。 IO が一番小さい時定数を持っており、 C_F , C_S , PC_S の順で大きくなっていく。時定数の違いによりニューロンの発火速度に差が出る。つまり、 IO は入力に対して内部値が変化する速度が速く、 C_S や PC_S は遅い。従って C_F が入力時系列の中で変化の速いプリミティブを自己組織化し、 C_S がプリミティブのシーケンスを自己組織化する。つまり MTRNN は入力時系列を階層構造に分解して学習することができる。このような特徴により、従来の RNNPB モデルに比べて、より多くの時系列を学習することが可能となっている。

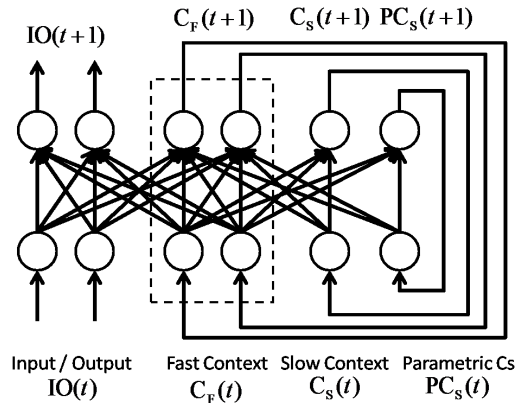


図 1: MTRNN の構造

MTRNN の学習には Back Propagation Through Time (BPTT) [6] と呼ばれる前進計算と重み更新に基づく手法が利用される。

2.2 スパース構造の導入

MTRNN の一部結合重みを 0 固定とすることでスパース構造を導入した。今回は図 1 の点線で囲われた部分の結合をスパース構造とした。これは C_F の内部結合に相当する部分である。 C_F はプリミティブを自己組織化するとされている。一般的にプリミティブ同士の相関は小さいと考えられるため、スパース構造は適していると考えられる。一方、 C_S はプリミティブの並びをコーディングしており、こちらをスパース構造にすると汎化性能に悪影響が出ると考えたため、 C_S は全結合としている。重みを 0 に固定する結合は、結合割合に応じてランダムで選択した。例えば、結合割合を 20% とするときには、対象となる結合本数のうち、結合重みを 0 に固定する 80% をランダムに選択し、残りの 20% に初期重みを割り当てた。

3. 評価実験

我々は評価のためにアルファベット列からなるデータを用いることにした。具体的には、英単語を正規文法に従って配置し、文章を作った。この文章から、アルファベットの出現位置でピークを持つような時系列に変換し、データを生成した。入力ノードはアルファベットの各 1 文字と対応しており、対応するアルファベットが出現した時点で発火している。今回は 100 文を用意し、80 文を学習データ、20 文を評価データとした。なお、評価に用いた MTRNN は、入力 30 ノード、Fast-Context 200 ノード、Slow-Context 6 ノード、Parametric-Slow 6 ノードである。評価した項目は以下の通りである。

汎化性能

未学習データを使用し、そのデータを再現できるか調査する。まず、未学習データを入力し、誤差信号から適

Learning Performance Improvement of Recurrent Neural Network by Sparse Structures: Hiromitsu Awano (Kyoto Univ.), Hiroaki Arie (RIKEN), Jun Tani (RIKEN), Tetsuya Ogata (Kyoto Univ.), Toru Takahashi (Kyoto Univ.), and Hiroshi G. Okuno (Kyoto Univ.)

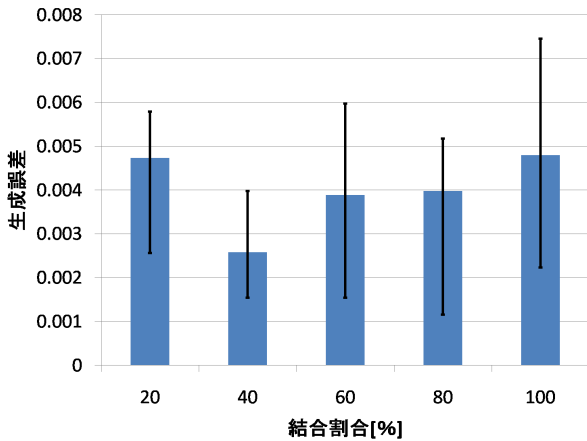


図 2: 汎化性能の変化

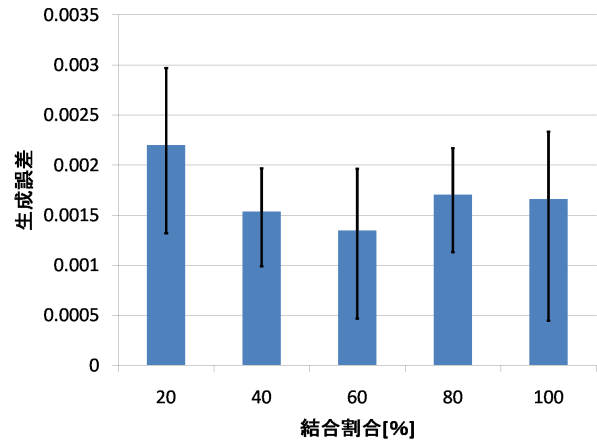


図 3: 引き込み性能の変化

切なパラメータ値を計算する．その後，計算されたパラメータのみから連想出力を行い，再現率を計算する．

ノイズに対する頑健性

既学習データにノイズを加え，ノイズの入っていない物を再現できるか調査する．まず，既学習データから1文字入れ替えることによってノイズを混入し，このデータを用いて誤差信号からパラメータ値を計算する．その後，連想出力を行い，ノイズの入っていないデータとの誤差を調べる．これはアトラクタの頑健性を調べていることに対応する．

4. 結果

実験では，結合割合を 20% から 100% まで 20% ずつ変化させ，汎化性能・ノイズに対する頑健性を調べた．MTRNN の結合重みは，はじめ乱数によって初期化されるが，乱数シードによって学習が収束する値が異なる．そのため，実験では各結合割合に対して 20 種類の乱数シードを用いて学習を行った．

汎化性能に関しては，未学習データ 20 パターンを用いて認識を行い，生成誤差の比較を行った．各結合率における生成誤差の平均と最大値・最小値を図 2 に示す．汎化性能の比較では，結合割合を 40% としたものが一番性能が良くなった．逆に，これより結合割合を下げてしまうと，汎化性能が悪くなる傾向にあることが分かった．

意図的にノイズを加えたデータに関しても，MTRNN で認識した後に，生成を行い，クリーンなデータと比較した．結果を図 3 に示す．横軸は結合割合，縦軸はノイズの入っていない元データとの誤差である．ノイズを混合したシーケンスの生成誤差に関しては，60% 結合の物が一番誤差が小さくなった．汎化性能のときよりも，高い結合率で誤差が小さくなるという結果になった．

5. 考察

結合割合を下げていくと，学習しているシーケンスの重なりが少なくなるため，シーケンスの構造をとらえた汎化が困難になることが予想され，汎化性能の結果もそれに沿ったものとなっている．逆に，結合率を上げていくと重み空間が広くなり，学習が局所解に収束してしまう例が増えてくる．例えば，全結合の生成誤差では，誤差が最大の物と最小の物でかなりばらつきがあるが，結

合率を 40% としたものは，ばらつきが少なくなっている．従って，結合率を適度に制限することで，局所解に陥ってしまうことを回避することが可能となる．

6. おわりに

本研究では MTRNN にスパース結合を導入し，結合率によって性能がどのように変化するかをアルファベット列からなるデータを用いて調べた．その結果，汎化性能においては 40% の結合率，ノイズデータの訂正能力では 60% の結合率で平均誤差が最も小さくなることが確認された．

残された課題としては，獲得された構造の解析と追加学習である．結合をスパースにしたことで，MTRNN 内部でのシーケンスがどのような構造で表現されているかを調べる必要がある．また，スパース構造を導入したことで，追加学習に対するロバスト性が向上していると期待できる．全結合の物では一旦獲得された構造の上に新しいデータで学習を行うと，初めの構造を壊してしまう．スパース結合では，新しく追加したデータに対する重みの修正が，局所的な修正で済むため，既存の構造を壊しにくいのではないかと考えられ，今後実験によって確かめる必要がある．

謝辞

本研究の一部は，JST さきがけ，科研費基盤 (B)，科研費学術創成，理研の支援を受けた．

参考文献

- [1] B.A. Olshausen and D.J. Field. Sparse coding of sensory inputs. *Current opinion in neurobiology*, Vol. 14, No. 4, pp. 481–487, 2004.
- [2] P. Kanerva and Research Institute for Advanced Computer Science (US). *Sparse distributed memory and related models*. Citeseer, 1992.
- [3] H. Jaeger and H. Haas. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science*, Vol. 304, No. 5667, p. 78, 2004.
- [4] Lazar Andreea, Pipa Gordon, and Triesch Jochen. SORN: a self-organizing recurrent neural network. *Frontiers in Computational Neuroscience*, 2009.
- [5] Yuichi Yamashita and Jun Tani. Emergence of functional hierarchy in a multiple timescale neural network model: A humanoid robot experiment. *PLoS Comput Biol*, Vol. 4, No. 11, p. e1000220, 11 2008.
- [6] G.E.H. De Rumelhart and RJ Williams. Learning internal representations by error propagation. *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. 1, p. 319.