

令和 4 年 6 月 2 日現在

機関番号：11501

研究種目：基盤研究(C)（一般）

研究期間：2019～2021

課題番号：19K11817

研究課題名（和文）パラメータの定まったニューラルネットワークの性質を調べるアルゴリズム基盤の構築

研究課題名（英文）Foundation of algorithm designs for artificial neural networks

研究代表者

内沢 啓 (Uchizawa, Kei)

山形大学・大学院理工学研究科・准教授

研究者番号：90510248

交付決定額（研究期間全体）：（直接経費） 2,800,000円

研究成果の概要（和文）：機械学習によってパラメータが定まったニューラルネットワークが、予め定められた特定の性質を持つかどうかを計算機を用いて判定する情報処理タスクが、一般に容易に計算できるか、あるいは計算困難となるかについて調査した。様々な性質を対象としてタスクを設定し解析を行った結果、判定したい性質の種類や、段数などのニューラルネットワークの構造によって、タスクの難しさが幅広く変化することを示した。特に、段数が1段のニューラルネットワークと2段のニューラルネットワークについては、その乖離が大きくなることを明らかにした。

研究成果の学術的意義や社会的意義

本研究で得られた成果により、パラメータの定まったニューラルネットワークの性質を問う判定問題は、入力として与えられるニューラルネットワークの構造の違いや、判定問題として問う性質の違いによって、多項式時間で解ける場合から、現実的な時間では解けないと考えられるほど難しい問題となる場合まで、非常に幅広く変化することを明らかにすることができた。特に、段数の大きいニューラルネットワークが深層学習の分野で高い能力を示す一方で、段数の大きな学習済みのニューラルネットワークから情報を取り出すタスクが計算困難になりやすいことを、理論的に明らかにすることができた。

研究成果の概要（英文）：We consider computational tasks of deciding if a given neural network possesses various predefined mathematical properties, and investigate how many computational resources are required to compute them. We then show that there exists a property for which it can be computationally very hard to check even if a given neural network is extremely simple (i.e., a neural network is of a single neuron). We also show that another property is computationally hard to check when a given neural network has two layers, while the property is easy to check (solvable in polynomial time) when a given neural network consists of a single neuron. Our results theoretically confirm that extracting information from multi-layer neural network can be computationally very hard.

研究分野：計算量理論

キーワード：計算の複雑さ ニューラルネットワーク しきい値回路

## 1. 研究開始当初の背景

神経細胞の理論モデルであるニューロンを基本構成要素とする情報処理システムをニューラルネットワークと呼ぶ。ニューラルネットワークは、ネットワークを構成するニューロンの個数、およびニューロン間の結合の強さを表す重みなどのパラメータを適切に設定することで、様々なタスクを効率よく処理できることが知られている。しかし、こうしたパラメータを手動で厳密に設定する作業は一般に困難であり、非自明な設計のアイデアが求められる場合が多かった。近年になり、深層学習と呼ばれる機械学習の手法により、目的の情報処理タスクに対して卓越した性能を持つニューラルネットワークのパラメータを自動的に定める手法が発達してきた。機械学習によってパラメータの定まったニューラルネットワークは、その性能により実用的な運用が実際に行われているが、手作業で構成されたニューラルネットワークのように、背後にあるアルゴリズムが明示的に提示されることはなく、そのニューラルネットワークが目的の情報処理タスクをどのように処理しているかは、深層学習を運用している技術者にとっても不透明である。そのため、機械学習によって得られた学習済みのニューラルネットワークが、どのような判断基準を内部で構築しているかという情報を取り出す技術が求められていた。

一方で充足可能性判定問題は、与えられた論理回路の出力を真とする入力が存在するかを問う判定問題であり、理論計算機科学分野の発展の中で中心的な役割を果たしている。入力としてどのような論理回路が与えられるかによって、充足可能性判定問題の計算複雑さは大きく異なることが知られている。機械学習によって構築されたニューラルネットワークが、どのような判断基準を内部で構築しているか理解するという上記の観点から、この充足可能性判定問題を捉えなおすと、学習によって各種のパラメータが定まったニューラルネットワークにおいて、ある特定の出力を誘導する入力は存在するかを問う判定問題と解釈できる。

## 2. 研究の目的

本研究では、上記の観点に基づいて、充足可能性判定問題に限らず、パラメータが定まったニューラルネットワークを入力として与え、そのニューラルネットワークがある特定の性質を持つかを問う判定問題の計算複雑さを調査することにより、パラメータの定まったニューラルネットワークに対するアルゴリズム開発の基盤を構築することを目指す。

## 3. 研究の方法

入力として与えるニューラルネットワークは、理論計算機科学の分野で古くから研究がなされているしきい値回路、および離散力学系のネットワークを採用する。しきい値回路とは、線形しきい値関数を計算することができるしきい値素子を基本構成要素とする論理回路である。また離散力学系のネットワークとは、しきい値素子などを基本構成要素としつつも、ネットワーク内にループ構造を許容し、さらに計算が離散時間に沿ってニューロンの発火パターンが変動するモデルである。

さらに入力として与えられたニューラルネットワークに対して問う性質は、既存の研究が存在する(1) 回路の不変性に基づく性質(文献[4])、(2) ニューラルネットワークの発火パターンに基づく性質(文献[1])、および(3) グラフの簡略表現に基づいた性質(文献[2, 3])の3つを考えた。

## 4. 研究成果

まず初年度の結果として、パラメータの定まったしきい値回路が入力として与えられ、その回

問うグラフの性質	既知の結果			本研究の結果	
	CIR[3]	CNF[2]	DNF[2]	THR	THR $\circ$ THR
辺は存在するか	NP 完全	NP 完全	P	P	NP 完全
最大次数が $k$ 以上か	NP 完全	-	-	NP 完全	NP 完全
最小次数が $k$ 以下か	$_2$ 完全	-	-	P	NP 完全
2 点間パスはあるか	NP 完全	-	-	P	$_2$ 完全
支配集合問題	NEXP 完全	NEXP 完全	PP 完全	P	NEXP 完全

表 1 : グラフの簡略表現に関する計算複雑さの結果 . ただし , CIR は一般の論理回路 , THR は 1 段のしきい値回路 , THR $\circ$ THR は 2 段のしきい値回路を表す .

路が特定の入力ベクトルの並び替えに対して出力の不変性を持つかを問う判定問題が、素子 1 個からなるしきい値回路であったとしても、多対一還元の下で NP 困難となることを示した。すなわちこの結果は、仮にしきい値回路を構成する素子の数が定数個であったとしても、入力ベクトルに含まれる 1 の個数に対する不変性についての情報を回路から取り出すことは計算困難となる場合があることを示唆している。一方で、入力ベクトルを単純に回転させる操作を適用した際の不変性について問う場合は、素子 1 つのしきい値回路については多項式時間で解けることを明らかにした。これらの結果から、しきい値回路に対して問う不変性の種類によって、計算複雑さが変わりうることが明らかになった。

さらに 2 年目の結果として、入力として離散力学系のニューラルネットワークが与えられたとき、そのニューラルネットワークに特定の発火パターンから別の発火パターンへ移行可能かどうかを問う判定問題は、ネットワークの基本構成要素となるニューロンの検知できる処理が、周辺に発火しているニューロンが存在するか否か、あるいは全てのニューロンが発火しているか否かである場合には、PSPACE 完全になる一方で、隣接するニューロンの発火状態をコピーすることだけができる場合は、多項式時間で解けることを示した。また、ある発火パターンから始まり、再度そのパターンに戻ってくるまでに、どの程度の離散時間が経過するかを問う判定問題についても、一般には PSPACE 完全になる一方で、周辺に発火しているニューロンが存在するかのみを個々のニューロンが検知できる場合には NP 完全となることを示した。

これらの結果に加え、さらに機械学習の媒体として非常によく用いられるスパース性を担保したしきい値の表現能力について、段数やニューロンの個数に制限があると、非常に単純な情報処理タスクであっても表現することができなくなることを示した。

3 年目の結果として、入力としてしきい値回路が与えられたとき、そのしきい値回路が簡略表現しているグラフが、様々なグラフの性質を持つかどうかを問う判定問題について研究を行った。その結果、同じグラフの性質を問う判定問題を対象とした場合でも、入力として 1 段のしきい値回路 (すなわち、1 つのしきい値素子) を与えた場合は多項式時間でその性質の有無を判定できるアルゴリズムが設計できるのに対し、入力として 2 段のしきい値回路を与えた場合はその決定問題の難しさが NP 完全、あるいは NEXP 完全などになることを明らかにした (表 1)。さらに、この結果に付随して、入力として与えられる計算モデルによって、グラフの簡略表現に基づいた同じ決定問題の複雑さが異なるならば、その違いが計算モデルの計算能力の違いにつな

がることを示す数学的な命題を得ることができた。この知見は回路計算量理論の成果として非常に興味深い。

本研究で得られた成果により、パラメータの定まったニューラルネットワークの性質を問う判定問題は、入力として与えられるニューラルネットワークの構造の違いや、判定問題として問う性質の違いによって、多項式時間で解ける場合から、NEXP に所属する問題となる場合まで、非常に幅広く変化することを明らかにすることができた。特に最終年度の結果により、段数の大きいニューラルネットワークが深層学習の分野で高い能力を示す一方で、段数の大きな学習済みのニューラルネットワークから情報を取り出すタスクが、理論的にも計算困難になりやすいことが明らかになった。

#### <引用文献>

- [1] C. L. Barrett, H. S. Mortveit, C. M. Reidys. Elements of a theory of simulation II: Sequential dynamical systems. *Applied Mathematics and Computation*, 107(2-3):121-136, 2000
- [2] Das B. Scharpfenecker, P. Torán, J.: CNF and DNF succinct graph encodings. *Information and Computation* 253, 436–447 (2017)
- [3] H. Galperin and A. Wigderson. Succinct representations of graphs, *Information and Control*, 56(3), 183 – 198, 1983.
- [4] D. Ensign , S. Nevill , A. Paul and S. Venkatasubramaian , ``The Complexity of Explaining Neural Networks Through (group) Invariants,`` *Proceedings of Machine Learning Research* 76:1-19 , 2017.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 1件 / うちオープンアクセス 2件）

1. 著者名 Mitsunori Ogihara, Kei Uchizawa	4. 巻 170
2. 論文標題 Synchronous Boolean Finite Dynamical Systems on Directed Graphs over XOR Functions	5. 発行年 2020年
3. 雑誌名 Proceedings of 45th International Symposium on Mathematical Foundations of Computer Science	6. 最初と最後の頁 76:1-76:13
掲載論文のDOI（デジタルオブジェクト識別子） 10.4230/LIPIcs.MFCS.2020.76	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

1. 著者名 Kei Uchizawa	4. 巻 181
2. 論文標題 Size, Depth and Energy of Threshold Circuits Computing Parity Function	5. 発行年 2020年
3. 雑誌名 Proceedings of 31st International Symposium on Algorithms and Computation	6. 最初と最後の頁 54:1-54:13
掲載論文のDOI（デジタルオブジェクト識別子） 10.4230/LIPIcs.ISAAC.2020.54	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------