

令和 3 年 6 月 24 日現在

機関番号：32682

研究種目：基盤研究(B) (一般)

研究期間：2017～2019

課題番号：17H01846

研究課題名(和文) 複数の評価手法を利用可能な大規模分散モンテカルロ木探索

研究課題名(英文) Large-scale distributed Monte-Carlo game-tree search algorithm that can employ different evaluation strategies

研究代表者

横山 大作 (Yokoyama, Daisaku)

明治大学・理工学部・専任准教授

研究者番号：80345272

交付決定額(研究期間全体)：(直接経費) 12,600,000円

研究成果の概要(和文)：現実世界の大規模な探索問題は全解探索は困難であり、一部を選択的に探索するアルゴリズムが用いられる。ゲーム木の探索はこのような問題の一例であり、モンテカルロ木探索が有効とされているが、将棋など「細い正解の一本道をたどらねばならない」ような問題領域においては収束が遅く、適用が難しい。我々が以前提案したアルゴリズムを改善するために、性質の異なる複数の評価手法を活用できるモンテカルロ木探索手法を検討した。検証の結果、適用の可能性が示されるとともに、効率的な実装方法の検討などが課題として明らかになった。また、実用的な探索アプリケーションへの応用可能性についても検討を行った。

研究成果の学術的意義や社会的意義

現実世界には複雑な制約のもとで適切な解を見つけることを要求される、大規模な探索問題が多く存在する。例えば、不完全な情報のもとで最適な戦略を見つける問題などがあり、ゲームをプレイするアルゴリズムを研究することでその問題のエッセンスを考えることが可能になる。本研究は、そのような問題を、現在の高性能な計算機を効率よく活用し、精度よく解くことを可能にするための基礎的な技術確立することを目指したものであり、広い応用範囲を持つ。

研究成果の概要(英文)：Large-scale search problems in the real world are not applicable to exhaustive search; randomized search algorithms have great ability to explore such problems. Game tree search is an example of such a problem; the Monte-Carlo Tree Search algorithm (MCTS) has been widely used. However, this great advance does not help to achieve good performance in Shogi that has a long-narrow path of 'correct' play. We try to evaluate an algorithm that can employ several different evaluation strategies to improve our previously proposed method. We evaluate the applicability of our method and found several difficulties, such as implementing issues. We also research the applicability for large-scale realistic problems.

研究分野：ゲーム情報学

キーワード：人工知能 アルゴリズム ゲーム情報学

### 1. 研究開始当初の背景

現実世界の大規模な探索問題は全探索は困難であり、一部を選択的に探索するアルゴリズムが用いられる。ゲーム木の探索はこのような問題の一例であり、モンテカルロ木探索が有効とされているが、将棋などに代表される「細かい正解の一本道をたどらねばならない」ような問題領域においては収束が遅く、適用が難しい。我々は、乱数を組み込んだ探索を利用するアルゴリズムを提案し、将棋での有効性を確認したが、単一の評価関数を用いることによる効率の悪さも明らかになった。

### 2. 研究の目的

本研究では、ごく限られた一本道の正解経路をたどっていかなければならないゲーム的特性を持つ将棋を対象に、性質の異なる複数の評価手法を活用できるモンテカルロ木探索手法を検討し、将棋を題材に実用的なアプリケーションとして実装を行い、大規模分散計算環境でその有効性を評価する、ことを目指す。

### 3. 研究の方法

複数の異なる評価手法を利用可能な分散モンテカルロ木探索のためのプログラミングフレームワークを試験実装し、並列探索アルゴリズムの設計と性能に関する調査を進めるとともに、将棋を題材としたゲームプレイヤを用いて分散計算アプリケーションを試験実装する。この試験実装結果を元に並列探索アルゴリズムの詳細設計と本格実装を行ない、性能面の評価と改善を行う。さらに、シミュレーションなどを用いた予備実験による評価を行って、性能面でのボトルネックなどの改善手法検討を行う。また、この手法を用いるときの実用性の評価を行うとともに、さらなる性能向上を図る。

### 4. 研究成果

複数の評価手法を利用するためには、その評価手法から得られる結果をどのように統合して探索結果を得るかという、解の表現方法の確立と、異なるリソース要求を持つ評価手法を適切に実行制御し、その結果を探索順序にいかに対応して反映していくか、という探索制御方法の確立の2つが必要となる。解の表現方法については、複数の手法の評価結果が全て確率分布を持つものとして取り扱うように考えることで、問題全体としては確率分布を持つ探索木の解を求めることになり、研究代表者がこれまでに提案してきたベイジアンアプローチを用いたモンテカルロ木探索アルゴリズムが適用可能となることが確認された。一方、探索順序への反映については、複数の評価手法が独立に確率分布を持つことで、分布がスパースになり、何らかの集約を加えることで探索順序制御への適用が可能になると期待されることが示された。実用化のために、適切な集約方法の確立や効率的な実装方法の検討などが課題として示され、今後の研究につながる知見が得られた。

本研究では、また、実用的な探索アプリケーションへの応用を検討するため、交渉ゲームやRPG、大規模データとシミュレーションを用いた人流推定など、より複雑な状況を扱う必要がある困難な問題に関する研究を行った。以下、研究成果について簡潔に示す。

#### (1) 交通スマートカードの乗降履歴データを用いた人流推定手法の検討と、シミュレーションを用いた経路推薦システムの検討

交通スマートカードから得られる乗降履歴データをもとに、鉄道網での人の流れを推定する問題に取り組んだ。乗降履歴には、乗車駅と降車駅、降車時間、の3つの情報しか記録されておらず、途中の移動経路を推定する必要がある。それぞれの乗客が合理的な行動をとったと仮定し、移動時間を最短にするような経路をすべての乗降履歴について探索することで、経路の途中も含めたあらゆる場所で、それぞれの時刻ごとの移動人数を推定することが可能になると示された。

この経路探索において、乗り換えにかかった時間などを見積もる際、平均運行時間間隔のみを用いて推定した場合、列車のダイヤ情報と突き合わせた場合、列車の実際の運行記録と突き合わせた場合の3通りの方法で推定を行い、計算に必要な時間と得られた推定結果の観点から比較を行った（図1）。それぞれの手法で推定移動人数に差異があり、運行記録を用いることで推定精度が向上できる可能性が示されたとともに、計算時間の増大を軽減するような手法の確立が必要であることが明らかになった。

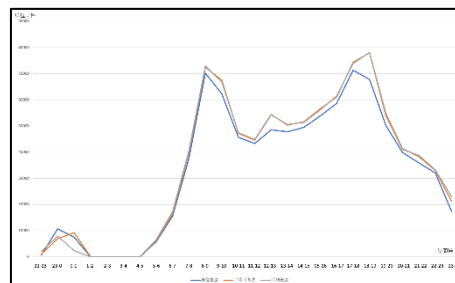


図1 人流推定手法による時刻ごとの推定移動人数の差異

この知見を基に、合理的に経路を選択する人間を模した多数のエージェントによる歩行者人流シミュレータを構築した(図2)。それぞれのエージェントは、道路の距離と混み具合から得られるリンクの推定移動時間に基づいた経路探索を行い、確率的な経路選択ポリシーに従って行動する。このシミュレータ上で、移動時に混雑が発生する状況を再現し、一部の人に混雑を緩和するような経路推薦を行ったときにどの程度集団全体の効用が改善するかを検証した(図3)。適切な経路探索アルゴリズムを用いることで集団全体の平均移動完了時間が短縮できること、また、経路推薦に従う人が7割を超えると、経路推薦を行っていない残りの人たちの移動完了時間も短縮され、恩恵があること、などが示された。

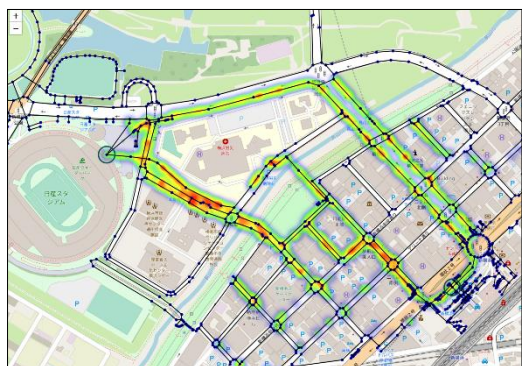


図2 現実の道路ネットワーク上で1940人が一斉に移動する状況のシミュレーション

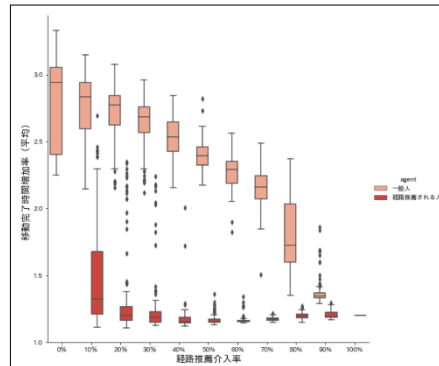


図3 経路推薦を行う割合を変化させたときの推薦群・非推薦群の移動完了時間分布変化

さらに、歩行者が混雑の中で過ごす時間をなるべく少なくすることを目指し、制御することを目指した経路推薦アルゴリズムを検討した。混雑状況を考慮して迂回路のリンク時間コストを設定することで、単純な方法と比較して、高密度空間にさらされた累積時間を短くすることができることが示された(図4)。特に、歩行者集団の後半に位置している人たちについて、単純な方法では過密空間に居た時間が長かったが、提案アルゴリズムで大幅に軽減されていることが示された。

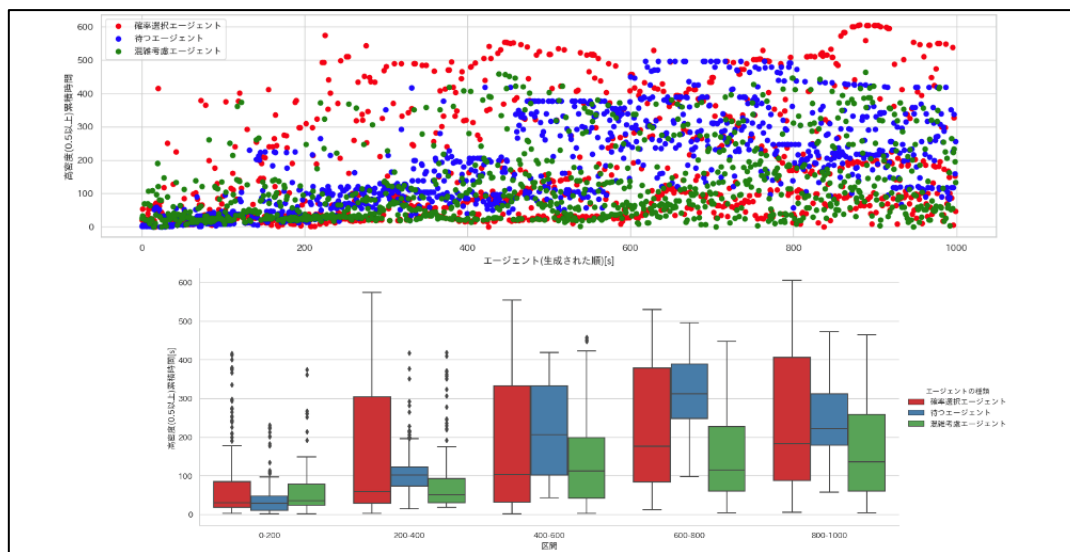


図4 異なる方法で迂回路推薦を行った場合の歩行者1000人が高密度空間に滞在した累積時間のプロットと、集団を20%ずつに分割した時の分布

これらの問題は、大量の探索を必要とする実世界応用問題であり、本研究課題が対象とする問題領域を広げ、実用性を高めるために役立つ問題領域であると考えている。今後も引き続き、研究課題として取り組んでいく。

## (2) 複雑なルールを持つゲームを対象とした探索アルゴリズムの検討と評価

複数の観点を持つ評価関数や、単純に最大化・最小化を目指すのみではない評価関数を持つような応用分野について、本研究を応用することを目指すため、そのような複雑な構造を持つゲームを題材に探索アルゴリズムの研究を行った。

完全情報ゲームの1つであるボードゲーム「マンカラ」を題材に、コンピュータプレイヤーの強さを対戦相手に合わせて自動的に変更できるような枠組みを検討した。方策として、評価関数の

示す次状態の評価値の中で最善のものではなく、割り引いた値を目指すようなプレイを行うことで、コンピュータプレイヤーの強さを調整できる可能性が示された。

3人で行う不完全情報カードゲームの「最中限」を題材に、複雑な評価構造を持つゲームでのプレイヤーの構築方法を検討した。最中限はトリックテイキングゲームの一種であり、トリックを得ることができるのは「カードの値が中央となったプレイヤー」、最終的に得られたトリックの値が「中央」となるプレイヤーが勝者となる、というルールである。つまり、相手プレイヤーの行動に合わせて中央となるような行動を選びたい、というゲームが再帰的に繰り返される、複雑な構造を持つ。ここでは、ヒューリスティックを組み合わせたプレイヤーに、ゲーム途中での目標得点を設定する枠組みを追加することで、ランダムプレイヤーとの対戦で勝率が向上することが確認できた。(図5)

また、より多様な問題領域での応用可能性を検討するために、General Game Playing (GGP) における戦略の研究を行った。GGPは、ゲーム記述のためのドメイン特化言語 GDL によって記述されたゲームルールを受け取り、事前の知識がない状態でゲームを理解してプレイする、という問題である。Tic-Tac-Toe や Connect Four など、多数のゲームが GDL で記述されており、同じ探索手法を性質の大きく異なるゲームに対して適用することが求められる。

ゲーム固有の知識をなるべく必要とせず、汎用で性能の良いプレイを行うため、ここでは Monte-Carlo Tree Search アルゴリズムを用いてプレイヤーを作成した。また、探索の中で必要となるシミュレーション(ランダムな行動選択を行ってゲームの結果がわかるところまで進行させる手法)の過程について、合理的な行動選択を選ぶ確率がより高まるよう、すでに行ったシミュレーションの状態遷移記録を利用した、状態の bi-gram を用いる戦略を検討した。シミュレーションの過程において、既知の状態の bi-gram は  $\epsilon$ -greedy 法により確率的に採用されるものとする。この手法を用い、シミュレーションで bi-gram を用いないプレイヤーと、 $\epsilon$  を変化させたプレイヤーとの対戦実験を行って勝率を評価した(表1)。ゲームには Connect Four を用いた。結果として、bi-gram を用いることで有意に強くなったという結論は得られなかったが、bi-gram の利用方法の工夫などを行えばシミュレーションの精度向上が図れるのではないかという結果が観察された。また、計算量が増大するため実装の工夫が必要となることも確認できた。

これらのゲームは、相対的に変化する評価基準を利用する必要があるものや、汎用の評価基準を利用する必要があるものなど、本研究課題が目標とする、複数の評価基準を利用する枠組みの応用範囲を広げる対象領域を持つ。今後も、提案手法の有用性を高めるために研究課題としていくことにする。

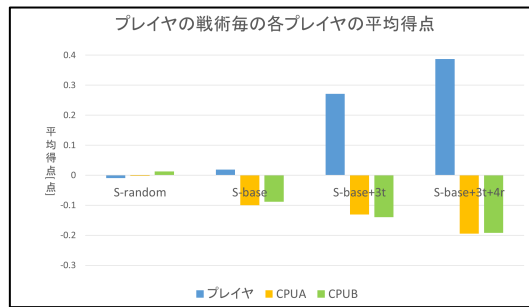


図5 ゲーム途中で目標得点を与える戦術の利用による勝率変化

表1 bi-gramを用いるシミュレーションで  $\epsilon$  を変化させたときの勝率

$\epsilon$	勝ち数	負け数	引き分け数
0.3	26	23	51
0.5	29	21	50
0.7	24	25	51

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 横山 大作	4. 巻 37-1
2. 論文標題 IEEE CoG2019会議報告	5. 発行年 2020年
3. 雑誌名 コンピュータソフトウェア	6. 最初と最後の頁 73-77
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 横山大作	4. 巻 35
2. 論文標題 KDD参加報告	5. 発行年 2018年
3. 雑誌名 コンピュータソフトウェア	6. 最初と最後の頁 86-89
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計9件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 wang xuanhao, 横山 大作
2. 発表標題 旅客乗降履歴を利用した人流推定の精度向上に関する考察
3. 学会等名 第11回データ工学と情報マネジメントに関するフォーラム（DEIM2019）
4. 発表年 2018年

1. 発表者名 藤本海右, 鶴岡慶雅
2. 発表標題 Counterfactual Regret Minimization による交渉ゲームの求解
3. 学会等名 第22回ゲームプログラミングワークショップ
4. 発表年 2017年

1. 発表者名 加納由希夫, 鶴岡 慶雅
2. 発表標題 内部報酬を自動生成する強化学習による一人用RPGの自動攻略
3. 学会等名 第22回ゲームプログラミングワークショップ
4. 発表年 2017年

1. 発表者名 Yang, Xu and Tsuruoka, Yoshimasa
2. 発表標題 Parametrized Control in Soccer Simulation with Deep Reinforcement Learning
3. 学会等名 第22回ゲームプログラミングワークショップ
4. 発表年 2017年

1. 発表者名 若泉 亮汰, 横山 大作
2. 発表標題 群衆移動時の混雑抑制を目指した経路推薦手法に関する考察
3. 学会等名 第13回データ工学と情報マネジメントに関するフォーラム (DEIM2021)
4. 発表年 2021年

1. 発表者名 大室 光, 横山 大作
2. 発表標題 最中限における戦略学習の試み
3. 学会等名 第62回プログラミング・シンポジウム
4. 発表年 2021年

1. 発表者名 若泉 亮汰, 横山 大作
2. 発表標題 全体最適化を目指した経路推薦システムのための混雑情報伝達に関する考察
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム (DEIM2020)
4. 発表年 2020年

1. 発表者名 リ チョウウ, 横山大作
2. 発表標題 プレイヤーに合わせて難易度を調整するカラハAIの研究
3. 学会等名 第83回情報処理学会全国大会
4. 発表年 2021年

1. 発表者名 上宮佳晃, 横山大作
2. 発表標題 General Game Playingにおけるモンテカルロ木探索のシミュレーション戦略改善に向けた検討
3. 学会等名 第83回情報処理学会全国大会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------