

共進化ジレンマゲームにおける平等ルールのオープンエンドな進化 Open-Ended Evolution of Impartial Rule in Co-evolutionary Dilemma Game

佐藤 尚[†] 松川 将也[‡]
Takashi Sato Shoya Matsukawa

1. はじめに

我々の社会には法や制度、規範などの様々なルールがある。これらのルールは、必ずしも万人に益をもたらすものばかりではない。社会が複雑化・多様化したために様々な制約や依存関係、選択困難なジレンマ状況などがその中で生じることにより、万人に対して等しく益をもたらすような「平等ルール」の形成を困難にしていると考えられる。また、社会やルールと社会成員は相互に作用・規定し合う循環的相互依存関係を作り出す。そのため、ルールは動的に変化しうるものであると捉えられる。このような状況の中で、どのような性質を持つルールが平等ルールとなりうるのか、また、どのような性質を持つ個体たちがそのルールの下で恩恵を受けられるのだろうか。

このような社会における様々な現象や人間同士の利害関係について、ゲーム論的枠組み[2]を用いて解析を行った研究が数多く存在する。囚人のジレンマゲームを題材に、ゲームの利得表と各プレーヤの戦略の両方を共進化させるモデルが山内ら[3]によって提案され、その後、4種類のジレンマゲームも考慮できるように拡張したモデルが山口ら[4]によって提案された。これらの研究では、ゲームの利得表をモラルや価値観として捉える。そして、プレーヤそれぞれが独立に自分の従うべきモラルや価値観を表す利得表を持ち、共進化ジレンマゲームを通して、どのような性質を持つ利得表が多くのプレーヤに支持されるのかが示された。これらのモデルでは、各プレーヤが独立に支持する利得表の下で自身が利益を獲得できるか否かが評価される。すなわち、あるプレーヤが支持していない利得表の下で、そのプレーヤがどれほどの利益を得られるのかについては確認できないモデルとなっている。しかし、社会において、我々は数多くのルールに従っており、常に1つのルールにのみ行動が規定されるというわけではない。

そこで本研究では、山口らのモデル[4]を参考に、全プレーヤが複数のルールそれぞれの下でゲームをプレイし、全てのルールを評価する共進化ジレンマゲームを提案する。本論の目的は、我々の社会と同様にジレンマ状況がある中で、どのようなルールが平等ルールとなりうるか、またその平等ルールが生まれる場合にはどのような過程を経て生まれるのか、更にそのルールが生まれるためにはどのような条件が必要なのかを明らかにすることである。

2. 共進化ジレンマゲーム

2.1 ジレンマゲーム

ゲーム理論[2]における最も単純な枠組みは、2種類の行動を持つ2人のプレーヤによるゲーム(2人ゲーム)であ

[†] 沖縄工業高等専門学校メディア情報工学科,
Department of Media Information Engineering, Okinawa
National College of Technology

[‡] (株) 創和ビジネスマシナズ, Sowa Business Machines Co., Ltd.

る。この2人ゲームでは、各プレーヤがそれぞれ選択した行動に応じて、予め用意された利得表から利得が決定される。各プレーヤの行動を「協調」と「裏切り」の2種類とした場合の利得表を表1に示す。なお、本研究では、このゲームにおける利得表を「ルール」と見なす。

表1: 2人ゲームの利得表と4種類のジレンマゲーム

プレーヤ1	協調 (C)	裏切り (D)
プレーヤ2	R	T
協調 (C)	R	S
裏切り (D)	T	P

- ・行き詰まりゲーム : $T > P > R > S$
- ・鹿狩りゲーム : $R > T > P > S$
- ・チキンゲーム : $T > R > S > P$
- ・囚人のジレンマゲーム : $T > R > P > S$

また、表1に示されるように、利得 T, P, R, S のそれぞれの大小関係を変えることによって、「行き詰まりゲーム」、「鹿狩りゲーム」、「チキンゲーム」、そして「囚人のジレンマゲーム」という4種類のジレンマゲームを考えることができる。

2.2 プレーヤの進化

プレーヤの進化モデルには Lindgren のモデル[1]を採用する。このモデルでは、自分と相手との過去の対戦履歴に応じてゲームにおける自分の行動を選択する。そのため、対戦履歴を記憶するためのメモリとそれに基づいて自分の行動を決めるための戦略を持つ。

対戦履歴 h_m は、 m ビットの2進数で表される。なお、本論では、0を裏切り、そして1を協調とする。

$$h_m = (a_{m-1}, \dots, a_1, a_0) \quad (2.1)$$

ここで a_0 は前回の相手の行動を表し、以下 a_1 は前回の自分の行動、 a_2 は前々回の相手の行動、 a_3 は前々回の自分の行動、というように表し、過去 $m/2$ 回分の対戦履歴を記憶する。

戦略 S も、対戦履歴 h_m と同様に2進数 (n ビット) で表される。

$$S = [A_0, A_1, \dots, A_k, \dots, A_{n-1}] \quad (2.2)$$

ここで A は次のゲームで選択される自分の行動を表す。対戦履歴 h_m を10進変換した値が A の添え字に対応する。すなわち、10進変換した h_m が k である場合、プレーヤは A_k の行動を選択する。なお、戦略 S は可変長であり、 n は2の m 乗となる。また、この戦略 S をプレーヤの遺伝子として捉え、遺伝的操作の対象とする。

プレーヤの遺伝子に対する遺伝的操作としては、以下に示される「一点突然変異」、「重複突然変異」、そして「遺伝子分割」の3つを用いる。

一点突然変異は、 n ビットの遺伝子からランダムに1つ

を選択し、ビットフリップ (0→1 or 1→0) する。

例) [01] → [00]

重複突然変異は、遺伝子をコピーして2倍の長さにする。

例) [01] → [0101]

遺伝子分割は、遺伝子を前半部分もしくは後半部分のどちらかを削除し、半分の長さにする。

例) [0110] → [01]

なお、本研究では、Lindgren のオリジナルモデル[1]とは異なり、どの戦略をどのくらい増減させるのかを戦略の適応度に応じて決定論的に決めるのではなく、ルーレット選択を用いて確率的に決定する。また、戦略から行動を決定する際、ある確率で戦略が指定する行動とは異なる行動を選択する、所謂「ノイズ」を採用しないものとする。

2.3 ルールの進化

本研究で提案するルールの進化モデルは、山口らのモデル[4]と異なり、各プレーヤの戦略を表す遺伝子の一部として表現するのではなく、ルール自体を表す遺伝子を別途用意し、独立に進化させる。

ルールを表す遺伝子は、「行き詰まりゲーム」、「鹿狩りゲーム」、「チキンゲーム」、そして「囚人のジレンマゲーム」の4種類のルールの型の中でどれを採用するのかを決めるための p ビットと、利得表の T, P, R, S のそれぞれの値を決めるための q ビットを合わせた計 r ビットで表現される。利得表の各値は、各ルールの型が規程する T, P, R, S の大小関係に従って、それらの中の最小値を基準値とし、遺伝子に従って加算させることで決定する。

p を2, q を12 とする計14ビットの場合のルールの遺伝子の例を図1に示す。まず、右から3ビットの2進数を10進変換した値をルールの基準値とし、次からの3ビットの値を先ほどの基準値に加算していくことで利得を決定する。

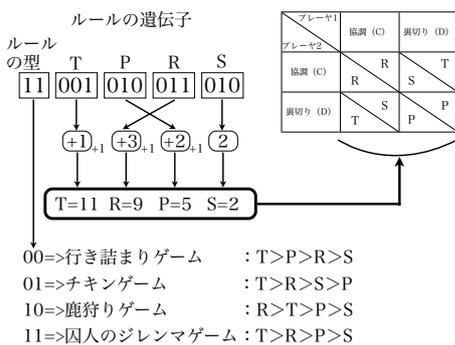


図1: ルールの遺伝子の例。ルールの型を決めるための p ビットと T, P, R, S のそれぞれの値を決めるための q ビットを合わせた計 r ビットの遺伝子で表される。各利得は次の手順で決定する。まず、 q ビットを4分割してそれぞれを10進変換する。次に、ルールの型が規定する各値の大小関係に従って並べ替えを行い、最小値を基準値とする。最後に遺伝子が表す値を基準値に加算していくことで決定する。

なお、基準値以外の3つの値に対して「1」を加算しているのは、遺伝子が「000」となった場合にも利得間の大小関係が成立するようにするためである。

ルールの遺伝子に対する遺伝的操作としては、以下に示される二点交叉と一点突然変異の2つを用いる。

二点交叉は、2つの親の遺伝子の一部を入れ替え、2つの子の遺伝子を作る。

例) 親1: [00000000000000] 子1: [00001111000000]
 ↓
 親2: [11111111111111] 子2: [11110000111111]

一点突然変異は、 r ビットの遺伝子からランダムに1つを選択し、ビットフリップ (0→1 or 1→0) する。

例) [11111111111111] → [11110111111111]

また、各ルールは式(2.3)~(2.5)を用いて評価する。

$$S_j = \sum_{i=0} x_{ij} \tag{2.3}$$

$$u_j = \sum_{i=0} \frac{E(x) - x_{ij}}{N - 1} \tag{2.4}$$

$$F_j = S_j - u_j \tag{2.5}$$

ここで、 x_{ij} はプレーヤ i がルール j の下で全プレーヤと総当たり戦を行って獲得した利得の合計、 S_j はルール j の下で総当たり戦を行った全プレーヤが獲得した利得合計の総和、 u_j はルール j における各プレーヤが獲得した利得の分散値、 $E(x)$ は総当たり戦を行った全プレーヤが獲得した利得合計の平均値、 N は全プレーヤ数、 F_j はルール j の適応度である。

分散値を用いることによって、それぞれのプレーヤが獲得した利得合計にばらつきが多いルールの適応度は低く評価される。これにより、どのルールが全プレーヤに対して最も偏りなく益(利得)をもたらすことができるかを確認することが出来る。

2.4 共進化ジレンマゲームの手順

複数のプレーヤ(の戦略)と複数のルールがそれぞれの進化モデルに従って共進化していく。全プレーヤは各ルールの下で総当たり戦を行う。この総当たり戦をまとめて1ステップとし、プレーヤは i ステップまで進化させる。一方、ルールは i ステップごとに進化させ、これを j 世代まで繰り返す。

3. シミュレーション実験

3.1 シミュレーション実験の設定

本研究で用いるシミュレーション実験の設定を以下に示す。世代数は10000, ステップ数は100とする。プレーヤの個体数は100とし、各プレーヤの遺伝子の初期状態は00(ALLD), 01(TFT), 10(ATFT), 11(ALLC)の4種類をそれ

ぞれ 25 個体 ($N / 4$) ずつとする。プレーヤの遺伝子を行う点突然変異率、重複突然変異率、遺伝子分割率はそれぞれ 0.00002, 0.00001, 0.00001 とする。また、プレーヤの遺伝子は最長 16 ビット、最短 2 ビットとする。ルール of 個体数は 20 とする。利得表の各得点 T, R, P, S を決めるための遺伝子長は 12 ビットとし、20 個のルールすべての初期値を 010010001000 とする。ルールの型を決める 2 ビットはランダムに与える。ルールの交叉率は 0.6, 点突然変異率は 0.01 とする。

3.2 シミュレーション実験の結果

まず、全ステップ (世代数 $10000 \times$ ステップ数 $100 = 100$ 万ステップ) を通して最も多く出現したプレーヤの戦略の上位 10 個体の個体数変化を図 2 に示す。この結果は次の 3 つのフェーズに分けることができる。

<フェーズ 1: 0~200000 ステップ>

裏切り戦略 (青, 赤) のプレーヤが支配している様子が示されている。その中で、TFT 戦略 (黄緑) のプレーヤが生き残っていることがわかる。

<フェーズ 2: 200000~400000 ステップ>

上位 10 個体以外の戦略のプレーヤが繁栄していたことがわかる。その中で、裏切り戦略 (青, 赤) を持つプレーヤが現れては消えている。

<フェーズ 3: 400000~1000000 ステップ>

裏切り戦略 (青, 赤) プレーヤ, TFT 戦略 (黄緑) プレーヤ, そして、協力戦略 (ピンク) プレーヤのそれぞれが入れ替わりながら支配している様子が示されている。また、どのステップでも TFT 戦略 (黄緑) プレーヤは少なからず生き残っていることがわかる。

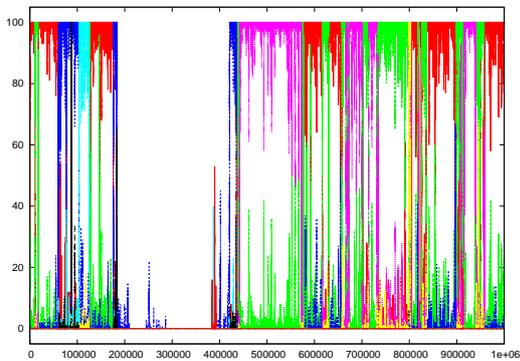


図 2: 全ステップを通して最も多く出現した上位 10 個体のプレーヤの個体数変化。縦軸がプレーヤの個体数、横軸はステップ数である。主に裏切り戦略 (赤, 青) と TFT 戦略 (黄緑), 協調戦略 (ピンク) が支配している。

次に、全ステップを通して最も多く出現したルール of 上位 10 個体の個体数変化 (図 3) およびルールの型別の個体数変化 (図 4) を示す。これらの結果は次の 2 つのフェーズに分けることができる。

<フェーズ 1: 0~400000 ステップ>

図 3 からは 1 つの戦略のルールが支配することはなく、行き詰まりゲーム型 (赤) のそれぞれの戦略がまばらに存在することがわかる。また、図 4 から行き詰まりゲーム型

(赤) のルールが支配していることが分かる。その中で、鹿狩りゲーム型 (青) とチキンゲーム型 (黄緑) のルールが生き残っていることがわかる。

<フェーズ 2: 400000~1000000 ステップ>

行き詰まりゲーム型 (赤) と鹿狩りゲーム型 (青) のルールが入れ替わりで支配している。行き詰まりゲーム型 (赤) のルールが支配している環境では、鹿狩りゲーム型 (青) とチキンゲーム型 (黄緑) のルールが生き残っている。鹿狩りゲーム型 (青) のルールが支配している環境では、行き詰まりゲーム型 (赤) と囚人のジレンマゲーム型 (ピンク) のルールが生き残っていることがわかる。また、鹿狩りゲーム型 (青) が支配している環境では、ほぼ一つのルールが支配していることがわかる (図 3)。

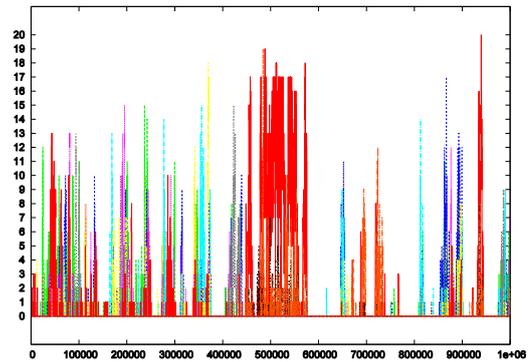


図 3: 全ステップを通して最も多く出現した上位 10 個体のルールの個体数変化。縦軸はルールの個体数、横軸はステップ数である。鹿狩りゲーム型 (赤) のルールが支配している状況が長く、他のルールはまばらに存在することが分かる。

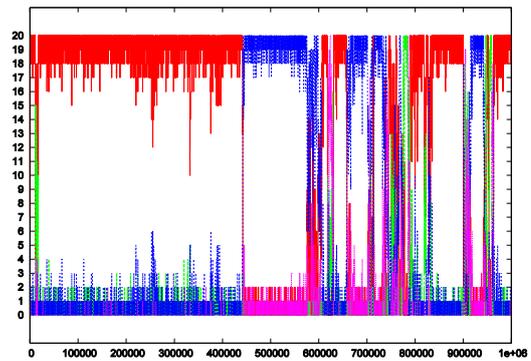


図 4: 4 種類のルール (チキン (黄緑), 行き詰まり (赤), 鹿狩り (青), 囚人のジレンマ (ピンク)) の型別の個体数変化。縦軸と横軸は図 3 と同じである。行き詰まりと鹿狩りがそれぞれ入れ替わるように支配している。

4. 議論

図 2 および 4 から分かるように、裏切り戦略 (青, 赤) のプレーヤが支配しているステップでは行き詰まりゲーム型のルールが支配しており、協力戦略 (ピンク) のプレーヤが支配しているステップでは鹿狩りゲーム型 (青) のルールが支配している。TFT 戦略 (黄緑) のプレーヤが支配しているステップでは、何らかのルールの型が長く支配す

ることがない。また、裏切り戦略(青, 赤)のプレーヤが支配している環境から協力戦略(ピンク)のプレーヤが支配している環境へ切り替わる間には, TFT 戦略(黄緑)のプレーヤが支配している環境が必ずあることが分かる。そこで, TFT 戦略(黄緑)のプレーヤが支配している環境において何が起きているのか明らかにしよう。

ここでは, 図4の430000ステップから450000ステップに注目して, 裏切り戦略(青, 赤)のプレーヤが支配する環境から TFT 戦略(黄緑)のプレーヤが支配する環境の切り替わる様子, そして, その後の協調戦略(ピンク)のプレーヤが支配する環境へと切り替わる様子を見てみよう。図4の430000ステップから450000ステップまでを拡大したものを図5に示す。

図2から分かるように, はじめに裏切り戦略(青, 赤)のプレーヤが支配する環境から TFT 戦略(黄緑)のプレーヤが支配する環境への切り替わりが起こる。協力と裏切りのどちらにも合わせることでできる TFT 戦略(黄緑)のプレーヤは, 行き詰まりゲーム型(赤)のルール他に鹿狩りゲーム型のルールでも得点を獲得できる。そのため, 行き詰まりゲーム型(赤)のルールが支配している中で鹿狩りゲーム型(青)のルールが個体数を増やすことができる。そのとき, 鹿狩りゲーム型(青)のルールにおいて多くの得点を獲得できる協力戦略(ピンク)のプレーヤが生まれ, その個体数を増やす。その後, 鹿狩りゲーム型ルール(青)と協力戦略(ピンク)プレーヤは個体数を増やし続け, 環境を支配した。

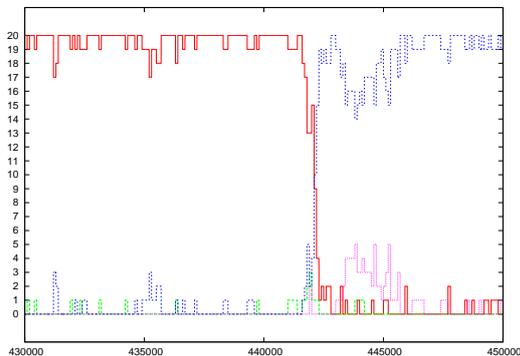


図5: 430000~450000ステップの上位10個体のルールの個体数変化。縦軸と横軸は図3と同様である。440000ステップあたりから行き詰まりゲーム型ルール(赤)が個体数減り始め, その代わりに鹿狩りゲーム型ルール(青)の個体数が増える。

次に, 550000~700000ステップ間を拡大したものを図6に示す。この区間では, 図2に示されるように, 協調戦略(ピンク)のプレーヤが支配する環境から TFT 戦略(黄緑)のプレーヤが支配する環境へと切り替わり, さらに裏切り戦略(青, 赤)のプレーヤの支配する環境へと切り替わる。

はじめに, 協調戦略(ピンク)のプレーヤの支配環境から TFT 戦略(黄緑)のプレーヤの支配環境へ切り替わりが起こる。先ほど説明したように TFT 戦略(黄緑)は行き詰まりゲーム型(赤)と鹿狩りゲーム型(青)のどちらでも得点を獲得することができるため, 鹿狩りゲーム型(青)のルールが支配している中で行き詰まりゲーム型(赤)のルールが個体数を増やし始める。そのとき, 行き詰まりゲ

ーム型(赤)で多くの点数を獲得できる裏切り戦略(青, 赤)のプレーヤが生まれ, 個体数を増やしている。その後, 裏切り戦略(青, 赤)のプレーヤが環境を支配し, 行き詰まりゲーム型(赤)のルールが環境を支配した。

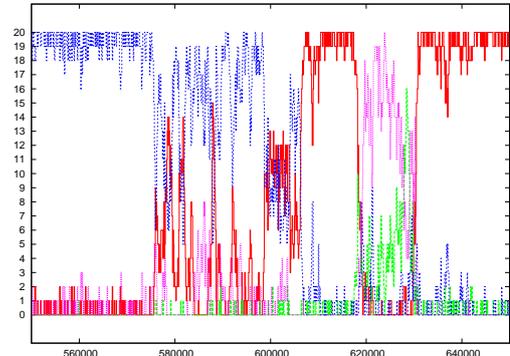


図6: 550000~650000ステップの上位10個体のルールの個体数変化。縦軸と横軸は図3と同様である。570000ステップあたりから鹿狩り型ルール(青)が個体数減り始め, その代わりに行き詰まりゲーム型ルール(赤)の個体数が増える。

5. おわりに

本研究では, どのような性質を持つルールが偏りなく益を与えられる平等ルールとなりうるのかを調べるために, ルールとプレーヤが相互作用しながら互いに進化していく共進化ジレンマゲームモデルを提案した。この共進化ジレンマゲームを用いたシミュレーション実験の結果, 全プレーヤが偏りなく点を得られる平等ルールとして, 裏切り戦略を助長する行き詰まり型ルールと, 裏切り戦略および協調戦略のどちらも助長する2つの均衡を持つ鹿狩り型ルールのそれぞれが1つの平等ルールに収束することなく, ダイナミックにルールが切り替わり続けるオープンエンドな進化ダイナミクスを確認した。また, TFT 戦略のプレーヤがそのような平等ルール及びプレーヤの戦略の性質を変えるために重要な役割を果たしていることを明らかにした。

参考文献

- [1] Lindgren, K., "Evolutionary phenomena in simple dynamics," *Artificial Life II*, pp.295-312, Addison-Wesley, (1991).
- [2] 生天目章, "ゲーム理論と進化ダイナミクス - 人間関係に潜む複雑系," 森北出版, (2004).
- [3] 山内敦, 星野力, "囚人のディレンマゲームにおける社会行動の進化," 情報処理学会研究報告「知能と複雑系」, Vol.98, No.24, pp.7-13, (1998).
- [4] 山口佳樹, 山内敦, 丸山勉, 星野力, "ジレンマゲームにおける価値観の発生とその解析," 計測自動制御学会論文集, Vol.38, No.3, pp.299-308, (2002).