

進化的手法を用いた繰り返しジャンケンにおける戦略モデルの獲得および評価

田上 嗣能 (福井大学大学院工学研究科)

白井 治彦・高橋 勇・黒岩 丈介・小高 知宏・小倉 久和 (福井大学)

1 はじめに

対戦型ゲームについて、様々な研究がなされている。ゲームの研究においてその目的は、ある局面におけるパターン認識の側面と論理的な推論とについて研究されている [1][2]。また、コンピュータを用いることで、より高度な戦略が得られるようになってきている。このように、ゲームにおいて、相手の戦略知識を獲得することは重要なことであり、そのための戦略表現と獲得が必要である。

本研究では、対戦型ゲームにおける相手の戦略知識の獲得手法として遺伝的プログラミング (以下 GP) を用いる。GP では、構造的表現を扱い学習することが可能であるため [3]、関数などにより表現した戦略そのものを環境に適応する形へと進化させることができる。そのため、より環境に適応した個体を得ることで、対戦相手の戦略知識の獲得を行うことができると考えられる。

戦略知識の獲得を行う対戦型ゲームとして、今回は繰り返しジャンケンゲームを設定する。ジャンケンは通常一回の勝負では勝敗の結果に対する期待値に変化はないため戦略がない。そこで、ジャンケンを繰り返し行うことで、過去の履歴から次の手を決めるといった戦略的要因を取り入れることができる。

本研究では、このような繰り返しジャンケンゲームにおける戦略知識の獲得として、まずその再現性の評価に固定戦略の戦略知識の獲得および評価を行った上で、逐次的な学習戦略に対する戦略知識の獲得・評価を行う。

2 戦略知識の獲得方法

2.1 繰り返しジャンケンゲーム

ジャンケンは一般によく知られているように、数人で勝負を行い勝敗を決するゲームである。ルールでは、3種類の手である、グー、チョキ、パーが3すくみ状態になっている。

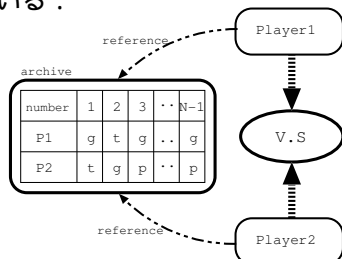


図 1: 対戦の様子

このような手から戦略的に出す手を決定するために、過去の情報を用いる形式として、繰り返しジャンケンゲームを設定する図 (1)。

繰り返しジャンケンゲームでは、対戦の形式を 1 対 1 で任意の回数を繰り返し最終的により多く勝利を上げた方を勝者とする。繰り返しジャンケンゲームでは、引き分けの場合も 1 回の対戦としてカウントする。

2.2 関数による戦略の表現

プレイヤーは、自身の持つ戦略に従い次に出す手を一意的に決める。戦略を獲得するには、その戦略知識を表現する必要がある。そこで、表現方法として以下のような関数式を用いて戦略知識を表現する。

$$n = F(o(t), o(t-1), \dots, m(t), m(t-1))$$

- $o(t)$: t 回前の相手の手
- $m(t)$: t 回前の自分の手

この関数式は、繰り返しジャンケンゲームの履歴から、自分の手と相手の手を変数として入力し、その結果を下記の関数の組合せにより演算して次の手を出力する。ただし、使用関数として、数式表現は用いず条件分岐型の関数のみを用いる。

表 1: 使用関数

operation	定義
$O(t)$	t 回前の自分の手を返す
$M(t)$	t 回前の相手の手を返す
$W(x)$	入力 x に勝つ手を返す
$L(x)$	入力 x に負ける手を返す
$E(x)$	入力 x と引き分ける手を返す
$IF(x,y,a,b,c)$	x と y の結果から 引き分けの場合 a を返す x が勝ちの場合 b を返す y が勝ちの場合 c を返す

以上の関数を用いて戦略知識を表現する。戦略知識については、S 式表現にで次のように表わされる。

$$F : IF(O(1), M(1), L(M(2)), GU, CHOKI)$$

この例では、獲得する側のプレイヤーにとって、一回前の対戦結果が引き分けの場合、二回前の手に負ける手を出し、勝ちの場合グーを出し、負けの場合にはチョキを出す。

2.3 戦略知識の進化的獲得方法

戦略知識の獲得手法として、GP を用いる。GP は GA の遺伝子型を拡張し、構造的な表現を扱えるようにしたものである。

本研究では、前述のように、戦略知識を関数式として表現した。GP では、このような関数式の組合せを直接扱い進化させることができる。

戦略の遺伝子コーディングは関数の S 式表現を元に木構造として扱う。戦略知識獲得の方法は図 2 に示すように行う。

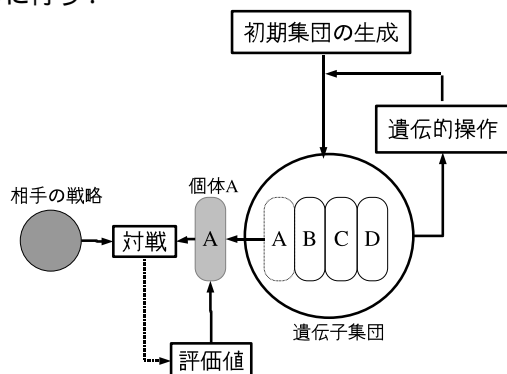


図 2: 戦略知識獲得の流れ

1. 戦略知識を獲得する戦略を用意する。GP 側はランダムに初期集団を生成する。
2. 相手の戦略と個体群の中の一つの個体の持つ戦略との間で繰り返しじゃんけんゲームを行う。
3. 各個体は、一試合で得られた勝敗結果をもとに個体の評価値を定める。
4. 全ての個体で終了したあと、遺伝子集団に対し遺伝的操作用意を行う。
5. 2～4 を繰り返し相手の戦略知識を獲得する。

3 戦略知識獲得実験

3.1 固定戦略の獲得

知識獲得の再現性の評価として、逐次的に局面で戦略の変化しない固定戦略を用いる。用いる戦略については以下に簡単に説明する

- 一筋戦略

グー、チョキ、パーのいずれかの手を出し続ける

- N 手前仕返し戦略

N 回前に出した相手の手をそのまま次の手として出す。

- 周期 N 繰り返し戦略

グー、チョキ、パーなど一定のパターンを繰り返し出し続ける。

- 場合分け戦略

履歴より、勝ち、負け、引き分けのそれぞれに応じて次に出す手を決める。今回は、勝ちの場合チョキ、引き分けの場合グー、負けの場合パーを出す手とした。

図 3 は 3 種類の初期集団に対する場合分け戦略における世代ごとのエリート戦略の評価値であり、最良とされる獲得した戦略は、

$$F: IF(M(1), O(1), CHOKI, GU, PA)$$

であり、実際の戦略と一致するものが得られた。

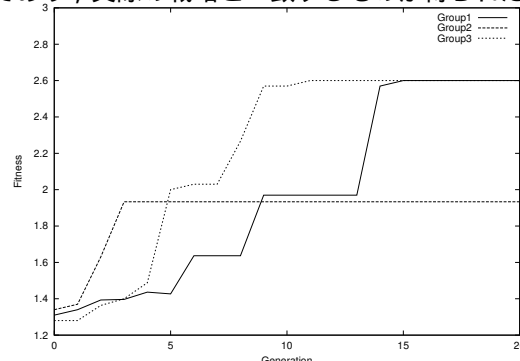


図 3: 場合分け戦略の結果

3.2 逐次的な学習戦略の獲得実験

プレイヤーが人間同士の繰り返しじゃんけんゲームにおいては、互いの戦略はそれまでの戦績によって変化する。そのため、全体を通して一意的に学習戦略を獲得できない。そこで、繰り返しじゃんけんゲームの過程における局面ごとの戦略知識の獲得を順次行い、それらの戦略知識の間での相関などにより、戦略知識の過程における特徴を獲得する。詳しくは当日の発表にて述べる。

4 考察とまとめ

本研究では、対戦型ゲームの戦略知識の獲得として、遺伝的プログラミングによる繰り返しじゃんけんゲームにおける戦略獲得を行った。まず、対戦相手の戦略知識の再現性の評価について、固定戦略における戦略知識の獲得を行った。その結果、今回定義した固定戦略である、一筋戦略、N手前仕返し戦略、周期N繰り返し戦略、場合分け戦略について、戦略表現に用いる分岐関数および入力変数により戦略知識の獲得ができた。

これにより、ある固定された戦略については、その戦略を再現することができる。このことから、逐次的に変化する場合においても、何らかのくせといった特徴を抽出することが可能となり、対戦型ゲームによらず、交渉などの駆け引きを行うような現場においても、相手の戦略知識の獲得が行えるのではないかと考えられる。

参考文献

- [1] モートン・D・デービス, ゲームの理論, 講談社, 1994.
- [2] 西山 賢一, 勝つためのゲーム理論, 講談社, 1995.
- [3] 伊庭 斉志, 遺伝的プログラミング, 東京電気大学出版, 1996.
- [4] 木村 将治, 牧野 泰裕, 小高 知宏, 小倉 久和, 繰り返しじゃんけんゲームを対象とした固定戦略による知識表現と遺伝的アルゴリズムによる知識表現, 電子情報通信学会論文誌 VOL.J85-D-I NO.5 pp.479-483, 2002.