

# More Effective Choice in the Prisoner's Dilemma

Robert Axelrod

Journal of Conflict Resolution, Vol.24 No.3 (1980) 379-403

## 1 The Tournament Approach

- 第1回に続き，囚人のジレンマゲームにおけるコンピュータトーナメントの第2回戦について報告する
  - － 囚人のジレンマが生じる状況で「良く振舞う方法」について更なる理解を得るため
- 囚人のジレンマゲームとは？
  - － 個人の合理性 (相手を出し抜きたい) とグループの合理性 (お互い協調したい) の間で葛藤が生じる
  - － 利得表は Table1
  - － 何度も相手と相互作用する時，裏切りを選択することは得策ではなく，効果的な方略が必要となる

TABLE 1  
Payoff Matrix for Each Move of the Prisoner's Dilemma

		Column Player	
		Cooperate	Defect
Row Player	Cooperate	3,3	0,5
	Defect	5,0	1,1

(The payoff to the row player is given first in each pair of numbers.)

- 繰り返し囚人のジレンマゲームにおいて，全ての対戦相手に対して優位な方略というのは存在し得ない。
  - － すなわち，結果は対戦相手の方略に依存してしまう。
    - \* 相手が協調を続けてくれるならば，こちらは裏切りを選んだ方が有利
    - \* 逆に，例えば相手が「こちらが1回でも裏切ったら，その後はずっと裏切りを選んでくる方略」の場合，ずっと協調すべきである。
- 対戦相手の方略が重要であることから...
  - － 対戦相手は出来るだけ良く振舞おう (高い利得を得よう) と試みる人である必要があり，
  - － それぞれのプレイヤーは相手が知的であると考え，高い利得を得ようと試みる人であると知っている必要がある
- 第1回目のトーナメントのレビュー
  - － エントリー：14 方略+RANDOM

- 優勝：TFT<sup>1</sup>
- 効果的な方略に必要とされる重要な概念は3つ
  - \* nice ルール
  - \* Forgiveness(寛容さ)
  - \* optimistic(楽観的)

## 2 The Second Round

- 結果として、第1回目よりも多くの効果的な方略における洞察を与えてくれた
  - その理由は参加者が第1回目の結果を取り入れ、数多くのルールを考えてくれたから
  - 加えて、参加者は相手が第1回目の結果を取り入れることを知っており、それがさらにレベルの高い戦いへと繋がったから
- 参加者
  - エントリー：62 方略
  - 参加者の専門分野：コンピュータサイエンス，経済学，心理学，数学，社会学，政治学，進化生物学
  - 国籍：アメリカ，カナダ，イギリス，ノルウェー，スウェーデン，オランダの6ヶ国

### 2.1 The Winner

- 優勝は1回目に続き、TFT だった
  - もちろん参加者は第1回目の優勝方略が TFT と知っており、それを打ち破ろうと方略を考えたはずである
  - 前回に続き、今回も提出された方略の中で一番シンプルな方略
  - 60 を超える方略が集まり、また参加者は前回なぜ TFT が優勝したのかについて知っていたにも関わらず

### 2.2 The Form Of The Tournament

- 対戦形式
  - 全方略総当り戦
  - 提出された62 方略に RANDOM を加えている
  - 自分自身の方略 (twin) との対戦も行う
  - 各方略同士の間は5 回行われる
- 利得表は Table1
- 試行数
  - 終了条件：0.00346 の確率<sup>2</sup>で毎試行終了抽選を行う
  - しかし、実際は5 回の平均で151 試行となるように調整された(らしい...)
  - 参加者はいつ終わるのが前もって分からないため、end-game-effect を消す効果がある

<sup>1</sup>初回は協調、その後は相手の前回の手を出す方略

<sup>2</sup>期待値が200 試行となるように設定された

## 2.3 The Contestants

- 方略のパフォーマンスは以下の点と相関は無かった
  - 作成者の属性 (国籍や役職など)
  - 使用言語
  - プログラムの長さ (TFT は除く)

## 2.4 The Representatives

- RANDOM を加えた 63 方略の総当り戦で詳細な結果は Table3<sup>3</sup>
- 方略数が多く、詳細な検討をするために重回帰分析の手法を取り入れる
  - ある方略の得点は特定の 5 つの方略 (以後、5 つをまとめて「代表方略」と呼ぶ) との対戦で得た得点で 96 %説明できる

T:方略の予測得点  $S_j$ :j 番目の方略との対戦で得た得点

$$T = 120.0 + (.202)S_6 + (.198)S_{30} + (.110)S_{35} + (.072)S_{46} + (.086)S_{27} \dots (1)$$

- TFT: $S_6 = 453, S_{30} = 453, S_{35} = 453, S_{46} = 452, S_{27} = 446$
- Table4 の 3 列目 ~ 7 列目は TFT と比較して、各方略が代表方略との対戦で得た得点の違いを表す
  - 0 : 違いなし
  - プラス値 : TFT と比べて、低い得点
  - マイナス値 : TFT と比べて、高い得点
- Table4 の最後の列 (Residual) は予想得点と実際の得点との差を表す
  - ほとんどの場合、差が 10 点より小さい範囲にあるため、代表方略との結果はその方略の全体的なパフォーマンスを表すのにふさわしいことが示された
- 代表方略との対戦結果が今後の議論の中心となる

## 3 Properties Of Successful Rules

### 3.1 Niceness

第 1 回目に引き続き「nice ルール」と順位の間で相関があった。

- nice ルール : 相手より先に裏切らないこと
  - 62 方略の内、半数以上が nice ルールを用いていた。
    - \* これは第 1 回目の教訓によるものだろう
  - 上位 15 位の内、8 位の方略を除き、全て nice ルールを用いている
  - 下位 15 位全てが NOTnice であった
- 代表方略の内、3 つが nice ルールを用いていた (Table4 の列が早い 3 つが nice ルールを用いている)
  - nice ルール同士の対戦では全試行で「お互い協調」になるため、得点に差が出ない
    - \* Table4 の 3 列目 ~ 5 列目で 0 が多く見受けられるのはそのためである
- NOTnice であるにも関わらず、8 位になった Harrington の方略

<sup>3</sup>表が見にくいいため割愛

- 方略の概要
  - \* TFT の改変版
  - \* 37 試行目に必ず裏切る .
  - \* その後 相手が裏切りを返すまでランダムに裏切り, こちらの裏切りに反応が無い場合, 裏切る割合を上昇させる
  - \* 相手が裏切りを返してきたら, ランダムに裏切るのをやめる. (TFT に戻る?)
- 2 番目の代表方略である REVISED STATE TRANSITION との対戦では TFT と比べて, 37.2 点も低かった
  - \* REVISED STATE TRANSITION は 1 次マルコフ過程を用いた方略
  - \* Harrington の方略が裏切り始めても, REVISED STATE TRANSITION は最初, 裏切りに反応しない
  - \* しかし, 相手が裏切りの度合いが増すと, 徐々に裏切り出し, 「お互い協調」後でも裏切りを選択するようになる
  - \* その結果, TFT よりも低い結果となった
- 相手がこちらの裏切りを幾度か見逃してくれる方略であっても, 限度を超すと相手が報復するようになり, 関係の修復が難しくなる .
  - REVISED STATE TRANSITION との対戦結果において, TFT に勝る方略が多くある
  - しかし, そのような方略は他の代表方略との対戦では振るわなかった

## 3.2 Provocability

nice ルールを持つ方略内での順位を決めたものは「怒り (provocability)」であった .

- NOTnice な代表方略は 2 つあり, それが nice ルールを持つ方略内の順位を決める要因となった
  - TESTER:総合順位 46 位 . Table6 列目
  - TRANQUILIZER:総合順位 27 位 . Table7 列目
- TESTER
  - 方略の概要
    - \* 第 1 試行目で裏切る
    - \* 相手が裏切りで返してきたら, 次から協調を選び, その後は TFT
    - \* 相手が裏切りで返さなかったら, 4 試行目, 6 試行目と偶数試行目で裏切る . 言い換えれば, 協調と裏切りを交互に繰り返す
  - TESTER は第 1 回目トーナメントに出場していれば好成績を収めたいだろう
    - \* 例えば, 第 1 回目トーナメントに出場していたら優勝していた TF2T<sup>4</sup>相手だと, 高得点を得ることが出来る
    - \* TF2T は今回のトーナメントに出場したが, 結果は 24 位と振るわなかった .
      - ・ TESTER は 2 回連続裏切らないため, TF2T は裏切りを返さないから
  - TESTER に苦しめられた方略
    - \* 第 1 回目に出場し kingmaker の役割を果たした DOWNING が用いた方法<sup>5</sup>を踏襲した方略
      - ・ これに該当する方略は 3 つあった
    - \* なぜ苦しめられたかという点, DOWNING の方法は自分が協調した後で相手が協調した場合, 大方続けて協調を選ぶから
  - TFT の対応
    - \* TFT は TESTER 相手でも好成績を残した理由は, 相手の裏切りに対して即座に裏切りで返したから
    - \* その結果, TESTER から謝罪 (協調) を引き出し, 良い結果へと繋がった

<sup>4</sup>相手に 2 回連続裏切られたら裏切りを返す方略 . 今まで HalfTFT と呼んでいたもの .

<sup>5</sup>前回の自分の選択別に相手が協調する確率を算出し, 行動を決定する方法

- TRANQUILIZER

- － 方略の概要

- \* 最初の何試行かは基本的に協調．相手の裏切りがあまりにも多い場合に限り，裏切りを返す
    - \* 裏切りを選ぶ
    - \* 相手が裏切りを返してこなかったら，より頻繁に裏切りを選択するようになる．
    - \* しかし，今まで獲得した得点が高ければ，2回連続して裏切らない．

- － TRANQUILIZER に苦しめられた方略

- \* 14位の Eatherley の方略
      - ・ TRANQUILIZER との結果にのみ，TFT と差が出てしまい，その差は 50.4 であった．
    - \* Eatherley の方略は「(相手が裏切った回数)/(今までの試行数)」を基に，裏切るかどうかを判定する．
    - \* その結果，TRANQUILIZER に裏切られてもなかなか裏切りを選択しないから搾取されてしまった．
    - \* Table5 は TRANQUILIZER と Eatherley の方略との行動パターンの例

TABLE 5  
Illustrative Game Between EATHERLEY and TRANQUILIZER

moves	1- 20	11111	11111	11111	11111
moves	21- 40	11212	11121	11111	11211
moves	41- 60	11211	11112	11112	11232
moves	61- 80	11121	21121	11111	12111
moves	81-100	21121	11111	23111	11112
moves	101-120	11111	11111	21112	11111
moves	121-140	23111	12121	11111	21111
moves	141-151	12111	11111	1	

Score in this game: EATHERLEY 387, TRANQUILIZER 492

Legend: 1 both cooperated  
2 EATHERLEY only cooperated  
3 TRANQUILIZER only cooperated  
4 neither cooperated

- TESTER と TRANQUILIZER を相手にする時に上手く振舞うために大事なこと

- － 怒り：相手からの「理由無き裏切り」に対して，すぐに裏切りを返すこと
    - \* TESTER, TRANQUILIZER 共に相手が裏切りを即座に返さなければ，調子に乗る，より頻繁に裏切ってくる
    - \* その結果，のんきな(罰を与えない)方略だと搾取されてしまう
  - － 好成績を取るためには「nice ルールを持ち，かつ相手の裏切りを罰すること」が必要であった

- 怒りと寛容さ<sup>6</sup>の関係は？

- － 寛容さは「お互い裏切り」状態を回避するため，「片方協調，片方裏切り」が反復している状態から「お互い協調」へ導くために必要
  - － 怒りは相手の「意味無き裏切り」に罰を与え，相手に裏切りを抑制させるために必要
    - \* 今回のトーナメントでは相手の裏切りに寛容な方略が目立った．
      - ・ 恐らく，多くの参加者が第1回目の教訓を踏まえたからだろう
    - \* そこを狙った方略 (TESTER や TRANQUILIZER など) によって，搾取されてしまった．
    - \* その結果，怒りが方略の順位を決める鍵となった．

<sup>6</sup>第1回のトーナメントで明らかになった概念．相手の裏切りに対して，協調を返す傾向．

## 4 The Top Three Rules

今まで見てきたように、良い順位を取るためには「nice」、「寛容さ」、「怒り」が必要だと分かった。ではトップ3の方略はどのような方略だったのか？

- 1位：TFT
  - nice ルール
  - 寛容さ：裏切りに対して、1回だけ裏切る。以前の裏切りについては考慮しない
  - 怒り：過去にどれだけ相手が協調しようとして、相手の裏切りに対してすぐ裏切りを返す
- 2位：Championの方略
  - 方略の概要
    1. 最初の10試行協調
    2. 11試行目から25試行目までTFT
    3. 26試行目から下の条件を満たした場合裏切り、そうでなければ協調
      - \* 前回、相手が裏切った場合
      - \* 現時点で、相手の協調率が60%未満の場合
      - \* 0~1のランダム数を生成し、それが相手の協調割合よりも高かった場合
  - 欠点
    - \* 代表方略5つ目のTESTERとの対戦で最初の10試行の内、5試行だけ搾取されてしまった
      - ・ TESTERは相手に裏切られない限り交互に裏切り、協調を選択する
      - ・ しかし、Championの方略は最初の10試行は無条件で協調を選択するため。
    - \* その結果、TESTERとの対戦結果において、TFTよりも12ポイント低い得点となった
- 3位：Ottoの方略
  - 方略の概要
    - \* 基本的にTFT。ただしチェック機能を付与。
    - \* チェック機能は、1. 相手がランダム方略か、2. 相手があまりにも協調的でないか、3. 「片方協調、片方裏切り」が反復していないかの3つをチェックする。
    - \* 3の反復は3試行前までの履歴をチェックする。もし判定されれば協調を選ぶ。
  - 欠点
    - \* 3の反復チェック機能が原因となり、TRANQUILIZERとの対戦結果が振るわなかった。その理由は以下の流れによるものである。
      1. TRANQUILIZERはたまに $n$ 試行目と $n+2$ 試行目に裏切るという行動を取るらしい
      2. その現象により、Ottoの方略の反復チェック機能が利いて、協調を選ぶ
      3. その試行でTRANQUILIZERは裏切りを選んだ場合、Ottoの方略は搾取されてしまう。さらに、TRANQUILIZERが持つ特徴により、その後も裏切る確率が上昇してしまう。

## 5 The Role Of The Environment

### 5.1 The Lessons Drawn From Round One

第1回トーナメントで優勝したTFTの代わり出場していれば、TFTよりも高い得点を得た方略が2つあった。TF2TとREVISED DOWNINGである。これらの方略が今大会ではどのような結果を残したのか？

- TF2T
  - 今大会での成績は24位であった。

- 原因は TESTER に搾取されたから
- REVISED DOWNING<sup>7</sup>
  - この方略も今大会に出場したが、下位に沈んでしまった。
  - 原因はやはり TESTER に搾取されたから
- 今大会で見られた面白い相互作用
  1. 「nice ルール」と「寛容さ」をどのように取り入れるかを考えて作られた方略
  2. 相手が 1 を持っているならば、どうすればアドバンテージを取れるかを考えて作られた方略
    - 1 は 2 によって搾取されてしまった。
    - しかし、代表方略の TESTER や TRANQUILIZER のような 2 は上位に入ることが出来なかった。
      - \* 理由は、搾取しようと裏切りを仕掛けると、「お互い裏切り」状態になってしまったから

## 6 Is Tit For Tat The Best Decision Rule?

前回に引き続き優勝した TFT は最も有効な方略と言えるのか？

- 答えは NO . 理由は以下の 3 点
  1. 「完全なランダム方略チェック機能」を持つ方略には勝てないと考えられるから
    - 完全にミス無くランダムを判定する機能は極めて難しいが、実現出来るはず
  2. 今大会の上位半分の方略のみで大会を行うと、TFT は 4 位になってしまうから
    - 1 位：今大会 25 位の方略，2 位：16 位の方略，3 位：8 位の方略
    - 恐らく、搾取しようとする方略が有利となったのではないが
  3. この論文の最初で述べたように、相手の行動に依存して結果が決まるため、最も有効な方略は一意に定まらないから
- TFT との対戦に勝とうと裏切りを選ぶと、「お互い裏切り」状態に陥り、結果的に自分の首を絞めることになる
- 相手から搾取しようとするすると 3 つの問題が生じてしまう
  1. 仕返しを受けるリスクが伴う
  2. 一旦相手を怒らせると (相手の許容範囲を超えて裏切ると)、修復するのが難しくなる
  3. 相手を RANDOM, または極端に協調的でないと判断し、そこから裏切りを始めようとする試みは判断を誤った時のリスクが高い。
    - TFT のような辛抱強い方略が有利になる場合がある
- 今大会において、高いリスクを伴わずに搾取に成功した方略は一つも無かった。
- TFT は以下の 3 つの要素が組み込まれた方略である
  1. nice : 不必要な争いを避ける
  2. provocability : 相手の裏切りを抑制させる
  3. forgiveness : 「お互い協調」状態へ修復する

<sup>7</sup>前回の自分の手が協調の時に相手が協調する確率、前回の自分の手が裏切りの時に相手が協調する確率を過去の履歴から計算し、それを基に行動を決定する方略。第 1 回トーナメントで DOWNING が犯した「最初の 2 試行で必ず裏切る」というミスを改良した方略

TABLE 4  
Performance of the Rules

Rank	Tournament Score	Performance with Representatives (Points lost relative to TIT FOR TAT)					Residual
		Rule 6	Rev. State Transition (30)	Rule 35	Tester (46)	Tran- quilizer (27)	
1	434.73	0	0	0	0	0	13.3
2	433.88	0	0	0	12.0	2.0	13.4
3	431.77	0	0	0	0	6.6	10.9
4	427.76	0	0	0	1.2	25.0	8.5
5	427.10	0	0	0	15.0	16.6	8.1
6	425.60	0	0	0	0	1.0	4.2
7	425.48	0	0	0	0	3.6	4.3
8	425.46	1.0	37.2	16.6	1.0	1.6	13.6
9	425.07	0	0	0	0	11.2	4.5
10	424.94	0	0	0	26.4	10.6	6.3
11	422.83	0	0	0	84.8	10.2	8.3
12	422.66	0	0	0	5.8	-1.2	1.5
13	419.67	0	0	0	27.0	61.4	5.4
14	418.77	0	0	0	0	50.4	1.6
15	414.11	0	0	0	9.4	52.0	-2.2
16	411.75	3.6	-26.8	41.2	3.4	-22.4	-11.5
17	411.59	0	0	0	4.0	61.4	-4.3
18	411.08	1.0	-2.0	-.8	7.0	-7.8	-10.9
19	410.45	3.0	-19.6	171.8	3.0	-14.2	3.5
20	410.31	0	0	0	18.0	68.0	-4.0
21	410.28	0	0	0	20.0	57.2	-4.9
22	408.55	0	0	0	154.6	31.8	.9
23	408.11	0	0	0	0	67.4	-7.6
24	407.79	0	0	0	224.6	56.0	7.2
25	407.01	1.0	2.2	113.4	15.0	33.6	2.5
26	406.95	0	0	0	0	59.6	-9.4
27	405.90	8.0	-18.6	227.8	5.6	14.0	8.9
28	403.97	0	0	0	3.0	1.4	-17.2
29	403.13	4.0	-24.8	245.0	4.0	-3.0	4.4
30	402.90	0	0	0	74.0	54.4	-8.6
31	402.16	0	0	0	147.4	-10.0	-9.6
32	400.75	0	0	0	264.2	52.4	2.7
33	400.52	0	0	0	183.6	157.4	5.7
34	399.98	0	0	0	224.6	41.6	-1.9
35	399.60	0	0	0	291.0	204.8	16.5
36	399.31	0	0	0	288.0	61.4	3.7
37	398.13	0	0	0	294.0	58.4	2.7
38	397.70	0	0	0	224.6	84.8	-.4
39	397.66	1.0	2.6	54.4	2.0	46.6	-13.0
40	397.13	0	0	0	224.6	72.8	-2.0



Table 4 (Continued)

Rank	Tournament Score	Performance with Representatives (Points lost relative to TIT FOR TAT)					Residual
		Rule 6	Rev. State Transition (30)	Rule 35	Tester (46)	Tran- quilizer (27)	
41	395.33	0	0	0	289.0	-5.6	-6.0
42	394.02	0	0	0	224.6	74.0	-5.0
43	393.01	0	0	0	282.0	55.8	-3.5
44	392.54	0	0	0	151.4	159.2	-4.4
45	392.41	0	0	0	252.6	44.6	-7.2
46	390.89	1.0	73.0	292.0	1.0	-.4	16.1
47	389.44	0	0	0	291.0	156.8	2.2
48	388.92	7.8	-15.6	216.0	29.8	55.2	-3.5
49	385.00	2.0	-90.0	189.0	2.8	101.0	-24.3
50	383.17	1.0	-38.4	278.0	1.0	61.8	-9.9
51	380.95	135.6	-22.0	265.4	26.8	29.8	16.1
52	380.49	0	0	0	294.0	205.2	-2.3
53	344.17	1.0	199.4	117.2	3.0	88.4	-17.0
54	342.89	167.6	-30.8	385.0	42.4	29.4	-3.1
55	327.64	241.0	-32.6	230.2	102.2	181.6	-3.4
56	326.94	305.0	-74.4	285.2	73.4	42.0	-7.5
56	309.03	334.8	74.0	270.2	73.0	42.2	8.4
58	304.62	274.0	-6.4	290.4	294.0	6.0	-9.3
59	303.52	302.0	142.2	271.4	13.0	-1.0	1.8
60	296.89	293.0	34.2	292.2	291.0	286.0	18.8
61	277.70	277.0	262.4	293.0	76.0	178.8	17.0
62	237.22	359.2	261.8	286.0	114.4	90.2	-12.6
63	220.50	311.6	249.0	293.6	259.0	254.0	-16.2