

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 1 日現在

機関番号：12601

研究種目：基盤研究(B) (一般)

研究期間：2014～2016

課題番号：26280130

研究課題名(和文) 確率分布を反映したゲーム木の大规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 based on probability distribution

研究代表者

横山 大作 (Yokoyama, Daisaku)

東京大学・生産技術研究所・助教

研究者番号：80345272

交付決定額(研究期間全体)：(直接経費) 12,200,000円

研究成果の概要(和文)：現実世界の大规模な探索問題は全解探索は困難であり、有用そうな部分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、近年提案されたモンテカルロ木探索がその有用性から広まっているが、将棋など「細い正解の一本道をたどらなければならない」ような問題領域においては終息が遅く、適用が難しい。我々は、局面評価値の確率分布をゲーム木でそのまま取り扱えるベイジアンアプローチに基づいた探索アルゴリズムについて、分散計算の適用による効率化の可能性を検証した。

研究成果の概要(英文)：Large-scale search problems in real world are not applicable exhaustive search; randomized search algorithms have great ability to explore such problems. Game tree search is an example of such problem; Monte-Carlo Tree Search algorithm (MCTS) has been developed and is getting widely used with good performance. This great advance, however, does not help to achieve good performance in Shogi that has long narrow path of 'correct' play. We propose a new randomized game-tree search algorithm based on Bayesian Approach that can treat evaluated values as probability distributions. In this research we evaluate the effectiveness of our approach using distributed computing method.

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

組合せ最適化や経路探索など、実世界で大規模な探索問題を解きたいという要求は多い。複雑な実問題を解くためには、全ての解の可能性をしらみつぶしに調べるのではなく、様々な技法を用いて探索範囲を限定し、有用な解が存在する可能性の高い部分を選択的に探索して、質の良い解をできるだけ早く求めるような、確率的探索などのアルゴリズムを用いることが多い。ゲーム木探索はこのような探索問題の一例であり、様々な枝刈り技法を用いた探索が発展してきたが、逐次探索に特化した枝刈り技法が特に効果があるため、大規模な問題を解くために必須である分散計算の適用が困難な状況にある。

一方、大規模並列探索に向けたアルゴリズムとして、UCT という制御手法を用いたモンテカルロ木探索(図 1)が近年提案され、囲碁への適用では特に高い性能を示しており、大規模分散計算での性能向上も進んでいる。しかし、将棋への適用は現在のところ従来の木探索手法より良好な性能は得られていない。これは、将棋においては、ある局面においての正解手順がごく少数に限られ、それ以外の手順は大きく形勢を損ねるような場合が多い、というゲーム的な特性%があることに起因していると考えられる。このような、狭い一本道の正解手順をたどることが要求される探索問題は将棋以外にも存在しているが、そのような問題では現在のモンテカルロ木探索手法では効率よく探索空間を絞ることができない、という問題点が知られている。

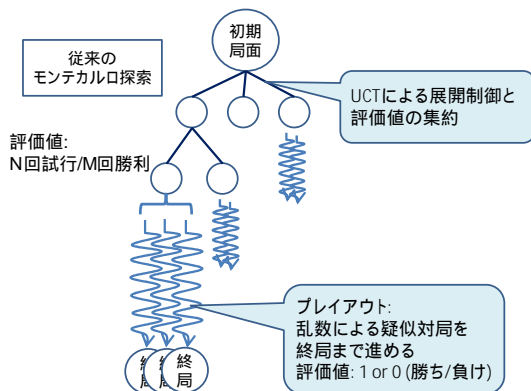


図 1 従来のモンテカルロ木探索

我々はこの問題に対し、探索における葉局面評価値の確率分布を探索木に伝搬させルート局面での最善手選択に反映させる、ベイジアンアプローチを利用したモンテカルロ木探索(図 2)を提案し、将棋を対象にその有効性の検証を試みている。提案アルゴリズムは極めて大量に計算リソースを用いた場合には有効であるが、効率に関して検討を加えていく必要があることが確認されている。また、大量のリソースを利用するための分散計算手法が検討されておらず、大規模問題を解くうえでの障害となっている。

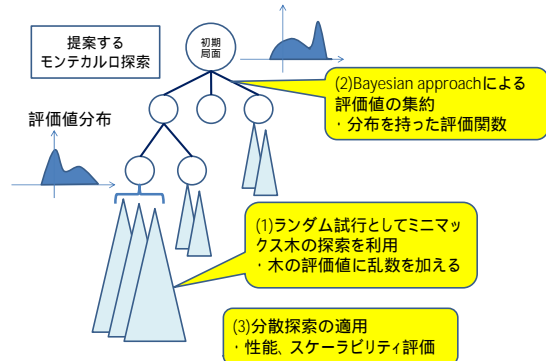


図 2 ベイジアンアプローチに基づくモンテカルロ木探索

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、大規模分散計算の適用が容易な、確率分布に基づいたモンテカルロ木探索アルゴリズムを構築し、実際の大規模環境でその有用性を明らかにする。具体的には、(1)ベイジアンアプローチに基づいたモンテカルロ木探索手法と、そのスケール可能な並列化手法を検討し、(2)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、(3)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

(1)我々は従来のUCTを用いた探索に対し、(a)プレイアウト評価の木探索による置き換え(b)評価値分布の集約手法の改良、の2点を変更した、ベイジアンアプローチに基づいた探索アルゴリズムを提案し、将棋での有効性を検証している。このアルゴリズムにおいて、大規模分散計算を想定した並列化を行うため、どのリーフを計算するのかという探索スケジューリング、どのタイミングで評価値分布を集約し配布するのかという情報更新手法などについて、適切な設計を行う。また、並列化に伴って現在の探索アルゴリズムに探索性能と実行効率のバランスを考慮した改良を加えていく。

(2)現在知り得る問題領域依存の工夫が十分に施された将棋プレイヤー「激指」を用いて分散探索アプリケーションを作成し、強いプレイヤー実現を目指すことを通して、実用的なアプリケーションを構築する上で有効な実装上の工夫、性能上のトレードオフなどについて検討を行い、適切な実装方法を明らかにする。

(3)大規模分散計算環境での評価を通して、設計及び実装の適切さと提案手法の有効性を検証する。

4. 研究成果

(1)手法の検討

我々は、モンテカルロ探索のシミュレーションとして乱数を付加した探索を用いること、Bayesian Approach に基づいて探索木中

の評価値伝搬を行うこと、の2点から構成されるモンテカルロ木探索アルゴリズムを構築し、将棋プレイヤーに応用した。

提案手法では、ランダムな指し手を生成するのではなく、評価関数に乱数を加えた従来のミニマックス探索を複数回行い、その探索木で得られる指し手と評価値をシミュレーション結果とする。複数のミニマックス探索により、モンテカルロ探索木のリーフでは複数個の評価値が得られる。これをそのリーフの「確率分布を持つ評価値」として扱うことにする。

乱数付加においては、合議方式と同様に、シミュレーション探索開始局面(すなわち、モンテカルロ木のリーフ局面)におけるシミュレーション回数と評価局面とをキーとした疑似乱数を生成して利用する。将棋の探索空間は合流が多いため、シミュレーション探索中に同一局面に至ったときに同じ乱数値を加えた評価値が得られるようにするためである。さらに、シミュレーション回数を乱数のキーに加えることで、モンテカルロ木の複数のリーフノードで同一局面が存在するとき、同じシミュレーション回数の時は同じシミュレーション探索結果(分布)になるような一貫性も実現できる。ただし、シミュレーション回数が異なる時は同一局面でも異なる分布となり、モンテカルロ木内で完全な一貫性を保っているわけではない。

提案手法では、モンテカルロ木での評価値伝搬において Bayesian Approach を利用し、リーフの評価値分布を考慮した探索木を構築する。評価値分布の伝搬を効率よく計算するために、リーフの評価値分布は離散的な確率分布であるとして複数のピンで表されるものとする。

図3はシミュレーションの回数とそこで得られる評価値分布を示している。シミュレーション回数が1回の際には、得られた評価値 v_1 に だけずれを加えた値にも小さな確率を与えた擬似的な分布を生成する。またシミュレーション数が2以上の時は、 $v_1 \pm \delta$ の点の擬似的な確率は削除し、それぞれのシミュレーションで得られた評価値(v_1, v_2 など)が、出現回数に従った確率で得られる確率分布であるとする。今回の実装では、1回目のシミュレーション(v_1)についてはオリジナルの評価値に乱数誤差を加えない探索を使用し、2回目以降のシミュレーション(v_2 など)にはオリジナル評価値に乱数誤差を加えて探索を行っている。

なお、将棋の場合、探索を通して局面の勝ち負けが確定していることを判定することが可能である。これを反映するため、シミュレーションの結果、詰みだと判断され勝ち負けが確定した場合には、評価値が誤差を持たない、1本のピンで表現される分布になるとした。

展開制御方式としては、ルートノードの確率分布に対して最も影響を与える度合いが

大きいリーフノードを選択して展開する、というQSS(Q Step Size)による探索制御を採用する。

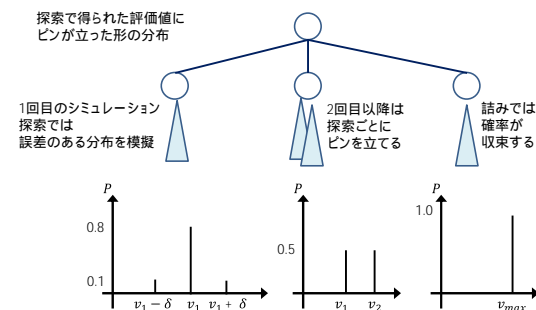


図3 シミュレーション結果による評価値分布の作成方法

(2) 評価実験

提案手法を実装し、オリジナルのプレイヤーとの対局を多数行うことで手法の性質と有効性を評価した。実装においては、コンピュータ将棋プレイヤー「激指」を利用し、その評価関数に模擬正規分布乱数を加えてシミュレーションを作成した。激指は実現確率打ち切り探索など既存の探索技法を多数導入した実用的なプログラムであり、世界コンピュータ将棋選手権などで優秀な成績を収めている。

実装したシステムを用いて、基礎的なパラメタ設定についてそれぞれ評価実験を行い、代表的なパラメタ設定を用いて提案手法の勝率と所要時間の関係を調べた。

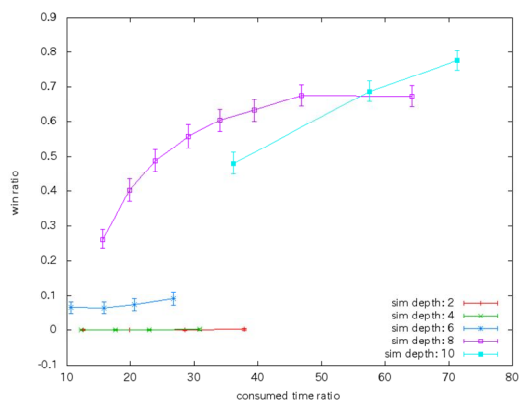


図4 PV長変更時の消費時間比率と勝率

Simdepth(シミュレーション探索の深さ)を2から10までの値で固定し、PVth(シミュレーションを呼ぶ末端ノードの深さ)を変化させたときの勝率変化を図4に示す。横軸はオリジナルプレイヤーと比較した消費時間の比である。

Simdepthが6より小さい場合には、PVthを4から16まで変化させてもほぼ勝率は横ばいであり、モンテカルロ木の深さを深くしていても効果がないことがわかるが、Simdepthを8まで大きくすると、モンテカルロ木の深さを深くすることで勝率が向上し、

提案手法で効果が得られることが確認できる。ただし、木の深さを深くしたときの勝率向上の度合いは漸減していき、深さ 12 程度で向上が頭打ちになることも見て取れる。Simdepth を 10 まで大きくすると勝率はさらに高くなり、シミュレーションサイズは大きいほど強くなるという結果が得られた。

一方、消費時間を考えると、Simdepth が 10 の場合は、8 の場合と比較して同一の PVth 設定での消費時間が 2.5~3 倍程度延びており、同一の消費時間を要する領域で比較すると、Simdepth が 8 の方が高い勝率を得られる場合もあることがわかる。つまり、使用リソースに対する効率の良さの観点では、必ずしもシミュレーションサイズを大きくした方が良いとは言えない。また、シミュレーションサイズを大きくするという事は、逐次実行部分の粒度を大きくするため、並列計算の適用を考える場合にはより不利になることが考えられる。

次に、 α を 100、Simdepth を 8 とし、PVth を変化させたときの勝率を図 5 に示す。各系列は Simnum(末端ノードのシミュレーション数)を 1 から 7 まで変えた場合である。横軸は対局で消費した時間をオリジナルプレイヤとの比率で示している。

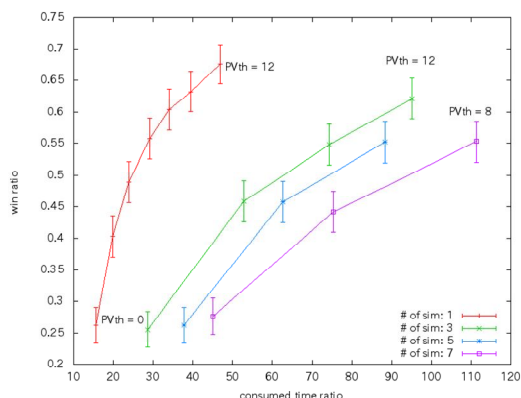


図 5 勝率と所要時間の関係

いずれの設定においても、PVth を延ばしていくことで勝率は向上し、実験の範囲ではまだ限界には達していないように見受けられる。前述したように、Simnum を増やしたときには勝率がむしろ低下しており、消費時間が伸びることも考慮すると、現在の手法で Simnum を増やすことはあまり有効ではないと結論づけられる。

シミュレーション回数を増やしたときに性能が低下する原因を探るために、シミュレーション探索で得られる探索結果値分布の性質を調査した。実戦棋譜において 20 手から 40 手目までの局面から 494 局面をランダムに抽出し、その一手後の局面と合わせた合計 988 局面を評価対象とした。それぞれの局面で、 $\alpha = 200$ 、Simdepth = 8 のシミュレーション探索を 100 回ずつ行い、得られたシミュレーション結果(乱数が付加された評価関

数を用いた探索木の値)とオリジナルの探索結果(乱数が付加されない評価関数を用いた探索木の値)との差を求めた結果を図 6 に示す。横軸はオリジナルの探索結果の値であり、正の値の時に初期局面の手番側が有利であることを示すように正規化されている。初期局面をオリジナルの探索結果値 25 点ごとの刻みで分けて集計し、縦軸に木の値の平均と標準偏差を示している。

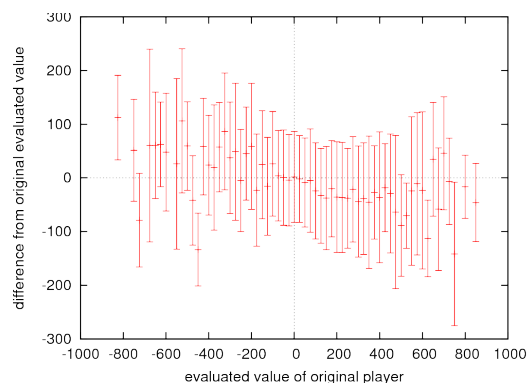


図 6 シミュレーション探索とオリジナル評価関数の探索結果のずれ

シミュレーション結果は、オリジナルの探索結果が正のときには負の誤差を、オリジナルが負だと正の誤差を持つような分布になっていることが見て取れる。つまり、自分に有利な結果が得られる手についてシミュレーションはその有利さをやや割り引いて考えがちであり、不利なときには逆にやや有利に考えがちである、という傾向にあるといえる。この誤差は、最善手と次善手の値の差を小さくする方向に働き、最善手を選びにくくする結果をもたらしている可能性がある。この誤差の大きさがどの程度探索結果に影響を与えているかは検証の必要があるが、この分布は Bayesian Approach の想定に沿っているとは言いがたい。

この結果から、単一のシミュレーション探索を利用するようなモンテカルロ木探索では、想定される勝率の確率分布を十分な精度で推定できるような確率分布を作成することができないのではないかと、との予想に達した。このことから、複数の評価観点を利用できるようにモンテカルロ木探索の手法を検討する必要があると考えられる。今後の研究課題としたい。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 1 件)

横山 大作, 喜連川 優, ベイジアンアプローチに基づくモンテカルロ木探索アルゴリズムの将棋への適用と評価, 情報処理学会論文誌, 査読あり, Vol. 55, No. 11, pp. 2389-2398, 2014

〔学会発表〕(計 9 件)

水上直紀, 鶴岡慶雅. 強化学習を用いた効率的な和了を行う麻雀プレイヤー, 第21回ゲームプログラミングワークショップ, pp.81-88, 2016.

山内智晴, 鶴岡慶雅. 将棋における個人に適応した着手推定モデルの構築, 第21回ゲームプログラミングワークショップ, pp.112-118, 2016.

河村圭悟, 水上直紀, 鶴岡慶雅. 多人数不完全情報ゲームにおける仮想自己対戦を用いた強化学習, 第21回ゲームプログラミングワークショップ, pp.188-195, 2016. (研究奨励賞受賞)

齋藤雄太, 鶴岡慶雅. 線形関数近似によるトリックテイキングゲームのQ学習, 第21回ゲームプログラミングワークショップ, pp.196-200, 2016. (研究奨励賞受賞)

川上裕生, 鶴岡慶雅. 静止探索を用いたArimaa 評価関数の比較学習, 第20回ゲームプログラミングワークショップ, 2015

水上直紀, 鶴岡慶雅. 期待最終順位の推定に基づくコンピュータ麻雀プレイヤーの構築, 第20回ゲームプログラミングワークショップ, 2015

Daisaku Yokoyama, Masaru Kitsuregawa, A Randomized Game-Tree Search Algorithm for Shogi Based on Bayesian Approach, The 13th Pacific Rim International Conference on Artificial Intelligence (PRICAI 2014), pp. 937-944, (2014)

Keisuke Suzuki, Yuto Hayamizu, Daisaku Yokoyama, Miyuki Nakano, Masaru Kitsuregawa, Comprehensive Analytics of Large Data Query Processing on Relational Database with SSDs, The 25th Australasian Database Conference (ADC 2014), pp. 135-146, (2014).

Daisaku Yokoyama, Masahiko Itoh, Masashi Toyoda, Yoshimitsu Tomita, Satoshi Kawamura, Masaru Kitsuregawa, A Framework for Large-Scale Train Trip Record Analysis and Its Application to Passengers' Flow Prediction after Train Accidents, The 18th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2014). PAKDD(1), pp. 533-544, (2014).

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

6. 研究組織

(1)研究代表者

横山 大作 (YOKOYAMA, Daisaku)
東京大学・生産技術研究所・特任助教
研究者番号: 80345272

(2)研究分担者

なし

(3)連携研究者

鶴岡 慶雅 (TSURUOKA, Yoshimasa)
東京大学・工学(系)研究科(研究院)・
准教授
研究者番号: 50566362

(4)研究協力者

なし