

平成 20 年度

学士学位論文

じゃんけんゲームに対する
遺伝的アプローチ

Genetic approach to Janken game

1090396 山下 将臣

指導教員 坂本 明雄

2009 年 3 月 5 日

高知工科大学 情報システム工学科

要 旨

じゃんけんゲームに対する 遺伝的アプローチ

山下 将臣

近年，コンピュータは急速な発展を遂げている．しかし，人と同じように思考し，振る舞うことのできるコンピュータは存在していない．例えば，コンピュータが人の心を読んで行動するなどということは，現在の技術ではなかなか困難なようである．

一方，遺伝アルゴリズムとは自然淘汰に基づく遺伝・進化のメカニズムを最適解探索の枠組みととらえてアルゴリズムを構成しようとするものである．遺伝アルゴリズムでは，親から子へとその形質を伝えつつ環境に適応していくという，生物における遺伝・進化のメカニズムを簡略化した形で利用している．

本研究では，コンピュータが苦手とする分野である人の心を読む，という行動をコンピュータに実現させることを目的とする．その例として，2人対戦ゲームであるじゃんけんゲームを取り上げ，過去のデータを元に人が次に出す手を読み，それによりじゃんけんゲームにおける戦略知識を獲得することを目的とする．具体的には，人が普段じゃんけんゲームを行うときにとっていると思われる戦略や癖をコンピュータに学習させる方法を，遺伝アルゴリズムの考え方をを用いて実現する．

キーワード 遺伝アルゴリズム，じゃんけん

Abstract

Genetic approach to Janken game

Masaomi YAMASHITA

Recently, information processing technology develops rapidly. However, a computer that can react like a person doesn't exist. For instance, it looks difficult that the computer reads person's mind by present technology.

On the other hand, the genetic algorithm is an approximation procedure because it solves optimization problems based on the mechanic of natural selection and natural genetics.

In this paper, we propose a strategy to read person's mind with the computer. The Janken game is adopted as an example. It aims to acquire an excellent strategy in the Janken game by estimating player's hands from past data. It is achieved in using the idea of the genetic algorithms.

key words genetic algorithm, janken game

目次

第 1 章	序論	1
第 2 章	じゃんけんゲーム	3
2.1	じゃんけんゲームの定義	3
2.2	じゃんけんゲームにおける問題点	4
2.3	じゃんけんゲームの遺伝的表現と評価方法	6
第 3 章	遺伝的アプローチ	10
3.1	遺伝アルゴリズムの概要	10
3.2	遺伝アルゴリズムの流れ	10
3.3	じゃんけんゲームへの適用	12
3.3.1	初期染色体集団の生成	12
3.3.2	適応度の計算	13
3.3.3	選択	14
	エリート戦略	14
	トーナメント選択	15
3.3.4	交叉	16
	2 点交叉	16
3.3.5	突然変異	17
3.3.6	終了条件	17
第 4 章	実験内容	19
4.1	パラメータの設定と実験環境	19
4.2	実験結果	20
4.2.1	単純パターンとの対戦	20

4.2.2	途中で戦術を変更する単純パターンとの対戦	20
4.2.3	人との対戦	22
第 5 章	考察	23
5.1	過去の研究と本研究の手法の違いについて	23
5.2	単純パターンとの対戦について	23
5.3	人との対戦について	24
第 6 章	結論	27
	謝辞	28
	参考文献	29

図目次

2.1 対戦相手とコンピュータがじゃんけんゲームを行う流れ	9
3.1 遺伝アルゴリズムの全体の流れ	11
3.2 初期染色体集団	13
3.3 エリート戦略	15
3.4 トーナメント選択	16
3.5 2点交叉	17
3.6 突然変異	18
4.1 途中で単純パターンを変更した場合の対戦結果による得点の推移	21
5.1 プレイヤー1との対戦による得点の推移	26
5.2 プレイヤー2との対戦による得点の推移	26
5.3 プレイヤー3との対戦による得点の推移	26
5.4 プレイヤー4との対戦による得点の推移	26

表目次

2.1	繰り返しじゃんけんゲームにおける勝敗表	4
2.2	あるプレイヤーが 100 回のじゃんけんゲームで出した手の回数	5
2.3	プレイヤー 1 がじゃんけんゲームで出した手の流れ	6
4.1	遺伝的パラメータ	19
4.2	単純パターンとの対戦結果	20
4.3	途中で単純パターンを変更した場合の対戦結果	21
4.4	人との対戦結果	22

第 1 章

序論

人は利害の対立する状況では、相手の心を読んで少しでも自分に有利になるように行動している。それは、一種の交渉のモデルとしても見ることができ、コンピュータが苦手とする分野である [1].

そこで、本研究ではコンピュータが交渉を行うための手始めとして、対戦相手に勝利するためには相手の心を読むことが必要となるじゃんけんゲームについて、対戦相手の出す手を予測して次に出す手を考えるような戦略知識の獲得を目指す。じゃんけんゲームは、複数の人が相互に影響する状況の中で、ゲームに参加する全員の利益の総和が零になる零和ゲームとして、我々にとって最もなじみの深いものの一つである。また、人がじゃんけんゲームを行うときには、相手が過去に出した手を考慮しながら次に相手が出す手を予測していると考えられるため、コンピュータに交渉をさせるために使用するモデルとして適しているのではないかと考えた。

一方、ダーウィンの進化論に基づき、自然淘汰・交叉・突然変異からなる生物の進化のメカニズムを簡略化した形で模倣した最適化の手法の一つに、遺伝アルゴリズムがある。遺伝アルゴリズムは、システムの観点からは遺伝子の並びである染色体の表現方法、遺伝子の複写による増殖、突然変異や交叉による新しい遺伝子型や表現型の生成、環境への対応の度合いに基づく個体間の競合による子孫数の決定に特徴がある。具体的には、遺伝アルゴリズムは問題の解となる候補を染色体で表現したものを複数用意し、適応度の高い個体を優先的に選択して突然変異・交叉の操作を繰り返しながら最適解を探索する。環境に適応している個体は、交叉によってその優れた遺伝子を子孫に残す。また、染色体の一部に突然変異が発生することで、偶発的に突然良い個体が発生することがある。なお、特定の染色体が他の染色

体より優れているかどうかを示す適応度は適応度関数によって与えられ、これにより、環境に適したより良い次世代集団を生み出していく。

本研究では、人とコンピュータとで心の読み合いを行うゲームとしてじゃんけんゲームを取り上げ、遺伝アルゴリズムの考え方をを用いて人相手にもある程度勝利することのできる戦略知識を獲得することを目指した。具体的には、人が普段じゃんけんゲームを行うときにとっていると思われる戦略や癖をコンピュータに学習させ、相手の特徴をつかむことで次に出す手を決定する方法を獲得する。これにより、人の出す手の特徴をうまくつかみ取ることができれば、対戦相手の戦略の癖を読み取れ、人相手にも勝ち越すような戦略が獲得できると考えられる。

第 2 章

じゃんけんゲーム

2.1 じゃんけんゲームの定義

じゃんけんゲームとは、ときには大勢で行われることもあるが、一般的には2人のプレイヤーの間で行われるゲームである。じゃんけんゲームには3通りの手が存在し、それぞれのプレイヤーは3通りの手の中から一つを選択し、互いに自分の選択した手を同時に見せ合うことで勝敗を決定する。じゃんけんゲームにおける3つの手は、それぞれ一つの手には優位であり、もう一方の手には不利であり、同じ手とは互角となる。このように、じゃんけんゲームには絶対的に優位な手というものが存在しておらず、それぞれの手は3すくみの関係になっている。3通りの手の個々の名称は、それぞれグー・チョキ・パーであり、互いの手に対して以下の特性をもっている。

- グーはチョキに対して勝ち、パーに対して負ける
- チョキはパーに対して勝ち、グーに対して負ける
- パーはグーに対して勝ち、チョキに対して負ける

以上のような特性を持つ手を互いのプレイヤーが出し合った場合について、勝った方をW、負けた場合をL、あいこの場合をDで表すと、じゃんけんゲームにおける勝敗の結果は表 2.1 のようになる。なお、表 2.1 において、p1 はプレイヤー 1、p2 はプレイヤー 2 を指す。

表 2.1 繰り返しじゃんけんゲームにおける勝敗表

	グー (p2)	チョキ (p2)	パー (p2)
グー (p1)	(p1 : D) (p2 : D)	(p1 : W) (p2 : L)	(p1 : L) (p2 : W)
チョキ (p1)	(p1 : L) (p2 : W)	(p1 : D) (p2 : D)	(p1 : W) (p2 : L)
パー (p1)	(p1 : W) (p2 : L)	(p1 : L) (p2 : W)	(p1 : D) (p2 : D)

2.2 じゃんけんゲームにおける問題点

じゃんけんゲームは、非常に単純なゲームながら対戦相手に勝つための決まった作戦というものは存在していない。

じゃんけんゲーム以外の2人対戦ゲームの代表的なものとして、将棋を例に挙げる。将棋は、じゃんけんゲームよりはるかに難しいルールや知識が必要であるにもかかわらず、2008年現在、コンピュータはプロ棋士とも対等に戦えるほどの力を獲得している。これは、2005年に日本将棋連盟がプロ棋士とコンピュータ将棋との公の場での無断での試合を禁止したことからもうかがい知れるものである [2] [3]。

また、将棋のほかに囲碁の場合では、将棋と比較して広い盤面であるため選択肢が多いことや、自分が有利な状況にもっていくための有効な評価関数の作成が困難であることが相まって、プロのレベルに届くのはまだまだ先だとされている。

これら二つの2人対戦ゲームと比較すると、有効な戦略知識の獲得が容易そうなじゃんけんゲームではあるが、人間相手に5割以上安定して勝てるというコンピュータはおそらく存在していない。そこには、じゃんけんゲームにおける「勝てる戦略」が明確に存在していない点と、ある程度、運が勝敗を左右してしまうことに原因があるように思う。

しかし、人がじゃんけんゲームを行うときには、過去に互いが出した手によって次に出す手を決定しているものと考えられる。つまり、本人にも分からないところで癖や特徴が出ている可能性がある。このことから、やはりじゃんけんゲームにおいて勝てる戦略知識を獲得する上で問題となるのは、人の出す手の癖をどのように読み取るかであるように思う。

表 2.2 は、あるプレイヤーが合計 100 回のじゃんけんゲームを行った際に出した手の総数である。この表 2.2 から、このプレイヤーはグーやパーよりも、チョキの手を出しやすいということが分かる。これは、このプレイヤーの一つの癖であると言え、このような癖を読み取ることはコンピュータに有効な戦略知識を獲得することができる方法の一つであると予想される。

表 2.2 あるプレイヤーが 100 回のじゃんけんゲームで出した手の回数

グーを出した回数	チョキを出した回数	パーを出した回数
29 回	38 回	33 回

次に、対戦相手となる人の出した手一つ一つの回数ではなく、対戦相手が過去に出した手と自分が過去に出した手の流れについても考慮する方法を考える。表 2.3 は、あるじゃんけんゲームの対戦結果について、双方のプレイヤーが出した手の流れを表したものである。これを見ると、プレイヤー 1 (p1) は、グーを出したときに相手にグーを出されると、次にパーを出しやすいことが分かる。同様に、プレイヤー 1 がグーを出したときに相手にチョキを出された場合には、プレイヤー 1 は次にチョキを出しやすいことが分かる。このように、単に対戦相手が出しやすい一つ一つの手を見るのではなく、自分の過去に出した手と対戦相手の過去に出した手の結果からもプレイヤーの癖をある程度読み取ることができそうである。

このような方法で人の癖をコンピュータに学習させることができれば、コンピュータでも人相手に一定以上の勝率を得ることができるという予測が立つ。これは、じゃんけんゲームの話ではあるが、コンピュータに普段人が無意識のうちにやっている相手の行動を読むということを学習させることにも繋がってくる。

そこで、人が過去に出した手を元にコンピュータが次に出す手を決定するような動作をコンピュータに適用する方法を考える。このための戦略知識の獲得方法には、遺伝アルゴリズムの考えを用いることにする。じゃんけんゲームに勝つための戦略の組み合わせの選択につ

表 2.3 プレイヤー 1 がじゃんけんゲームで出した手の流れ

p1 が二手前に出した手	p2 が二手前に出した手	p1 が 1 手前に出した手の回数
グー	グー	グー:2 回, チョキ:4 回, パー:8 回
グー	チョキ	グー:3 回, チョキ:6 回, パー:2 回
グー	パー	グー:1 回, チョキ:3 回, パー:4 回
チョキ	グー	グー:5 回, チョキ:0 回, パー:2 回
.....
パー	パー	グー:2 回, チョキ:5 回, パー:4 回

いての問題は、1 手ごとに最適な解が変化してしまうが、1 手単位で見ると、一種の組み合わせ最適化問題に近いものであると言える。そこで、組み合わせ最適化問題を解くことに適している遺伝アルゴリズムの考えを用いることで、じゃんけんゲームについて人相手にも勝てる戦略をコンピュータ上で進化的に獲得することを目指す。

2.3 じゃんけんゲームの遺伝的表現と評価方法

じゃんけんゲームにおいて出される 3 通りの手、グー・チョキ・パーをそれぞれ $G \cdot C \cdot P$ で表し、その集合を $H = \{G, C, P\}$ とする。また、集合 H に含まれる要素の順序は G, C, P であるとする。戦術の読みの深さを d とし、じゃんけんゲームの一つの戦術を H^d から H への関数 α で表す。

じゃんけんゲームの対戦相手（人側）を記号 U ，コンピュータ側を記号 I とする。また、対戦相手が次に出す手を U_0 ，1 手前の手を U_1 ，2 手前の手を U_2 ，一般に k 回前の手を U_k で表す。なお、対戦相手である人が次に出す手である U_0 は、現在の時点ではもちろん分かってはいないが、過去に K 回の対戦が終了していれば、 $U_k (1 \leq k \leq K)$ は分かっている。同様に、コンピュータ側の手については $I_k (0 \leq k \leq K)$ で表す。

ここで、例えば読みの深さを $d = 2$ とし、 U_1 と I_1 の手からコンピュータが次に出す手

を決定するような H^2 から H への関数 α が、以下のようなものであったとする.

$$\alpha(G, G) = P, \alpha(G, C) = P, \alpha(G, P) = G$$

$$\alpha(C, G) = G, \alpha(C, C) = C, \alpha(C, P) = C$$

$$\alpha(P, G) = G, \alpha(P, C) = P, \alpha(P, P) = G$$

このとき $\alpha = PPGGCCGPG$ と表し、この表現を遺伝アルゴリズムの適用の際に用いる染色体とする. なお、コンピュータ上では $G = 0, C = 1, P = 2$ として扱っているため、 $\alpha = 220011020$ となる.

この戦術 α は、対戦相手である人が1手前にグーを出し、コンピュータが1手前にグーを出した状況であれば、コンピュータは次に出す手 I_0 にパーを選択する、というものである. 同様に、 $U_1 = G$ かつ $I_1 = C$ のとき $I_0 = P$, \dots , $U_1 = P$ かつ $I_1 = P$ のとき $I_0 = G$ という戦術をとる. すなわち、関数 α は、1手前の対戦相手の手と1手前のコンピュータの手から、次にコンピュータが出す手を決定していることになる. 例えば、過去に K 回の対戦を終えているとき、過去の対戦データから $K - 1$ 回の対戦について評価できることになる.

そこで、例えばこの戦術 α に対して、対戦相手の過去5回の手を $GGCGP$, コンピュータの手を $CGCGC$ であった場合に適用すると、以下のようになる.

1 戦目

戦術 α は、適用に関して過去1回分の対戦結果が必要なため、1戦目では適用できない.

2 戦目

1手前の1戦目では、 $U_1 = G$ かつ $I_1 = C$ なので、戦術 α を用いた場合にコンピュータが次に出す手は $\alpha(G, C) = P$ である. 実際には、対戦相手はここで G を出しているので、2戦目はコンピュータの勝ちとなる. なお、戦術 α とは異なり、コンピュータ側はここで G の手を選択しているが、これは α の評価には関係しない.

3 戦目

1 手前の 2 戦目では、 $U_1 = G$ かつ $I_1 = G$ なので、戦術 α を用いた場合にコンピュータが次に出す手は $\alpha(G, G) = P$ である。実際には、対戦相手はここで C を出している
ので、3 戦目はコンピュータの負けとなる。

4 戦目

1 手前の 3 戦目では、 $U_1 = C$ かつ $I_1 = C$ なので、戦術 α を用いた場合にコンピュータが次に出す手は $\alpha(C, C) = C$ である。実際には、対戦相手はここで G を出している
ので、4 戦目はコンピュータの負けとなる。

5 戦目

1 手前の 4 戦目では、 $U_1 = G$ かつ $I_1 = G$ なので、戦術 α を用いた場合にコンピュータが次に出す手は $\alpha(G, G) = P$ である。実際には、対戦相手はここで P を出している
ので、5 戦目は引き分けとなる。

以上のように、今回の例では戦術 α が適用されている場合の対戦結果はコンピュータ側の 1 勝 2 敗 1 引き分けとなる。なお、この戦術が 5 戦以降にも適用されたとすると、6 戦目にコンピュータが出す手は $\alpha(P, C) = P$ となる。

なお、読みの深さはここでは 2 として扱ったが、じゃんけんゲームを行う際の対戦回数によってこの値は変更する可能性がある。それは、読みの深さを大きくするとより詳しい対戦経歴が分かるようになり、精度が上昇すると考えられるが、対戦回数に対して読みの深さがあまりに大きいと、過去に対戦経験のない場合が数多く現れ、戦術の評価自体が困難になる恐れがあるためである。よって、この読みの深さ d については、対戦回数によって適当な値に変更すべきであると考えられる。実験では、対戦回数を 100 回として行うため、この読みの深さについては、 $d = 3$ が適当であると考えられる。

そして、適応度の計算に際しては、最近の対戦結果に大きな重みを与えるように工夫する。これは、最初の方の対戦結果によって対戦相手に有効な戦術を持つ染色体が収束してしまい、最近の対戦結果が染色体の評価にあまり影響を与えなくなることを抑えるための処

置である。また、対戦相手が途中で戦術を変更してくる可能性があり、その場合にもコンピュータが対応できるようにするために、このような処置を行う。これにより、もちろんコンピュータは対戦相手の新たな戦術への対応に手数はかかるものの、最近の手に重みを与えない通常の方法と比較すると、早い段階で対戦相手に有利な戦術に対応することができると予想される。

じゃんけんゲームの遺伝的表現や評価方法は以上のような方法で行い、これによりコンピュータに対戦相手に有効と思われる戦術を学習させる。実際に対戦する際は、対戦相手とコンピュータとが互いに手を出すごとに、複数個作成した染色体から新たな対戦結果を評価に反映させた次世代の染色体を次々と作成していく。そのようにすることで、対戦回数が増えれば増える程、コンピュータは対戦相手の出す手の癖を読み取ることができ、勝率は上昇していくものと思われる。

図 2.1 は、繰り返し行われるじゃんけんゲームにおけるコンピュータ側の戦術を適用していく流れを表したものである。なお、最初の数手についてはコンピュータは過去に対戦相手との対戦がない、もしくは過去に対戦相手との対戦回数が少なく有効と言える戦術を獲得することが難しいため、ランダムに出す手を選択する。

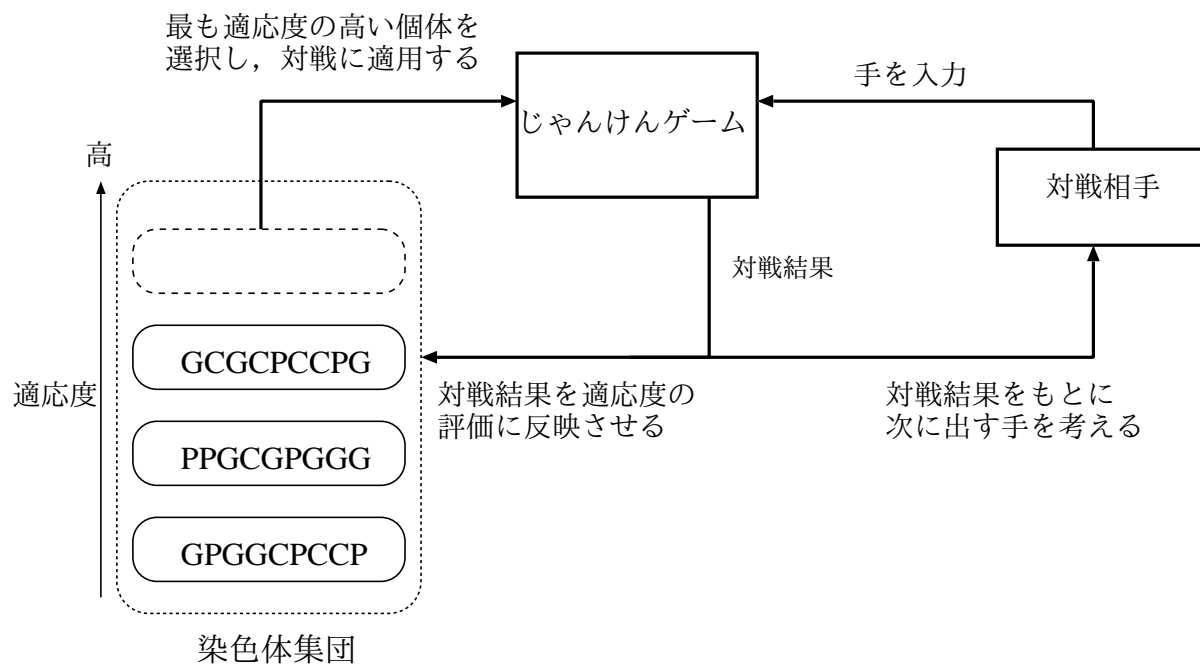


図 2.1 対戦相手とコンピュータがじゃんけんゲームを行う流れ

第 3 章

遺伝的アプローチ

3.1 遺伝アルゴリズムの概要

遺伝アルゴリズムとは、生物集団の進化のメカニズムを工学的にモデル化したアルゴリズムで、最適化問題に対する近似解法の一つである [4]。遺伝アルゴリズムはダーウィンの進化論に基づいており、自然淘汰・交叉・突然変異からなる生物の進化のメカニズムを簡略化した形で模倣している。

具体的には、問題の解となるものを染色体で表現し、染色体集団の中から適応度が高い染色体を親として設定し交叉することで、次世代により良い遺伝子を残そうとする。そして、これを繰り返すことで最終的には環境に最適、もしくはそれに近い個体が解として得られることが期待できる。

3.2 遺伝アルゴリズムの流れ

遺伝アルゴリズム全体の流れは、図 3.1 の通りである。

図 3.1 の各部分については、具体的には以下のような処理を行うものである。

1. 初期集団の生成

1 世代目となる染色体集団を、あらかじめ設定された個体数だけランダムに生成する。

2. 適応度の計算

各染色体に対し、環境（条件）に対する適応度を計算する。

3. 選択

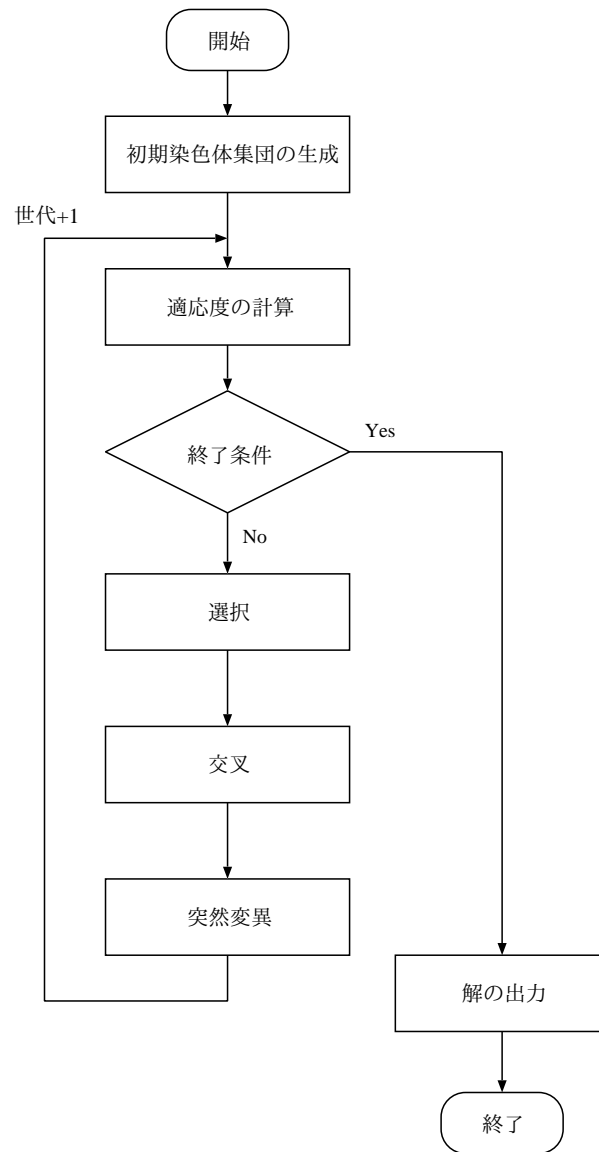


図 3.1 遺伝アルゴリズムの全体の流れ

染色体集団の内から，次世代に残す染色体を選択する．2 で求まった適応度が高いほど，親として選択される確率が高くなる．

4. 交叉

3 によって選択された染色体同士で交叉の処理を行う．交叉とは，親として選択された二つの染色体を元にして遺伝子を組み換え，子となる新しい染色体を二つ生成する処理である．

5. 突然変異

突然変異確率によって決められた確率で突然変異の処理を行う。これは、染色体の遺伝子をランダムに変更する処理であり、通常低い確率に設定される。

6. 終了条件

あらかじめ設定された終了条件を満たしていれば、この時点で適応度の高い解を出力し、処理を終了する。そうでなければ、2に戻り、再度適応度の計算を行う。

3.3 じゃんけんゲームへの適用

遺伝アルゴリズムは、適用する問題によって初期染色体集団の生成方法や適応度の評価の方法が異なっている。また、両親の選択方法や交叉の方法についても数多くの種類がある。本節では、じゃんけんゲームにおける遺伝アルゴリズムの適用について、初期染色体の構成や適応度の計算について述べる。

以下では、染色体集団の個体数を N_{pop} 、1つの染色体に含まれる遺伝子数、すなわち染色体の長さを L で表す。なお、読みの深さが d の戦略を用いるとき、染色体の長さは $L = 3^d$ である。

3.3.1 初期染色体集団の生成

遺伝アルゴリズムでは、染色体集団の第一世代として初期集団を生成する必要がある。ここでは、一般的に行われている方法と同じく初期集団の染色体の遺伝子についてはランダムに決定する。すなわち各染色体は、じゃんけんゲームにおける3通りの手である G, C, P を表す数字の 0, 1, 2 がランダムに L 個並んだものであり、3種類の数字は等確率で出現する。このような染色体を N_{pop} 個生成して、初期染色体集団とする。図 3.2 は、初期染色体集団を表したものである。

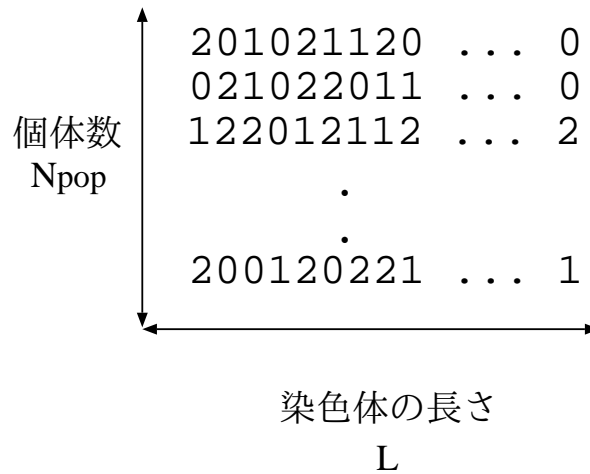


図 3.2 初期染色体集団

3.3.2 適応度の計算

適応度は、各染色体が環境にどれほど適応しているかを数値で表したものである。この数値は、遺伝アルゴリズムにおいて次世代に残す染色体を選択するために使用される。じゃんけんゲームにおける適応度は、コンピュータ側が過去に対戦相手の出した手から次に対戦相手が出すと思われる手に勝てる度合いを数値化したものになる。具体的には、コンピュータは過去に出した互いの手の履歴を参照することで、対戦の結果ごとに対戦相手が次に出しやすいと思われる手を遺伝子に持っている染色体の適応度を高くする。すなわち、過去の履歴から対戦相手が次に出す手を読み、よりその手に勝てる可能性が高い染色体の適応度が高くなる。この処理を1度の対戦が終了するごとに行う。

本研究では、じゃんけんゲームの適応度を式 3.1 を用いて求める。なお、式 3.1 について、対戦を数多くこなす場合を考えると、最近の対戦結果を古い対戦結果よりもより重要視するようにするため、対戦回数の値が増加する程、よりそのときの対戦結果が適応度の値に反映させる計算式を用いた。これにより、対戦回数が少ない段階での染色体の収束を抑えることができる。また、対戦相手が途中で戦術を変更した場合にも、比較的早く対応することができる。それぞれの対戦結果でコンピュータに与える得点は、実験を何度か繰り返し、その中で最も良い値を用いるものとする。

$$\sum_{i=1}^m \{(i \text{ 回戦目の対戦結果の得点}) \times i\} \quad (3.1)$$

コンピュータが勝利したときの得点 $P_w = 2$

コンピュータが引き分けたときの得点 $P_d = 1$

コンピュータが敗北したときの得点 $P_l = 0$

$m =$ 適応度の計算が行われている対戦回数の値

3.3.3 選択

選択とは、各染色体の適応度から次世代に残す染色体を選択する処理のことを指す。一般的に、遺伝アルゴリズムにおける両親の選択方法はいくつか種類があるが、本研究では、エリート戦略とトーナメント選択を用いて実験を行った。以下に、それぞれの選択方法の特徴と、本研究への適用方法を示す。

エリート戦略

エリート戦略とは、染色体集団の中で最も適応度の高い個体を次世代に残す選択方法である。これは、最も良い個体は必ず次世代に残されるため、他の選択方法と異なり現在の世代で最も良い解を確率的な操作によって消滅させてしまうことがない点において優れている。しかし、エリート戦略によって選択された個体の特徴が急速に染色体集団内に広がっていく可能性が高くなり、その結果、局所的な最適解に陥る可能性が高くなるという危険性も含んでいる。そこで、本研究ではエリート戦略と共にトーナメント選択を併用することで、局所的な最適解に陥る可能性を減少させる。

図 3.3 は、じゃんけんゲーム用に作成した染色体集団に対して、エリート戦略が行なわれている様子を表したものである。この図のように、最も適応度が高い染色体がエリートとして選択され、初期染色体集団のサイズに比例した数だけ適応度が低い染色体にコピーされる。これにより、集団内の適応度の高い染色体の数を増やすことができ、世代が進むごとに

より良い染色体集団が生成されていくことになる。

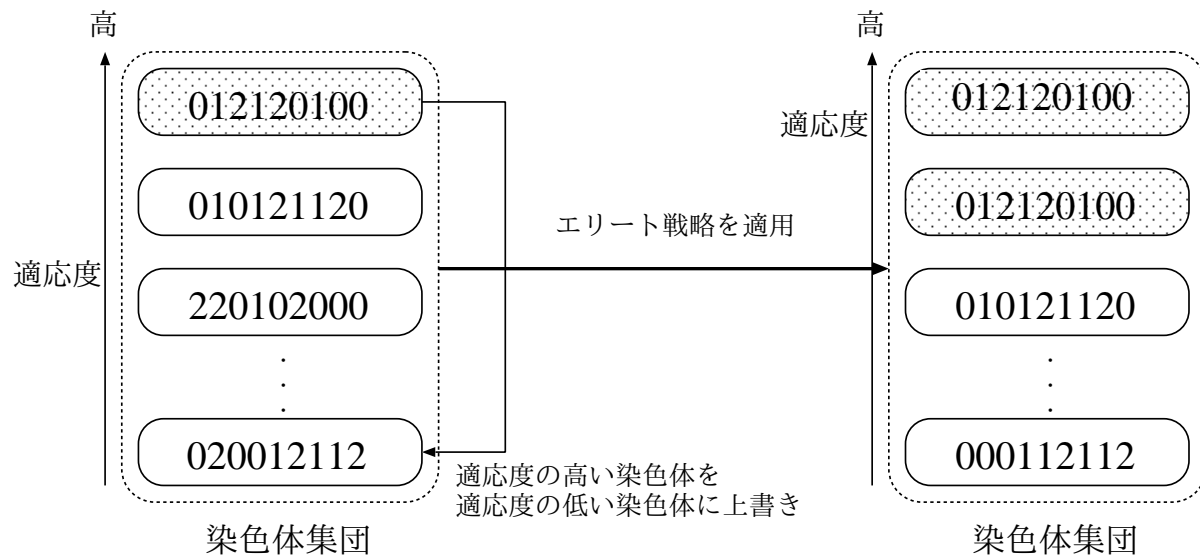


図 3.3 エリート戦略

トーナメント選択

トーナメント選択は、染色体集団の中からあらかじめ設定された数だけ個体をランダムに選択し、その中で最も適応度の高い個体を次世代に残す選択方法である。選択される個体数は一般的にトーナメントサイズと呼ばれ、この値を変更させることで選択圧をコントロールすることができる。具体的には、トーナメントサイズを大きくすれば適応度の低い個体が淘汰される可能性が高くなり、結果として局所的な最適解に陥る可能性が高くなるという特徴を持つ。

図 3.4 は、じゃんけんゲームのプログラム内の染色体集団に対して、トーナメント選択が行なわれている様子を表したものである。この図のように、まず集団からトーナメントサイズ分だけ染色体がランダムに選択される。そして、選択された染色体の中から最も適応度が高いものが残されるのである。この処理を初期染色体集団のサイズ N_{pop} 個分選択されるまで繰り返す。

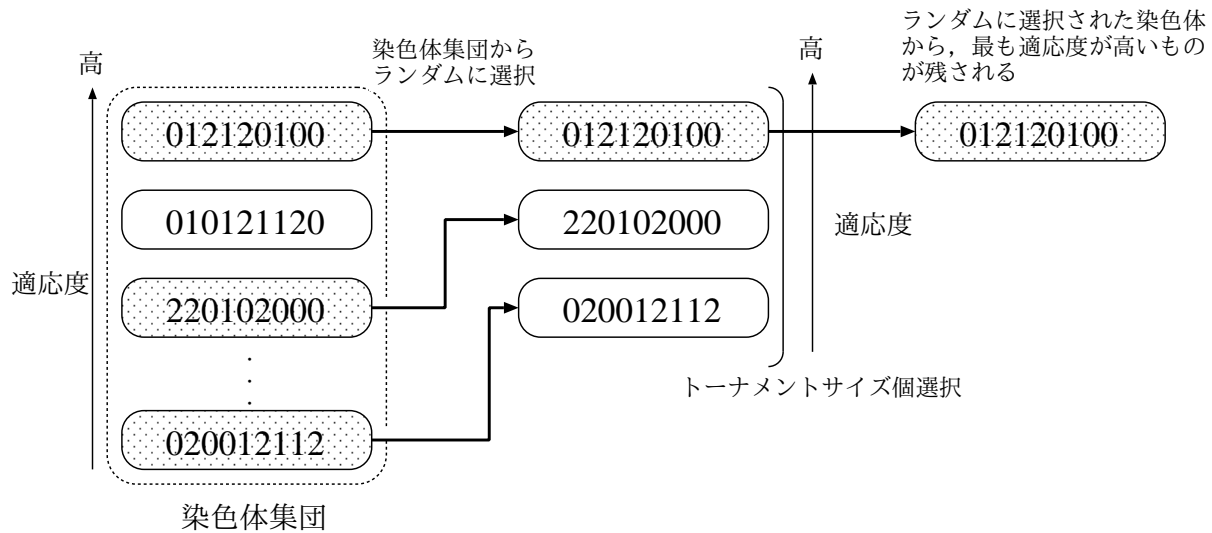


図 3.4 トーナメント選択

3.3.4 交叉

交叉とは、選択された個体2つに対してある交叉位置で双方の染色体の特徴を残しながら次世代の染色体を作り出す処理である。なお、交叉という処理はあらかじめ設定される交叉確率という値に基づき実行されるため、毎回必ず行なわれるものではない。交叉確率とは、その名の通り交叉の処理を行う確率のことである。交叉の方法には多くの種類があるが、以下では本研究で使用した2点交叉について述べる。

2点交叉

図 3.5 は、2点交叉が行われている様子を表したものである。まず、親となる染色体を選び、ランダムに交叉点を2点選択する。この交叉点は、両親共に同じ点を選択する。そして、交叉点の間の遺伝子以外については子にそのまま遺伝させ、交叉点の間の遺伝子は親1から子2へ、親2から子1へと遺伝させる。これにより、選択された2つの親の特徴を残しつつ、次世代の新たな染色体を作成することができる。

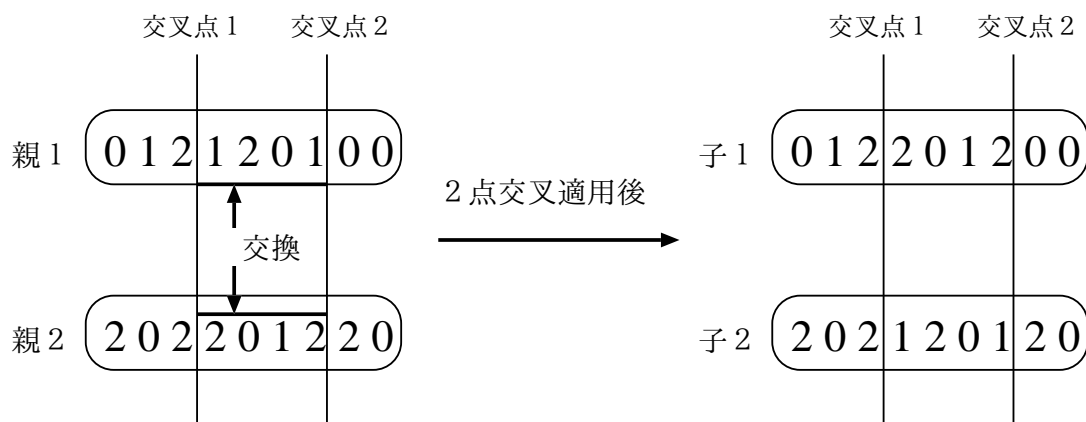


図 3.5 2点交叉

3.3.5 突然変異

突然変異とは、個体の遺伝子の一部の値を書き換える処理である。この突然変異という処理は、交叉という処理だけでは初期染色体集団内の個体に依存した、限られた範囲の特徴しか持たない個体しか生成することができないため行われる処理である。しかし、この突然変異という処理が頻繁に行われるようであれば、両親の持つ環境に適した良い特徴を破壊してしまう可能性がある。そこで、一般的に突然変異確率と呼ばれる、あらかじめ低い値に設定された確率でこの突然変異という処理を行う。この処理により、交叉だけでは生成することのできない、多様性のある個体を生成することを目的としている。

本研究では、一つ一つの遺伝子に対して、あらかじめ設定される突然変異確率によって値を別のものに変更する処理を突然変異として行う。具体例として、図 3.6 を示す。

3.3.6 終了条件

終了条件とは、与えられた問題に対して遺伝アルゴリズムの適用を終了させる条件のことである。本研究では、終了条件として染色体集団が設定された最大世代数に達した場合を終了条件として設定する。

各遺伝子の値について，別の値に強制的に変更する

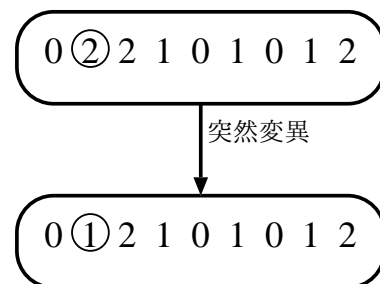


図 3.6 突然変異

第 4 章

実験内容

4.1 パラメータの設定と実験環境

人とコンピュータとのじゃんけんゲームでの対戦を実際に行なうにあたって使用した各パラメータは、表 4.1 のとおりである。これらは、一手ごとに最適な戦術を求めるための処理時間の問題や、どのような戦術を用いてもじゃんけんゲームはある程度の運が勝敗を決してしまうことを考慮して決定した。また、対戦相手となる人にストレスなく対戦を行ってもらうため、処理時間を人が気にならない程度に抑えるようなパラメータ設定とし、表 4.1 のような値を用いて実験を行うこととした。

なお、コンピュータがランダムに手を出すのをやめ、表 4.1 のパラメータを用い遺伝的アプローチのもと、対戦相手に勝てるような手を考えるようにするのは、4 戦目の対戦からとする。また、読みの深さを 3 として扱っているが、これは対戦相手の過去 1 手に加えてコンピュータの過去 2 手分を読むように設定した。

表 4.1 遺伝的パラメータ

個体数 (染色体集団のサイズ)	50
最大世代	10
交叉確率	0.95
突然変異確率	0.05
トーナメントサイズ	10
読みの深さ	3

4.2 実験結果

4.2.1 単純パターンとの対戦

実際に人との対戦を行う前に，ある特定の単純パターンの手を繰り返し出す仮想的なプレイヤーと対戦を行った．単純な繰り返しパターンとして，パターン 1 : GCP，パターン 2 : GPPCP，という手の流れを繰り返す 2 つのパターンを用意した．一つの対戦は，連続した 100 回のじゃんけんにおける勝敗または引き分けの数を決めるものとし，この対戦を異なる乱数の種を用いて 10 回行った．コンピュータ側から見た 10 回分の対戦結果の平均を表 4.2 に示す．この表 4.2 から，単純パターンとの対戦では，有効な戦術を獲得できていることが分かる．また，パターン 2 のようにパターンの手数が長くなると，パターン 1 と比較すると有効な戦術の獲得に時間がかかり，勝利数が減少していることが分かる．

表 4.2 単純パターンとの対戦結果

	勝利数	引き分け数	敗北数
GCP	89.0	4.7	6.3
GPPCP	54.3	23.1	22.6

4.2.2 途中で戦術を変更する単純パターンとの対戦

次に，対戦の途中で使用する単純パターンを変更する場合，コンピュータが新しいパターンに対応できるか実験を行った．対戦回数を 200 回とし，始めの 100 回の対戦ではパターン 1 : GPCP を繰り返し出し，101 回目から 200 回目までの対戦ではパターン 2 : CGPPC を繰り返し出すものとする．これを 10 回分繰り返した平均のコンピュータ側から見た 101 回目から 200 回目までの対戦結果を表 4.3 に示す．表 4.3 から，コンピュータは途中で対戦相手が出す手のパターンを変更した場合でも新たなパターンに対応できていることが分かる．

表 4.3 途中で単純パターンを変更した場合の対戦結果

勝利数	引き分け数	敗北数
60.1	22.2	17.7

また，図 4.1 は，先程の 10 回の実験の中の一試合についてのコンピュータ側の蓄積点の推移をグラフ化したものである．対戦結果によってコンピュータに与える得点を，それぞれ勝利したときに 1 点，敗北したときに -1 点，引き分けたときに 0 点としている．図 4.1 から，コンピュータはまずパターン 1 : GPCP に対しては対戦回数 30 回ほどまでは勝ち負けを繰り返し，その後の 100 回戦目までの対戦では得点を伸ばしていることが分かる．そして，対戦回数が 100 回を越えたところでまた得点の伸びが鈍化し，およそ対戦回数 150 回目から再び得点を伸ばしていることが分かる．

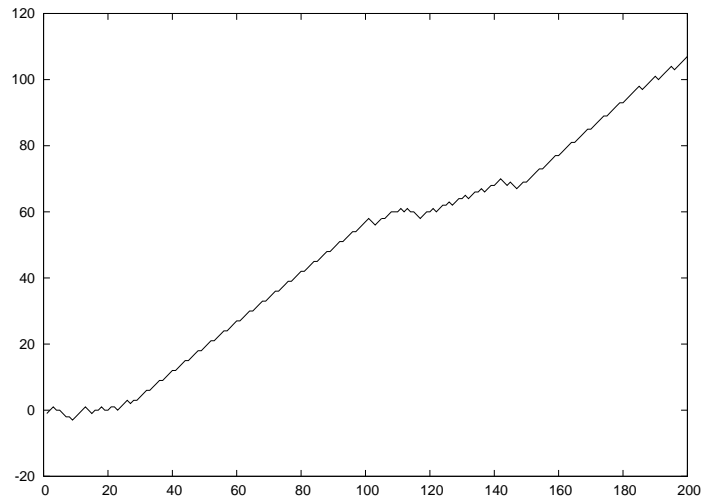


図 4.1 途中で単純パターンを変更した場合の対戦結果による得点の推移

4.2.3 人との対戦

次に，実際に人との対戦を行った．それぞれの対戦相手とコンピュータとのコンピュータ側から見た対戦結果を表 4.4 に示す．表 4.4 から，プレイヤー 1 からプレイヤー 4 までについては，僅かではあるがコンピュータは勝ち越すことができていることが分かる．プレイヤー 5 については，32 勝 34 敗と負け越してはいるが，全体的にはコンピュータは勝ち越すことができており，ある程度コンピュータは人の出す手の癖を読めていると言える．

表 4.4 人との対戦結果

	勝利数	引き分け数	敗北数
プレイヤー 1	36	30	34
プレイヤー 2	39	36	25
プレイヤー 3	39	30	31
プレイヤー 4	37	33	30
プレイヤー 5	32	34	34

第 5 章

考察

5.1 過去の研究と本研究の手法の違いについて

本研究を行う際に参照した木村将治氏らによる文献 [1] では，じゃんけんゲームにおけるいくつかの固定した戦術に対して，遺伝アルゴリズムを用いて有効な戦術の獲得をしている．固定戦術として，ある一つの手を繰り返し出すものや，一つの手の流れを繰り返すもの，相手が過去に出した手を累積しその中で最も相手が多く出している手に勝つ手を出す戦術等を用いている．また，染色体の長さは可変とし，例えば 00112 の表す染色体は，周期 5 のパターン GGCCP を対戦中ずっと繰り返し出す戦術として表現している．

一方，本研究では遺伝的アプローチによってじゃんけんゲームを行う際の対戦相手として，固定戦術ではなく人を想定している．また，染色体の表現方法は，コンピュータが出す手の流れをそのまま染色体の遺伝子として表すのではなく，相互のプレイヤーが過去に出した手によって出す手を変更する戦術を用意し，それを染色体として表現した．これにより，過去の対戦結果から次に対戦相手が出しやすい手を読むことで，人相手に有効な戦術の獲得を目指している．

5.2 単純パターンとの対戦について

単純パターンとの対戦では，本研究の手法は有効な戦術を獲得することができた．これは，読みの深さを 3 とし，対戦相手の過去 1 手とコンピュータ側の過去 2 手から次に出す手を決定するものとしたので，対戦相手の出す手のパターンを読むことができたと思われる．

当然、同じ読みの深さ 3 でも対戦相手の過去 3 手から次に出す手を決定する手法の方が単純パターンに対して早く有効な戦術を獲得することができ、勝率が高くなると予想される。しかし、人を対戦相手とした場合、この手法であるとコンピュータがどのようなアルゴリズムで出す手を決定しているかを人に読まれる恐れがあるため、本研究ではこの手法を用いなかった。

具体的な対戦結果の推移については、図 4.1 から、パターンの長さが 4 のときには 20 手から 30 手ほどかけて対戦相手に有効な戦術を探り、その後 100 回戦目までは大きく勝ち越していることが分かる。しかし、対戦相手の出す手のパターンを長くすると有効な戦術の獲得により手数がかかると予想される。

そして、図 4.1 から対戦回数が 100 回を越えたところで得点の伸びが止まっていることが分かる。これは、新たな長さ 5 のパターンに対して今まで用いてきた戦術が有効でないことが影響していると考えられる。これに対して、コンピュータ側は 40 手ほどかけて新たな相手の出す手のパターンに有効な戦術を探ることで、その後の対戦回数 140 回目から 200 回目までは大きく勝ち越せていると思われる。このことから、コンピュータは使用するパターンを変更した場合においても、対戦を何度か繰り返すことで新たに有効な戦術を獲得できているものと思われる。

また、木村将治氏らによる文献 [1] においても、単純パターンとの対戦では有効な戦術を獲得することができている。本研究と用いている手法に違いはあるが、対戦回数を 1000 回とした単純パターンとの対戦に対して最大世代数 (100) だけ繰り返し各染色体を評価できたという実験条件の違い、さらにコンピュータが出す手の流れをそのまま染色体の表現としている点が有効に働いたのだと思われる。

5.3 人との対戦について

人との対戦では、表 4.4 からある程度は有効な戦術が獲得できたと思われる。

木村将治氏らによる文献 [1] の方法は、本研究の対戦相手の過去数手からコンピュータが

次に出す手を決定する手法に性質が近いと思われるが、この方法だと人にコンピュータの出す手のアルゴリズムを読まれる可能性があるという意味で脆さがある。本研究では、読みの深さについて、対戦相手とコンピュータの双方の手からコンピュータが次に出す手を決定するような手法を用いているため、人に気づかれることなく人の癖を読み取ることができると考えられる。

図 5.1, 図 5.2, 図 5.3 および図 5.4 は、それぞれ表 4.4 にある人との対戦結果におけるプレイヤー 1 からプレイヤー 4 との対戦結果の得点の推移を表したものである。コンピュータが勝利すると +1 点, 引き分けると 0 点, 負けると -1 点としている。その中で、特にコンピュータの勝利数が多かった図 5.2 と図 5.3 については、前半では勝ち負けを繰り返し、中盤でコンピュータが勝率を伸ばし、後半ではまた勝ち負けを繰り返しているのが分かる。この結果から、前半から中盤にかけては対戦回数が増える程コンピュータは人の出す手の癖を読み取り、勝率を上げ、後半では人側もコンピュータに出す手を読まれないように工夫しているように考えられる。しかし、図 5.1 と図 5.4 についてはそのような傾向は見られず、前半から後半にかけて勝ち負けを繰り返していることが分かる。

このことから、人のプレイヤーと対戦を行う場合、対戦の後半になるほどコンピュータは人の癖を読み取ることで勝率が上昇していき、すなわち対戦回数が多い程コンピュータはより勝ち越せると当初は想定していたが、一概にそのようには言いにくい結果となった。

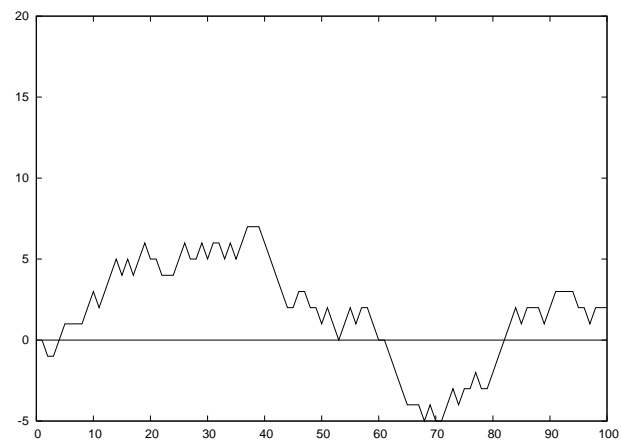


図 5.1 プレイヤー 1 との対戦による得点の推移

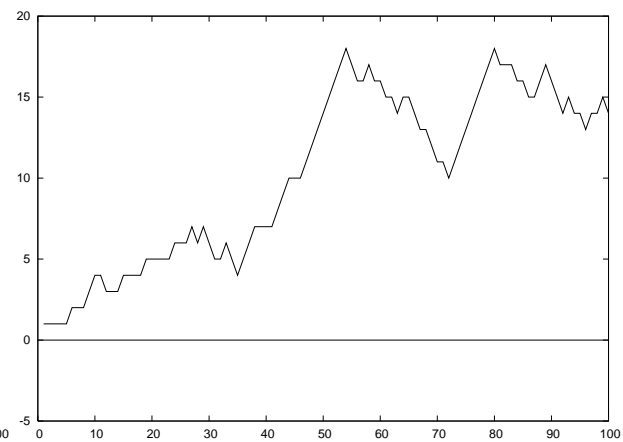


図 5.2 プレイヤー 2 との対戦による得点の推移

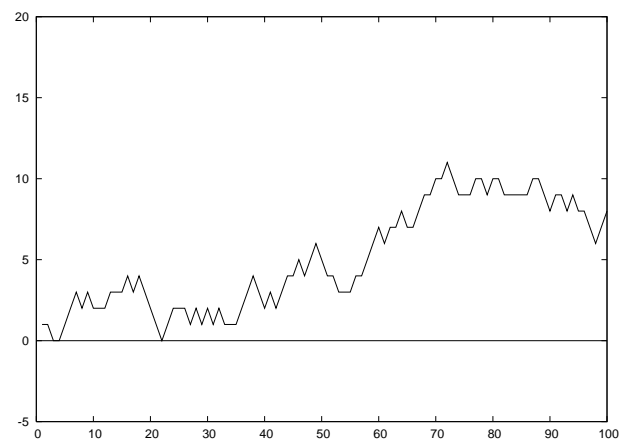


図 5.3 プレイヤー 3 との対戦による得点の推移

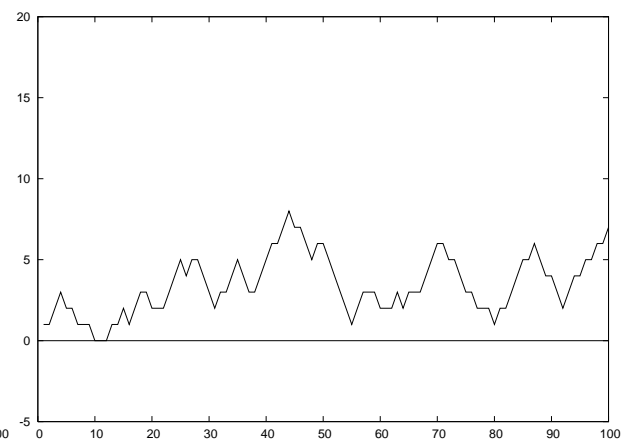


図 5.4 プレイヤー 4 との対戦による得点の推移

第 6 章

結論

本研究では、じゃんけんゲームに対して遺伝的アプローチにより人相手に有効な戦術の獲得を目指した。その結果、特定の手の流れを繰り返し出し続ける単純パターンとの対戦では非常に有効な戦術を獲得することができ、高い勝率を得ることができた。また、途中で使用する単純パターンを変更した場合においても、コンピュータは新たなパターンに有効な戦術を見つけ出すことができるという実験結果を得た。そして、実際に人を対戦相手として対戦した場合には、平均すると僅かではあるが勝ち越すことができ、人の出す手の癖を読み取ることができた。

今後の課題として、今回使用しなかった遺伝的操作を用いる方法が考えられる。本研究では、人相手に有効な戦術の獲得に遺伝的アルゴリズムの考え方を聞いたが、その中でも遺伝的操作の方法には選択・交叉・突然変異方法には様々なものがある。そのため、異なる操作方法、組み合わせを用いることで結果がよくなる可能性がある。従って、そのような方法を検討すべきである。

また、より良い遺伝的パラメータを求めることでコンピュータの勝率を高めることができるため、人相手の実験回数を増やし様々なパラメータでの勝率を確認することも課題として挙げられる。じゃんけんゲームはどうしても確率が絡んでくる問題であるため、各パラメータを用いた実験に対して、それぞれ繰り返して実験を行う方法が考えられる。

謝辞

本論文は、著者が2008年4月から2009年2月までの高知工科大学工学部情報システム工学科在学中に、同学科坂本研究室において行った研究活動の成果を記したものである。

本研究を行うにあたり、本論文やプログラム・就職活動等、多くのご指導を下さった坂本明雄教授に心から感謝致します。論文の執筆を開始するのが遅くなりましたが、何とか完成に至ったのはひとえに先生のおかげです。本当にありがとうございました。

私は4年生から坂本研究室へと配属になったのですが、その際研究室には私を含めて三人の学生しかいなかったため、不安に思うこともありましたが、先生を含めて全員で力を合わせることで、新しいメンバーとなる3年生を大勢研究室に呼び込むことができました。そして、それからは大勢で楽しく研究生活を過ごすことができました。みんなの努力があったからこそだと思います。特に、研究室で唯一の同期の森田和樹君にはプログラムやコンピュータのトラブル等、本当にお世話になりました。心から感謝致します。大学院に進むそうですが、これからは今まで以上に努力しなければならないと思います。健康第一でがんばってください。そして、坂本研究室唯一の大学院生として様々な面でお世話になった小林弘典さんに感謝致します。本当にありがとうございました。また、某カードゲームでの対戦や大会への参加等、勉学の面以外で楽しく研究生活を過ごすことができたのは、坂本研究室の後輩の方々のおかげだと思います。これからも腕に磨きをかけてください。

福本研究室の方々にはサーバー管理を初めとして大変お世話になりました。また、論文の梗概へのアドバイスを初めとして、本稿を記すにあたって貴重なご意見を頂いた福本昌弘准教授および高田喜朗講師に心から感謝致します。

そして、本学に入学して以降、情報システム工学科の先生方を初めとして多くの方々にお世話になりました。その全ての皆様に感謝致します。また、本研究を行うにあたってじゃんけんゲームの対戦相手としてプレイヤーに参加して下さった方に感謝致します。本当にありがとうございました。

参考文献

- [1] 木村将治, 牧野奉裕, 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
固定的戦略による知識表現と遺伝的アルゴリズムによる知識獲得”, 電子情報通信学
会論文誌, Vol. J85-D-I, No.5, pp.479-483, 2002 年 5 月.
- [2] <http://www.shogi.or.jp/osirase/news/2005-10.html> (2008 年 12 月 19 日)
- [3] “コンピュータ将棋は止まらない (ミニ特集)” 情報処理, Vol.49, No.12, 2008 年 8
月.
- [4] 伊庭齐志, 遺伝的アルゴリズムの基礎, オーム社, 2001.