繰り返しの囚人のジレンマゲームにおける断続平衡†

蜷 川 繁*・新 谷 憲*

Punctuated Equilibria in the Iterated Prisoner's Dilemma†

Shigeru NINAGAWA* and Akira SHINTANI*

In the evolution of strategies in the iterated Prisoner's Dilemma, the abrupt decrease of species and the following return to the stationary value have been observed. This phenomena called punctuated equilibrium always accompanies the collapse of the cooperative relationship.

Key Words: Prisoner's Dilemma, punctuated equilibrium, genetic algorithms

1. はじめに

囚人のジレンマゲームとは2人のプレイヤーによって行なわれるゲームであり、各プレイヤーはC(協調), D(裏切り)いずれかの手を同時に出し、Table 1に示す利得表にしたがって得点が与えられる。ただし、 $T > R > P > S$, $(S+T)/2 < R$ である。これを繰り返して行なうゲームを繰り返しの囚人のジレンマゲーム(Iterated Prisoner's Dilemma, IPD)と呼ぶ¹⁾。ここでは $R=3, S=0, T=5, P=1$ とする。

本研究では、多数のプレイヤー(戦略)がIPDを行なうシステムにおいて進化計算を行なった場合の、システムの大域的な振る舞いに注目する。これまでIPDでは、主に生態学的な観点から、各戦略が自分以外のすべての戦略と対戦する総当たり戦の形式がもっぱら用いられてきた²⁾。しかし、たとえばジレンマ状況にある1対のエージェントが協調して目的を達成するといった問題のように、総当たり戦よりも1対1対戦のほうが工学的にはより重要であると思われることから、本研究では1対1対戦の形式を採用した。

2. 実験方法

本研究では進化計算の手法として遺伝的アルゴリズムを用いる。一定数の戦略からなる戦略の集合から2つの戦略をランダムに選び、1対1対戦で試合を行なわせる。1試合は200回の対戦からなり、戦略間の適合度の差異を際立たせる

Table 1 Payoff table. x, y in (x, y) in the table mean the payoff to the row player and the column player respectively.

	C	D
C	(R,R)	(S,T)
D	(T,S)	(P,P)

ため、1試合の合計得点の2乗をその戦略の適合度とする。これをすべての戦略について行ない、適合度に比例した確率で戦略を選ぶルーレット選択方式で戦略を選択した後、1点交叉、突然変異を経て次世代の戦略を生成する。

戦略は記憶長3、すなわち、過去3手の対戦結果によってつぎの手を決定するものとする。そのための遺伝子表現としては文献³⁾と似た形式を用いた。すなわち、過去3手の対戦結果が取り得る場合の数は $4^3=64$ 通りあるので、それぞれの場合にとるべきC(=0)またはD(=1)を割り当てるための64ビットと、最初の3手を指定する3ビットを加えて染色体は67ビットとなる。

3. 実験結果

個体数を1000とし、染色体の初期集団として各ビットに独立に1/2の確率でCまたはDをランダムに割り当てる2項分布の場合と染色体におけるDの割合が0~1にわたり均等に分布している一様分布の2通りの初期集団の生成方法を用いた。ビットごとの突然変異確率を $30/(67 \times 1000) \approx 0.447 \times 10^{-3}$ とした。これは各世代の戦略の集団において平均30ビットに突然変異が発生することを意味する。2項分布および一様分布の場合についてそれぞれ60通りの初期集団を生成し、3000世代まで計算機によるシミュレーションを行なった。

Fig. 1は初期集団が一様分布の場合の世代ごとの種数、すなわち異なる戦略の数の変化を示している。ほぼすべての試行において、種数は過渡的な状態の後、安定値(約420)に落ち着き、それが長期にわたり続く。しかし、Fig. 1の試行では約2200世代目から種数が急激に減少し、2255世代で種数は最小値89となった後、徐々に種数は安定値に回復し約2450世代ではほぼ元の値に戻っている。本研究では短期間で種数が安定種数の約2/3以下になる現象を大量絶滅とよび、大量絶滅後に安定種数に回復する現象を放散とよぶことにする。初期集団が2項分布および一様分布の場合について、それぞれ60回中、ともに33回の試行において大量絶滅、およびそれに続く放散が観測された。いずれの場合も大量絶滅は5~100世代にわたって発生し、その後の放散は20~300世代にわたって続いた。ただし、大量絶滅が発生する世代は

† 平成12年度電気関係学会北陸支部連合大会で発表(2000・9)

* 金沢工業大学情報系 石川県石川郡野々市町扇が丘7-1

* Division of Information and Computer Science, Kanazawa Institute of Technology, Nonoichi, Ishikawa
(Received February 12, 2002)

(Revised May 1, 2002)

Fig. 1 The number of species at each generation. Top: 0 ~ 999 generations, Middle: 1000 ~ 1999 generations, Bottom: 2000 ~ 2999 generations.

各試行ごとに不規則だった。

つぎに、大量絶滅が発生した時の試合内容の変化を調べる。1試合は200手からなり、1世代で500試合が行なわれるので、各世代で100,000回中に発生する対戦結果におけるR,PおよびSTの度数分布を調べた。Fig. 2はFig. 1に示した試行における世代ごとの対戦結果の度数分布である。ほとんどの試行において、長期的にみた場合、対戦結果はRが優勢であった。しかし、Fig. 2の場合、大量絶滅が始まった約2200世代からRが急激に減少する一方でSTが急増し、2233世代で最大値47929となるがすぐに減少し、それに代わってPが増加する。2244世代では、Pが優勢となるが、その後再びRが急増する一方でPは減少し、約2260世代には再びRが優勢となる。ほかの試行においても、大量絶滅に際して常に対戦内容の急激な変化が観測された。それらのほとんどはこの例のようにRが一時的に減少し、それに代わりSTやPが

Fig. 2 Composition of competition at each generation. Top: 1 ~ 1000 generations, Middle: 1001 ~ 2000 generations, Bottom: 2001 ~ 3000 generations.

急増するが、再びRが優勢になるというものだった。一般にRが優勢に戻る際に要する世代は数十世代であり、放散に要する世代にくらべて短い。

つぎに大量絶滅前後で集団のGTYPEがどのように変化したかを見るために、ある世代における1000個体中のビットごとのCの度数を求め、絶滅前後でのそれらの差を求めた。Fig. 3の上図は種数が最小となった2255世代におけるCの数から大量絶滅が発生する直前の2150世代におけるCの数を引いたものであり、下図は放散が終了し種数が安定した2600世代におけるCの数から、2255世代におけるCの数を引いたものである。この結果を見ると、放散にくらべ大量絶滅に要する世代は少ないにもかかわらず、GTYPEの変動は大きいことがわかる。

初期集団が2項分布と一様分布の場合に実験結果に違いがみられなかった原因としては、一様分布の場合、常にCまた

Fig. 3 Difference of frequency of C's in all population in distinct generations. Top: the difference between $t = 2255$ and $t = 2150$, Bottom: the difference between $t = 2600$ and $t = 2255$.

はDを出すような極端な戦略が早期に淘汰されるため、2項分布の場合と同じような結果になったと思われる。

また、初期集団が2項分布および一様分布の場合について、それぞれ60回中、ともに27回の試行において大量絶滅は発生しなかった。しかし、それらの試行においても、種数については安定値を取る一方、対戦内容では長期的にはRが優勢となるが、前述の大量絶滅の定義には合致しないものの、突発的な種の減少が観測された。

本実験では1試合の合計得点の2乗を適合度としたが、大量絶滅時には戦略の間に大きな得点差が生まれるため、このようなスケールとは無関係に上記のような現象が発生すると思われる。

4. おわりに

1対1対戦形式で記憶長3のIPDの戦略の進化シミュレーションを行なった。その結果、種数はほぼ一定値をとり、対戦結果はRが大勢を占める協調社会が長期にわたって続いた後、大量絶滅とともにGTYPEは大きく変化し、協調社会が突発的に崩壊するが、その後、放散とともに協調社会は急速に復帰し、GTYPEおよび種数は緩やかに変化するという進化のパターンが多くみられた。従来より、生物界においては進化は少しずつ継続的に起こるとする漸進進化 (gradual evolution) が唱えられてきた。それに対して、長期にわたる進化の停滞期の間に突発的に発生する大量絶滅とその後の放

散による進化を断続平衡 (punctuated equilibrium)⁴⁾とよぶ。本研究の結果はIPDにおいて断続平衡が進化の要因のひとつであることを示唆している。

文献⁵⁾では遺伝的アルゴリズムにおける断続平衡の発生を数学的に定式化しており、一点交叉と突然変異からなる遺伝的アルゴリズムを用いた場合、断続平衡進化はある意味では当然の結果といえる。本研究においても、大量絶滅の定義を緩めたり、測定する世代数を長くすることによって、ほとんどすべての試行において大量絶滅が観測されると思われる。

IPD以外の人工的なシステムにおける断続平衡はネオダーウィンモデル (Neo-darwinian model)⁶⁾や計算機生態系 (computational ecosystem)⁷⁾等において報告されているほか、文献⁸⁾では、染色体の長さが変化する総当たり戦のIPDにおける断続平衡について簡単に触れられている。

大量絶滅の際には常に対戦結果の変化が観測されたが、逆に対戦結果の変化は、常に大量絶滅を伴うわけではない。このことからIPDにおいては断続平衡以外にも進化の要因があると思われる。また、生物界においては、大量絶滅の原因として環境の急激な変化等が挙げられているが、本研究において、外因がない場合でも絶滅が起こることが明らかとなった。今後は種々のパラメータを変化させた場合やノイズを導入した場合でのIPDの進化について研究していく予定である。

参考文献

- 1) R. Axelrod: The Evolution of Cooperation, Basic Books (1984)
- 2) J. Maynard-Smith: Evolution and the Theory of Games, Cambridge University Press (1982)
- 3) R. Axelrod: The evolution of strategies in the iterated prisoner's dilemma, Genetic Algorithms and Simulated Annealing, Pitman, 32/41 (1987)
- 4) N. Eldredge and S.J. Gould: Punctuated equilibria: an alternative to phyletic gradualism, Models in Paleobiology, Freeman, 82/115 (1972)
- 5) M.D. Vose and G.E. Liepins: Punctuated equilibria in genetic search, Complex Systems, 5, 31/44 (1991)
- 6) C.M. Newman, J.E. Cohen and C. Kipnis: Neo-darwinian Evolution Implies Punctuated Equilibria, Nature, 315-30, 400/401 (1985)
- 7) H.A. Ceccatto and B.A. Huberman: Persistence of nonoptimal strategies, Proc. Natl. Acad. Sci. USA, 86, 3443/3446 (1989)
- 8) K. Lindgren: Evolutionary phenomena in simple dynamics, Artificial Life II, Addison-Wesley, 295/312 (1991)