

木村将治* 西野順二** 小高知宏** 小倉久和**

* 福井大学大学院工学研究科 ** 福井大学工学部

1 はじめに

対戦型ゲームにおいて、相手の戦略知識を獲得することは有用であり、そのために相手の戦略などの知識表現と、その獲得の方法が必要である。

本研究では、対戦型ゲームの一つとして、じゃんけんを取り上げる。じゃんけんを繰り返して行う、繰り返しじゃんけんゲーム（以下 RJG）を対象にする。じゃんけんを繰り返すことで、相手が前に出した手を参考に次の手を考えることで、じゃんけんに戦略的要因を取り入れることができる。

対象プレーヤーはあらかじめプログラムで用意し、その戦略知識の獲得を目的とする。その手法には、遺伝的アルゴリズム（以下 GA）を用いた。GA の戦略の表現方法は、遺伝子コードとして G、T、P を持つ適当な長さの個体を生成し、その長さを繰り返し長とする、繰り返し戦略表現とした。

2 GA の戦略知識の表現方法

じゃんけんは、グー、チョキ、パーの手が、3すくみの強弱関係にある（以下グーは G、チョキは T、パーは P と表す）。本研究では、プレーヤーを 2 人に限定し、引き分けも 1 回とカウントして、1000 回の対戦で行う。

GA で獲得する戦略のモデルは、繰り返し戦略表現を用いる。遺伝子コードとしてグー、チョキ、パーをそれぞれ 1、2、3 に対応させ、それらを遺伝子とする適当な長さの個体を生成する。例えば遺伝子表現された戦略の例として、112233 という遺伝子の個体は、GGTTPP という、周期 6 の手を繰り返す戦略を表す。そして、その遺伝子長を繰り返し長とする、繰り返し戦略によって自分の手を出す。生成したそれぞれの個体と、RJG の戦略を対戦させ、得られた評価値をもとに、選択淘汰、交叉、突然変異を行いながら、RJG の戦略に勝る戦略の獲得を試みる。

繰り返し戦略は、固定的で単純な戦略であり、相手の過去の手を参照することもない。しかし、相手の出す手に規則性のようなものがあるとすれば、うまく繰り返しの長さを出す手を工夫すれば、相手の戦略に勝ちやすいパターンにおとしいれるような戦略が獲得できると考える。

3 繰り返しじゃんけんゲームの戦略

以下に、本研究で相手として用意した繰り返しじゃんけんゲームの自動プレイヤーの戦略を説明する。

- 一筋戦略
3 種類の手の中で一つだけを選択し、試合中ずっとその手を出し続けて、他の 2 種類の手は出さない戦略である。G 一筋戦略、T 一筋戦略、P 一筋戦略の 3 種類がある。
- 繰り返し戦略
GTPGTP... といった、あらかじめ用意したパターンを繰り返す戦略である。
- 仕返し戦略
相手の手と同じ手を、自分の手として出す戦略である。
- 場合分け戦略
1 回前の勝負の結果が、勝ち、引き分け、負けの場合に応じて、その次の手を 3 種類のいずれかに決めておく戦略である。
- 累積戦略
相手が過去に出した手を累積し、最も多く出している手に勝つ手を出す戦略である。過去何手分を覚えるかによって、異なる戦略となる。
- 履歴戦略
相手の過去の手を履歴として記憶しておき、それから相手の手を予測して出す手を決める戦略である。記憶するパターンの長さ n により、履歴 n 戦略と呼ぶ。
- 相互履歴戦略
相手の手と同時に、自分の手も記憶していく履歴戦略である。相互履歴戦略は、履歴 1 で相手と自分の 1 手前の手と、それに続く手を記憶するため、全部で (手の種類)³ (3 手分) のパターン、つまり $3^3 = 27$ 種類がある。履歴 2 では、 $3^5 = 243$ 種類にもなり、記憶する履歴を長くすると同じパターンが出てくるまでに時間がかかるのが欠点である。

これら RJG の対象プレーヤー戦略と、GA の戦略を対戦させ、それぞれの RJG の対象プレーヤー戦略に対して有効な戦略の獲得を目指した。

4 GAによる戦略知識の獲得方法

初期集団は、最大遺伝子長以内の個体を G、T、P にそれぞれ対応する、1、2、3 をランダムに決定し、初期遺伝子として生成する。

こうしてできた各個体と、RJG の戦略を対戦させ (図 1)、そこで得られた得点から評価値を与え、その評価値に応じて選択、交叉、突然変異を行い次の世代を作る。

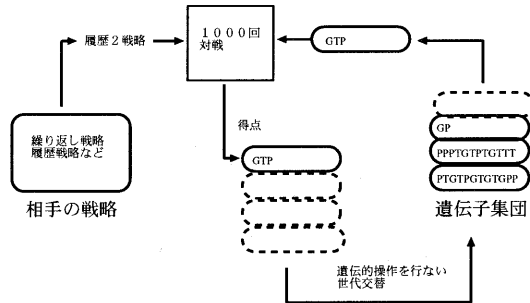


図 1: 対戦の流れ

評価値は、次の3通りの評価関数で与える。

1. 勝ち * 2 + 引き分け
2. (勝ち * 2 + 引き分け) / 遺伝子長
3. (勝ち * 2 + 引き分け) - 遺伝子長

式1は、勝ちを2点、引き分けを1点としてその合計を評価値とする方式、式2は、長い遺伝子長の個体に淘汰圧力を加えるために、遺伝子長で割るものである。これにより、遺伝子座1つあたりの貢献度評価が出るようになる。式3は、遺伝子長で引く方法である。式2に比べて、長い遺伝子への圧力は小さく、早い世代で短い個体に収束してしまうのを防ぐことを期待できる。

5 実験

各戦略に対して、GA の戦略による獲得実験を行なった。実験対象を、3章に挙げた戦略から選び、その戦略に対して4章に示した3つの方式で評価値を与えそれぞれ実験を行なった。実験に用いたパラメータは表1の通りである。最も良かった方式についての結果と式の番号を表2に示す。

表から分かる通り、単純な戦略に対しては、全勝できるような戦略を獲得できている。また履歴2戦略に対しても、大きく勝ち越す戦略を獲得できている。しかし、相互履歴1戦略に対しては、僅かに勝ち越してはいるものの、有効な戦略を獲得することはできなかった。

表 1: 遺伝的パラメータ

個体数	100
最大世代	100
最大遺伝子長	100
交叉率	0.6
突然変異率	0.05

表 2: GA と各戦略を対戦させた結果

相手の戦略	勝ち	負け	引き分け	方式
G 一筋	1000	0	0	2
繰り返す PCCPGCGGP	1000	0	0	3
1 手前仕返し	1000	0	0	1
場合分け GTP	1000	0	0	2
1 0 手累積	1000	0	0	1
履歴 2	430	143	427	2
相互履歴 1	323	305	372	1

6 まとめと今後の課題

今回の実験では、単に勝ち数が多いだけでなく、できるだけ簡単な形で戦略を獲得することを目指し、3種類の式で評価値を与え、それぞれ実験を行なった。結果として、ほとんどの戦略に対していずれかの方式を用いることで有効な戦略を獲得する事ができ、RJG の戦略に対する戦略知識の表現方法として、繰り返す戦略表現は有効である事が分かった。また、獲得の手法として用いた GA も、予想しなかったような短い個体で有効な戦略が得られるなど、期待した以上の効果が得られた。

今後は、これを人間の系列に用いて、有効な戦略を獲得できるか実験する。それにあたり、対戦の回数を人間が集中して続けられるように1000回から30回にすること、対戦の途中までで知識獲得ができるようにし、その結果をリアルタイムに対戦に反映できるようにすることなどが課題である。

参考文献

- [1] 北野宏明編 "遺伝的アルゴリズム" 産業図書 (1993)
- [2] メラニー・ミッチェル著、伊庭斉志監訳 "遺伝的アルゴリズムの方法" 東京電機大学出版局 (1997)