

審査の結果の要旨

氏名 井上 康介

井上 康介 (いのうえ こうすけ) 提出の本論文は、「部分観測環境におけるエージェントの自律的行動獲得」と題し、全7章からなる。本論文では、身体性を有するエージェントがその身体性と環境・タスクとの相互作用に立脚した状況識別様式を自律的に獲得し、これに基づいて複数タスクを実現する行動を獲得する手法の提案を行っている。

第1章では、エージェントの身体性から帰結される本質的な問題として、(1) 環境の部分観測性、(2) 観測入力 of 解釈方法の身体・環境・タスクへの立脚性および(3) 複数タスク実現における身体性に立脚したタスク識別機構の必要性について議論した。これらの問題を、外的に付与される即時報酬に基づいて解決する学習手法の意義について議論し、これを研究の目的として示した。

第2章では、提案する手法が扱う問題領域を定式化し、本研究を通じて想定する離散的時間について議論を行った。また、提案手法を教示システムとしてみた場合の提案手法の優位性について議論した。

第3章では、単一のタスクに対して、エージェントの状況認識を身体・環境・タスク間の相互作用に基づいて実現する状況認識機構の獲得、およびこれに基づく行動決定機構の獲得を実現する学習手法の提案を行った。提案手法では、環境の部分観測性に起因する知覚騙し問題を解決するために、エージェントの観測入力・動作出力の短期記憶を利用した状況識別を表現する決定木構造の内的状態表現を利用し、これに対して観測の識別に基づく状況識別に対応する分岐を逐次的に追加するという方法で、身体・環境・タスク間相互作用に立脚した観測入力の解釈様式の獲得を実現する。ここで内的状態構成は、最初単一の内的状態からなる状態表現を身体・環境・タスク間の相互作用に即して適応的に分割することにより行われるが、ここでは(1) 観測情報をより詳細に利用するための観測空間の分割、および(2) 部分観測性に起因する知覚騙し問題への対処のためにより過去の履歴を利用するための状態分割の2つの分割方法があり、これを適切に実行する必要がある。提案手法では、過去の経験データに対する統計的処理に基づいて、どちらの分割をどのように行うべきかを判別することでこの問題に対処する。

第4章では、第3章で提案した手法を複数タスクに拡張する手法を提案した。具体的には、個別のタスクに対応する行動獲得を第3章で示す手法により実現した上で、未知のタスクに対して

投入されたエージェントが、現在扱っているタスクの識別を身体・環境・タスク間の相互作用に立脚して実現するために、現在各タスクを扱っている可能性を表現する確信度を更新し、適切なタスク依存知識を適用するという方法を採用。このとき、(1)各タスクに対する学習過程において、タスク識別に十分な経験データが得られていない、(2)タスクを識別するのに十分な行動を行った時点から、各タスクにおけるゴールを達成するのに必要な行動が獲得されていない、という2点の問題に対応するため、適切な追加学習を行う学習スケジューリング法を適用する。

第5章では、以上に提案した手法を実移動ロボットのナビゲーションタスクに適用して評価・考察を行った。まず、複数のスタート点から特定のゴール点に到達する行動を、初期のスタート位置をロボットに知らせずに実現する行動の獲得を、実ロボットを再現した実際的なシミュレーションにより実現し、獲得された行動を実ロボットに適用することにより、提案手法が実移動ロボットによるナビゲーションタスクに対して有効に機能することを確認した。ただし、ここではロボットに与える即時報酬は、各時点におけるロボットの位置・姿勢からゴール状態に到達するまでに必要な動作ステップ数の増減に基づいて与えた。次に、ロボットに与えられる即時報酬がロボットの動作原理に立脚していない場合として、Wavefront法に基づくポテンシャルにより計算される即時報酬に基づく単一タスクに対するナビゲーション行動の獲得をシミュレーションにより行った。これにより、ロボットの行動原理とは独立した視点から与えられる即時報酬のみに基づいて提案手法により行動獲得が可能であることが示された。

第6章では、提案手法に関する考察・評価を行った。具体的には、単一タスクに対する学習手法に関して、計算量・記憶量および環境の性質・学習パラメータに対する学習性能の依存性、内的状態表現における観測情報の利用の意義について議論した。また、付与する即時報酬の満たすべき条件について議論した。次に、実環境への適用における誤差の影響について議論した。これらの議論から、提案手法の適用可能範囲とその拡張についての議論を行った。

以上を要約するに、本研究では、身体性を有するエージェントが、部分観測環境下において、身体・環境・タスクに立脚した状況認識様式を獲得し、これに基づいて複数タスクに対して適切な行動を実現する行動獲得を、エージェントの各時点における状況・動作に対して与えられる即時報酬に基づいて実現する行動獲得手法の提案が行われ、実機実験・シミュレーションにより実環境への適用可能性が示された。これは精密機械工学のみならず工学全体に寄与するところが大きい。

よって本論文は博士（工学）学位請求論文として合格と認められる。