

囚人のジレンマにおける GA の利用

T01K041F 清水 崇行

1. 囚人のジレンマ (Prisoner's Dilemma Game :PD)

囚人のジレンマは2人非ゼロ和ゲームである。非ゼロ和ゲームの場合はそれぞれの利得の大きさが評価基準になる。PD では2人のプレイヤーが協調、裏切りの手を選択し、利得行列と呼ばれる表に基づいて利得の計算を行う。

<利得表>

P1 \ P2	協力	裏切り
協力	R,R	S,T
裏切り	T,S	P,P

($T > R > P > S$ $2R > T + S$)

この条件は共に協調の手をとるほうがと共に裏切りの手を取るより利得が高いが、自分が協調して相手が裏切れば利得が低くなることを示している。またこのゲームでは1回の対戦なら共に裏切ったほうが合理的といわれている。

2. 戦略について

<代表的な戦略>

AllC(all cooperate)すべて協力

AllD(all defect)すべて裏切り

TFT(tit for tat)しっぺ返し

最初のゲームでは協力し、次のゲームでは前のゲームで相手が取った動作を行うというものである。

3. GA 適用の目的

戦略を個体群として交叉、突然変異を使うことによって戦略を進化させ、どの戦略がもっと利得を稼げるかを調べるために使用する。

利点・・・ゲームの利得により遺伝的操作を使うことによって戦略を変化させられる。

欠点・・・存在する戦略数が多くなるとすべての解を見るのは不可能

4. GA による戦略の進化

まず、進化させる戦略の個体の集団を初期集団の生成によって設定する。その生成された戦略個体を対戦させることによって、それぞれの個体は利得を得る。そして、その利得によってそれぞれの個体の適合度が決定する。それぞれの個体は、その適合度をもとに交叉や突然変異などの GA オペレータを行って新しい個体に生まれ変わる。このループを GA の終了条件に達するまで繰り返し、終了条件に達すれば終了となる。

6. 戦略のコード化

進化モデルを用いての PD の戦略を進化させて行くには、戦略をコード化する方法を決めなければならない。

過去2、3回目の双方の行動を参照する戦略でコード化を行う。

過去の双方の2手を参照する場合

DD DD(case 1) D(0) OR C(1)

DD DC(case 2)

DD CD(case 3)

⋮

CC CC(case 16)

AllC(0000000000000000)と表現できる。

戦略は $2^{16} = 64$ 通りある

過去の双方の3手を参照する場合

戦略は 2^{64} 通りある

7. 実験

<2回前を参照>

遺伝子表現：過去2回分の履歴を所有する16ビットの遺伝子+1回目、2回目の行動を決める2ビットの遺伝子

選択：ルーレット選択

交叉：1点交叉

突然変異率：0.01

個体数：20

世代数：50

適合度：対戦によって得られたスコア

利得行列：

P1 \ P2	協調	裏切り
協調	3,3	0,5
裏切り	5,0	1,1

(1)対戦相手：(AllC,AllD,TFT,TFTの逆)の

[結果]

戦略	獲得利得
000000101000100100	1391
000000101010101100	1391
000100111000110100	1391
000100111000111100	1391
000110101000101100	1391
001100101000001100	1391
001110111010100100	1391

1391の利得を稼ぐ戦略を7つ発見した。

1、2回目の行動は裏切りで共通したものだったが、それ以外には特に目立った共通点は見当たらなかった。

(2)対戦相手：1で得られた戦略

[結果]

戦略	獲得利得
000001000100000111	2824
000001000100001111	2824
000100000100001111	2824
000101000100000111	2824
000100001100000111	2824
001000000100001111	2824

2824の利得を稼ぐ戦略を6つ発見した。

この場合も同じく1、2回目の行動は協調で共通したものだった。

(3)利得行列の値を変えて同様の実験を行ってみる

P1 \ P2	協調	裏切り
協調	5,5	1,7
裏切り	7,1	2,2

[結果]合計利得の値に変化はできるが、得られる戦略には変化がなかった。

<3回前を参照>

遺伝子表現：過去3回分の履歴を所有する64ビットの遺伝子+最初の1回、2回、3回の行動を決める3ビットの遺伝子

(1)対戦相手：(AllC,AllD,TFT,TFTの逆)

[結果]

獲得利得1385の戦略を2つ見つけた。この戦略は29ビット目の値が違うだけで残りの値は全く同じものだった。

(2)対戦相手：(1)で得られた戦略

[結果]

獲得利得586を稼ぐ戦略を9個発見した

9. 考察

囚人のジレンマは対戦相手に依存するゲームで利得表の値を変化させても得られる戦略は変化しない。また参照する回数を増やしていくと同じ対戦相手でも得られる戦略は減少することがわかった。

獲得利得586を稼ぐ戦略を9個発見した

10. 参考文献

囚人のジレンマは対戦相手に依存するゲームで利得表の値を変化させても得られる戦略は変化しない。また参照する回数を増やしていくと同じ対戦相手でも得られる戦略は減少することがわかった。

獲得利得586を稼ぐ戦略を9個発見した

10. 参考文献

遺伝的アルゴリズムの方法
メラニー・ミッチェル著
遺伝的アルゴリズム①
北野宏明 編