

無限繰り返し囚人のジレンマの戦略進化における 記憶長との関係の分析

○宇野巧馬 高橋真吾(早稲田大学)

Analysis of the relationship with memory length in the strategic evolution of the infinitely repeated prisoner's dilemma

*T.Uno and S.Takahashi(waseda university)

概要—無限繰り返し囚人のジレンマにおいて戦略を1記憶から4記憶まで進化させたとき, [1001 0*0* 0*0* 0001]戦略が長寿戦略であると分かっている. また, 記憶長の短い集団が記憶長の長い集団と対戦すると記憶長の短い集団の方が高い平均利得を獲得する. つまり, 無限繰り返し囚人のジレンマの戦略進化における適切な記憶長について分かっていない. 協調状態に進化するには記憶長は長い方がよいのか. また, 記憶長が戦略進化にどのような影響を与えるのか. 進化型 IPD モデルを採用して, 分析・考察していこうと考える.

キーワード:囚人のジレンマ 記憶長 戦略進化

1. 研究背景

ゲーム理論はこれまで多くの分野で研究が行われており, 現実の多くの問題状況に対して深い理解や分析の助けとなる知見を与えてきた. 本研究では無限繰り返し囚人のジレンマに深く関係する. ここでは, ゲーム理論と無限繰り返し囚人のジレンマについての概要を説明する.

1.1 ゲーム理論

ゲーム的状况を記述する基本的枠組みである. ゲーム的状况とは, 意思決定の相互作用が主体のとりうる戦略とその結果に対する利得の関係によって認識された状況. 通常のゲーム理論では, ゲーム的状况を記述する際に, 各プレイヤーが認識しているゲーム的状况

は全員に共通であるという, 共通知識を前提とし, また, 各プレイヤーが合理性をもって推論し, 自己の利得の最大化を目指して意思決定する, という仮定を置いている.

1.2 無限繰り返し囚人のジレンマ

囚人のジレンマとは, 表1のような利得構造においてパレート最適である(協力, 協力)ではなく, (裏切り, 裏切り)がナッシュ均衡になることである.

表1 利得表

自分, 相手	協力	裏切り
協力	R, R	S, T
裏切り	T, S	P, P

行列の要素はプレイヤー1とプレイヤー2の獲得する利得を表す. ただし,

$S < P < R < T$. $(S+T)/2 < R$ を満たす. 囚人のジレンマを無限回繰り返す無限繰り返しの囚人のジレンマでは, 戦略進化に伴い囚人のジレンマは解消されることが知られている.

2 研究目的

Lindgren は戦略を $\{0, 1\}$ 列からなる遺伝子として, 遺伝的アルゴリズムを用いて自動進化させることによって, 強い戦略を自動生成するモデル(進化型 IPD モデル)を導入した [1]. また, 糸井らは進化型 IPD モデルに改良を加えて長寿戦略が $[1001 \ 0*0* \ 0*0* \ 0001]$ であることを発見した [2]. しかし, このモデルは戦略長の増減が対応する戦略をそのまま 2 倍または半減するだけである. つまり, 記憶長の長い戦略が短い戦略を淘汰するのは当然である. 本研究では, 進化型 IPD モデルをベースとして戦略進化と記憶長の関係を分析する.

3 モデル

本研究では強い戦略が自動生成される進化型 IPD モデルをベースとする. 具体的なモデルについて構築が出来ていないため, 進化型 IPD モデルについて説明する.

3.1 戦略, 記憶の表現方法

戦略を 1(協調)または 0(裏切り)の 2 ビットで表現する. 記憶長が n の時, 記憶を $(a_{2n}, a_{2n-1}, \dots, a_1)$ で表現する. この場合 $a_{2n} : n$ 回前の自分の戦略, $a_{2n-1} : n$ 回前の相手の戦略, $a_1 : 1$ 回前の相手の戦略である.

3.2 戦略淘汰

戦略は人口という属性を持ち, 自信を含めたすべての戦略と総当たりで戦い, 獲得した利得の平均値と戦略全体での利得の平均値の差を評価値として人口を変動させる. t 世代目の戦略 i の獲得平均利得 S_i は, 戦略 i の人口比率 x_i (戦略 i の人口/全体人口), 戦略 i

が戦略 j より獲得した利得を $g_{i,j}$, $N(t)$ を t 世代目に存在する戦略の種類の数, 戦略 i と戦略 j が対戦した回数を $c_{i,j}$ と定義すると

$$S_i(t) = \sum_{j=1}^{N(t)} \frac{g_{i,j}(t)x_j(t)}{c_{i,j}(t)}$$

また, 全体での利得の平均 \bar{S} は以下で表す.

$$\bar{S}(t) = \sum_{i=1}^{N(t)} S_i(t) x_i(t)$$

戦略 i の人口動態は以下で表せる.

$$x_i(t+1) - x_i(t) = \alpha(S_i(t) - \bar{S}(t)) x_i(t)$$

\bar{S} 以上の利得を得ている戦略は増加し, \bar{S} 未満の戦略は人口を減らせる.

4 今後のアプローチ

研究目的で述べたように戦略の増やし方を考える必要がある. そこで, 戦略長をその時の記憶長で期待利得の高い戦略を確率的に付加して伸ばす, また期待利得の低い方のビットを減らす突然変異を導入したい.

今後のアプローチをいかに示す.

①記憶長 1 から 4 の集団それぞれの戦略の平均獲得利得の算出

②①で求めた利得をもとに確率的な突然変異を進化型 IPD モデルに組み込み実験を行う.

5 参考文献

- [1] Kristian Lindgren Evolutionary Phenomena in Simple Dynamics Artificial Life II pp.295-312, Addison-Wesley(1990)
- [2] 糸井良太, 田中美栄子 ノイズ付き進化型繰り返しの囚人のジレンマにおける長寿戦略の探求 情報処理学会論文誌 Vol.6 No.1 31-37(March. 2013)
- [3] 安田俊一 異集団による囚人のジレンマ 松山大学論集 第15巻 第6号 2002年