

論文審査の結果の要旨

論文提出者氏名 今川 孝久

本論文は、二人ゲームを対象としたモンテカルロ木探索について、既存手法の性質と限界を様々な角度から明らかにするとともに、新たに分かった問題点に対応する改善手法を提案したものである。モンテカルロ木探索では、計算機資源の許す範囲で試行を行い、その結果をもとに最善手を推定する。また、最善手の推定のためには、各選択枝の利得の推定が行われる。このようにモンテカルロ木探索は汎用的な枠組みであるため、アルファ碁をはじめとして様々な応用で成果をあげているが、一方で苦手とするドメインの存在も知られている。しかし、ドメインと木探索の様々な性質を踏まえた性能の分析については、これまで知見がほとんどなかった。

本論文は 8 章からなる。第 1 章では、研究の背景と目的が述べられている。第 2 章では、これ以降の議論で必要となるモンテカルロ木探索と、その周辺分野の関連研究が導入されている。モンテカルロ木探索は、木構造を持たない多腕バンディット問題の手法を、木探索に応用したものである。両者は、繰り返し試行を行い推定を改善してゆくという点で、共通の枠組みを持つ。しかし、多腕バンディット問題で成立していた性質の多くは、現実の木探索では成り立たない。そのため、多腕バンディット問題に近い理想的な木と応用問題との間をつなぐモデルの必要性が指摘されている。

第 3 章では、モンテカルロ木探索の性能が落ち込む状況について、新しい条件を示している。それは、最善手を逃した場合の不利益の大きさの分布の偏りである。この点について理論的な分析に基づく対応方法を提案し、その効果を計算機実験で確認している。これまで囲碁の攻め合いなど個別に指摘されていた事象の背景を一般化し、対応方法を示したことは高く評価される。

第 4 章では、試行回数に制約がある前提で最善手を判別するために、累積リグレットに替えて単純リグレットを最小化する戦略について扱う。これまでは、単純リグレットを直接最小化することは難しいため、累積リグレットを最小化する手法が主流であった。今川氏は、報酬の分布を計量することと多少の仮定を導入することで、単純リグレットを最小化する、具体的かつシンプルなアルゴリズムを提案している。提案手法の性能は、幅と深さが一様な木のモデルでは既存手法を越えないものの、現実の探索問題に近い非一様な木

のモデルでは上回り、有力な手法であることが確認された。

第5章では、ゲーム木の探索中にある節点で勝敗が確定した際の、既存のモンテカルロ木探索の振る舞いの問題点を指摘している。勝敗の確定は、多腕バンディット問題では生じない、木探索特有の問題の一つである。さらに今川氏は、勝敗の確定と木探索における試行結果の管理を整合させる新たな手法を提案し、計算機実験で効果を確認した。

第6章では、第7章の準備として、確率変数の期待値の最大値を取り扱う。今川氏は、収束性と計算コストの両面で扱いやすい性質を持つ、新しい推定量 Simplified Weighted Estimator (SWE) を提案した。有限の試行回数での推定の正確さを評価するために、多腕バンディット問題の計算機実験を行い、様々な推定量との比較を行っている。その結果、一部の設定では同程度に良い推定量はあったものの、総合すると SWE が安定して良いことが示された。

第7章では、第6章で提案した SWE をモンテカルロ木探索に統合する方法を新しく提案している。SWE が木探索でも同様に効果的であれば、各選択肢の推定が正確になり、そのことを通じて最善手の判別が正確になると期待される。性能評価には、確率的に報酬を返す終端節点を持つ木と、確定的に報酬を返す終端節点を持つ木の、2種類の人工的な木が用いられた。最善手の判別に関して、SWE を統合したモンテカルロ木探索の性能は、前者の木で既存手法を上回り後者でも同等程度であった。また、共通して、各節点の推定値が改善したことが確認された。総合して、提案手法は有力な手法と評価できる。

第8章では、全体のまとめと、今後の研究方向についての考察を行なっている。

以上のように本論文は、モンテカルロ木探索という研究上も実用上も重要なテーマについて、理論的な基礎を持つ手法を提案し実用的な性能向上を実現したもので、木探索を扱う研究分野全体に対する学問上の貢献が大きいと評価できる。したがって、本審査委員会は博士(学術)の学位を授与するにふさわしいものと認定した。