

称号及び氏名	博士（理学）松本 裕二
学位授与の日付	平成 19 年 9 月 30 日
論文名	「知識表現を切り替える学習に関する研究」
論文審査委員	主査 馬野 元秀 副査 加茂 静夫 副査 佐藤 優子

## 論文要旨

人工知能(artificial intelligence)は、提唱者である John McCarthy によると「知的機械をつくるための科学や工学」であり、人間の知的活動をコンピュータ上で実現することを目標としている。学習は人間のもっとも知的な活動のひとつであるので、人間が行っている学習を計算機上で実現することは非常に重要である。人工知能の重要な分野のひとつである機械学習(machine learning)では、人間の学習と同様の機能を計算機上で実現させることを目標としており、機械学習のアルゴリズムとして、ニューラルネットワーク、決定木、遺伝的アルゴリズム、サポートベクターマシンといったさまざまな手法が提案されている。これらは一度に大量のデータに対して適用することが多いが、人間が実際に学習するときには、データを一度に全部見ることはできず、データを少しずつ受け取りながら知識をつくりあげていくと考えられる。

人間の学習では、始めの段階では問題に対する知識を全く持っていないので、与えられたデータをそのまま覚えておく。そして、推論時にはそれまでに与えられたデータと照らし合わせて、もっともよく似ているデータの答の部分を推論結果にすると考えられる。データの数が多くなると、すべてのデータをそのまま覚えられなくなり、整理された知識をつくって、その知識に基づいて推論を行うようになると考えられる。つまり、データ数が少ないときに適した知識の表現方法と多いときに適した表現方法は異なっており、データが増えるにつれて人間の知識の表現は具体的なものから抽象的なものに変化していくと考えられる。このような学習として、始めの段階では与えられたデータをそのまま覚えておき、学習が進むにつれて if-then ルールや決定木などの形で知識を持つような、知識の表現

形式が切り換わっていく学習が考えられる。また、人間はそれ以外にも知識の表現形式を変えずに、与えられたデータを基にして知識を再構成することや、知識を変えずに推論の方法だけを変更することもあると考えられる。

そこで、本研究では、データがひとつずつ与えられるごとに推論を行いながら学習も行う状況において、抽象度が異なるいくつかのルール形式の知識表現を用いて、推論法の切り換え、知識の再構成、知識表現の切り換えを行うことにより、データの量に応じて、知識の量を抑えながらある程度の推論性能を保つような推論と学習を進めていく方法を提案し、その有効性を示す。また、知識の生成と検証のときに保存データをふたつに分けることで生成された知識が過剰適合することを防いでノイズデータを多く含む問題に対して適切に対処できるようにし、最近のデータを中心に記憶しておくことで、時間的にデータの傾向が変化するようなデータにも対応することができることを示す。

具体的には、知識表現(ある表現形式での知識の集合と生成法と複数の推論法を組み合わせたもの)をあらかじめ複数用意しておき、推論/学習時にはこのうちのひとつを用いる。また、与えられた入力データとその正解を蓄えておく領域(保存データ集合)を持つ。そして、不正解が連続したことにより、推論法の切り換えと知識集合の再構成と知識表現の切り換えを行う。

推論法の切り換えでは、現在の知識表現が持つすべての推論法を用いて、保存データ集合に対して推論を行い、その中で正解率が最大となる推論法に切り換える。知識集合の再構成では、現在の知識表現が持つ知識生成法を用いて、保存データ集合から知識集合をつくり直す。知識表現の切り換えでは、各知識表現の知識集合を再構成してから、保存データ集合に対して推論を行って正解率を求めて、評価値(正解率が高くて知識集合の量が小さいほどよい)が最大となる知識表現に切り換える。計算機シミュレーションを行い、不正解が連続することでデータ数が少ないときに適したものからデータ数が多いときに適したものへと知識表現が切り換わっていることを示す。

さらには、人間のように、ノイズデータを多く含む問題やデータの傾向が時間によって変化するような問題にも適切に対処することができるように、与えられたデータに関して2つの拡張を行う。

ひとつめは保存データ集合を知識生成用と知識評価用に分割することで、これによって、生成された知識が過剰適合することを防ぎ、ノイズデータを多く含む問題に対して適切に対処できるようになる。計算機シミュレーションによって、分割を行う場合のほうがよい結果が得られることを示す。ふたつめは保存データ集合の容量を制限することで、新しいデータだけを残すことで、クラスの傾向が時間によって変化するような学習データに対して適切に対処できるようになる。また、知識生成と知識評価に必要なデータだけを残すことができれば、推論の性能を下げずに知識の生成と評価に必要な時間を短縮することができる。計算機シミュレーションによって、クラスが時間によって変化するデータに対しては保存データ集合の容量を制限することが有効であることを示す。また、この2つの拡張

を組み合わせることでよりよい結果が得られることを示す。

本論文は 5 章で構成されており、第 1 章では、本研究の背景と目的について述べる。

第 2 章では、本研究に必要となる前提知識となる、ファジィ理論、クラスタ分析、ID3、インスタンススペースの学習について述べる。

第 3 章では、本論文の枠組みについて述べ、学習データと同じ形式、ファジィ if-then ルール、決定木、ファジィ決定木の 4 つの知識の表現形式から、シミュレーションに使用した 6 つの知識の生成法と 6 つの推論法について述べる。そして、これらを組み合わせた 6 つの知識表現を用いて、UCI (University of California, Irvine) の機械学習用データセットの手書き数字識別データを使用したシミュレーションを行い、知識表現が適切に切り換わっていくことを示す。そして、関連研究として、本論文のように複数の学習アルゴリズムを使用する研究をいくつか述べ、それらをいくつかの点で分類することによって、提案法との違いを明らかにする。

第 4 章では、保存データに関する 2 つの拡張について述べる。シミュレーションによって、提案法の拡張前と拡張後との性能を比較し、これらの拡張の有用性を示す。

第 5 章では、本研究で得られた結果をまとめ、今後の課題を示す。

## 審査結果の要旨

学習は人間の最も知的な活動のひとつであり、人間のような学習を計算機に行なわせようという研究が盛んに行なわれている。これらの研究ではデータ集合全体から知識を学習するというものがほとんどである。しかし、人間の場合、データを少しずつ受け取りながら知識をつくりあげていくと考えられ、さらにデータ数が少ないときの知識の表現方法と多いときの表現方法が異なっており、データが増えるにつれて具体的なものから抽象的なものに変化していくと考えられる。このような方法をモデル化し、計算機シミュレーションを行ない、その有効性を示している。

本学位論文の主要な成果は次の通りである。

- (1) 学習の進行に従って、知識の表現方法を切り換えていく学習方法をモデル化している。知識表現の切り換え以外に、推論法の切り換えと保存データ集合からの新たな知識の再構成を考え、連続して推論を間違ったときにいずれかを行なう方法を提案している。具体的な知識表現を設定して計算機シミュレーションを行なった結果、推論法の切り換え、知識の再構成、知識表現の切り換えが起こり、データの量に応じて、知識の量を抑えながらある程度の推論性能を保つことが可能であることを示している。

- (2) 保存データ集合を知識生成用と知識評価用に分割することにより、生成された知識の過剰適合を防ぐことができ、これにより推論性能が上がるとともに、ノイズをより多く含むデータに対しても対処できるようになることを計算機シミュレーションによって示している。
- (3) 保存データの量を制限し、新しいデータだけを残すことにより、傾向が時間的に変化するようなデータに対しても対処できるようになることを計算機シミュレーションによって示している。
- (4) 項目(2)、(3)を組み合わせ、保存データの知識生成用と知識評価用への分割と量の制限を行なった場合の計算機シミュレーション結果も示している。

本研究の成果は、まず、人間が行なっているような学習の進行に従って知識の表現方法を切り換えていく学習方法のモデル化にある。このような学習は人間ならば明らかに行なっているように思えるが、従来はこのような考え方はまったくされていなかった。そして、計算機シミュレーションの結果、連続して推論を間違った回数を用いるという簡単な方法でも、知識の量を抑えながらある程度の推論性能を保つことが可能であることが分かったということは大きな成果である。

さらに、保存データ集合を知識生成用と知識評価用に分けることにより学習性能がよくなることや、保存データの量の制限によりノイズを含むデータや傾向が時間的に変わるデータに対するシミュレーションの結果も非常に興味深いものである。

本研究は、学習の進行に従って知識の表現方法を切り換えていく学習方法を提案し、その計算機シミュレーションにより有効性を示した。本研究は人工知能分野の学習に関する理論と実践の発展に大きく貢献するものである。

本委員会は、学位論文の審査および最終試験の結果から、申請者に対して博士(理学)の学位を授与することを適当と認める。