

称号及び氏名	博士（理学）山村 忠義
学位授与の日付	平成 19 年 3 月 31 日
論文名	「段階的な視覚をもつマルチエージェントにおける強化学習に関する研究」
論文審査委員	主査 馬野 元秀 副査 石井 伸郎 副査 佐藤 優子 副査 寺岡 義伸

## 論文要旨

複数の自律的なエージェントの行動により複雑かつ動的な問題を解決しようとするマルチエージェントの研究が 1990 年代以降、さかんに進められている。

行動主体であるエージェントは、(1) 環境から情報を得て、(2) その情報を基にして判断を行い、(3) 行動をとる、という 3 つのプロセスを繰り返す。(1) でどのような情報を得られるかは、解く問題によって異なる。例えば、マルチエージェントの標準問題によく用いられる追跡問題では、自分の周りの一定範囲の情報が得られる。次に (2) では学習された知識に基づいて判断を行う。そのための知識の学習法として強化学習などが挙げられる。このとき、観測された情報から行動を決めるための知識を獲得する。最後に (3) で判断に基づいて行動を行う。

これら一連の行動をくり返した後に、結果として報酬が与えられる。報酬に至る各行動の適否についての情報は存在しない。報酬を基にして、エージェントは最適または合理的な行動選択を学習する。

このようなエージェントを用いた通常の追跡問題において、各エージェントは絶対座標（地図上の座標）を基にした視覚をもつ。その視覚範囲（情報を獲得できる範囲）の広さは一定で、視覚範囲内では環境に対する正確な情報が得られるが、視覚範囲外では情報がまっ

たく得られないという設定であることが多い。そして、エージェントの視界を広くすると観測情報が増え、うまく利用すると適切な判断ができるようになるが、知識の数が多くなり、学習に時間がかかるという問題がある。そこで、状態数を抑制する方法として、観測情報の状態のいくつかを同じ状態であるとみなすことにより状態数の削減を図る「粗視化 (coarse graining)」が挙げられる。

本論文では、自分を中心とする相対座標に基づく視覚を用い、距離と方向に応じて段階的に大雑把な情報しか得られない「段階的視覚」を提案する。これは、実際の人間の視覚に近いだけでなく、粗視化の一方法でもある。そして、追跡問題において、このような視覚をもつエージェントでも強化学習により効果的に学習が行われることを実験により示す。段階的視覚は、正確な距離や角度を測定できなくなるという意味で不完全な観測情報を扱っている。

第 1 章では、本論文の概要として、エージェント学習の背景と目的について述べた。

第 2 章では、マルチエージェントとマルチエージェントにおける学習でよく用いられる強化学習について述べた後に、その問題点について述べた。

第 3 章では、人間の視覚がもつ「距離が大きくなるにつれて見えにくくなる」という距離感と、「正面はよく見えるが、左右の方向は見えにくい」という人間の視野の性質を反映させた「段階的視覚」を提案する。状態の表現が距離と方向に応じて異なり、近傍では正確な情報が入るが、距離が遠くなると情報が雑多となり、記号化された「近距離」や「中距離」であることしか分からず、見えている対象物の区別もつきにくくなり、方向も「左」「左前」「前」「右前」「右」と記号化されている。そして、中距離では味方エージェントと相手エージェントの区別がつかないために、5 種類の状態表現の方法を提案し、これらについて比較実験を行う。その結果は、いずれの方法でも段階的視覚の方が段階的でない視覚よりもより早く学習が行われていることを示す。なお、エージェントの強化学習アルゴリズムには、環境同定型の代表的な手法である Q-learning を基にした方法を用いる。

強化学習のもう一つの代表的な手法として、経験強化型の Profit Sharing がある。第 4 章では、Q-learning と Profit Sharing の比較実験を行う。荒井が 1998 年に発表した文献では、2 つの学習法を比較した結果、Profit Sharing の方が Q-learning よりも有効であるとしていたが、自分を中心とする相対座標に基づく視覚と段階的視覚を用いたエージェントでは、逆の結果となり、Q-learning の方が Profit Sharing よりもかなりよい性能を示すことが分かった。これは、エージェントの設定と学習パラメータの値に加えて、状態数に対する 2 つの学習法の特性が結果に大きく影響していることが分かった。

ここまでは、基本的に相手の各エージェントに対する観測情報のみを用いる段階的視覚について考えてきたが、第 5 章では、味方エージェントに対する観測情報を用いる場合について考える。味方エージェントとの関係についてのルールも新たに構築され、より効果的な学習が可能であると予想できるが、状態数が大幅に増え、膨大な学習時間が必要になると考えられる。エージェントの視覚範囲内について距離区分と方向ごとに味方エージェントと相手エージェントに対するどのような観測情報を用いるかで 5 種類の視覚を提案し比較実験を行う。このとき、相手エージェントに対する情報に応じて、味方エージェントに対する観測情報の扱い方が異なっている。その結果、状態数が大幅に増え、学習時間も大幅に増えるが、味方に対する情報を用いる方がよい結果となっている。

また、本論文の追跡問題の環境において、3 章で提案した段階的視覚を用いたエージェントと、本章で提案した段階的視覚を用いたエージェントとの比較実験を行った結果、限られた時間内では状態数が少なく、一定の性能のある方が効率がよい場合もあった。

第 6 章では、他の生物がもつ視覚として、人間より広い視角をもっているが、距離感が乏しい「馬」と、全方向を見ることができ、距離感がさらに乏しい「昆虫」の視覚を取り上げ、人間の視覚を含めた 3 種類の段階的視覚についての比較を行う。次に、対象物との距離感と見える方向の異なる数種類の視覚を用いて比較実験を行う。その結果、学習時間を多くすると昆虫がもっともよい性能を示した。

第 7 章では、本論文のまとめと、今後の課題について述べた。

本論文では、相対座標に基づき、距離と方向に応じて見え方が変わり、段階的に大雑把な情報しか得られない設定で、現実の人間に近い視覚である段階的視覚の提案を行った。そして、マルチエージェントの標準問題によく用いられる追跡問題において、段階的視覚をもつエージェントでも効果的に学習が行われることを実験により示した。

## 審査結果の要旨

複数の自律的なエージェントにより複雑かつ動的な問題を解決しようとするマルチエージェントの研究がさかんに進められている。従来のエージェントでは、視覚を用いて環境情報を得るとき、絶対座標に基づき、すべての方向の一定範囲が見えるという設定が普通であった。この論文では、自分を中心とする相対座標に基づき、距離と方向に応じて段階的に大雑把な情報しか得られない「段階的視覚」を提案して、追跡問題において、このような視覚をもつエージェントでも強化学習により効果的に学習が行なえることを示している。さらに、人間の視覚だけでなく、馬と昆虫の段階的視覚を定式化し、人間の視覚との比較を行なっている。

本学位論文の主要な成果は次の通りである。

- (1) 人間の視覚がもつ距離感と視角の性質を反映させた「段階的視覚」を提案している。これは、自分を中心とする相対座標に基づき、「距離が大きくなるほど見えにくくなる」という距離感と「正面はよく見えるが、左右の方向は見えにくい」という視角の性質を反映させている。学習には、Q-learning と呼ばれる強化学習の手法を用いているが、その状態の表現が近傍、近距離、中距離ごとに異なる方法を提案している。さらに、中距離において、5 種類の状態表現法を提案し、段階的ではない視覚も含めて、計算機シミュレーションにより比較実験を行なっている。その結果、5 種類のいずれの方法でも、段階的視覚の方が段階的でない視覚よりも学習が早く行なわれることを示している。
- (2) 強化学習に Profit Sharing を用いた場合について検討している。この場合でも、Q-learning と同様に、段階的視覚の方が段階的でない視覚よりも学習が早く行なわれることを計算機シミュレーションにより示している。さらに、Q-learning と Profit Sharing の比較を行なった結果、Q-learning の方が圧倒的によい結果が得られることが分かった。これは従来の結果とは逆で、その理由として状態数に関係しているだろうと予想している。
- (3) 上記 (1)、(2) では、相手に対する情報のみで状態をつくっていたが、味方に対する情報も用いる場合について考えている。状態数が非常に増えて学習に時間がかかるが、最終的には、味方に対する情報を用いる方がよい結果が得られることを計算機シミュレーションにより示している。

- (4) さらに、人間の視覚だけではなく、馬と昆虫の段階的視覚を考え、比較を行なっている。距離感が人間より乏しいが視角が少し広い馬の段階的視覚と、すべての方向が見えるが距離感がさらに乏しい昆虫の段階的視覚を定式化し、人間の段階的視覚と比較している。学習回数を多くすると昆虫がもっともよい結果が得られることが分かった。

本研究の成果は、まず段階的視覚の定式化にある。このような現実の人間に近い視覚に基づくシミュレーションは、従来、行なわれていなかった。そして、このような視覚でも、追跡問題において十分な学習性能を示すことが分かった。これは状態の大幅な削減を可能にし、学習可能な範囲を大きく拡大するものであり、この研究の大きな成果である。また、人間以外の馬と昆虫の段階的視覚の定式化とそれを用いたシミュレーションの結果は非常に興味深いものである。