

論文の内容の要旨

論文題目 棋譜データに基づく、 ゲーム木探索の性能評価

氏名 竹内 聖悟

将棋やチェス，囲碁などのゲームは，ルール及びゴールが明確であるため，研究の対象として扱いやすいという利点がある．ルールとゴールが明確である一方，探索空間が広いため，単純な力任せの探索では最善手を求めることが不可能で，ゲームの知識を利用した探索や評価が必要である．そのため，人工知能やコンピュータサイエンスなどのテストベッドとして発展してきた．三目並べのように単純なゲームから囲碁のように複雑なゲームまで様々な難易度のゲームがあるため，適切な難易度の問題を選択することが可能である．また，ゲーム自体が面白く人々の興味を惹くことも魅力の1つである．

ゲームプログラミングの分野では，プログラムの強さが人間のチャンピオンレベルに達することが一つの目標として研究され，そのレベルに達したゲームも多い．強いゲームプログラムでは，先を読む「探索」と局面の特徴により優劣を判断する「評価関数」が重要である．ゲームプログラムは，局面を与えられると，探索と評価関数によって指し手に優劣をつけ，その局面における最善手を出力する．なお，探索も評価関数も与えられた局面に対して評価値と呼ばれる局面の優劣を示す数値をつけるものであり，評価関数は深さ0の探索とみなすことも可能であるため，以降，両者をまとめてゲーム木探索と呼ぶ．

ゲーム木探索の改良についての研究も数多くされてきており，近年，コンピュータ将棋では，プロ棋士の棋譜から評価関数の重みを自動学習させる手法 (Bonanza Method) が成功をおさめ，多くの開発者が利用し，成果を挙げている．また，コンピュータ囲碁では，モンテカルロ木探索と呼ばれる探索手法が多くのプログラムに用いられ，大きな成果を挙げている．ゲーム木探索の改良には新たな特徴の追加やパラメータの調整などがあり，それらが有効であるか改良前後での性能評価が必要となる．

性能評価の従来手法としては、改良前後のプログラムを用意して互いに対戦 (自己対戦) させる、他のプログラムと対戦させる、問題集を解かせるなどを行い、結果の比較により性能を評価する手法がある。これらの手法には、有意な結果を得るのに必要な時間が長時間となることや、対戦相手に特化しすぎたプログラムになる可能性があること、問題集に現れる局面に偏りがあること、性能の評価はできても次の改良に有用な情報が得られないことなどの問題点がある。

本論文は棋譜データに着目し、棋譜データを用いた効率的なゲーム木探索の性能評価手法の構築を課題とし、性能評価手法の提案とその評価を行った。ゲーム木探索の改良において性能評価は重要な要素であるが、性能評価そのものを対象とした研究はなく、本研究は独創的な研究である。

性能評価の時間の短縮や特徴の有用性の評価が可能となれば、効率的にゲーム木探索の改良を行うことができ、強さの向上へとつながる。Bonanza Method により評価関数の重みの調整を自動で行うことが可能になったが、未だにどの特徴を用いるべきかを判断する方法はなく、人間の知識によって判断されている。よって、どの特徴が有用であるかが判定できるのであれば、プログラムの強さの向上に非常に有用であると考えられる。

本論文では、評価値と棋譜の勝敗との単調性と一貫性から評価を目視で行う Evaluation Curve や機械学習の分野で用いられている分類器の性能評価手法をゲームに応用した勝敗予測の正確さの指標、ケンドールの順位相関係数を評価値と勝敗の順位に応用した単調性の数値指標、評価値の分布を利用した期待値指標を提案した。

棋譜データには様々なプレイヤーが存在するため、従来手法の問題点の1つである対戦相手への特化を避けられる他、現れる局面が実際のゲームの局面であるので現れる局面の偏りを解消できるという利点がある。また、棋譜の各局面の評価値が必要であるが、評価関数による評価はほとんど時間がかからず、探索を行う場合であっても短い時間の探索であるため、既存の手法に比べ短い時間で評価が可能である。

はじめに、Evaluation Curve について説明する。まず、何らかの手段によって全局面について勝率が分かっている場合を考える。この場合、評価値の大小と勝率の大小が一致するか、つまり評価値と勝率との間に単調性があるかによって性能評価ができる。また、ある特徴について正しく評価できているかを調べたいとすると、その特徴を持った局面だけを集め、評価値と勝率の関係が他の場合と変わらないか、つまり一貫性を見ることでその特徴についての性能評価が可能となる。しかし、実際には局面の勝率を知ることはできず、全局面を網羅することも不可能である。そこで、勝率を求めるために多数の棋譜データを用意し、棋譜の勝敗を利用して勝率の近似を行う。まず、棋譜に現れる局面にその棋譜の勝敗をつけ、評価値を求める。同じ評価値を持った局面を集め、それらの勝敗から勝率を近似する。つまり、局面毎に勝率を求めるのではなく、同じ評価値を持つ局面に対して勝率を求めている。そして、これらの関係を利用する手法として、評価値に対して勝率をプロットする Evaluation Curve を提案した。単調性は得られるグラフが単調増加かによって確認が可能であり、一貫性はある特徴を持つ局面だけを対象としてプロットした場合にグラフが変化するかによって確認が可能である。

続いて、Evaluation Curve で説明した単調性についての数値指標を説明する。ケンドールの順位相関係数とは2つの順位間の相関を測る手法であり、評価値と勝敗の間の相関を計測することで、評価値と勝敗の単調性の数値指標となる。そこで、ケンドールの順位相関係数を単調性の数値指標として利用することを提案した。

Evaluation Curve には勝率の情報はあるが、その勝率の計算にどれだけの局面が現れたという情報はない。別々の棋譜から作った2つの Evaluation Curve があり、それぞれ、勝率 0.5 付近、

勝率 0.0 と 1.0 付近にデータが集まっているグラフであるとする．この場合後者の方が良いが，Evaluation Curve では区別がつかない．そこで，先手勝ち，負けの時それぞれの評価値の期待値を利用した指標として，期待値指標を提案した．これにより，評価値の分布を考慮した性能評価が可能となり，Evaluation Curve の補完が可能になった．

最後に，機械学習の分野において分類器の性能評価を行う手法のうち，ゲーム木探索に適した 6 つの手法をゲーム木探索の性能評価手法として利用した．局面の評価値が正なら先手の勝ち，負なら先手の負けとして分類器による分類，そして棋譜の勝敗を実際の分類と見て，機械学習における分類器の性能評価手法を勝敗予測の正確さの指標として提案した．

提案手法による性能評価の有効性を示すため，チェスや将棋，囲碁を対象としてゲーム木探索の性能評価を行い，自己対戦や経験的知見との比較を行った．

まず，チェスと将棋において，評価関数を対象として実験を行った．チェスにおいては，ある特徴を正しく評価できていないプログラムとして，全特徴の中からその特徴を除去したプログラムを使うことにして，除去されたプログラムと元のプログラム，その特徴の重みについて学習を行ったプログラムの 3 つの性能を評価した．特徴を除去したプログラムは弱くなると予想されるが，自己対戦でも他の 2 つのプログラムに有意に負け越した．また，Evaluation Curve では，その特徴についてのグラフが他のグラフから分離することが観察され，その特徴について正しく評価できていないことと符合した．Evaluation Curve による性能評価と自己対戦による評価は一致しており，ある特徴について問題があるかの判定にも成功したと言え，提案手法の有効性を確認できた．将棋においては，終盤において有効な特徴である王の危険度の差を対象とし，その特徴について人間が経験に基づいて調整したプログラム，自動調整を行ったプログラム，その特徴について評価を行わないプログラムの 3 つの性能を評価した．チェス同様に，その特徴について評価を行わないプログラムでは，グラフが分離することと，自己対戦で有意に負け越すことが確認され，将棋においても Evaluation Curve による性能評価の有効性を確認できた．

さらに，囲碁において，モンテカルロ木探索を対象として実験を行った．モンテカルロ木探索の強さについて経験的に知られている知見として，シミュレーションの数が多いと強い，単純なモンテカルロ法よりもゲーム木探索との組み合わせの方が強い，ゲームの知識を利用した拡張を行う方が強い，などがある．これらの強さに関する知見について，モンテカルロ木探索の囲碁プログラムを使い，Evaluation Curve，勝敗予測の正確さの指標，単調性の数値指標，期待値指標による性能評価を行った．実験結果では提案手法と強さに関する経験的な知見とが一致し，提案手法によるゲーム木探索の性能評価の有効性を示すことができた．また，プレイアウト数が異なる同じ評価値でも勝率が異なることを示し，異なるプレイアウト数の勝率の比較を行っている手法に危険があることを発見した．

これらの実験結果から，提案した性能評価手法が様々なゲーム，ゲーム木探索において有効であることを示すことができた．