

Title	繰り返し囚人のジレンマゲームを対象とした競合共進化モデルに関する研究
Author(s)	横山, 智生
Citation	
Issue Date	2001-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1477
Rights	
Description	Supervisor:平石 邦彦, 情報科学研究科, 修士

修 士 論 文

繰り返し囚人のジレンマゲームを対象とした 競合共進化モデルに関する研究

指導教官 平石邦彦 助教授

北陸先端科学技術大学院大学
情報科学研究科情報システム学専攻

横山智生

2001 年 2 月 15 日

目 次

1	はじめに	1
2	繰り返し囚人のジレンマゲーム	4
3	競合共進化モデル	7
4	戦略の進化	9
4.1	戦略のコード化	9
4.2	GA による戦略の進化	11
4.3	CCE による戦略の進化	13
5	シミュレーション	19
5.1	交叉	19
5.2	突然変異	20
5.3	再生	20
5.4	初期集団の設定	21
5.5	各設定	22
6	結果	23
6.1	収束速度	23
6.2	生成集団の特徴	29
6.3	生成個体の特徴	30
6.4	集団と個体の推移	32
6.5	“その他” の集団の特性	40
7	まとめ	45
7.1	考察	45

7.2 おわりに	49
謝辞	51

第 1 章

はじめに

人間の文明は協調関係のうえに成り立っているといっても過言ではない。しかし、この協調関係がどのようにして成り立っているかという疑問に対しては様々な議論が成されている。エゴイストとは常に他人を押しつけて生きてゆくものだろうか。エゴイストが、自発的に他人と協調することはありえないのだろうか。彼らは、中央の権力に強制されなければ、協調などしないのだろうか。このような問題が、長い間人々の興味をそそってきた。

今から 300 年以上も前においては、中央の権力による統治なくしては協調社会などありえないというのが定説であった [3]。だからこそ強い政治組織が必要であるという考えが定着し、それ以降、政治組織の本来あるべき姿について議論するときには、権力による政治を抜きにしても協調関係が生まれることを期待できるかどうかしばしば焦点となった。しかしながら、現代社会においては様々な場面で中央権力の存在しない協調関係が多く見られる。このような問題を議論するうえで特に問題になるのがジレンマの発生である。ことに、国と国との間の協調関係における国際問題では、ジレンマの発生がしばしば焦点となってきた。

このようなジレンマの発生している複雑な問題を単純化するために、近年ではゲーム理論による解析が行われてきている。ゲーム理論は、意思決定主体であるプレイヤーの間に存在する利害の競合を分析することを目的とする応用数学の一分野であり、経済学、生物学、政治学、情報科学など様々な研究分野を対象とする意思決定状況の分析の枠組みとして考えられるようになってきた [4]。その中でもジレンマ状況の集約モデルとして知られ

ているのが囚人のジレンマゲームである。囚人のジレンマゲームは 1950 年に二人の科学者によって発見されたもので、様々な分野のジレンマ問題で利用されている。

その繰り返しゲームである繰り返し囚人のジレンマゲーム (Iterated Prisoner's Dilemma Game: IPD) を用いてこのようなジレンマ問題を解明しようとする研究が多く報告されていて、それらの研究の多くが IPD の戦略を遺伝的アルゴリズム (Genetic Algorithm: GA) を用いて進化させるといった方法を用いている。中でも最も有名なものが、R.Axelrod による研究である。Axelrod の研究では協調関係がどのような場面に生じるかを理解することにより、ある特定の状況で協調関係を進展させていくのにふさわしい行動を明らかにするのが目的とされ、様々な戦略の総当たり戦とそれにおける特徴的な戦略の性質の分析を行っている [5]。彼は IPD の様々な戦略の総当たり戦を行うために、ゲーム理論家などに戦略の計算プログラムを持ち寄ってのトーナメント戦の参加を呼びかけた。そのトーナメント戦の結果、しっぺ返し戦略 (Tit for Tat : TFT) が総合優勝した。しかし、この TFT は個々の対戦では勝利することなく、ほとんどの対戦において同点で引き分けている。すなわち TFT は、相手を打ち負かすことを目的とする戦略ではなく、協調関係を維持するための戦略であることがいえる。以降、このような、IPD の戦略を進化させることにより協調関係を導き出すような研究においては、この TFT が常に注目されてきた。

このように IPD の戦略を GA を用いて進化させると協調関係の成り立つ集団構成へと進化する可能性があることが分かっているが、ここで、GA による社会形態のモデル化が適切であるのかが疑問点となってくる。近年では社会形態の複雑化が進み、様々な集団が互いに影響しあって共存しているのが現状である。GA の進化のメカニズムは、幾つかの個体の集まりからなる 1 つの集団を設け、集団内の個体を GA オペレータによって進化させていくことにより集団そのものを変化させていくものである。個々のエージェントがどのように影響しあっているのかを解明することはできても、集団同士がどのように影響しあっているかは解明できない。そこで、近代社会の複雑な形態をより忠実にモデル化するために、GA のような単純な集団構成の進化モデルではなく、より複雑な集団構成の進化モデルを IPD の戦略の進化に適用することを考える。本研究では、その進化モデルとして競合共進化モデル (Competitive Co-Evolution Model : CCE) の適用を試みる。

CCE は主に対戦ゲームにおける最適戦略を探索する手法として提案された進化モデルで、最適戦略の存在するゲームの最適解探索において、有効である場合があることが報告

されている [2]。CCE の進化のメカニズムは、幾つかの個体の集まりである集団を 2 つ設け、それぞれの集団内の個体は、互いにもう一方の集団内の個体を意識しながら共に進化していくというものであり、集団同士が互いに影響しあう構成になっている。このような複雑な集団構成を持つ CCE を IPD の戦略の進化に適用すると、どのような結果を生じるのかという疑問点が焦点となってくる。しかしながら、IPD のように一般的に最適戦略が存在しないとされている問題に対して、その戦略を CCE によって進化させるとどのような挙動をとるのかという研究はあまり報告されてはいない。そこで、本研究では、IPD の戦略を同じ条件下で GA と CCE のそれぞれで進化させ、GA と CCE が示す特徴の違いを明らかにする。ここでは、主に解の集団全体の振る舞いに着目し、その収束速度や収束解の性質を調べることにより GA と CCE の差別化を図る。GA を用いた従来研究で報告されているように、互いに協調的な戦略を取り合うような集団構成への進化が CCE でも可能なのかが一番の焦点となる。また、それらの結果が GA と CCE のどのようなメカニズムの違いから生じているのかについて考察し、IPD における CCE の特性を明らかにする。

第 2 章

繰り返し囚人のジレンマゲーム

囚人のジレンマゲーム (Prisoner 's Dilemma Game: PD) は、1950 年に Rand 研究所の M.Flood と M.Dresher によって発見され A.W.Tucker によって定式化されたものであり、合理的な手の組み合わせがパレート最適をもたらさない二人非ゼロ和ゲームである。非ゼロ和ゲームの場合は、ゼロ和ゲームのような勝ち負けではなく、それぞれの利得の大きさが評価基準となる。PD では 2 人のプレイヤーが“ 協調 ”か“ 裏切り ”の手を選択し、表 2.1 の利得行列に基づいて利得を得て、その利得の大小でどちらが勝利したのかが決まる。表 2.1 では、 X が行プレイヤー、 Y が列プレイヤーを表す。

表 2.1: 利得行列

X, Y	協調	裏切り
協調	R, R	S, T
裏切り	T, S	P, P

この利得表には以下の条件が付け加えられる。

$$\begin{aligned} T &> R > P > S \\ 2R &> T + S \end{aligned}$$

これらの条件の意味は、共に“ 協調 ”の手をとる方が共に“ 裏切り ”の手をとるよりも利得が高いが、自分が“ 協調 ”の手をとり、相手が“ 裏切り ”の手をとると相手の利得が高くなり、自分の利得が低くなるというものである。PD では両プレイヤーが“ 裏切り ”をとるのが合理的となり、両プレイヤーが“ 協調 ”をとると 2 人の合計利得が最大になる。よって支配的な手の組み合わせが存在することになり、その支配的な手というのが非協調を意味するため、パレート最適をもたらす協調関係を意味する手をとるのが最も困難なゲームだと考えられている。このように相手がどのような手を選択しようと、自分の方は協調するより裏切った方が必ず得である。けれども、両方とも裏切りを選べば、両方が協調しあう場合に比べ損をしてしまう。ここにジレンマが発生するわけである。

この PD を繰り返し行うゲームが IPD であり、一般的にその繰り返し回数は未知数とする。これは、繰り返し回数が既知数ならば最終回から逆に分析することにより、理論的に協調関係を引き出すことができないことが明らかなためである。具体的には、まず、最終回ではその後のゲームが存在しないから一回限りのゲームと全く同じ状況になる。この状況において、裏切りの手をとるという合理的なエージェントの論理は何の欠陥も持たないから、双方のエージェントとも裏切りの手をとることになる。次に最後から 2 番目の回では最終回で双方のエージェントが裏切りの手をとることがわかっているから、双方のエージェントとも将来を考慮する必要がなくなる。再び裏切りの手が支配的になり、双方のエージェントとも裏切りの手をとることになる。次に最後から 3 番目の回では、最後から 2 番目の回以降で双方のエージェントが裏切りの手をとることがわかっているから、双方のエージェントとも裏切りの手をとることになる。以下、このように帰納することにより繰り返し回数が既知ならば、双方のエージェントとも初回から常に裏切りの手をとることを予測し得る。しかし繰り返し回数が未知ならば、協調関係を発現させ得る可能性がでてくる。なぜならば、繰り返し回数が既知の場合の帰納が出発点を持たなくなり、常に将来ゲームが行なわれ得るという可能性が現在の手の選択に影響を与えることになるからである。

このゲームには様々な戦略が存在するが、その中でも代表的なものが、*AllC*(all cooperate、全て協調)、*AllD*(all defect、全て裏切り)、*TFT*(tit for tat、しつぺ返し) である。Axelrod はこの IPD で相手の過去の手に応じて自分の手を決定するプログラムを募集し、2 度のトーナメントを行なったが 2 度とも TFT が優勝したと報告している [1]。しかし

ながら、 TFT は個々の対戦で勝利することによって利得を得たのではなく、むしろ引き分けに近い形で、敗戦を避けて高い総利得を得ているのであり戦略として最強というわけではない。しかし、 TFT が集団内を占めると、全ての個体が“協調”の手をとり合う協調集団が生成され、合計利得が高くなる。この場合、いかなる戦略もこの集団内に侵略することは不可能となり、 TFT は集団的安定な解の1つであると定義されている [6]。

第 3 章

競合共進化モデル

CCE は 2 つの集団を共に競い合わせながら進化させていく進化モデルである。例として、ある生物（被食者）とそれを捕食するもの（捕食者）の関係を考える。捕食者はより捕食能力を高めるように進化が進み、被食者はより捕食されにくくなるように進化を深めていく。自然に、互いに相手を意識しながら自分自信を進化させていくことになる。つまり、その相手が、自分自信が進化していく際の評価規範となるわけであり、それぞれの生物は、互いに自分が相手に対して有利になるように進化していくことになる。CCE はこのような自然界の進化のメカニズムを模倣した進化モデルであり、対戦ゲームの解探索の手法として 1995 年に D.Rosin らによって提案された [2]。一般的な GA では、進化させる集団は 1 つである。それに対して、CCE では、2 つの集団を用意し、互いに相手集団に対して優位に立とうとしながら進化していく。2 つの集団を互いに競い合わせるという点において、ゲームのような対戦型の問題に適しているのではないかと考えられている。

具体的には、GA 操作によって実際に進化していく進化集団と、その進化集団の進化の際の評価規範となる環境集団との 2 つの集団を設けるわけだが、ある一定の条件の下で、それら 2 つの集団は互いに役割を交代することになっている。つまり、それまで進化集団だった集団が環境集団になり、環境集団だった集団が進化集団になるわけである。ここで、その集団の交代の条件が問題となってくる。最適戦略の存在するような対戦ゲームにおいては、進化集団が環境集団に勝利すれば、つまり進化集団が利得で環境集団に上回れば集団の交代を行うものとしている [2]。しかし、この方法が最適戦略の存在しないような対戦ゲームにおいても有効なのかは報告例がない。本研究では、この CCE の交代条件

に着目しながら、IPD の戦略の進化に適用することを試みる。

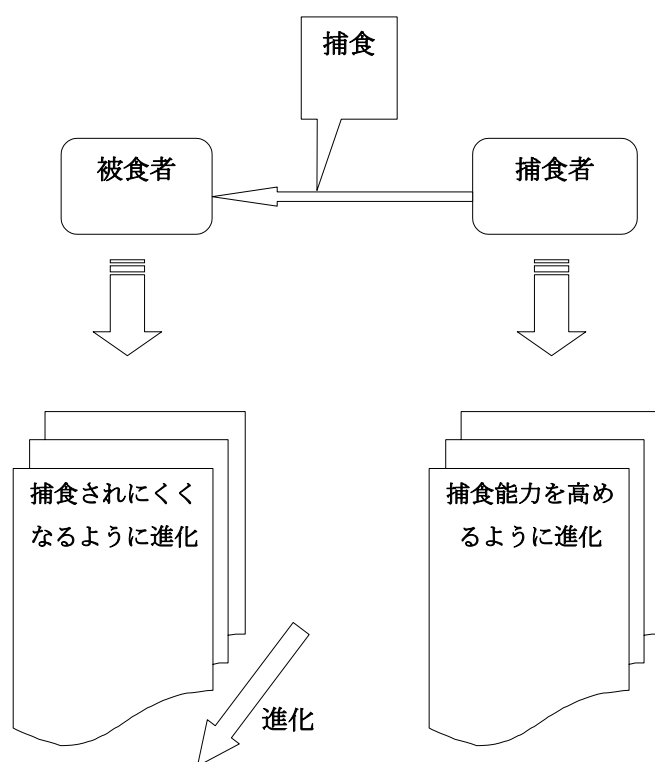


図 3.1: concept of CCE

第 4 章

戦略の進化

4.1 戦略のコード化

進化モデルを用いて IPD の戦略を進化させていく際には、戦略をコード化する方法を決めなければならない。ここでは過去の双方の 2 手を参照する戦略でコード化を行なう。図 4.1 のように、1 ビット目が自分の初手、2、3 ビット目が相手の初手に対する自分の 2 手目、4 ビット目以降は 3 手目からの自分の手を決めるコードとなり、全部で 19 ビットの表現となる。この 19 ビットのコードで、1 つの個体の戦略を表す。

3 手目以降の手は、表 1 と表 2 のコード表にその戦略のコードを照らし合わせることでより決まる。“ C ”は“ 協調 ”、“ D ”は“ 裏切り ”のコード化をそれぞれ表している。表 4.1 のコード化の欄には例として TFT の戦略のコード化を記している。例えば、表 4.2 より 1 手前の自分の手が“ 協調 ”、相手の手が“ 裏切り ”であれば記述は“ S ”となり、2 手前の自分の手が“ 裏切り ”相手の手も“ 裏切り ”であれば記述は“ P ”となる。この記述“ S-P ”を表 4.1 に照らし合わせると、コード化の欄に“ D ”と書かれている。よって自分の次の一手は“ 裏切り ”を出すことになる。このように 1 手前の自分と相手の手、2 手前の自分と相手の手を参照することにより次の自分の一手が決定される。

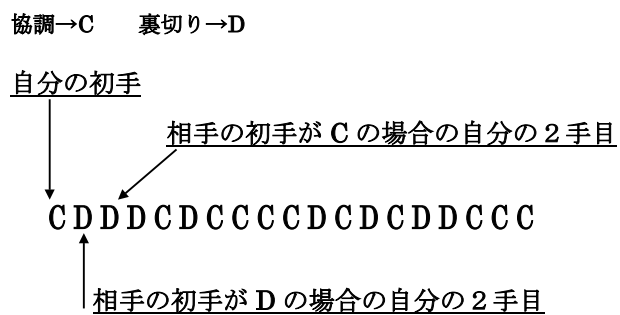


図 4.1: coding

表 4.1: 16 ビットのコード化

R	R	R	R	T	T	T	T	S	S	S	S	P	P	P	P	1 手前
R	T	S	P	R	T	S	P	R	T	S	P	R	T	S	P	2 手前
C	C	C	C	C	C	C	C	D	D	D	D	D	D	D	D	コード

表 4.2:

自分	相手	記述
C	C	R
D	C	T
C	D	S
D	D	P

4.2 GA による戦略の進化

図 4.2 に GA による IPD の戦略の進化のアルゴリズムを示す。まず、進化させる戦略個体の集団を初期集団の生成によって設定する。その生成された戦略個体同士を集団内で IPD の総当り対戦させることによって、それぞれの個体は利得を得る。そして、その利得によってそれぞれの個体の適応度が決定される。それぞれの個体は、その適応度をもとに交叉や突然変異等の GA オペレータを行って新しい個体に生まれ変わり、そしてまた対戦ゲームを繰り返す。そして、そのゲームによって得た適応度の評価をもとに選択によって再生される。このループを GA の終了条件に達するまで繰り返し、終了条件に達すれば終了となる。ここでの終了条件は世代数としている。つまりあらかじめ設定している世代数に達するとアルゴリズムは終了する。

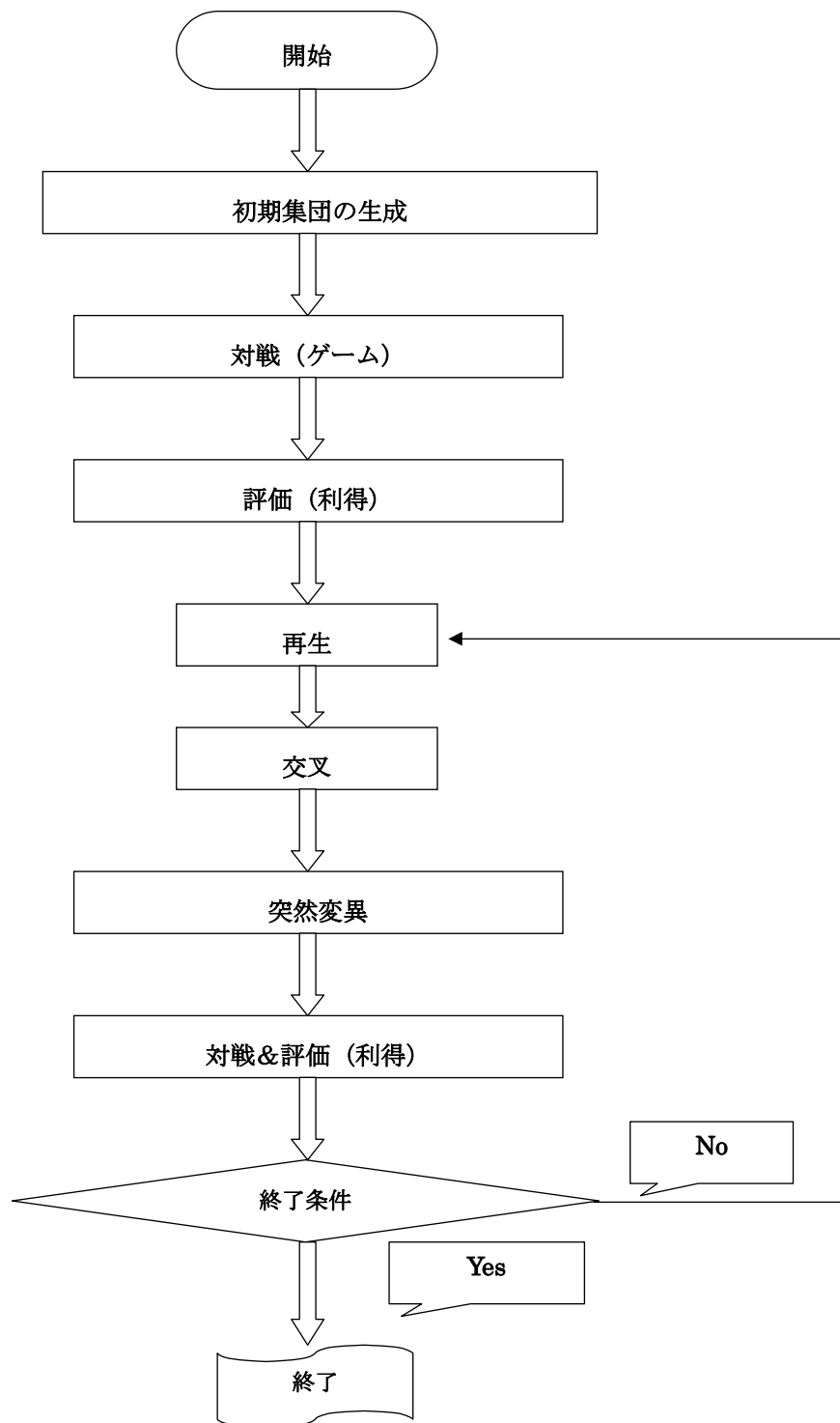


図 4.2: flowchart-GA

4.3 CCE による戦略の進化

図 4.3 に CCE のアルゴリズムを示す。進化させていく集団が 1 つである GA に対して、CCE では実際に進化していく進化集団と、適応度の評価規範となる環境集団の 2 つの集団を設ける。

まず、初期集団の生成で進化集団の個体群と、環境集団の個体群を設定する。進化集団内の個体は環境集団内の個体と IPD による総当りゲームを行い、それによって利得を得て、それぞれの個体の適応度を決める。そして、その適応度をもとに進化集団内の個体に対して GA 操作を行い、新しい個体群からなる進化集団が形成される。その生成された新しい進化集団の個体を、また環境集団の個体と IPD 対戦させ、それぞれの個体に再び適応度を与える。このように進化集団は GA の終了条件を満たすまで環境集団との IPD 対戦で得た適応度をもとに GA 操作を繰り返すことになる。ここでの GA の終了条件も、GA による戦略の進化と同じように設定世代数とする。この終了条件を満たせば進化集団は環境集団とその役割を交代する。つまり、GA の終了条件に達するとそれまでの進化集団が環境集団に、環境集団が進化集団に変化する。そして、同様な進化を繰り返し、CCE の終了条件を満たせば全ての行程が終了する。

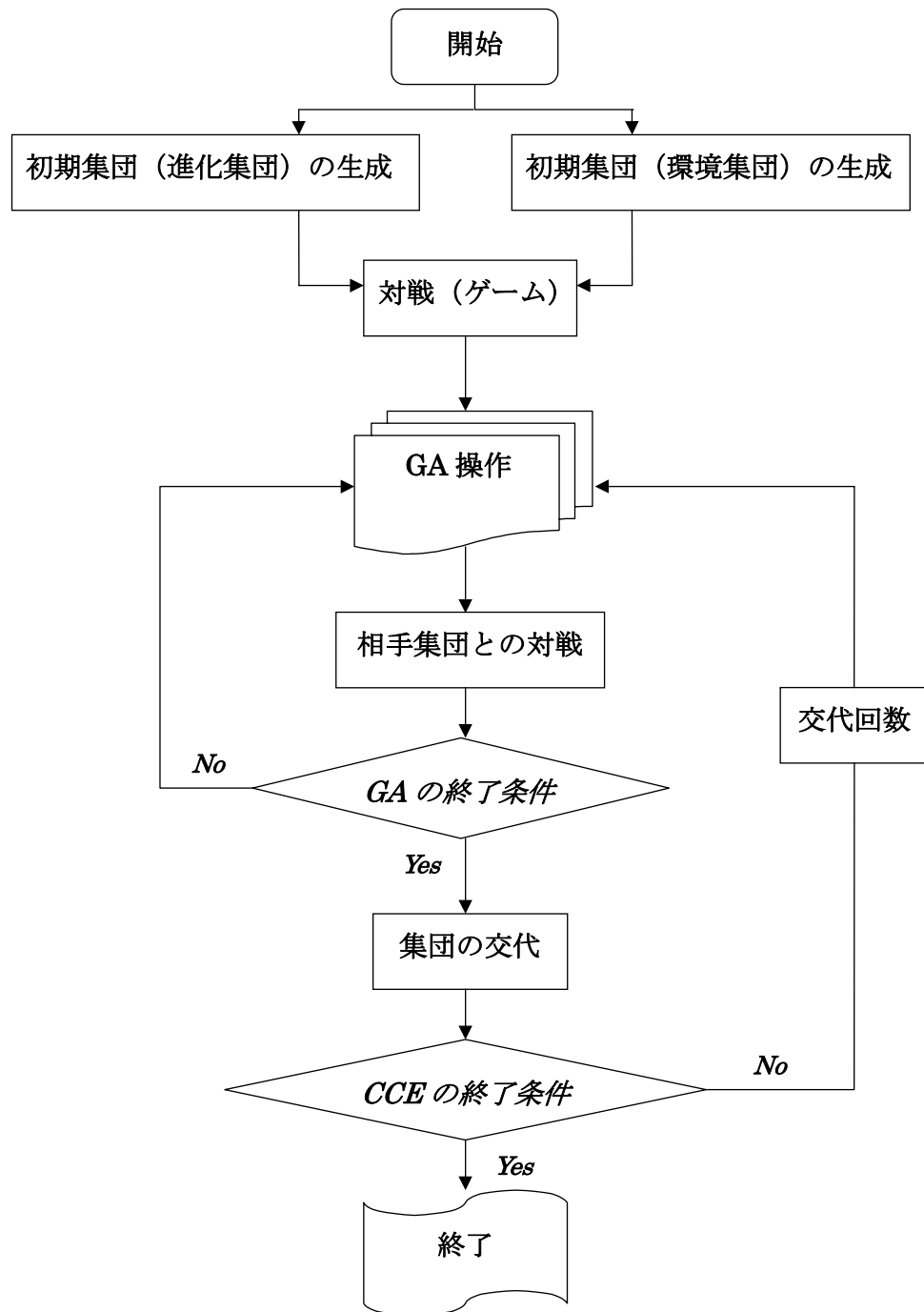


図 4.3: flowchart-CCE

しかしながら、ここで CCE の終了条件をどうするかが問題となってくる。最適戦略が存在するような対戦ゲームにおいて既に報告されているように [2]、進化集団内の個体の総利得が対戦相手である環境集団内の個体の総利得を上回れば CCE を終了して集団の交代を行う方法について検証してみる。表 2.1 の利得表にあるように、両者が“協調”を出し合えば両者の利得は共に“ R ”で、両者が“裏切り”を出し合えば両者の利得は共に“ P ”となる。そして、一方が“協調”、もう一方が“裏切り”を出した場合、“協調”を出した方の利得は“ S ”となり“裏切り”を出した方の利得は“ T ”となる。ここでは、 TFT と $AllD$ からなる集団を考える。まず最初の世代では GA と CCE の集団内の TFT と $AllD$ の比率は全く同じものであると仮定する。集団内の個体数を N とし初期集団における TFT の割合を n 、1 個体と 1 個体の対戦ゲーム数を G とすると、

TFT 1 個体と $AllD$ 1 個体の対戦

$$TFTの総得点 = S + P(G - 1)$$

$$AllDの総得点 = T + P(G - 1)$$

TFT 1 個体と TFT 1 個体の対戦

$$TFTの総得点 = GR$$

$AllD$ 1 個体と $AllD$ 1 個体の対戦

$$AllDの総得点 = GP$$

TFT 1 個体の 1 世代における総得点

$$GRnN + (S + PG - P)(1 - n)N = \alpha \quad (4.3.1)$$

$AllD$ 1 個体の 1 世代における総得点

$$(T + PG - P)nN + (1 - n)NGP = \beta \quad (4.3.2)$$

集団内における全 TFT の得点合計

$$nN\alpha$$

集団内における全 $AllD$ の得点合計

$$(1 - n)N\beta$$

集団内における全個体の得点合計

$$nN\alpha + (1 - n)N\beta$$

となる。これは、進化集団、環境集団共通の得点である。

ここで、世代が1世代進み、進化集団内の TFT の割合が e 増加した場合を考える。すると、先に示した得点合計の式は以下のように変化する。

(進化集団)

TFT 1 個体の 1 世代における総得点

$$GRnN + (S + PG - P)(1 - n)N = \alpha \quad (4.3.3)$$

$AllD$ 1 個体の 1 世代における総得点

$$(T + PG - P)nN + (1 - n)NGP = \beta \quad (4.3.4)$$

集団内における全 TFT の得点合計

$$(n + e)N\alpha$$

集団内における全 $AllD$ の得点合計

$$(1 - n - e)N\beta$$

集団内における全個体の得点合計

$$(n + e)N\alpha + (1 - n - e)N\beta \quad (4.3.5)$$

(環境集団)

TFT 1 個体の 1 世代における総得点

$$= GR(n + e)N + (S + PG - P)(1 - n - e)N = \bar{\alpha} \quad (4.3.6)$$

AllD 1 個体の 1 世代における総得点

$$= (T + PG - P)(n + e)N + (1 - n - e)NGP = \bar{\beta} \quad (4.3.7)$$

集団内における全 *TFT* の得点合計

$$nN\bar{\alpha}$$

集団内における全 *AllD* の得点合計

$$(1 - n)N\bar{\beta}$$

集団内における全個体の得点合計

$$nN\bar{\alpha} + (1 - n)N\bar{\beta} \quad (4.3.8)$$

ここで、進化集団の全個体の得点合計と環境集団の全個体の得点合計を比較してみる。

式 (4.3.5) と式 (4.3.8) より

環境集団の全個体の得点合計－進化集団の全個体の得点合計

$$\begin{aligned} & nN\bar{\alpha} + (1 - n)N\bar{\beta} - (n + e)N\alpha + (1 - n - e)N\beta \\ &= \bar{\beta}(1 - n) + \beta(n - 1) + n(\bar{\alpha} - \alpha) + e(\beta - \alpha) \end{aligned} \quad (4.3.9)$$

式 (4.3.4) と式 (4.3.7) より

$$\bar{\beta} - \beta > 0$$

$n \leq 1$ なので、

$$\bar{\beta}(1 - n) \geq \beta(n - 1) \quad (4.3.10)$$

また、 $n + e \leq 1$ 、式 (4.3.3)、式 (4.3.4)、式 (4.3.6) より、

$$n(\bar{\alpha} - \alpha) \geq (\beta - \alpha) \quad (4.3.11)$$

式 (4.3.3) と式 (4.3.6) より,

$$\bar{\alpha} - \alpha \geq 0 \quad (4.3.12)$$

式 (4.3.10)、式 (4.3.11)、式 (4.3.12) より,

$$\bar{\beta}(1 - n) + \beta(n - 1) + n(\bar{\alpha} - \alpha) + e(\beta - \alpha) \geq 0 \quad (4.3.13)$$

式 (4.3.13) より、環境集団の全個体の得点合計の方が進化集団の全個体の得点合計より高くなっていることがわかる。このことにより、進化集団の TFT の割合が高くなる限り、進化集団の全個体の得点合計が環境集団の全個体の得点合計を上回ることがないことが証明された。つまり、進化集団の TFT の割合が低くならない限り集団の交替が行なわれないということになる。これでは、進化集団が常に”非協調集団”へと進化していき、限り CCE の動作が正常に行なわれない。仮に進化集団が”協調集団”へと進化し続けると一切集団の交替が行なわれず、無限ループに陥ってしまう。このように、進化集団内の個体の総利得が対戦相手である環境集団内の個体の総利得を上回れば CCE を終了して集団の交代を行う方法は、IPD 戦略の進化では適用することができないことがわかった。よって、ここでは CCE が正常に作動するために、CCE の終了条件を設定世代数とする。

また、GA による戦略の進化による結果と CCE による戦略の進化の結果を比較し易くするために、2つのモデルのその他の GA オペレータは全く同じ設定とする。このように、GA と CCE はアルゴリズムとしての違いはそれほど大きいものではない。しかしながら、GA では同じ集団内で対戦を行いながら進化していくため、評価規範である自分以外の個体も常に変化していくが、CCE では、同じ集団内では一切対戦ゲームは行なわず、環境集団の個体と総当りゲームを行なうことにより利得を評価するため、評価規範はある一定の間変化しない。この評価関数の違いが結果にどのように影響してくるのかを実験により調べる。

第 5 章

シミュレーション

5.1 交叉

ここでは、 n 点交叉 (multi-point crossover) を用いる。図 5.1 のように $n = 2$ として、2 点の交叉点を個体別に設定し戦略のコードに交叉を施す。まず、交叉を行なう 2 つの親個体をランダムに選びだし、その 2 つの個体に対して共通の交叉点をランダムに 2 箇所設定する。それぞれの交叉点に沿って、個体のコードを 3 つに分解し、2 つの個体をつなぎ合わせることによって全く異なる個体 2 つが生成される。ここでは、交叉率を 100% に設定する。つまり、集団内の全個体数を N とすると、交叉回数は $\frac{N}{2}$ となる。

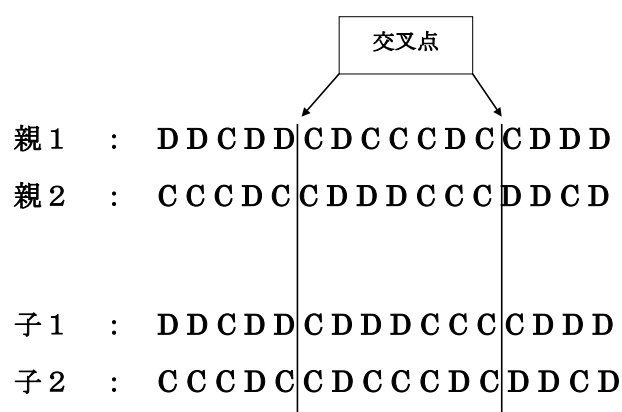


図 5.1: 2-point crossover

5.2 突然変異

突然変異は図 5.2 のように、個体の遺伝子 1 ビットを反転することによって行なう。突然変異を行なう個体は設定された突然変異率に従ってランダムに選び出される。そしてその選び出した個体の中で、突然変異を施す遺伝子をランダムに 1 つ選びだしコードを反転する。具体的には、選び出された遺伝子コードが”C” なら”D” へ、”D” なら”C” へと変換する。

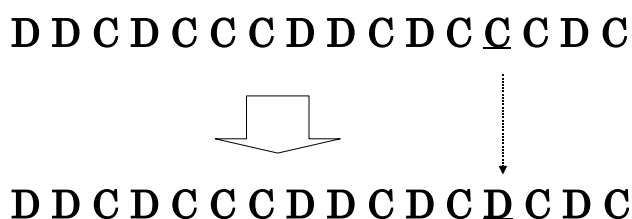


図 5.2: mutation

5.3 再生

ここでは、再生にルーレット選択を用いる。適応度が高い個体ほど選択される可能性が高くなり、適応度の低い個体ほど選択確率が低くなっている。つまり、環境に適応している生物ほどより多くの子孫を残し、適応していない生物は淘汰されていくという自然界の進化のメカニズムを模倣しているわけである。具体的には、各個体の適応度に応じて、各々の個体にランク（ルーレット幅）を設定し、ランダムにルーレットを回して個体の選択を行なう。その際、適応度の高い個体ほどルーレット幅が広く、選び出される可能性が高くなっている。このルーレット選択は、確率的に個体を選び出すので適応度の高い個体は確実に再生され、適応度の低い個体は確実に淘汰されるというわけではない。

各個体 s_i の適応度を $f(s_i)$ 、集団内の総個体数を N とすると、選択確率 P は以下のようになる。

$$P(i) = \frac{f(s_i)}{\sum_{i=1}^N f(s_i)}$$

ここで、適応度 $f(s_i)$ は、各個体の 1 ゲーム当たりの平均利得であり、対 1 個体当たりの対戦ゲーム数を G 、各個体の総利得を V とすると、以下のようになる。

$$f(s_i) = \frac{V(s_i)}{GN}$$

5.4 初期集団の設定

ここでは、多数の利己的なエージェントからなる集団が協調集団へと進化することが可能なかを考察するため、初期集団は多数の非協調的な個体と少数の協調的な個体とを混ぜ合わせたものにする。非協調的な戦略には全て裏切りの All D を用い、協調的な戦略には TFT を用いる。All C は集団的安定 [6] な個体ではない為、早い段階で All D のような非協調的な個体によって淘汰される可能性が非常に高い。よって本研究では All C を初期集団には用いない。

GA、CCE とともに表 5.1 のように初期集団を 4 つのパターンに設定してシミュレーションを行う。4 パターン全て、多数の All D と少数の TFT の組合せになっており、パターン”4”に近付くにつれ、TFT の割合が低くなり All D の割合が高くなっている。

表 5.1: 初期集団の設定

パターン	<i>TFT</i> の割合	<i>AllD</i> の割合
1	30%	70%
2	20%	80%
3	10%	90%
4	4%	96%

5.5 各設定

シミュレーションでは、1 個体対 1 個体の IPD のゲーム対戦回数は 100 回行なう。GA オペレーションの諸設定では、突然変異率 3%、エリート保存率 3%、交叉率 100%とする。以上は GA、CCE とともに共通の設定である。個体サイズにおいては、GA は 100 個体、CCE は進化集団 50 個体と環境集団 50 個体の計 100 個体とする。終了条件となる設定世代数は、GA は 100 世代、CCE は 10 世代×20 交替の計 200 世代とする。つまり、CCE においては、20 回の集団の交替を行ない、その 1 つの交替の中で進化集団を 10 世代進化させることになり、2 つの集団はそれぞれ計 100 世代ずつ進化することになる。これらの設定により、GA と CCE の IPD のゲーム対戦回数はほぼ等しくなる。以上の設定で、初期集団の 4 パターンにおいてそれぞれ 100 回ずつシミュレーションを行なう。また、IPD の利得は表 5.2 のように設定する。

表 5.2: IPD 利得

X, Y	協調	裏切り
協調	3, 3	1, 4
裏切り	4, 1	2, 2

このとき、集団内のほとんどの個体が“ 協調 ”の手を出し合えば集団の平均利得は限りなく“ 3 ”に近づき、“ 裏切り ”の手を出し合えば平均利得は限りなく“ 2 ”に近づくことになる。

第 6 章

結果

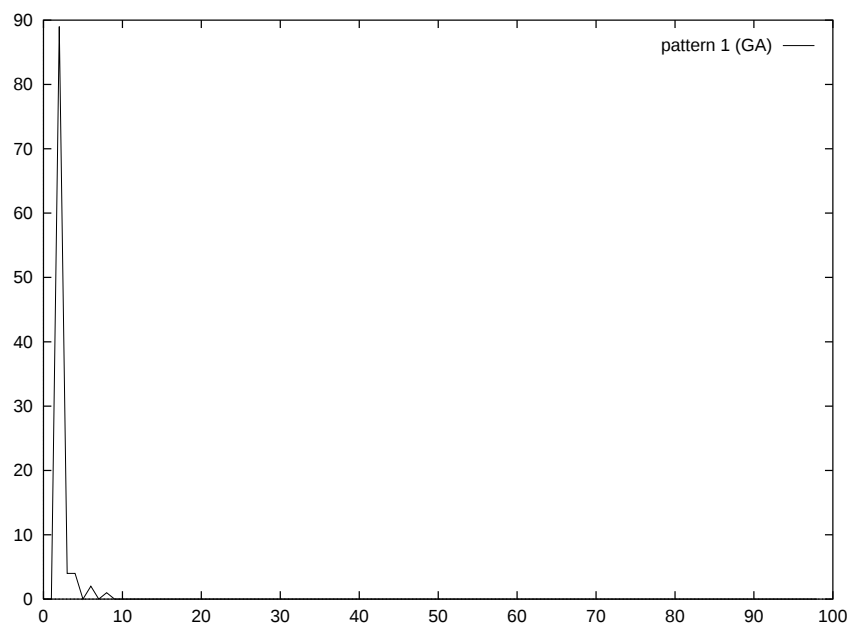
6.1 収束速度

ここでの収束とは、集団内の全ての個体の平均応度がある程度一定値に収まり、終了するまでそのままその値を保ち続ける場合を指す。表 6.1 はそれぞれの初期集団のパターンにおいて、GA と CCE のシミュレーション 100 回分の収束世代の平均をとったものである。数値は収束した世代数を表している。

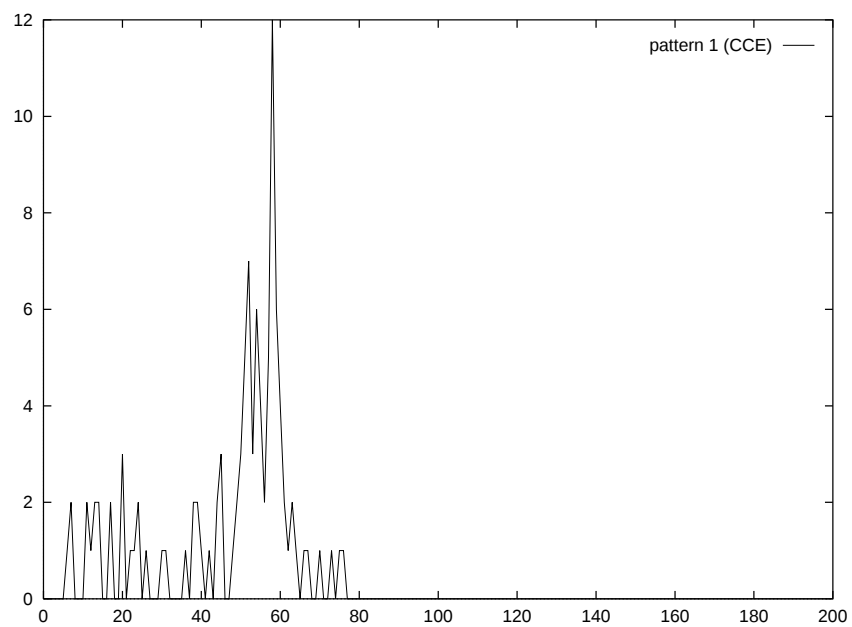
表 6.1: 収束速度

パターン	GA	CCE
1	3.28	53.77
2	3.45	63.80
3	4.23	89.21
4	4.12	92.94

図 6.1 と図 6.2 のグラフはそれぞれパターン” 1 ”における GA と CCE の収束状況の分布図である。縦軸がデータ数を表し、横軸が収束した世代数を表している。



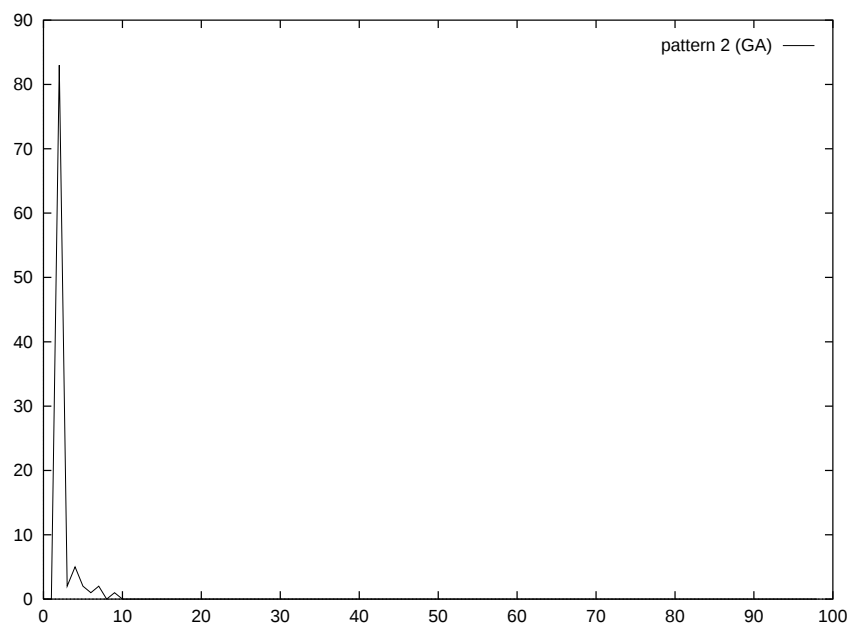
⊠ 6.1: pattern 1 (GA)



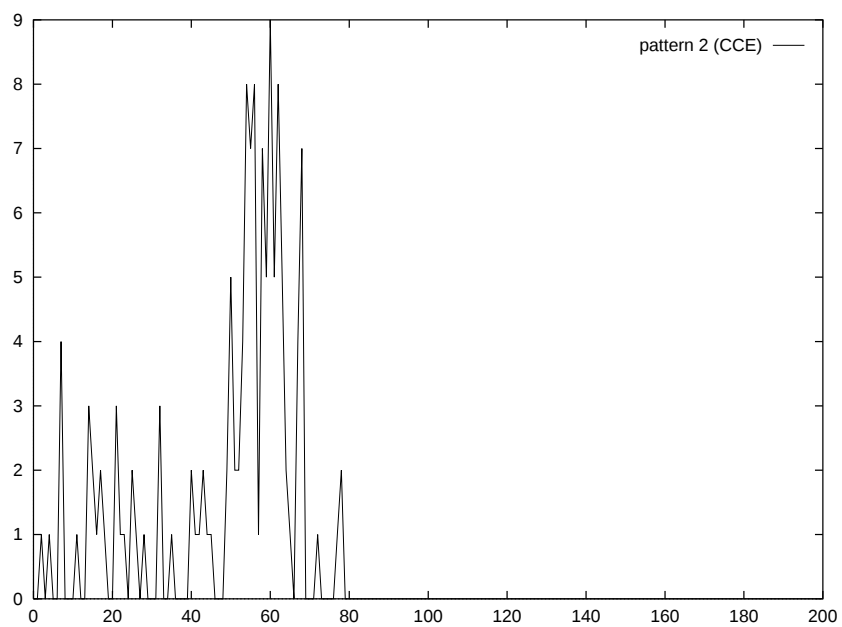
⊠ 6.2: pattern 1 (CCE)

図 6.3 と図 6.4 のグラフはそれぞれパターン” 2 ”における GA と CCE の収束状況の分布図、図 6.5 と図 6.6 のグラフはそれぞれパターン” 3 ”における GA と CCE の収束状況の分布図、図 6.7 と図 6.8 のグラフはそれぞれパターン” 4 ”における GA と CCE の収束状況の分布図である。

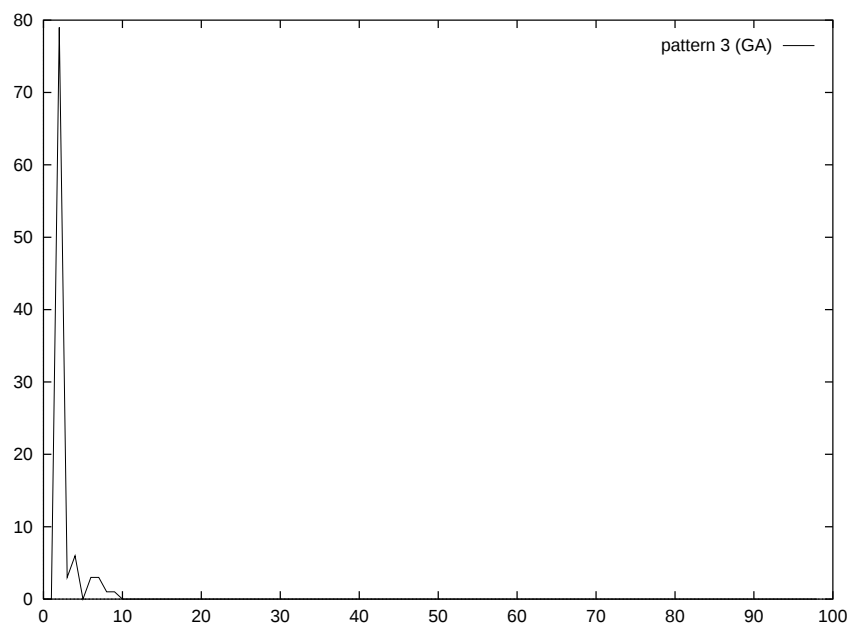
全てのパターンにおける GA、CCE それぞれの収束状況の分布図を見ても分かるように、GA は CCE に比べかなり速い世代で収束してしまっているが、CCE は遅い世代で収束している。また、GA の場合、パターンを変えても収束状況にそれほど変化は現れないが、CCE の場合、パターン” 4 ”に近付くにつれ、つまり初期集団における *TFT* の割合が低くなるとそれに比例して収束速度が遅くなっている。



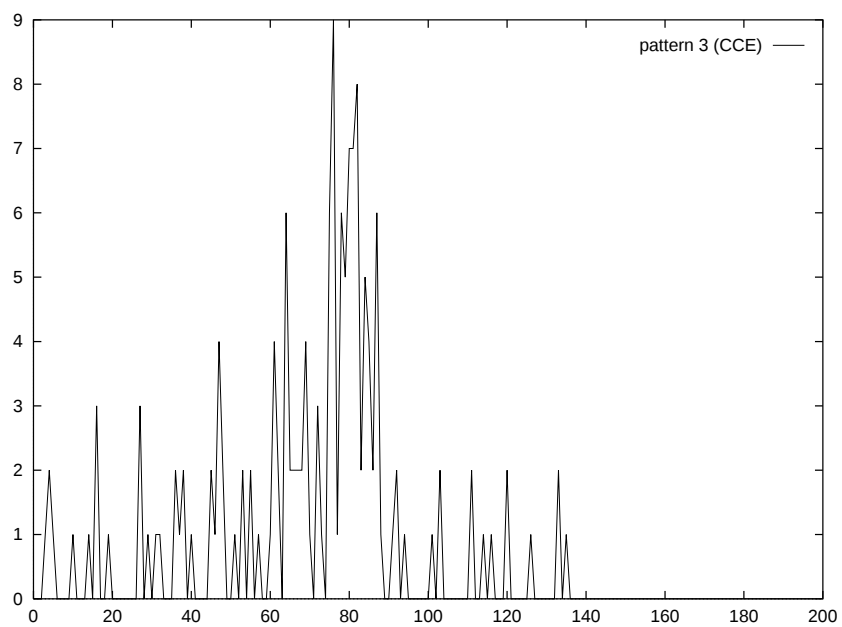
⊠ 6.3: pattern 2 (GA)



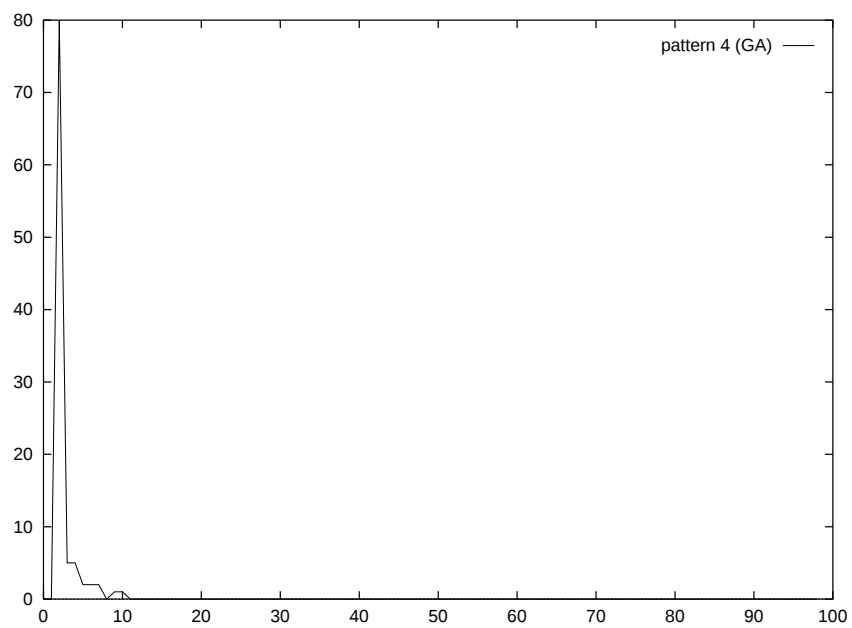
⊠ 6.4: pattern 2 (CCE)



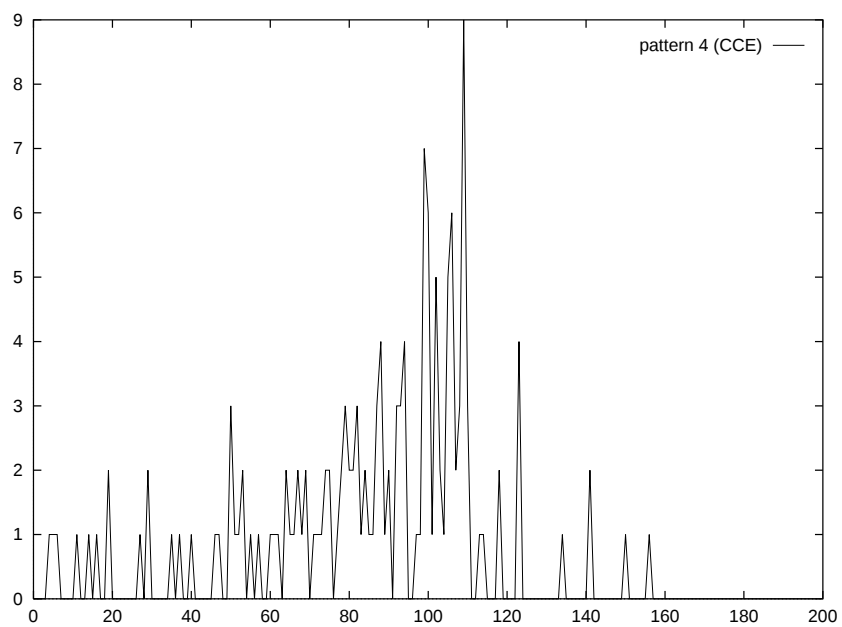
⊠ 6.5: pattern 3 (GA)



⊠ 6.6: pattern 3 (CCE)



⊠ 6.7: pattern 4 (GA)



⊠ 6.8: pattern 4 (CCE)

6.2 生成集団の特徴

本研究では、進化モデルが協調的な手をとるような戦略個体の集団に進化させることが可能なのかということが1つの大きな目的であるので、生成された集団が協調的な集団なのかを判断する必要がある。そのため、ここでは、生成された集団を3種類に分類することによって、その集団の特徴の分析を行なう。表 5.2 の IPD 利得より、集団内の個体が全て”協調”の手をとり合っている場合、集団の平均利得は”3”となり、集団内の個体が全て”非協調”の手をとり合っている場合、集団の平均利得は”2”となる。このことを利用して、集団内の個体の平均適応度が 2 ± 0.1 に収束した場合を”非協調集団”とし、 3 ± 0.1 に収束した場合を”協調集団”とし、それ以外の値を示すものは”その他”としている。

表 6.2: 集団の分類の定義

集団の分類	集団内の全個体の平均適応度
協調集団	3 ± 0.1
非協調集団	2 ± 0.1
その他	その他の値

表 6.3 に各パターンにおける GA、CCE それぞれの最終平均利得による集団の分類の割合を示す。全体的に CCE より GA の方が”非協調集団”に進化する割合が高く、特に、パターン”3”、”4”のように初期集団における TFT の割合が低い場合においてはその差が明確である。そして、その分”その他”の集団に進化している割合が GA より CCE の方が高くなっている。

表 6.3: 集団の分類の割合

パターン	集団の分類	GA	CCE
1	協調集団	91.1%	87.2%
	非協調集団	8.9%	0.8%
	その他	0%	12.0%
2	協調集団	44.8%	48.3%
	非協調集団	55.2%	36.8%
	その他	0%	14.9%
3	協調集団	26.6%	26.2%
	非協調集団	71.8%	53.7%
	その他	1.6%	21.1%
4	協調集団	3.8%	14.6%
	非協調集団	95.1%	63.4%
	その他	1.1%	22.0%

6.3 生成個体の特徴

集団そのものがどのような特性を持っているのかという点と同時に、集団内の個体そのものがどのような特性を持っているのかという点についても調べる。個体の特性の分別は以下のように行う。各世代ごとに生成された個体全てに対して、あらかじめ用意した“ All D ”の戦略個体と“ All C ”の戦略個体を対戦させる。その結果によって各個体を“ All C 類 ”、“ All D 類 ”、“ TFT 類 ”、“ その他 ”の 4 種類の特徴に分別する。各特徴の定義は表 6.4 より定める。表内の数値は適応度（1 ゲーム当りの平均利得）を表している。

表 6.5 に世代ごとの、各種類の個体の占める割合の平均値を示す。GA の方は、パターン “1” においては All C 類や TFT 類の割合が高く、All D 類の割合が低い。それがパターン “4” に近づくにつれてその割合が段々と入れ替わり、最終的には All D 類の割合が非常に高くなってきている。それに対して CCE の方は GA ほど割合の差が大きくはなく、パターン “4” においても比較的 All D 類の割合は低く、その分その他の個体の割合が高くなっている。

表 6.4: 個体分類の定義

個体の分類	<i>AllC</i> と対戦	<i>AllD</i> と対戦
<i>AllC</i> 類	3 ± 0.1	1 ± 0.1
<i>AllD</i> 類	4 ± 0.1	2 ± 0.1
<i>TFT</i> 類	3 ± 0.1	2 ± 0.1
その他	その他の値	その他の値

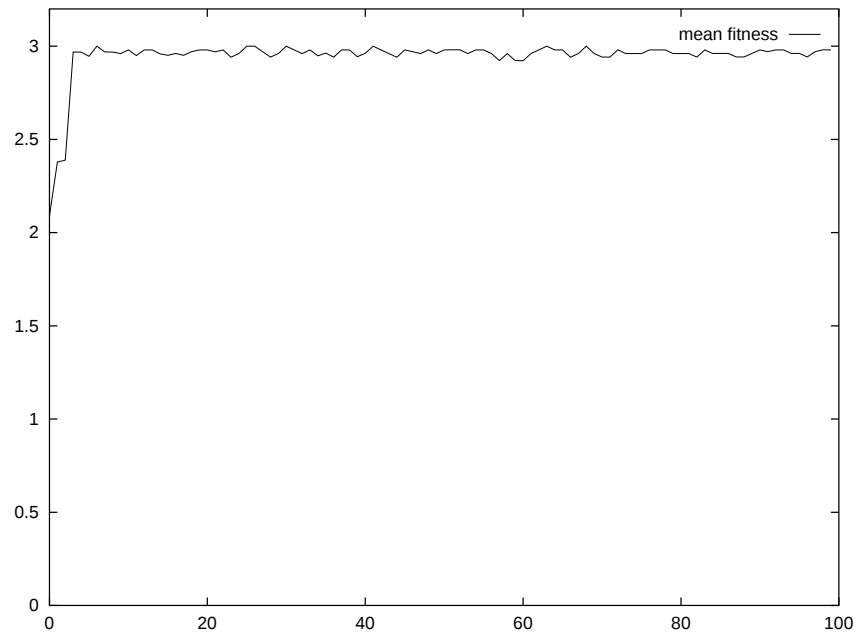
表 6.5: 集団の分類の割合

パターン	集団の分類	GA	CCE
1	<i>AllC</i> 類	26.9%	21.2%
	<i>AllD</i> 類	5.1%	7.2%
	<i>TFT</i> 類	61.3%	56.7%
	その他の個体	6.7%	14.9%
2	<i>AllC</i> 類	12.5%	16.6%
	<i>AllD</i> 類	19.6%	17.2%
	<i>TFT</i> 類	59.7%	42.3%
	その他の個体	8.2%	23.9%
3	<i>AllC</i> 類	9.4%	7.8%
	<i>AllD</i> 類	76.6%	61.7%
	<i>TFT</i> 類	10.9%	17.8%
	その他の個体	3.1%	12.9%
4	<i>AllC</i> 類	2.4%	5.1%
	<i>AllD</i> 類	65.4%	27.4%
	<i>TFT</i> 類	5.1%	12.1%
	その他の個体	27.1%	55.4%

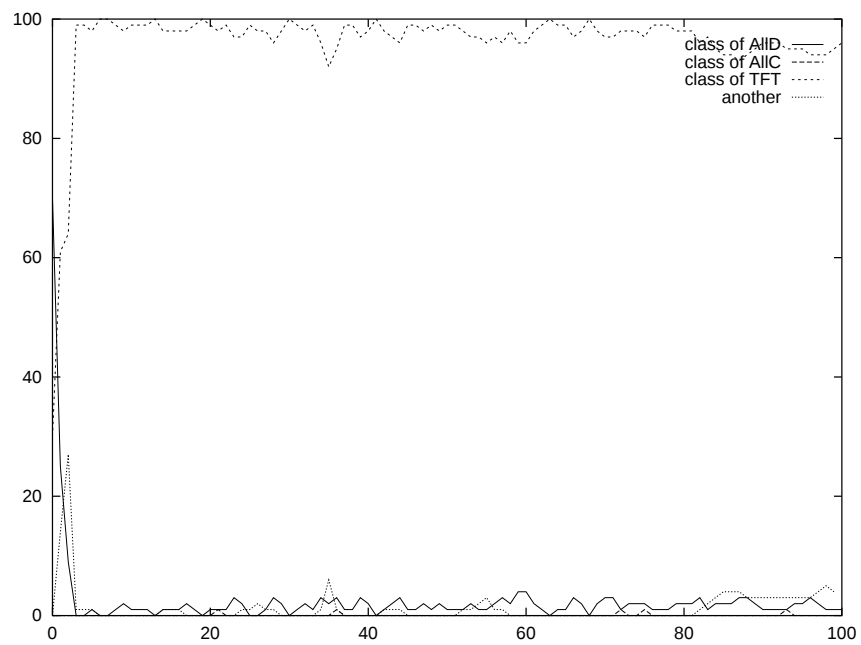
6.4 集団と個体の推移

図 6.9 と図 6.10 は GA のパターン” 1 ”におけるあるデータ結果である。図 6.9 は集団の特性図である。これは、集団の特性が世代毎にどのように変化しているのかを調べたもので、縦軸が集団内の全個体の平均適応度を表しており、横軸は世代数を表している。図 6.10 は個体の特性図である。これは、分類したそれぞれの個体が世代毎にどのような割合で存在しているのかを調べたもので、縦軸は個体数、横軸は世代数を表している。図 6.9 を見るとわかるように、集団内の平均利得が” 3 ”に収束しているため、この場合” 協調集団 ”に収束したということになる。この 2 つのグラフを見比べると、平均適応度が 3 世代目くらいでほぼ収束しているのに対し、それと同時期に集団内のほとんどは” TFT 類 ”によって占められ、その状態が終了まで続いている。図 6.11 と図 6.12 は同じく GA のパターン” 3 ”におけるあるデータ結果である。集団内の平均利得が” 2 ”に収束しているため、この場合” 非協調集団 ”に収束したということになる。平均適応度が早い世代でほぼ収束しているのに対し、集団内のほとんどは” AllD 類 ”によって終始占められている。このように、GA の場合” 協調集団 ”に収束する場合も” 非協調集団 ”に収束する場合も同じように早い世代で収束し、” 協調集団 ”に収束する場合は、集団内の個体のほとんどは” TFT 類 ”によって占められ、” 非協調集団 ”に収束する場合は、集団内の個体のほとんどは” AllD 類 ”によって占められるといった単純な構成になっているケースが多い。以上のデータ結果に対して、次は CCE の結果を比較する。図 6.13 は CCE のパターン” 1 ”におけるあるデータ結果である。CCE の場合、このように集団 A と集団 B のそれぞれの集団における平均適応度が表示されることになる。この場合も図 6.9 の GA の場合と同じように” 協調集団 ”に収束しているケースである。しかしながら、GA の場合に比べ、非常に遅い世代で収束しているのがわかる。図 6.14 と図 6.15 は集団 A と集団 B のそれぞれの個体の特性図である。集団 A、B とともに比較的” TFT 類 ”の割合が多い状態で推移しているが、GA の場合に比べ圧倒的に” その他 ”の個体の割合が高くなっているのがわかる。図 6.16 は CCE のパターン” 3 ”におけるあるデータ結果である。集団内の平均利得が” 2 ”に収束しているため、この場合” 非協調集団 ”に収束したということになる。集団 A、B とともに、最終的には集団内をほぼ” AllD 類 ”が占めてしまっているが、その状態に至るまでは、” その他 ”の個体の割合が高くなっている。このように、CCE の場合、GA に

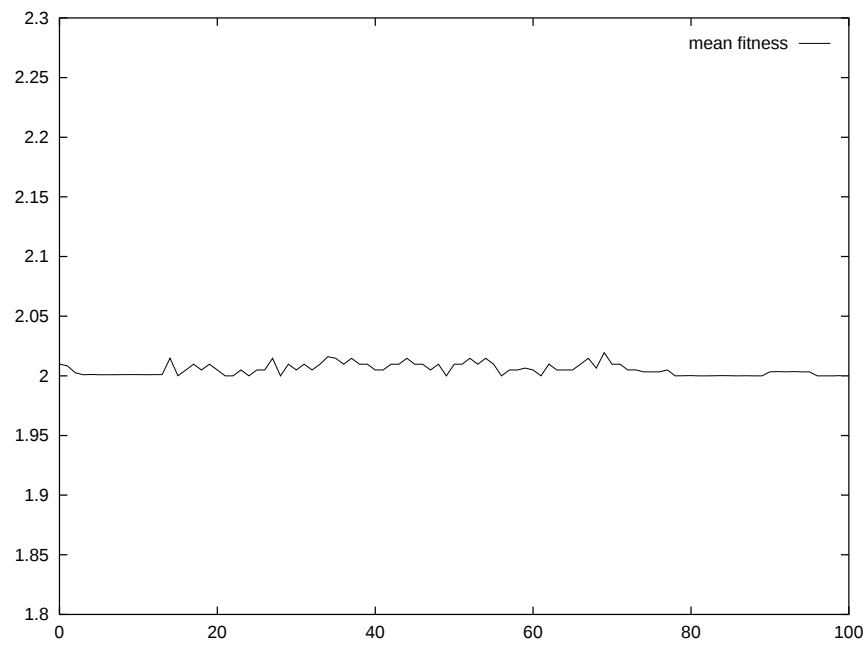
比べ収束する非常に速度が遅く、その推移も複雑な形になっているものが多い。それに、GA の場合は集団内の各種の個体の割合が分かり易く単純な構成になっているケースが多いのに対し、CCE の場合は全体的に” その他” の個体の割合が高く、複雑で解析の難しい構成になっているケースが多く見られる。



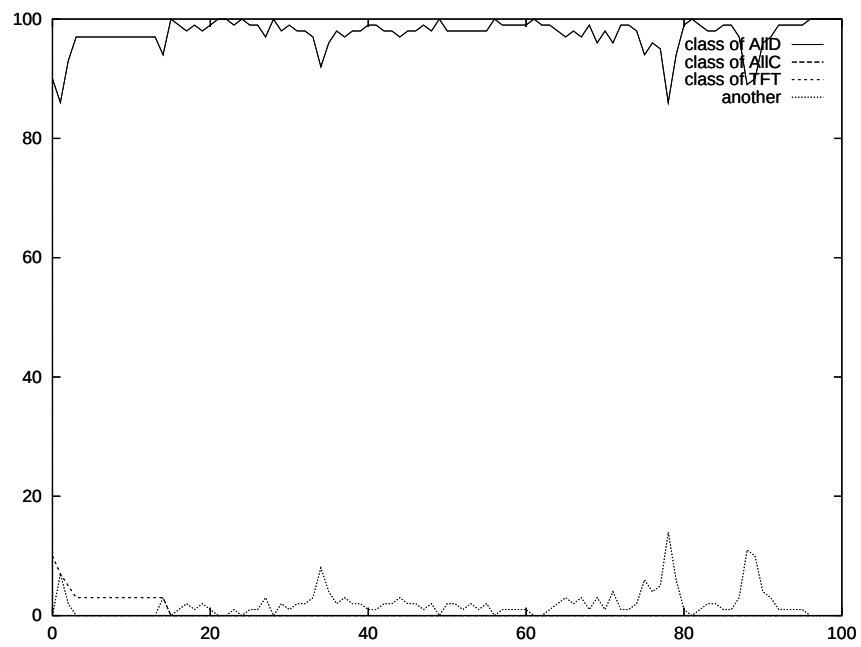
☒ 6.9: characteristic of population - pattern 1 (GA)



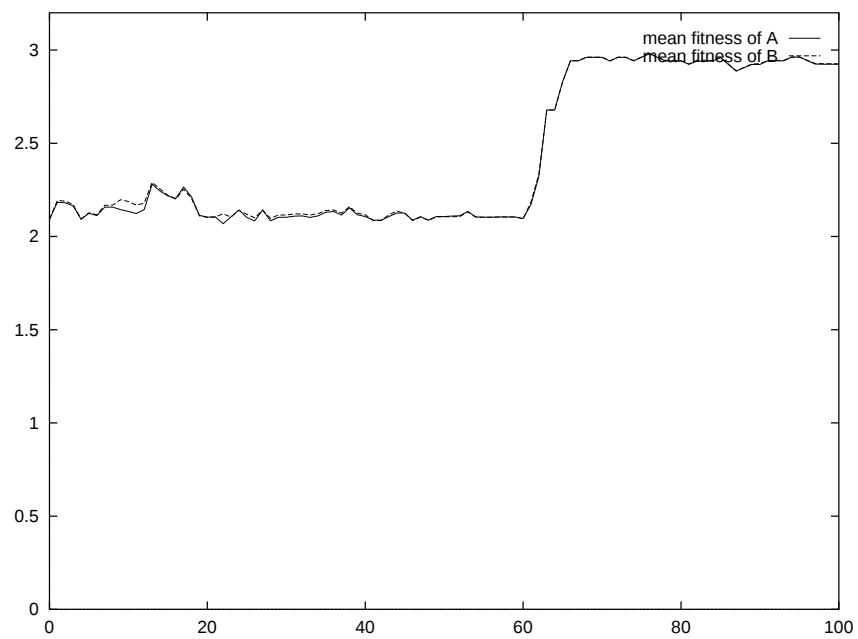
☒ 6.10: characteristic of individual - pattern 1 (GA)



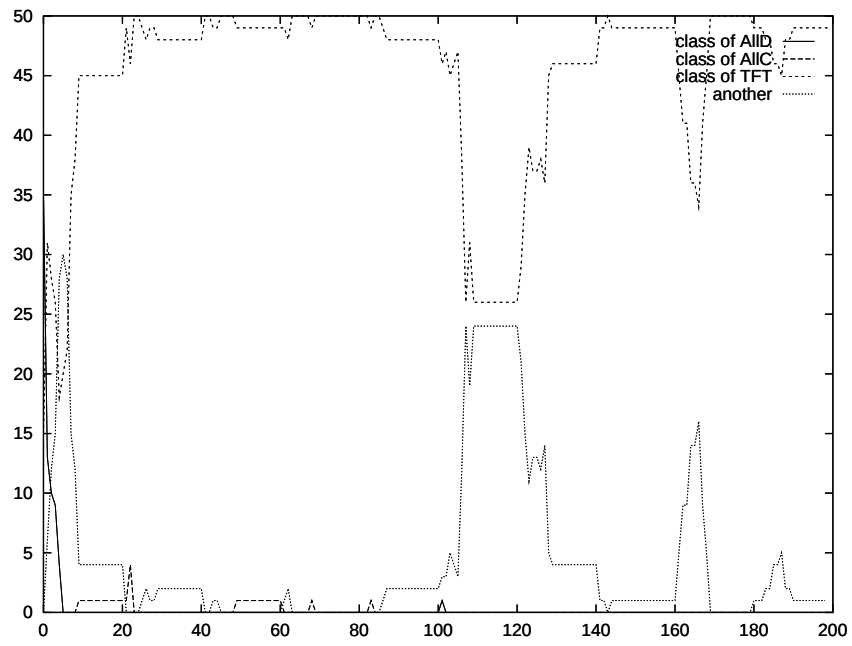
6.11: characteristic of population - pattern 3 (GA)



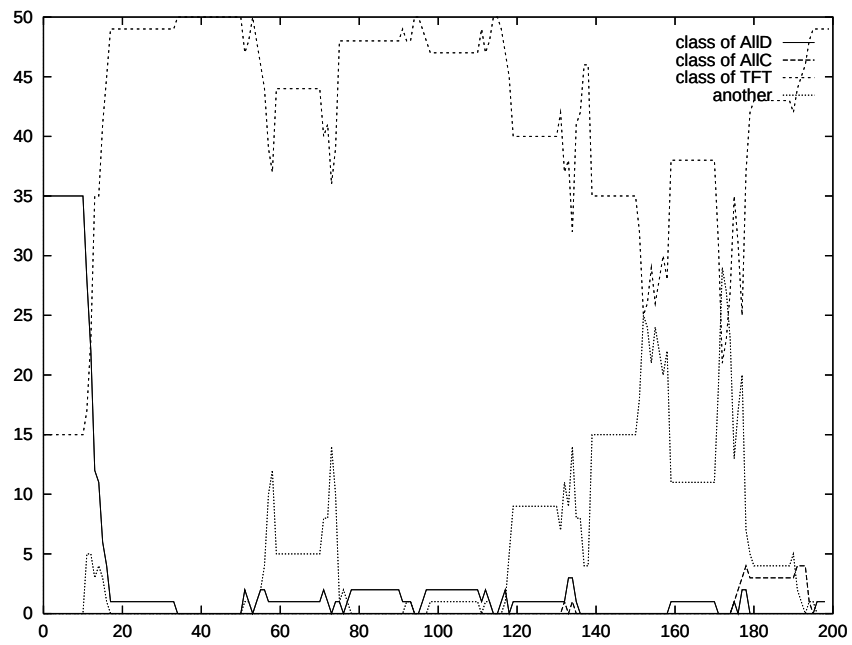
6.12: characteristic of individual - pattern 3 (GA)



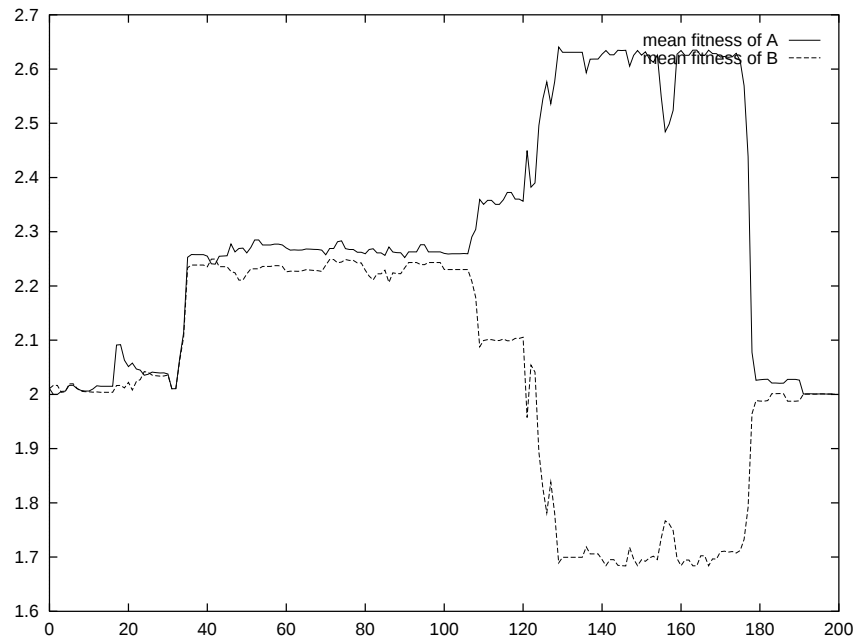
✕ 6.13: characteristic of population - pattern 1 (CCE)



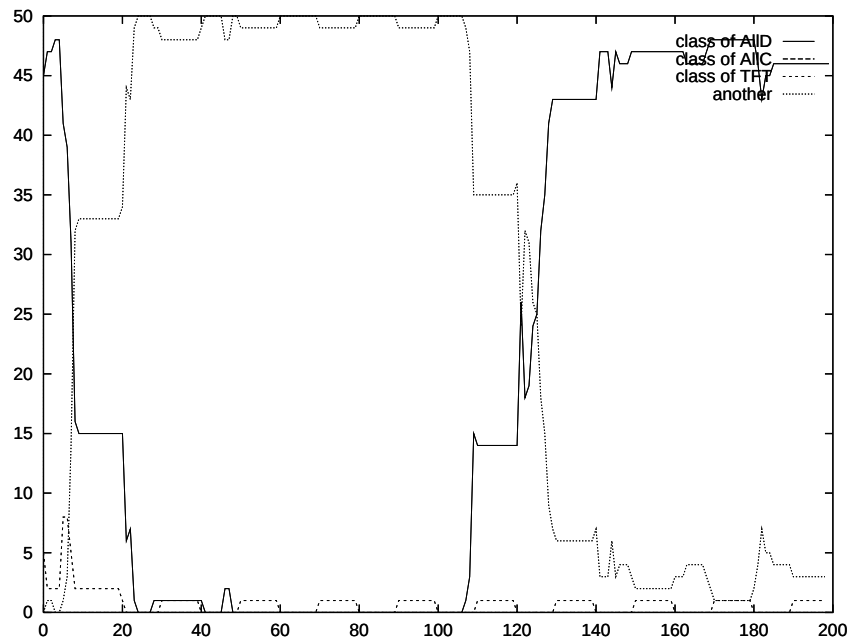
6.14: characteristic of individual - pattern 1 (CCE) -A-



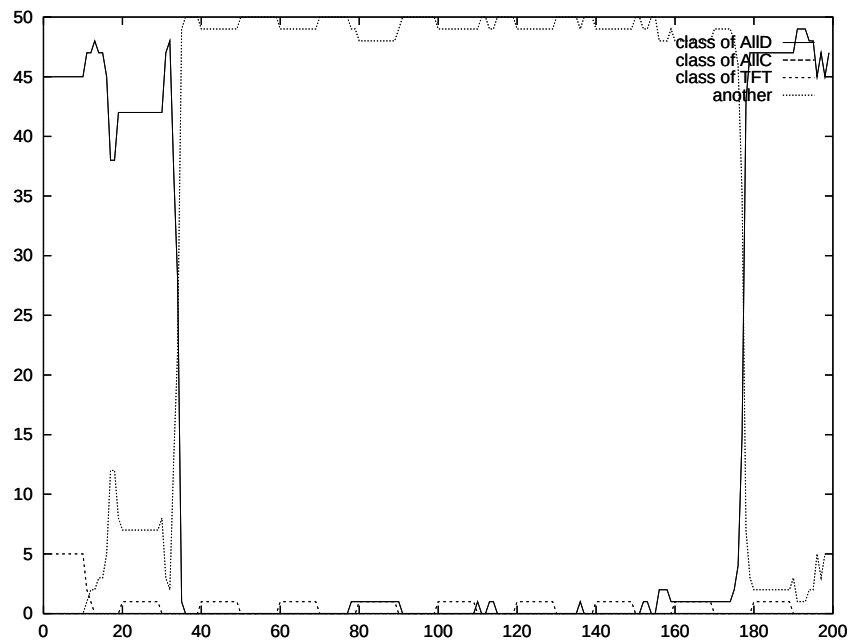
6.15: characteristic of individual - pattern 1 (CCE) -B-



⊠ 6.16: characteristic of population - pattern 3 (CCE)



⊠ 6.17: characteristic of individual - pattern 3 (CCE) -A-



⊠ 6.18: characteristic of individual - pattern 3 (CCE) -B-

6.5 “その他”の集団の特性

次に、CCE の場合に比較的多く見られた”その他”の集団がどのような特性を持っているのかを調べる。”その他”の集団に分類された集団群を見てみると、その 64.4% が集団内の個体の平均適応度が”2.5”に収束しているものであった。図 (6.19) に平均適応度が”2.5”に収束した場合の集団特性の推移グラフを示す。図 (6.20) と図 (6.21) はそれぞれの集団内のそれぞれの分類個体がどのような割合で推移しているのかを表したグラフである。収束時のそれぞれの集団内の個体の戦略コードは以下のようになっている。”協調”が”C”で、”裏切り”が”D”を表している。

集団 A の個体 a D D C D D C D C C C C D D D D D C D D

個体 B の個体 b C C D D C C C C C C C D D D D D C C D

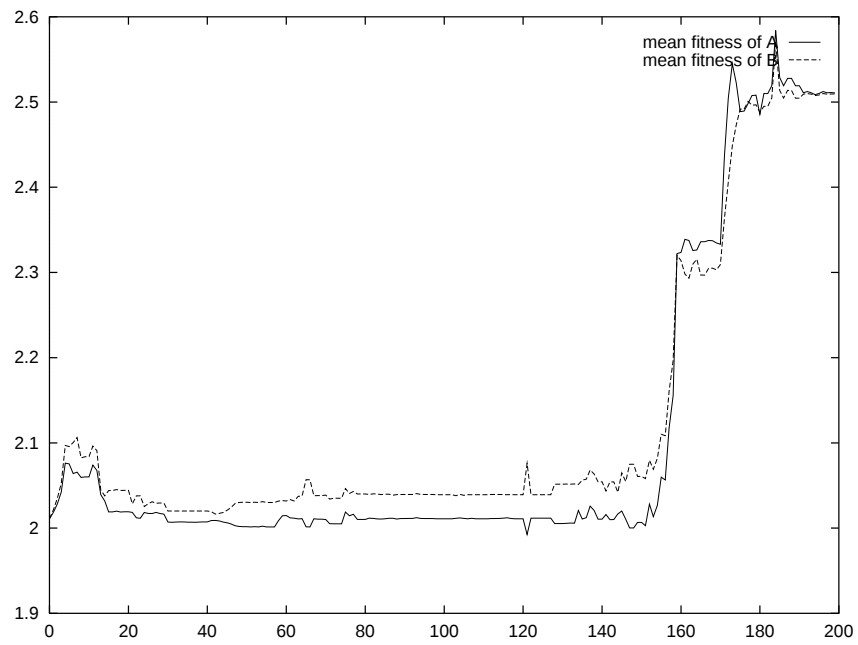
この 2 つの個体を IPD 対戦させると、以下のようになる。

a D C D C D C D C D C D C ...

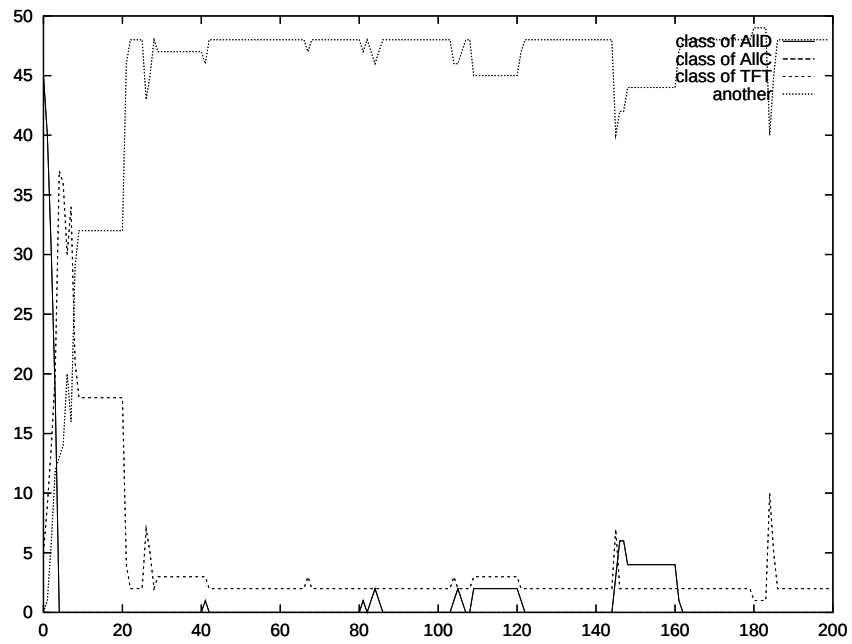
b C C C D C D C D C D C D ...

つまり、それぞれの個体は交互に”C”と”D”を出し合っていることになる。”C”と”D”を出し合えば、利得”1”と利得”4”を交互に得ることになり、平均すると”2.5”になるわけである。

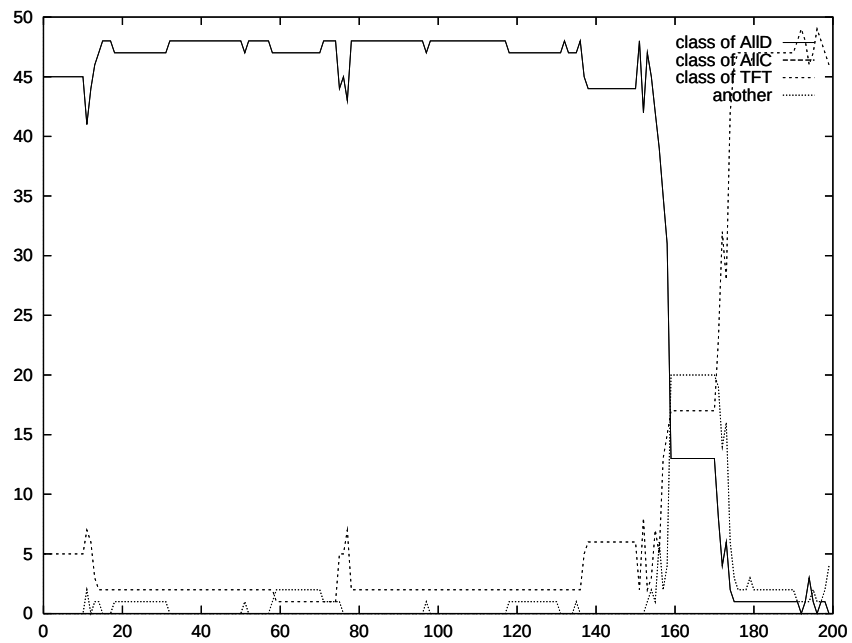
次に比較的多かった集団は、平均適応度が”2.3”に収束している場合である。図 (6.22)、図 (6.23)、図 (6.24) に平均適応度が”2.3”に収束した場合の集団特性、個体特性の推移グラフを示す。この場合、実際の IPD 対戦では、それぞれの個体が”D”を 2 回連続して出した後、1 回だけ”C”を出している。この動作を繰り返すことにより、それぞれの個体は、”2”、”1”、”4”の利得を順に繰り返し得ることになる。よって平均は”2.3”になるわけである。



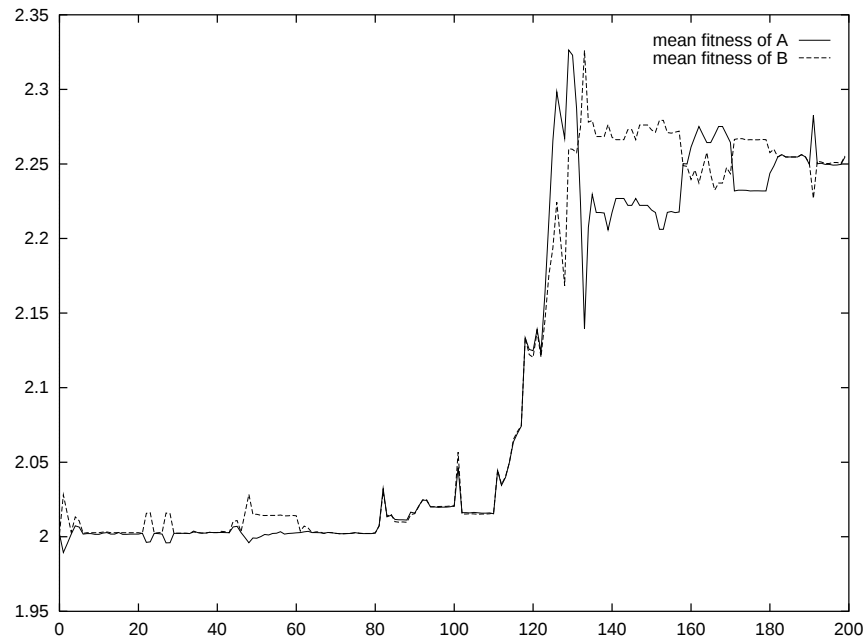
✕ 6.19: characteristic of population - pattern 3 (CCE)



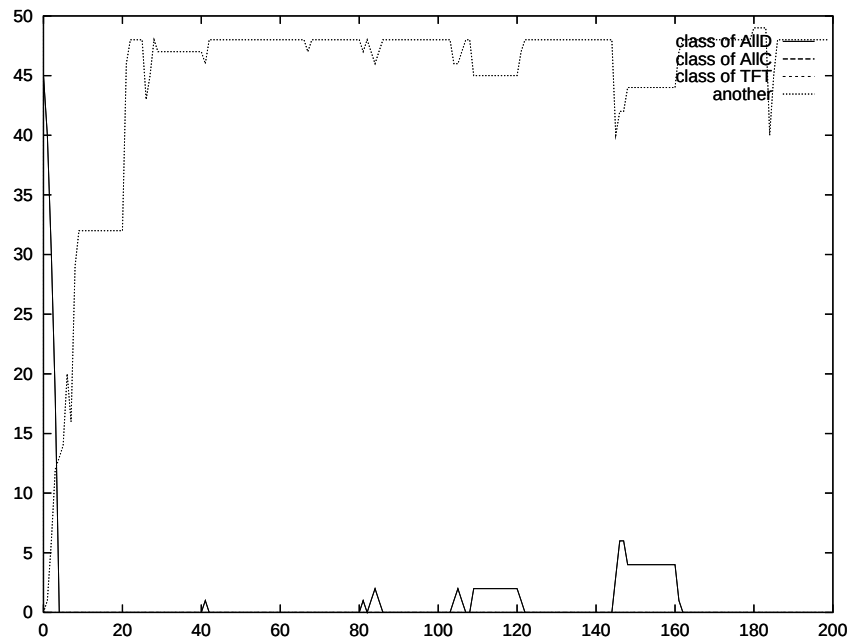
⊠ 6.20: characteristic of individual - pattern 3 (CCE) -A-



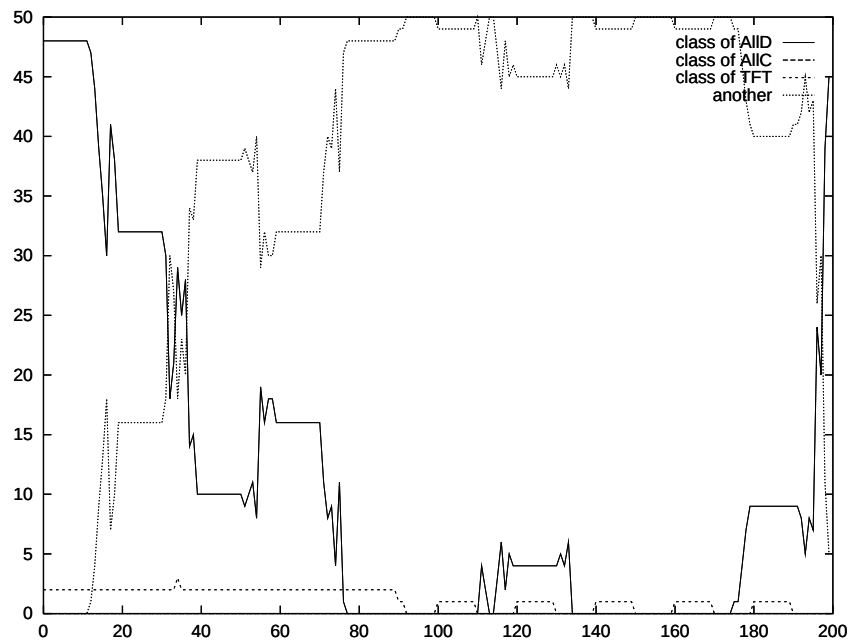
⊠ 6.21: characteristic of individual - pattern 3 (CCE) -B-



⊠ 6.22: characteristic of population - pattern 4 (CCE)



⊠ 6.23: characteristic of individual - pattern 4 (CCE) -A-



⊠ 6.24: characteristic of individual - pattern 4 (CCE) -B-

第 7 章

まとめ

7.1 考察

シミュレーションの結果より、IPD の戦略を CCE を用いて進化させた場合でも GA と同じように協調集団が得られる可能性があることがわかった。それと同時に、GA と CCE とでは収束速度、収束集団の特徴、集団内の個体の特徴などに大きな違いがあることも判明した。表 6.3 から分かるように、初期集団の *TFT* の割合が高い場合においては GA と CCE の分類したそれぞれの集団の比率の差はそれほど大きくはないが、*TFT* の割合が低い場合においては GA より CCE の方が協調集団になる割合が高くなっている。逆に CCE より GA の方が非協調集団になる割合が高くなっている。初期集団における *TFT* の比率が低いほど、GA に対して CCE の方が協調集団を生成し易い。

その理由を明らかにするために、集団内の個体がどのような特徴を持つ解なのかを調べる必要がある。表 6.5 より、GA に比べ CCE の方は全てのパターンにおいて“その他の個体”の割合が高く、特に初期集団における *TFT* の割合が低くなるほどその差は大きくなっている。つまり、CCE は GA より集団内の個体の多様性を保ちながら進化していく。これは、GA と CCE のメカニズムの違いが影響しているものと考えられる。評価規範となる相手個体も同じような方向性で常に変化していく GA に対して、CCE では一定期間、評価規範となる相手個体は全く変化しないからである。このことは表 6.1 のように、GA に比べ CCE の収束速度が極端に遅いことにも影響している。対戦相手が進化と同時に常に変化する GA では、集団内の *TFT* の割合が低くなるほど *TFT* 戦略が得る利得が低く

なり、適応度そのものの値が低くなる。それに対して CCE では評価規範となる環境集団は変化しないので TFT 戦略が得る利得は変化しない。よって、進化集団の TFT の割合が進化によって低くなったとしても、環境集団内の個体が全て $AllD$ でない限り、再び TFT が進化集団内に増えてくる可能性がある。よって、CCE は戦略の多様性を保つことで非協調集団へと進化する可能性を低くしている。GA では、進化によって集団内の TFT が減り $AllD$ が増えてくると、加速的に非協調集団へと進化してしまう。

これらのことを数学モデルを用いて以下のように証明する。

集団内の個体数を N とし初期集団における TFT の割合を n 、1 個体と 1 個体の対戦ゲーム数を G とするとし、以下の利得表より利得を決める。

表 7.1: 利得行列

X, Y	協調	裏切り
協調	R, R	S, T
裏切り	T, S	P, P

集団内の個体数を N とし初期集団における TFT の割合を n 、1 個体と 1 個体の対戦ゲーム数を G とすると、

第 1 世代における、それぞれの対戦による各個体の得点は以下ようになる。

TFT 1 個体と $AllD$ 1 個体の対戦

$$TFT \text{ の総得点} = S + P(G - 1)$$

$$AllD \text{ の総得点} = T + P(G - 1)$$

TFT 1 個体と TFT 1 個体の対戦

$$TFT \text{ の総得点} = GR$$

$AllD$ 1 個体と $AllD$ 1 個体の対戦

$$AllD \text{ の総得点} = GP$$

TFT 1 個体の 1 世代における総得点

$$GRnN + (S + PG - P)(1 - n)N = N\alpha \quad (7.1)$$

AllD 1 個体の 1 世代における総得点

$$(T - P)nN + NGP = N\beta \quad (7.2)$$

ここで、ある世代 x における集団内の *TFT* 個体数の割合を $W(x)$ とし、その次世代での *TFT* 個体が集団を占める割合の確率を $W(x + 1)$ とすると、
GA の場合、

世代 x における *TFT* 1 個体の総得点

$$GRNW(x) + (S + PG - P)(1 - W(x))N = N\alpha(x) \quad (7.3)$$

世代 x における *AllD* 1 個体の総得点

$$(T - P)NW(x) + NGP = N\beta(x) \quad (7.4)$$

世代 x における集団内の全 *TFT* の得点合計

$$N^2\alpha(x)W(x)$$

世代 x における集団内の全 *AllD* の得点合計

$$(1 - W(x))N^2\beta(x)$$

世代 x における集団内の全個体の得点合計

$$N^2\alpha(x)W(x) + (1 - W(x))N^2\beta(x)$$

となり、次世代での *TFT* 個体が集団を占める割合の確率 $W(x + 1)$ は、以下のようになる。

$$W(x + 1) = \frac{N^2\alpha(x)W(x)}{N^2\alpha(x)W(x) + (1 - W(x))N^2\beta(x)} \quad (7.5)$$

一方、CCE の場合は、対戦相手である評価規範が変化しないので、ある世代 x における集団内の TFT 個体数の割合は初期集団の TFT の割合 n と同じ値となることから、
世代 x における TFT 1 個体の総得点

$$GRnN + (S + PG - P)(1 - n)N = N\alpha \quad (7.6)$$

世代 x における $AllD$ 1 個体の総得点

$$(T - P)nN + NGP = N\beta \quad (7.7)$$

と、式 (7.1)、式 (7.2) と全く同じ定式になる。それより、
世代 x における集団内の全 TFT の得点合計

$$N^2\alpha W(x)$$

世代 x における集団内の全 $AllD$ の得点合計

$$(1 - W(x))N^2\beta$$

世代 x における集団内の全個体の得点合計

$$N^2\alpha W(x) + (1 - W(x))N^2\beta$$

となり、次世代での TFT 個体が集団を占める割合の確率 $W(x + 1)$ は、以下のようになる。

$$W(x + 1) = \frac{N^2\alpha W(x)}{N^2\alpha W(x) + (1 - W(x))N^2\beta} \quad (7.8)$$

式 (7.5) と式 (7.8) を整理すると、
GA の $W(x + 1)$ は、

$$W(x + 1) = \frac{1}{1 + (1 - W(x))\frac{\beta(x)}{\alpha(x)W(x)}} \quad (7.9)$$

CCE の $W(x+1)$ は、

$$W(x+1) = \frac{1}{1 + (1 - W(x)) \frac{\beta}{\alpha W(x)}} \quad (7.10)$$

ここで、式 (7.9) の $\frac{\beta(x)}{\alpha(x)}$ と式 (7.10) の $\frac{\beta}{\alpha}$ を比較する。式 (7.6) と式 (7.7) より、 $\frac{\beta}{\alpha}$ には $W(x)$ が含まれていないため、 $W(x)$ の値の変動によって $\frac{\beta}{\alpha}$ の値が変化することはない。一方、 $\frac{\beta(x)}{\alpha(x)}$ には、 $W(x)$ が含まれているため、 $W(x)$ の値の変動によって値が左右することになる。

$$\frac{\beta(x)}{\alpha(x)} = \frac{(T - P)W(x) + GP}{(GR - PG + P - S)W(x) + PG - S - P} \quad (7.11)$$

ここで、

$$\begin{aligned} T &> R > P > S \\ 2R &> T + S \end{aligned}$$

の条件より、 $W(x)$ の値が小さくなるほど $\frac{\beta(x)}{\alpha(x)}$ の値は大きくなる。よって、式 (7.9) の $W(x+1)$ は式 (7.10) の $W(x+1)$ に比べ、 $W(x)$ の値が変化すればするほどその値が大きく変化することになる。つまり GA では、 $W(x)$ の値が低ければ低いほど次世代での $W(x)$ の値 ($W(x+1)$) が低くなることになる。これらより、ある世代の TFT の割合が低ければ低いほど、GA は CCE に比べ次世代での TFT の割合が更に低くなる可能性が高いということが言える。

7.2 おわりに

考察より、IPD の戦略の進化に CCE を適用した場合の特徴としては、GA を適用した場合に比べ初期集団の TFT の割合が低ければ低いほど協調集団を生成する確率が高くなるが、収束するまでに多くの世代を必要とする、ということが言える。今後の課題として

は、コード化の方法や CCE の終了条件の見直し、“その他”の集団のより具体的な解析などが考えられる。

謝辞

本研究の遂行、本論文の構成において終始御指導頂いた平石邦彦助教授に深く感謝致します。

また、研究を進めるにあたって日頃から有益な御助言を頂いていた、当講座の宋少秋助手、高島康裕助手に深く感謝致します。

参考文献

- [1] Axelrod,R.,The evolution of Cooperation, Basic Books,1984
- [2] Christopher, D. Rosin and Richard, K. Belew. ,Methods for competitive co-evolution: Finding opponents worth beating. In Larry J. Eshelman, editor, Proceedings of the Sixth International Conference on Genetic Algorithms, pp.373-380, San Francisco, California, July 1995. Morgan Kaufmann Publishers, Inc. ISBN 1-55860-370-0
- [3] Hobbes,T.,1651.Leviathan.New York,Collier Books edition,1962
- [4] 鈴木 光男, 新ゲーム理論, 1994
- [5] Axelrod,R.and W.D.Hamilton,The evolution of Cooperation,Science,Vol211,pp.1390-1396,1981
- [6] J.Maynard Smith,Evolution and Theory of Games,Cambridge University Press,1982