

## 評判を考慮した繰り返し囚人のジレンマにおける戦略の進化

鳥取大学大学院工学研究科情報エレクトロニクス専攻

村上拓, 田中美栄子

### 1. はじめに

囚人のジレンマは相互協力が全体の利益であると解っていてもなかなか協力できない利己的な個人からなる社会のモデルである。本稿では評判を気にすることによって相互協力が自発的に創発するモデルを考えた。

### 2. 評判付き囚人のジレンマ・モデル

#### 2-1. 評判の定義

協力的な行動には正の、裏切りには負の評判がつくと考える。そこで、 $i$  と  $j$  が対戦した場合に、 $j$  が  $i$  につける評判を、 $i$  の  $j$  に対する協力の度合いとして以下のように定義する。

$$R_{i,j} = \frac{C_{i,j} - D_{i,j}}{C_{i,j} + D_{i,j}} \quad (1)$$

ここで  $C_{i,j}$  ( $D_{i,j}$ ) は  $i$  が  $j$  に対して協力 (裏切り) の手を出した回数とする。

#### 2-2. 一般的な評判 (立札モデル)

各プレイヤーはそれぞれ、自分の評判を表示するための立札を持ち、相手と対戦すると、式(1)で計算した評判を相手の立札に貼り付ける。立札には、今まで対戦した相手から受けた評判が貼り付けてある。プレイヤーは相手と対戦する際、相手の立札に貼り付けてある過去の評判の平均値

$$R_i = \frac{1}{n} \sum_{j=1}^n R_{i,j} \quad (2)$$

を相手の評判として対戦するかどうか判断する。

立札は常に全員に公開されているわけではなく、対戦しようと話しかけてきた相手にのみ公開する。つまり、対戦候補者を乱数で選んだ後、評判を見ることができる。これは、全員が全員の評判を知っているという状況を防ぐためである。

### 2-3. 評判による対戦相手の選択

各プレイヤーは、この程度の評判の相手なら対戦してもいい、という基準値を持つ。この基準値は初期設定時に乱数で与えられ (0.1 刻みで -1 から 1 までの 21 種類を一樣乱数で与える)、世代交代時に他の誰かのものを模倣する以外で変化することはない。基準値は相手と対戦するかどうかの判断のみに使われ、対戦時の行動には影響しない。もし、相手の評判が自分の基準値を下回っていれば、対戦を拒否することができる。同様に、相手も基準値を持ち、自分の評判が相手の基準値を下回っていれば対戦を拒否される。

基準値によって、必ず協力する相手でなければ対戦しないプレイヤーや、評判を気にせず誰とでも対戦するプレイヤーなどが生まれる。

### 3. 戦略

#### 3-1. 行動パターンについて

人の行動は、その人特有の「行動パターン」があると考え、裏切りと協力をそれぞれ 0 と 1 としたビット列で表現する。このパターンは、過去の相手や自分の行動に影響を受けると考え、以前の自分と相手の行動によって、複数の行動パターンを使い分けるものとする。例えば、前回の相手の行動を考慮する場合、プレイヤーは相手が裏切った時に使用する行動パターン  $a_1$  と相手が協力した時に使用する行動パターン  $a_2$  の 2 種類を持つこととなる。この 2 成分からなる戦略  $\langle a_1, a_2 \rangle$  を初期設定時に乱数で各プレイヤーに与え、世代交代時に変更する場合を除いては、変更しないものとする。

#### 3-2. 戦略による行動の決定

パターンは最初から順に使用し、最後まで使い

終わったら最初に戻って繰り返し使用する. 例えば<0010>の場合, このパターンを 3 回目や7回目に使用する時は, ビット列の 3 番目に当たる 1, つまり協力が次の行動となる. 初回の行動決定には戦略を使えないので, 乱数で決める.

### 3-3. 世代交代

全てのペアが(対戦拒否も含めて)対戦し終わると世代交代を行う. プレイヤーを得点の降順にソートして平均値より低いプレイヤーは元の戦略を棄て, 上位の戦略を順に模倣する(ノイズは入る).

## 4. 結果

全員同士が対戦するとき協力度体制が生成されないが, 対戦可能な相手を2割以下とすると多くの初期状態に対し協力度体制が実現する. Fig.1 に示す例では 30 世代位でpayoff(右軸)が3に飽和した. 同図に戦略<math>a\_1, a\_2</math>の強度変化(左軸)を各成分中の 1 の数(<math>\langle 0100, 1110 \rangle</math>は(1,3)と表示した. 戦略群(<math>\*, 3</math>)等が増加すると平均利得が急上昇し, 代わって(<math>\*, 4</math>)という戦略群が成長して安定する.

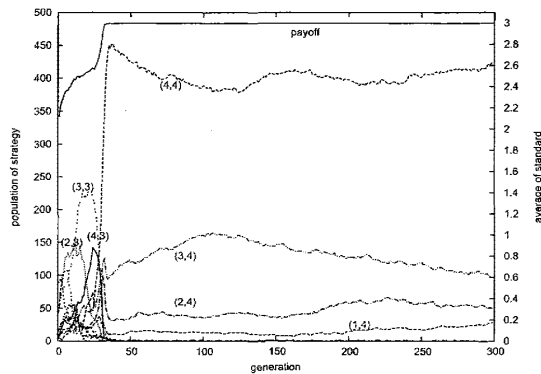


Fig.1 戦略の生成消滅と平均利得の変化

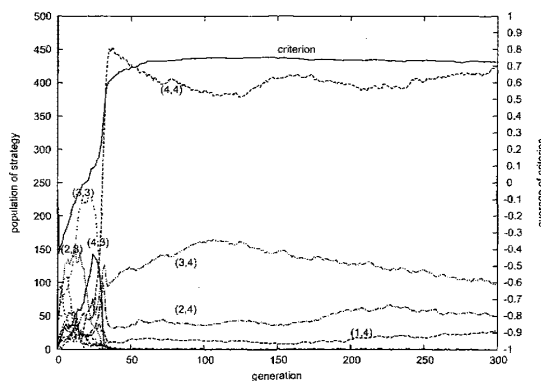


Fig.2 戦略の生成消滅と平均基準値の変化

協力的な戦略の成長に伴い, プレイヤーの判断基準(criterion)の平均値も上昇するが, Fig.2 に示すように最大値1には達せず0.7程度で安定する.

## 5. 結果の考察と今後の課題

進化アルゴリズムにより, 各プレイヤーが決めた基準以上の評判を有する相手とだけ対戦することで, 協力度体制を構築できることを確認した. 世代交代により, 最終的には全てのプレイヤーが協力の手を出すようになるが, 必ずしも全面協力(4,4)の手ばかりではなく, しっぺ返し(0,4)(Tit-For-Tat: TFT)に近い戦略群, 例えば(1,4)や(2,4)等が支配的となっていることが観察される.

本稿では, 乱数で選択されて来た対戦相手の評判が自分の要求する基準値を下回る場合には対戦を拒否できる, というモデルを考えた. この背景には, 現実社会においても, 相手を選ぶという行為は日常的に行われており, 裏切りやすい相手と付き合い続けることは, 資源の無駄になるため, 裏切りやすい相手とは付き合いを拒否し, 協力的な相手を求めるのは自然だと考えた事による. しかし実際には, 損な相手との関係を拒否できるとは限らない. 地理的な問題, 社会的地位や立場, 世間体などにより, 自分の損失になる場合でも付き合いざるを得ない状況は存在する. 卑近な例では, 北朝鮮の問題がこれに該当する. 北朝鮮の拉致問題や核問題は, 日本を始め各国の利益になる可能性は低い, 無視することはできない. 今回は問題が複雑化しすぎることを避けるため, 対戦を拒否されれば, 必ずそれに従うものとしているが, その拡張については今後の課題としたい.

## 6. 参考文献

[1] Martin A. Nowak and Karl Sigmund: Evolution of indirect reciprocity by image scoring, Nature 393, 11 June 1998.