

審 査 の 結 果 の 要 旨

氏 名 劉 雲青

本論文は、「Application of Simple Regret Bandit Algorithms on Monte-Carlo Tree Search (和題：単純リグレットバンディットアルゴリズムを利用したモンテカルロ木探索)」と題し、全6章から成る。モンテカルロ木探索 (Monte-Carlo Tree Search, MCTS) は、囲碁をはじめとする、評価関数の設計が難しい複雑な完全情報ゲームのための探索手法として、ゲーム AI 分野において広く利用されているアルゴリズムである。本論文は、モンテカルロ木探索における各ノードの着選選択の手法として、単純リグレットを基準とするアルゴリズムを導入することを提案し、その有効性を実験的に検証した論文である。

第1章は「Introduction」と題し、ゲーム AI における探索アルゴリズムの概要を述べ、本研究が対象とするタスクであるモンテカルロ木探索において、各ノードでの着手選択の問題が、多腕バンディット問題として定式化できることを述べている。多腕バンディット問題においては、これまで得られた情報をもとに利得を獲得することを目指すこと (exploitation) と、新たな選択肢を試すことで利得に関する情報を得ること (exploration) を両立させる必要があるが、その問題に対する解法として、累積リグレット (cumulative regret) と単純リグレット (simple regret) に基づく手法があることを紹介し、本論文において後者を利用する理由を説明している。

第2章は「Preliminaries」と題し、ミニマックス法とモンテカルロ木探索の概要、多腕バンディット問題に対する代表的な解法である UCB アルゴリズムを紹介している。また、UCB アルゴリズムが累積リグレットに基づくものであることを説明するとともに、より累積リグレットの上限が小さい Improved UCB アルゴリズムを紹介している。さらに、単純リグレットに基づく UCB アルゴリズムの変種について述べている。

第3章は、「Adaptation of the Improved UCB Algorithm」と題し、最初に Improved UCB アルゴリズムをモンテカルロ木探索に適用するうえでの問題点を議論している。具体的には、Improved UCB アルゴリズムが、最適ではない候補手に計算リソースを費やしがちであること、探索を途中で打ち切った場合

に最善手が確定しないことを述べている。これらの問題点をふまえ、CCB (Combined Confidence Bounds) バンディットアルゴリズムを提案し、Improved UCB アルゴリズムからの改善点を説明している。提案手法の有効性を検証するための予備的な実験として、人工的な多腕バンディット問題によるアルゴリズムの性能評価を行っている。

第4章は、「Regulation of Exploration in Simple Regret Minimization」と題し、CCB バンディットアルゴリズムをモンテカルロ木探索に適用した CCB-MCTS アルゴリズムを提案している。最初に、本論文が提案する CCB バンディットアルゴリズムと、単純リグレット基準に基づく既存のバンディットアルゴリズムの違いを論じ、提案手法は最善手以外に費やされる計算リソースを削減できる可能性があることを述べるとともに、提案手法を用いた場合の単純リグレットの上限を数学的に示している。また、CCB-MCTS の性能を検証するための実験として、9 路盤の囲碁と 9 路盤の NoGo を用いた実験を行い、CCB-MCTS がどちらのゲームにおいても既存手法を上回る性能をあげること、特に、プレイアウトの回数が増えるにしたがってその差が顕著になることを述べている。

第5章は、「Asymmetric Move Selection Strategies」と題し、Max ノードと Min ノードで異なる着手選択基準を用いるモンテカルロ木探索アルゴリズムである Asymmetric MCTS を提案している。具体的には、Max ノードにおいては単純リグレットに基づいて着手を選択し、Min ノードにおいては累積リグレットに基づいて着手を選択するという手法になっている。前章と同様、9 路盤の囲碁と 9 路盤の NoGo を用いた実験を行い、どちらのゲームにおいても提案手法が既存手法の性能を上回ることを述べている。

第6章は、「Conclusion」と題し、本論文を総括し、今後の課題について述べている。

以上これを要するに、本論文はゲーム AI における重要な探索手法であるモンテカルロ木探索アルゴリズムに関して、単純リグレットを評価基準とした着手選択手法を複数提案し、それらの有効性を実際のゲーム木探索の実験によって明らかにしたものであり、情報処理工学に貢献するところが少なくない。

よって本論文は博士（工学）の学位請求論文として合格と認められる。