

令和元年6月10日現在

機関番号：32665

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K16121

研究課題名(和文) 学習する対象から情報収集するアルゴリズムの構築とその応用

研究課題名(英文) Development and application of algorithms for information collection from learning targets

研究代表者

早川 隆 (HAYAKAWA, Takashi)

日本大学・医学部・助教

研究者番号：30756789

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：研究代表者は自身の先行研究の結果を拡張し、学習対象が学習し変化する場合にもその対象から効率的に情報収集するための学習理論・計算機アルゴリズムを導いた。導かれたアルゴリズムの性能を、理論的な解析や計算機シミュレーションによって評価し、情報収集の効用について新たな知見を得た。また、学習時系列を記憶することが性能の鍵となることを見出し、情報収集に使う人工神経回路の動作についての新しい理論を導き、効率的に時系列を記憶する回路を設計できた。以上のように、学習する対象から人工神経回路等を用いて情報収集をするための理論上・応用上の基盤となる結果を得た。

研究成果の学術的意義や社会的意義

我々人間が日常的に経験しているように、どのような種類の学習であっても、学習対象について情報収集することが効率的な学習の鍵となる。学習対象自体が知能を持たず変化しない場合については、コンピューターが自動的に情報収集するためのアルゴリズムが研究されてきたが、学習対象自体が知能を持ち学習し変化する場合についてはあまり研究されてこなかった。研究代表者は後者の場合にも適用できるコンピューターアルゴリズムを導き、その有用性を示し、応用面での技術的なブレイクスルーを得た。この成果を基盤として、人工知能の様々な学習機能が加速され、絶えず変化する環境に適応できるものになっていくことが期待される。

研究成果の概要(英文)：In this project, the researcher developed a new algorithm that makes a learner actively collect information about the target of learning and that is applicable even if the target itself learns and changes its strategy. The performance of the new algorithm was tested by theoretically analyzing the behaviour of the algorithm or by implementing the algorithm with neural networks in computer simulations. From these analyses, novel utility of information collection was discovered. In the simulations, the performance of the algorithm turned out to largely depend on the ability of the neural network to remember the time series of events in learning. Then, the researcher developed a new theory about the behaviour of neural networks and found that the theory allows us to design neural networks that efficiently remember time series. These results laid theoretical and technical foundations for information collection from learning targets.

研究分野：ニューラルネットワーク

キーワード：情報量最大化 ニューラルネットワーク 統計力学 強化学習

1. 研究開始当初の背景

コンピューターが仮想的な人工神経回路を用いて学習し、様々な実社会の問題において性能を発揮し注目を集めていた。そのようなコンピューターの学習において、コンピューターが具体的に解きたい問題とは別に情報を収集するための方法が知られており、情報量最大化学習と呼ばれていた[1,2]。この方法は学習対象から入力を受動的に得られる場合においてのみ定式化されていたが(図1(a))、研究代表者は自身の先行研究において、コンピューターが入力を得つつ学習対象に対してフィードバックを与えるような双方向性の相互作用をする場合に情報量最大化学習を拡張した(図1(b)) [3]。これによって、コンピューターが情報を能動的に収集する場合に、どのようなアルゴリズムで行えば情報理論的に最適かが明らかになった。自律的なコンピューターシステムが何らかのスカラ値の報酬を最大限獲得しようとする学習を強化学習と呼ぶが、研究代表者が導いたアルゴリズムを強化学習を行うシステムに組み込み情報をたくさん収集させると、報酬を集めるのも速くなることが示せた[3]。



2. 研究の目的

「研究の背景」で説明したように、研究代表者は先行研究でコンピューターが情報を能動的に収集するための情報論的な学習理論・アルゴリズムを構築した。この理論は学習対象がそれ自体学習し変化していく場合(図1(c))に拡張できると予想された。コンピューターが通常の情報収集に加えて、学習対象が何を学習しようとしているかについて情報収集し、学習に役立つような理論ができると予想された。実社会の問題では、学習対象の背後に人がいて、その人自身が学習をしているケースは非常に多くあり、そのような問題に対して理論を拡張・適用できれば幅広い応用があると考えられた。具体的には以下の2つの例に最終的に応用することを想定し、そのための基礎理論と基盤技術を構築することを目的とした。

(1) 脳コンピューター・インターフェース(BCI)

BCIとは、患者の脳活動(脳波など)からインターフェースとなるシステムを介して直接義手などを動かそうとするものである[4]。この時、患者が自分の脳活動を調節してうまく動かすのは往々にして難しい。インターフェースとなるシステムが、患者が何を学習しようとしているのかについての情報を収集し、意図を汲み取ることができれば、学習を促進できると予想される。

(2) 社会ジレンマ・ゲーム理論

相手の意図を読み合ったり裏切ったりする状況はゲーム理論の分野で研究されている。特に人間関係の作るネットワーク上で協力または裏切りを選択するゲームモデルが注目されていた[5]。この分野では裏切りによる社会ジレンマを解消して協力関係を促進することが目標となっていたが、人々が政策などのゲームパラメータの変化に対してどのように適応・学習するかについて情報を収集できれば、ジレンマを解消しやすくなると予想された。

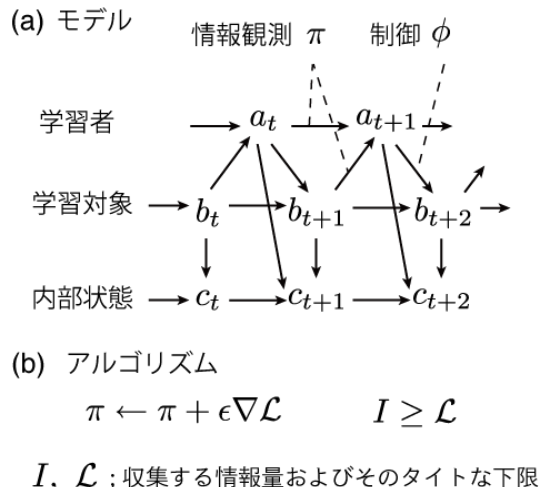
3. 研究の方法

(1) 確率モデルおよび理論の導出

研究代表者は先行研究において、神経回路が収集する情報量のタイトな下限が神経回路の内部変数のみを用いて表せ、その下限に関する最適化アルゴリズムが情報収集アルゴリズムを与えることを数学的に示した。この結果を改良して、「研究の目的」で説明したような学習する対象からの情報収集のための学習理論および学習アルゴリズムを導いた。

まず、先行研究ではマルコフ決定過程と呼ばれる確率過程を仮定して理論を導出した。しかし学習する対象を扱うためには、学習対象の内部状態が全部観測できない場合を考える必要があったので、部分観測マルコフ決定過程と呼ばれる確率過程およびその改変版を用いた(図2(a))。この一般的な確率過程に対して、収集する情報量のタイトな下限となる量を導き、この量の最適化アルゴリズムを抽象的モデルのレベルで導いた(図2(b))。

図2



具体的なモデルに関してはいくつかの単純で基本的な場合を、一般化されたランジュバン方程式などの比較的簡単な確率過程を用いて定式化し、近似的に理論解析を行った。またより非線形のシステムを用いて高精度の学習を可能にする目的で、問題設定に合った内部構造を持った確率的リカレント神経回路を用いて定式化し、アルゴリズムの挙動を数値的に調べた。この場合のリカレント神経回路システムは、対象をモデル化するための部分神経回路・記憶効果を持つリカレント部分神経回路・観測情報の符号化/対象へのアクションを行う部分神経回路など複数の部分回路から構成される。学習対象の部分は何かの観測不可能な目的関数に関する確率的勾配アルゴリズムによって変化する確率過程として表した。また、先行研究では離散の状態空間のみを考えていたが、問題の要請から先行研究の結果を連続状態空間の場合に拡張した。

(2)リカレント神経回路を用いた時系列の記憶・処理に関する理論

「研究成果」にて後述するが、研究をすすめるうちに、リカレント神経回路の記憶性能がアルゴリズム全体の性能のキーファクターとなることに気付いた。この点においてアルゴリズムの性能を改良するために、レザボワ計算と呼ばれる枠組みを組み込んだ。レザボワ計算とは、リカレント神経回路の全てのパラメータを最適化するのは困難で非効率だから、回路のほとんどの結合パラメータをランダムにとる方法である。しかし従来知られているレザボワ計算の枠組みは、そのまま組み込むことは困難であった。この点を克服するためにリカレント神経回路に動的な均衡が生じカオス性が抑制されるような構造を入れた。そして、回路の挙動を数値シミュレーションと後述するように研究代表者自身が導いた新しい理論を用いて解析した。

(3)具体的な応用とその性能評価のための計算機シミュレーションプログラムの開発

アルゴリズムの具体的な応用例における性能評価のために、ある程度の規模の計算機シミュレーションを行うプログラムの開発を行った。神経回路を用いた情報収集のアルゴリズムを実問題のコンピューターシミュレーションにおいて適用しようとすると、かなりの計算量が要求されることがわかった。このため、Intel Xeon プロセッサ上で 16-24CPU コアを用いて、BCI や社会ネットワーク上のゲームのシミュレーションを並行的に実行し、学習アルゴリズムも並列的に実行するようなプログラムコードを開発した。

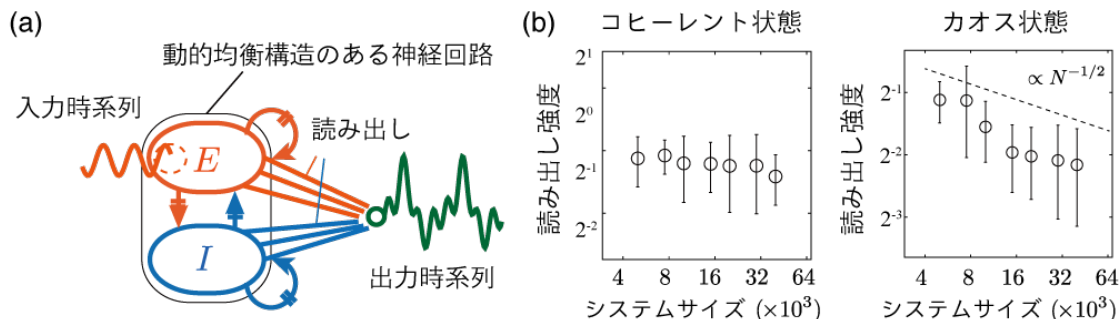
4. 研究成果

(1)情報収集の理論およびアルゴリズムの導出

「研究の方法」に述べた要領で、具体的に理論およびアルゴリズムを導いた。先行研究と同様に、学習対象の学習傾向を含めた情報量の下限となる量が構成できた。理論的にはこのアルゴリズムを適用すれば、学習対象の振る舞いについての情報が得られるが、その学習への効用は具体的な問題によって異なると予測された。例えば BCI のように学習対象の学習を学習者が補助するような場合は、学習対象が一種の局所解に陥っている際に、情報収集アルゴリズムはそれを抜け出させやすくする効用があると理論から予測された。一方、ゲーム理論のように学習対象と学習者の利益が拮抗する場合は、情報収集は必ずしもポジティブな効果ばかりでないことも理論から予測された。アルゴリズムのこれらの振る舞いは特別な問題設定に限ると解析的に示すことができた。一方、数値シミュレーションにおいて情報収集アルゴリズムの性能は、多分にモデルの設定の詳細に依存することが明らかになった。小規模な問題設定で見られた顕著な傾向は、リカレント神経回路部分の持つ記憶性能が全体の性能に最も反映されるというものであった。これらの結果の出版準備を行った。

(2)ランダムリカレント神経回路を用いた時系列学習とそのための理論

図 3



上記のように、研究をすすめるうちに、応用上はリカレント神経回路の記憶性能がアルゴリズム全体の性能のキーファクターとなることに気付いた。「研究の方法」で述べたように、アルゴリズムの性能を改良するために、レザボワ計算と呼ばれる枠組みを組み込んだ。従来知られているレザボワ計算の枠組みはあまり効果的ではなく、その原因は回路の動作のカオス性にあると考えられた。改良のためにリカレント神経回路に動的な均衡が生じカオス性が抑制される

ような構造を入れたモデルを用いた。このモデルを詳しく研究していたところ、構造のあるリカレント神経回路に関する新しい理論が導け、レザボワ計算の記憶効果の利用が容易となった。具体的には、構造の入った神経回路ダイナミクスを場の理論を用いて解析する方法を世界に先駆けて導出することができた。そしてその理論を用いて、コヒーレント状態とよばれる非カオス的でありながら学習に有用な回路状態を誘導できることを示した。図 3(a), (b)に示すように、コヒーレント状態からは入力時系列の履歴を反映した多彩な出力時系列の読み出しが安定に行われるのに対して、カオス状態からの出力時系列は神経回路のサイズの $1/2$ 乗に反比例して減弱してしまい不安定であることがわかった。この結果を論文にまとめ、プレプリントサーバーに公開した後、論文誌に投稿した。

(3)情報収集アルゴリズムの具体的な応用、およびその性能評価のためのシミュレーションプログラムの開発

具体的な応用例の数値シミュレーションによってアルゴリズムの性能を評価するために BCI と社会ネットワーク上のゲームのシミュレーションプログラムを開発した。学習対象の神経回路の出力によって脳活動信号パターンが決まるように設定し、インターフェースシステムではパラメータ値によって決まるフィルターを使って脳活動信号から運動パターンを決める。そして情報収集アルゴリズムを実装した神経回路をインターフェースシステムに組み込んだ。このように開発したシステムによって情報収集アルゴリズムの性能の数値的な評価がある程度可能となった。社会ネットワーク上のゲームモデルに関しては、従来のモデルは本研究のアルゴリズムの性能評価には用いることができないとわかった。そこで、各々のプレイヤーが局所的なゲーム結果の履歴によって確率的に戦略を変えていくモデルを新たに開発し、その性質を調べた。結果として、協力が促進されるゲームパラメータと裏切りが促進されるゲームパラメータが同定された。この新たなゲームの枠組みの上で、学習者がゲームパラメータを変更しながらゲームのプレイヤーの挙動を観察・統計的に処理し、そこから情報を得るようなシステムを設計することが可能となった。

<引用文献>

- [1]A.J. Bell & T.J. Sejnowski, Neural Computation, 1995
- [2]K. Torkkola, Journal of Machine Learning Research, 2004
- [3]T. Hayakawa & T. Aoyagi, Physical Review E, 2015
- [4]S.G. Mason et al., Ann. Biomed. Eng., 2007
- [5]H. Ohtsuki, Nature, 2006

5 . 主な発表論文等

〔雑誌論文〕(計 0 件)

〔学会発表〕(計 5 件)

(1) 早川 隆、深井 朋樹、動的均衡にある神経回路のコヒーレントダイナミクスの理論、日本物理学会 第 74 回年次大会, 2019 年

(2) 早川 隆、深井 朋樹、動的な興奮抑制均衡にある神経回路のコヒーレントダイナミクスの理論, 第 41 回 神経科学大会, 2018 年

(3) Takashi Hayakawa, Tomoki Fukai, Mean-field theory and coherent states of dynamically balanced neuronal networks, CNS2018, Seattle USA

(4) 早川 隆、動的均衡にある神経回路の平均場理論～LFP 様リズムの出現, 配位対称性の破れ, コヒーレント転移, 脳回路機能理論研究会, 2018 年

(5) Takashi Hayakawa, Tomoki Fukai, Interplay of microscopic and macroscopic dynamics in randomly connected neuronal networks, Bernstein Conference, 2017 年

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕

(オンラインプレプリント)

Spontaneous and stimulus-induced coherent states of dynamically balanced neuronal networks, arXiv:1711.09621 (<https://arxiv.org/abs/1711.09621>)

ホームページ等

(日本語) <https://sites.google.com/a/nihon-u.ac.jp/hayakawa-neuroinfophysics-jp/home>

(英語) <https://sites.google.com/a/nihon-u.ac.jp/hayakawa-neuroinfophysics/>

6 . 研究組織

特になし。

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。