

論文の内容の要旨

論文題目 Automatic Construction of State Evaluator
toward General Game Player

氏名 金子 知適

囲碁や将棋などの思考ゲームには、ルールが単純であるにもかかわらず状態空間が大きいいため、力任せの探索により最善手を求めることは不可能であるという特徴がある。従ってゲーム特有の知識を探索に組み込む研究が必要であり、人工知能の応用として魅力的な研究分野となっている。特に機械学習の観点からは、思考ゲームはルールが明確で扱いやすいことと訓練例が豊富に入手可能であることから、学習の良い題材である。

本研究では、汎用ゲームプログラム (general game player) を作ることを目標に、思考ゲームの局面評価器 (state evaluator) を自動生成する手法の提案とその評価を行った。汎用ゲームプログラムは、(未知の) ゲームのルールを与えられると、そのゲームの強いプレイヤーとしてふるまうプログラムである。汎用ゲームプログラムの作成のためには、ゲームのルールをもとにプログラムがゲームの指し方を学習する手法が必要である。プログラム中のゲームの指し方に相当する知識を、局面評価器と呼ぶ。局面評価器は、評価関数など探索で使われるヒューリスティクスを実装したものである。評価関数とは、局面を勝ちやすさの観点から数量化する関数である。たとえば、将棋のルールを覚えることは比較的容易でも将棋の勝ち方を習得するのは困難なように、正確な評価関数を定義することは大変困難である。他の重要なヒューリスティクスとして、終盤でこの局面が勝ち(負け)であることを証明するには探索でいくつの局面を読む必要があるかを推定する、証明数予測器がある。どんな探索アルゴリズムを用いる場合でも、強いゲームプレイヤーを作るためには、正確なヒューリスティクスとそれを効率よく実装した局面評価器が必要となる。

ヒューリスティクスは、局面を特徴ベクタに変換するための複数の特徴関数と、特徴ベクタから勝ちやすさを予測するモデル(線形モデルやニューラルネットワークなど)からなる。よく使われる特徴関数には、オセロにおける「打てるマスの数」「ひっくり返らない石の数」や将棋における「駒の損得」「玉の安全度」などがあり、ゲーム毎に適切な特徴関数は異なる。

本研究では、ゲームのルールから、そのゲームに特化した局面評価器を自動生成することを目的とし、そのために特徴関数の獲得、モデルのパラメータの調整、効率の良い処理プログラムの作成を自動で行う手法を開発・統合した。また、オセロにおける実験により、世界最高のプログラムの成績に近い優れた局面評価器を自動生成することに成功した。

本研究には以下の意義がある。第一に、強いゲームプレイヤーを作ること目標とする思考ゲームの研究において、良い評価関数を得るために重要である。これまでにモデルのパラメータ調整を自動化し成果をあげている研究は多いが、それらの研究では特徴関数自体は人間が与えていた。しかしながら、ゲームが複雑になるほど、人手によって有効な特徴関数を作成することは難しく、将来囲碁や将棋のような複雑なゲームで、または歴史が浅く人間の専門家が少ないような新しいゲームにおいて強いプログラムを作るためには特徴関数の自動獲得が不可欠である。第二に、汎用性は応用の点から重要である。一般にゲームに限らず機械学習やパターン認識などの分野では、モデルの学習が成功するためには、良い特徴関数により入力を性質の良い特徴空間に変換してからモデルに与えることが重要であることが知られている。しかしながら、良い特徴関数をどのようにすれば得られるかに関してはまだ手法が確立されていない。特定のゲームに関する特徴関数の研究ではなく、様々なゲームを扱える汎用的な手法の研究を進めたことで、将来、他の分野の特徴関数を設計する問題への応用が可能であると考えられる。

本手法の特徴関数生成の部分は、ゲームのルールを述語論理で記述し、その背景知識を手がかりに特徴関数を自動生成するという手法に基づいている。この生成手法は対象となるゲームの性質に依存しないため、ユーザがルールを述語論理で記述するだけで様々なゲームに適用できる汎用的なものである。特徴関数を述語論理で表現することで、統一的な特徴関数の扱いが可能になっている。この生成手法のアイデアは以前からあったが、以下の問題点により、これまで十分な実験・検証が行われてこなかった。それは、述語論理で表現した特徴関数は、他の図形的パターンのような表現方法と比較して、局面評価の効率が非常に悪いという点である。それに伴い(1)探索の際に十分な効率が得られない、(2)必要な数の特徴関数を扱えないという問題が生じ、大規模なデータを用いて学習を行うことは事実上不可能であった。本研究では、問題(1)について部分計算と論理最適化の技術による変換で解決し、問題(2)については、(1)で変換した論理特徴をさらに図形的パターンに変換することで解決した。

まず述語論理で表現された特徴関数を評価する効率を改善するために、特化した局面評価器を生成する手法を提案・実装した。特化した評価器は、部分計算・論理最適化の技術の組み合わせにより作成され、カウンタを用いた差分計算により高速に局面を評価する。実際に特化した局面評価器を生成するプログラムを実装して実験した結果、特徴関数の評価効率を飛躍的に改善した事

を確認している。オセロで実際の探索を用いた実験では、本手法は演繹データベースなど標準的な手法に対して4,000倍以上の効率向上を示し、ネットワークの最適化により、さらに最大約40倍の効率化が可能であることが示された。オセロだけでなく詰将棋でも効果を確認している。またメモリー効率も優れたものであった。

この効率の良い局面評価器の実現により、問題(1)の評価関数の局面の評価速度についてはかなり解決し。またこの効率の改善により問題(2)の特徴関数の数についても、約1万種類の特徴関数を扱うことに成功している。過去、述語論理を使用した研究においては約200種類程度であった。

さらにその手法をもとにして、論理特徴を図形的パターンに変換する手法と、そのようなパターンに特化した局面評価器を構築する手法を開発した、これにより、さらに30倍近くの効率の改善を実現した。オセロにおいて1秒間に10万局面程度の局面評価速度を実現し、効率については人手で書いたプログラムに近づいている。また開発した手法を総合的に評価するために、特徴関数獲得、パラメータ調整、局面評価器への変換を総合した大規模な実験を行い、評価関数と証明数予測器の2種類のヒューリスティクスの学習を行った結果、正確さについても大幅に向上し、世界最強のオセロプログラムで使われている評価関数の性能に近づきつつあることを確認した。現在までにオセロでもっとも良いとされている特徴関数は、約30万種類の図形的パターンを使うものである。しかし図形的パターンはどの形が重要であるかがゲームごとに異なり、人間がそれを与える必要があったため、ゲームに依存しない手法でのパターンの獲得は困難であった。本研究では、まず論理特徴を獲得し、それらをパターンに変換することでその問題を回避し、汎用ゲームプログラムが人間の助けなしにパターンを獲得できるようにしたと言える。

今後の課題は、本手法を様々なゲームに応用し、手法の汎用性を示すことである。特に、必要な特徴関数が作れるかどうか、また、不要な特徴関数の生成を避ける方法があるかについて、広い範囲のゲームで検証することが重要である。