

戦略パッケージ法を用いた競合共進化による 適応型戦略の獲得*

1 U-4

根路銘 もえ子 遠藤 聡志 山田 孝治 宮城 隼夫 †

琉球大学工学部‡

1 はじめに

ゲーム戦略獲得問題における適応アルゴリズムとして、陽に評価関数を設定せずに相対的な評価により解の探索を行う競合共進化アルゴリズムが有効であると考えられている。しかしながら、実ゲームにおいては、全ての戦略を打破する1戦略が存在することは稀であるため、最良戦略を決定することは困難である。そこで、本稿では、競合共進化アプローチに複数の戦略を1つの戦略セットとするパッケージの概念を導入したアルゴリズムを提案する。さらに、本アルゴリズムをTic-Tac-Toeゲームへ適用し、その有効性を検討する。

2 競合共進化

競合共進化とは、他の個体と直接競合することにより相対的な評価を受け、その評価を基に進化することで、全体としてより良い方向へと進化する自然のメカニズムのことである。このメカニズムは、工学的アルゴリズムとしても様々な問題への応用がなされている[1]。

競合共進化モデルは、まず、遺伝的に区別された2つの集団で構成され、一方の集団内の1個体の適応度は、他方の集団内の個体との競合により決定される。各々の集団の個体が交互に評価されることにより適応度は更新され、集団は相互作用的に進化する。このように、状況に適応した評価値に基づいて進化するため、対戦戦略に応じて戦略の評価が異なるゲーム戦略獲得問題に対して有効であると考えられる。

3 競合共進化によるゲーム戦略獲得

競合共進化の有効性を示すために、全順序および半順序戦略ゲームに適用し、これらのゲームにおいて最良解を獲得した[3]。そこで、本稿では、実ゲームにおける有効性を検証するために、Tic-Tac-Toeゲームに適用する。本稿で対象とするTic-Tac-Toe(TTT)ゲームは、最も基本的な二人零和有限完全情報確定ゲームである。ある局面に対する対処法をコード化したものをここで

は、戦略と呼び、競合共進化により最良戦略を求めることを目的とする。

3.1 TTT 戦略のコーディング

TTTゲームに競合共進化モデルを適用するために、戦略をコーディングする必要がある。1つの戦略は先手と後手の戦略を持っているため、ゲーム盤のマス数である9個を1つの配列としている。この配列の中には、ゲーム盤のどこに自分の手を置くかという判断の数字が入る。戦略のコーディング法を図1に示す。

strategy	1-9	1-8	1-7	1-6	1-5	1-4	1-3	1-2	1
----------	-----	-----	-----	-----	-----	-----	-----	-----	---

図1: 戦略のコーディング

このコーディング法を用いることで、全ての個体に対して実行不可能な対戦がなくなるという特徴がある。また、一般的遺伝子操作においても、実行不可能な戦略を生成しない。

3.2 シミュレーション

上記の様にコーディングされた戦略集団を2集団生成し、一方を先手戦略集団、もう一方を後手戦略集団としてそれぞれを進化させる。各世代で獲得された戦略をランダムに用意した評価戦略500と対戦させ、各世代中で最も高い勝率を図2に示す。評価戦略に対して、先手集団における最も勝率の高かった戦略は、第18世代の戦略であり、500の評価戦略に対して、勝ち...404, 引き分け...81, 負け...15という成績を治めている。しかしながら、次世代における最良戦略は極端に悪くなり、それ以降、評価値は振動を繰り返し、18世代目の戦略と同等以上の強さの戦略は獲得されていない。また、後手戦略においても評価値の振動現象が確認された。このような振動現象が起こる原因として、18世代目の戦略は、次世代における敵戦略に対して強さを発揮できなかったために淘汰され、別の戦略が生成されたためと考えられる。このように、1戦略のみでは、ある局面に対してだけの最良戦略となってしまう。従って、全ての局面における最良戦略の獲得には、複数の戦略を必要とするのではないかと考えられる。

*Competitive Co-evolution Based the Acquisition of Strategy with Packaging Method

†M.Nerome S.Endo K.Yamada and H.Miyagi

‡Faculty of Engineering, University of the Ryukyus

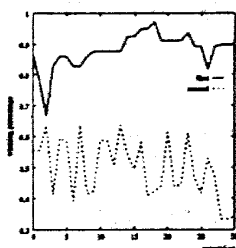


図 2: TTT における戦略の勝率

4 戦略のパッケージ化

最良戦略を生成するために複数個の戦略を必要とするゲームにおいては、最良戦略をただ1つ決定することはできない。そこで本稿では、複数個の戦略をパッケージ化し、パッケージ内の戦略数が最小かつ最良戦略であるパッケージを獲得するアルゴリズムを提案する。

4.1 パッケージ内の戦略

パッケージ内の戦略同士は、お互い補填的である必要がある。例えば、2つの戦略 S_1 , S_2 でパッケージが構成されている場合、敵 O_1 , O_2 に対して以下の条件を満たさなければならない。

condition 1: $S_1 > O_2$, $S_2 < O_2$

condition 2: $S_1 < O_1$, $S_2 > O_1$

このように、お互いを補填し合う戦略でパッケージを構成することにより、より強いパッケージの構成が可能となる。ここで、 $S_1 > O_2$ は、 S_1 が O_2 に勝つことを意味する。

4.2 戦略パッケージ法アルゴリズム

戦略パッケージ法を用いた競合共進化アルゴリズムを以下に示す。

step1: 2集団 ($Pop1$, $Pop2$) を生成 (1集団内のパッケージ数は n , 1パッケージ内の戦略数はランダムに設定)

step2: $Pop1$ の敵として $Pop2$ から戦略をサンプリング

step3: $Pop1$ の各戦略は敵と対戦し、結果から各評価値を計算

step4: $Pop1$ のパッケージの進化

step5: $Pop1$ の戦略の進化

step6: $Pop1$ が $Pop2$ に対して優位性を示すか終了条件を満たすまで、step3 から step5 までを反復

step7: $Pop2$ と $Pop1$ の役割を代えて、step2 から step6 までを反復

step8: 本アルゴリズムの終了条件を満たすまで、step2 から step7 までを反復

4.2.1 パッケージの進化

パッケージの進化アルゴリズムを以下に示す。

step1: 2パッケージをランダムに選択

step2: 2パッケージに交差、突然変異を適用

step3: 評価値の低いパッケージを淘汰し、新しいパッケージを集団に追加

4.2.2 戦略の進化

戦略の進化アルゴリズムを以下に示す。

step1: 2戦略をランダムに選択

step2: 2戦略に交差、突然変異を適用

step3: 新しい戦略をパッケージに追加

step4: 追加されたパッケージを敵と対戦させ、評価値を計算

step5: パッケージ内の各戦略の評価値を比較し、新しいパッケージを決定

4.3 シミュレーション

本アルゴリズムの有効性を示すために、TTT ゲームに適用する。各世代で獲得されたパッケージ戦略をランダムに用意した評価戦略 500 と対戦させ、各世代中のパッケージの平均勝率を図 3 左図に示す。また、各世代の最良パッケージ内の戦略数の変化を図 3 右図に示す。

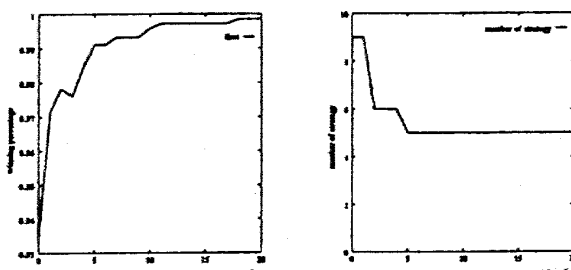


図 3: TTT におけるパッケージの勝率および戦略数

図 3 の左図より、勝率の上昇が確認できる。最終世代の最良パッケージは、500 の評価戦略に対して、5 戦略で勝ち...500 という成績を治めている。また、右図より、パッケージ内の戦略数が減少し、戦略数 5 に収束していることが確認できる。以上のことより、TTT ゲームにおいて本アルゴリズムは、最良戦略パッケージの獲得が可能であるといえる。

5 おわりに

本稿では、複数戦略を 1 セットとする戦略パッケージ法を導入した競合共進化を提案した。また、シミュレーションにより本アルゴリズムにより有効な戦略パッケージが獲得可能であることを示した。

参考文献

- [1] C. D. Rosin and R. K. Belew. (1995) Methods of Competitive Co-evolution (Finding Opponents Worth Beating). In Sixth International Conference on Genetic Algorithms.
- [2] Gerhard Weiß. (1994) Some Studies in Distributed Machine Learning and Organizational Design. Techn. Rep. FKI-189-94. Institut für Informatik, TU München.
- [3] M. Nerome, K. Yamada, S. Endo and H. Miyagi. (1996) Competitive Co-evolution Model on the Acquisition of Game Strategy. In First Asia-Pacific Conference on Simulated Evolution and Learning (SEAL'96), pp.357-364.