

# 囚人のジレンマゲームにおける戦略

非線型物理学教室 SP96142 渡辺基成

## 囚人のジレンマとは？

ある罪を犯した2人の囚人が、警察から司法取引を条件に互いの犯した罪を白状するように言われる（2人は互いに情報を交換できない。）ここで、2人の行動を裏切る（自白する）と、協力する（自白しない）に分けて、刑罰（得点も）との関係を表にする。

表1 囚人のジレンマの損得

| 自分—相手 | 協力する                     | 裏切る                      |
|-------|--------------------------|--------------------------|
| 協力する  | 自:2年, 相:2年<br>自:3点, 相:3点 | 自:5年, 相:1年<br>自:0点, 相:5点 |
| 裏切る   | 自:1年, 相:5年<br>自:5点, 相:0点 | 自:3年, 相:3年<br>自:1点, 相:1点 |

表より、相手が協力を選ぶ時、自分も協力を選んで2年の刑になるよりは、相手を裏切って1年の刑の方が良い。また、相手が裏切るのであれば自分だけが協力して5年の刑になるよりは、裏切って自分の刑を軽くした方が良い。

つまり、相手の手がどちらであろうと、裏切ることで、自らの利得をあげることができる。

しかし、相手も同じ事を考えているので、2人は裏切りあい、ともに3年の刑になる。これでは、2人とも協力した方が刑期が短くなってしまふ。つまり2人が理性的に行動すると逆に悪い結果をもたらす。ここにジレンマが生じるのです。

この囚人のジレンマを繰り返し行なった、繰り返し囚人のジレンマゲーム (IPD) について考えていきます。囚人のジレンマとの違いは、繰り返される段階において、どのような戦略が有効なのか、相手のとった前回以前の手が参考にできます。

## IPD の戦略の進化

(1) IPD の戦略を GA で進化させます。過去2回の結果をもとに自分の戦略を表現するため、ま

ず  $4 \times 4 = 16$  ビットの遺伝子を用意し、そのビットが1なら裏切り、0なら協力するとする。更に、最初の2手は過去の情報がなく決まらないのでその手も2ビットで表現し、計18ビットの遺伝子となる。

(2) ランダムな戦略をとる相手と何回か対戦を行ない、得た総得点を適合度とする。

(3) その適合度を元に、より良い遺伝子を選ぶ。

(4) 選ばれた遺伝子を交叉、突然変異させる。

(5) (2) から (4) を繰り返す。

## 結果

100個の遺伝子をランダムな相手と100回対戦させ、600世代繰り返し、世代ごとに適合度の最大値、平均値、最小値と最大時の裏切った数（1の数）のグラフを作成した。

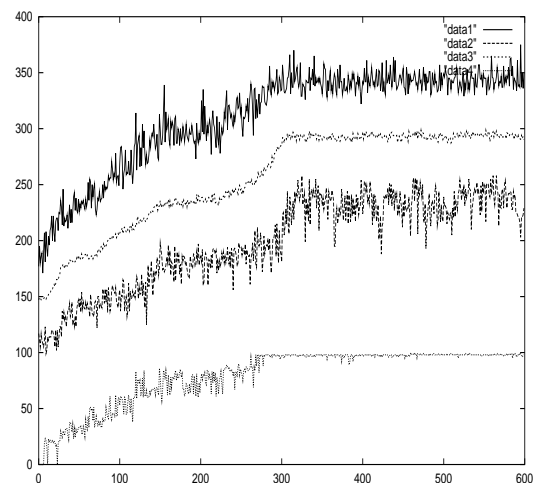


図1: IPD 実行結果 横軸: 世代数、縦軸: (上から) 最大、平均、最小の得点及び、最大時の裏切った数

最初のランダムな遺伝子作成において1と0の割合を1:1にすると、適合度の変動が見えにくいため、全て0にした。その結果、ランダムな相手に対しての対戦では、裏切り続けた方が得点が高いことが分かった。さらに、戦略を持った相手との対戦についても調べた。