ハーモニーサーチアルゴリズムとニューラルネットワークによるオセロ戦術の学習

A1178647 岡田 良健

指導教員 ゴンサルベス タッド

1. はじめに

　世の中には様々な目的に合わせて最適化問題が存在する．ゲームも最適化問題としてとらえることができる．ゲームにおいて各プレイヤーは自分の利益の最大化を目的として行動を決定する．本研究は，「大きな複雑度を持つ問題に対し専門知識をベースとせずに，機械学習によって状況評価を学習すること」を，オセロAIの状況評価の役割を持つニューラルネットワークとその重みを改善するハーモニーサーチアルゴリズムを組み合わせた機械学習により求めることを目的としている．

1. オセロ

　オセロ（あるいはリバーシ）は，二人零和完全情報確定非対称有限非協力交互進行収束ゲームに属する．オセロのプレイヤーの目的は勝利であり，各局面においては優勢になるように行動をする．単純なルールながらゲームの複雑度は人間がゲーム木の全展開を把握可能な程度を越えており，ゲーム探索空間は1060ほどである．一般にゲームAIは，状況の評価をする評価関数またはそれにゲーム木の探索アルゴリズムを追加したものに従って行動をとる．また，一般にゲームAIは専門知識をベースとした評価関数を持つ．

1. 研究方法

　本研究では, 評価関数の改善に焦点を置いた．ただし，専門知識をベースとせずに，個体群ベースのメタヒューリスティックアルゴリズムの一種であるハーモニーサーチアルゴリズムと，機械学習の手法であるニューラルネットワークを組み合せることにより，評価関数の改善を行った．また，探索アルゴリズムには一般的なαβ探索法を用いた．これらにより，機械学習を利用したオセロの状況評価の改善の実験に取り組んだ．

* 1. ハーモニーサーチアルゴリズム  
      (Harmony Search Algorithm)

　音楽家の調和の探し方をアイデアのベースにして作られた個体群ベースのメタヒューリスティックアルゴリズムである．組み合わせ最適化やパラメータ最適化に適している．本研究ではニューラルネットワークの重みの最適化のためにハーモニーサーチアルゴリズムを用いた．また，進化の過程は個体間の対戦の繰り返しと勝者の生き残りである．

以下に，ハーモニーサーチアルゴリズムの擬似コードを示す．

fitness function F(A), A = [ a1, a2, … adim ]

parameters (dim, hms, hmcr, par, mi)

parameters (fw, min, max) (: if continuous-value)

set of candidate values of decision variables  
 Xi = {xi(1), …, xi(Ki)} (: if discrete-value)

HM ← initial harmonics[ A1, A2, …, Ahms ]

**while** it < mi or other termination condition

**for each** aj new in Anew

**if** rand\_float(∈[0,1)) < hmcr

aj new = aj i (in Ai in HM, i = rand\_int(∈[1,hms]))

**if** rand\_float(∈[0,1)) < par

aj new = aj new + fw･rand\_float(∈[-1,1])

(: if continuous-value)

aj new = xj(k+m) (aj new was xj(k), m∈{-1,1})

(: if discrete-value)

**else**

aj new = rand\_float(∈[min, max]) (: if continuous-value)

aj new = xj(k) (k = rand\_int(∈[1, Kj])) (: if discrete-value)

**if** F(Anew) > F(Aworst in HM)

replace Aworst with Anew

best = Abest in HM

図1 ハーモニーサーチアルゴリズムの擬似コード

* 1. ニューラルネットワーク (Neural Network)

[神経](http://ja.wikipedia.org/wiki/%E7%A5%9E%E7%B5%8C)ネットワークの構造と機能を模した数学モデルであり，学習アルゴリズムである．本研究ではニューラルネットワークのモデルとして，代表的な人工ニューラルネットワークモデルである多層パーセプトロンを用いた．多層パーセプトロンは，入力に対し，入力層，中間層，出力層の順に信号を伝え，結果を出力するものである．機械学習アルゴリズムの分類としては，入力とそれに対応すべき出力が存在するため教師あり学習である．

　以下に多層パーセプトロンの構造を示す．

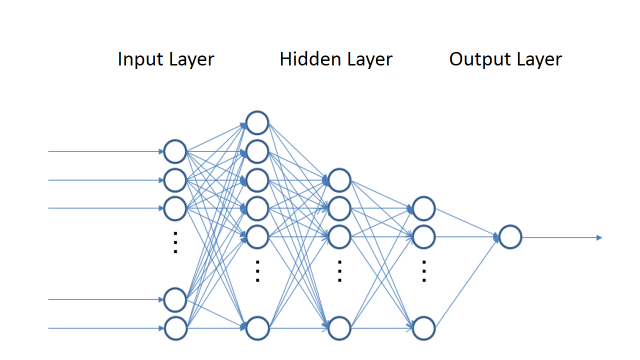


図2 多層パーセプトロンの構造

1. 改善度

　一般に複雑なゲームにおける“ゲームの強さ”を測る絶対的な指標はなく，正確に絶対値で表すことや単純な式によって算出することはできない．そのためにゲームAIにおいて改善度の確認は，実際に対戦を行うなどして強さの目安となるものを測定する必要がある．ただし，特定の対戦相手または戦術に対してではなく，あらゆる戦術に対して強い戦術へと改善しなければいけない．これは学習において，ゲームの戦術を本質的に学習するということである．

　本研究では，ハーモニーサーチアルゴリズムによる進化の各世代において複数個体間での対戦により強さの目安の測定をし，改善度を保証した．

1. 実験環境

　本研究では，プログラミング言語としてJavaを用いた．開発環境は Intel Core i7-3970X CPU 3.50 GHz プロセッサで行った．学習の実行環境はMicro Trend Server CPU 32 core RAM 1000 GBで行い，実行時間は8128.82分であった．

1. おわりに

　学習後に対戦成績をとった．対戦は，予め知識が組み込まれた2つのオセロAIに対して各50回と，学生36人に対して各約3回（計126回）である．結果の勝率は, 学習前に約25%だったAIに対して99%，学習前は約10%だったAIに対しては97%を超えた．一般人に対しては80.16%となった．以上から，ハーモニーサーチアルゴリズムとニューラルネットワークの組合せによって，専門知識をベースとせず，特別な対戦相手ではなく個体間の対戦の経験だけを経て，状況評価の学習をし, オセロの戦術を独自に理解したいえる．

　課題としては，パラメータ調整やアルゴリズムの改良による学習精度向上，探索の高速化による学習時間短縮，改善度の測定精度向上，特徴抽出の自動化や最適化などがある．今回の学習方法をベースとし，他分野への応用を行いたい．

1. 参考文献

[1] Zong Woo Geem, “State-of-the-Art in the Structure of Harmony Search Algorithm”, Recent Advances In Harmony Search Algorithm

Studies in Computational Intelligence, vol. 270, 2010, pp. 1-10.

[2] Siang Y. Chong, Mei K. Tan, and Jonathon D. White, “Observing the Evolution of Neural Networks Learning to Play the Game of Othello”, IEEE Transactions on Evolutionary Computation, vol.9, no.3, June 2005, pp. 240-251.