

Effective Choice in the Prisoner's Dilemma

Robert Axelrod

Journal of Conflict Resolution, Vol.24 (1980) 3-25

1 Introduction

1.1 Iterated Prisoner's Dilemma

- 国際政治では囚人のジレンマが生じる状況が数多く生まれている (*Snyder, 1971*)
 - 冷戦中、アメリカとソ連の間での軍事拡大、軍隊配置など
- 囚人のジレンマ
 - お互い協調：両方とも満足
 - 片方協調，片方裏切り：裏切った方は一番高い得点，裏切られた方は一番低い得点
 - お互い裏切り：両方とも不満足

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.)

- 囚人のジレンマの特徴
 - 相手の選択によらず裏切りを選択する方が得なので，両者ともに裏切る動機を持つ
 - しかし，お互い裏切りを選択するとお互い協調した場合に比べて，得られる利得が低い
 - よって，個人の合理性（裏切って高い利得を望むこと）とグループの合理性（お互い協調を目指すこと）の間で葛藤が生まれる
- 相手との関係
 - 相手との関係が今後も続く場合「全て裏切る」といった自分本位な行動を取ることは適切ではない
 - そのような状況で効果的な選択を行うためにはどのような方略が有効なのか？
- TFT(しっぺ返し方略)
 - 冷戦時，アメリカとソ連の両者ともに TFT という方略を取っていた

- * TFT：最初は協調を選び，2 試行目からは前回の相手の手を選択する方略
- TFT の改良 (?): TFT よりも高い利得が得られると期待できる改良をすると...
 - * 改良箇所：相手の協調に対して，90%の割合で裏切る
 - * 結果：TFT よりも利得が低くなった
 - * 理由：長い間，お互い裏切り状態が続いてしまった

1.2 The Study of Effective Choice

囚人のジレンマを用いた研究を 3 つに分けることができる。

1. 心理実験

- プレイヤーに囚人のジレンマゲームをやってもらい，その結果からプレイヤーの属性などの要因が協調度合いや行動パターンにどのような影響を与えるかを検討する研究
 - 中アフリカにおける西洋化 (Westernization) の効果 (*Bethlehem*, 1975)
 - 仕事に重きを置く (career-oriented) 女性の攻撃性 (*Baefsky and Berger*, 1974)
 - 思考スタイルが抽象的 vs. 具体的 (*Nydegger*, 1974)
- 10 年の間 (1970 ~ 1980) で学術誌 *Psychological Abstracts* の中だけでも，350 もの繰り返し囚人のジレンマゲームに関する論文が掲載されている

2. モデルに関する研究

- あるモデルを作成する時に囚人のジレンマが持つ構造を基盤として利用している研究
 - 囚人のジレンマが生じる状況での相互作用をベースとした腕相撲のモデル (*Richardson*, 1960)
 - 独占市場での競争を囚人のジレンマとしてモデル化 (*Samuelson*, 1973)
- 政治，社会，経済についての数多くのモデルは囚人のジレンマの構造を基に作成されている

3. 囚人のジレンマゲームが持つ根本的な特徴を明らかにする研究

- 実験室や現実世界を超えて，より抽象的なゲームとして囚人のジレンマゲームの特徴を分析する研究
 - 合理的行動の重要性 (*Luce and Raiffa*, 1957)
 - 協調行動が持つ影響 (*Taylor*, 1976)
- しかし，上述の 3 種類の研究からは「ゲームを上手にプレイする方法」についての知見を得ることができない
 - 繰り返し囚人のジレンマゲームについて効果的な選択の仕方について検討するには新たなアプローチが必要である
 - 新たなアプローチ：コンピュータトーナメントを行い，プログラム同士で対戦させる
 - なぜ，コンピュータプログラムを使うのか？
 - * 「効果的な方略」とは，ある特定の方略に効果的なだけではなく，様々な種類の方略に対して効果的でなくてはならない
 - ・ あらゆる対戦相手と対戦しなければ「効果的な方略」について検討できないため，人間同士の対戦では無理がある
 - * 効果的な選択を行うためには，今までの行動履歴を考慮する必要がある
 - ・ コンピュータでは全ての行動履歴を保持でき，それを誤りなく利用することが出来る
- 結果，囚人のジレンマという状況を上手く処理する方法に関する新たな洞察を得ることが出来た

2 The Tournament

2.1 The Winner

優勝した方略は TFT(Rapoport のプログラム)であった。

- 提出された方略：14 方略
- TFT は 14 方略中、もっともシンプルな方略であった
 - － プログラムのフォーマットと言語 (FORTORAN) は統一されており、TFT はたったの 4 行
- プログラムのエントリーを受け付ける前に行われた予選で、TFT が一位であったことが全参加者に知らされていた
 - － 要するに、参加者はこの TFT に打ち勝つ方略を考えずであり、それにも関わらず TFT が優勝した
- シンプルな方略である TFT が、複雑な方略に勝ったという結果はコンピュータチェストーナメントでの結果と矛盾している
 - － コンピュータチェストーナメントを勝ち抜くには、より複雑な方略が求められる
 - － そのチェストーナメントの優勝者である Joss もこの囚人のジレンマの大会に出たが、14 方略中 13 位であった
 - * Joss の方略は上述した、相手の協調後に対して 90%の確率で協調するように TFT を改変した方略であった

2.2 The Form of the Tournament

- 対戦形式
 - － 15 方略総当たり戦
 - * 提出された 14 方略に RANDOM 方略を加えている
 - * 自分自身の方略 (twin) との対戦も含んでいる
- 対戦回数、試行回数
 - － 方略同士の対戦は 200 試行を 1 セットとし、それを 5 セット行う
 - － すなわち、一つの方略は $15(\text{方略}) \times 200(\text{試行}) \times 5(\text{セット})$ 試行行う
- 利得表は Table1
- 指標
 - － 全 15 方略相手にした時に獲得した得点の平均
 - * 理論上では 0 点 ~ 1000 点の範囲を取るが、結果的には全ての試行で「お互い協調」の時に獲得する得点 (600 点) と全ての試行で「お互い裏切り」の時に獲得する得点 (200 点) の範囲に収まっている。(Table2 参照)
 - * 優勝した TFT は平均 504 点
 - － 各対戦の得点は 5 セットの平均
- 結果は Table2

2.3 The Contestants

出場者

- ゲーム理論、特に囚人のジレンマに関する論文の著者 16 名¹
- 経済学、心理学、数学、政治学、社会学と様々な専門領域から参加し、国別では 3ヶ国

¹共同作品が 2 つある

3 Properties of Successful Rules

3.1 Niceness

Table2の結果を見ると、高得点群(1位~8位)と低得点群(9位~15位)に大別できる。この違いが生まれた原因は高得点群が「niceルール」を用いていたからである。

- niceルールとは
 - 第一試行目に裏切らない、かつ最後の数試行以前で相手よりも先に裏切らないルール
- 高得点群の8方略は全て niceルールを用いており、niceルール同士の対戦ではほぼ全試行で「お互い協調」となり、お互い約600点の得点を手にする
- 低得点群は高得点群との対戦において、400点以上の得点を得ることがほとんど出来なかった(Table2の左下参照)
 - その結果、高得点群の得点は472点~504点の範囲にあるが、niceルールを用いていない低得点群は最高でも401点しか得られていない
 - niceルールは協調的な(niceルールである)相手に対して協調的に振舞うが、協調的でない(niceルールでない)相手には協調的に振舞わない

3.2 Effectiveness with the Kingmaker

高得点群と低得点群の違いは niceルールにあることが分かったが、高得点群の中での違いは何か？

- 2つの kingmaker：GRAASKAMP²と DOWNING
 - 高得点群の中での違いは2つの kingmaker との対戦成績が原因となっている
 - 1位の TFTは両 kingmaker との対戦成績が良かったが、2位の TIDEMAN AND CHIERUZZIは GRAASKAMP との対戦成績が悪かった
- DOWNING
 - 以下の2つの確率を毎試行更新し、その確率に従って行動を決定する方略。利得
 1. 前回の自分の手が協調の時、相手が協調する確率
 2. 前回の自分の手が裏切りの時、相手が協調する確率
 - * $1 = 2$ の時、裏切る
 - ・自分が協調しても裏切っても、相手が協調する確率が同じと判断している。その時は裏切る方が得
 - * $1 < 2$ の時、裏切る
 - ・こちらが裏切ると相手は協調し、こちらが協調すると相手は裏切ると判断している。その時も裏切る方が得
 - * $1 > 2$ の時、協調する
 - ・こちらが協調すれば相手が協調してくれ、逆にこちらが裏切れば相手も裏切ると判断している。その時は協調した方が得
- DOWNINGの欠点
 - DOWNINGの欠点は1と2の初期値を $1/2$ としているため、第一試行目と第二試行目は相手の行動に依らず、常に裏切りになってしまう点である
 - * この「初回から2回連続裏切り」に対して、対戦相手は報復として裏切りを選択する
 - * その結果、DOWNING自身の得点は振るわなかった
 - しかし、この欠点のおかげで kingmakerの役目を果たした
 - * 高得点群の8方略が DOWNING との対戦した時に得た得点には TIDEMAN AND CHIERUZZIの601点から NYDEGGERの158点と大きな差が生まれた

²TFT以外の方略は作成者の名前で呼び、大文字で表記する

- GRAASKAMP
 1. 初回から 50 試行目まで TFT
 2. 51 回目に裏切る
 3. 51 回目から 56 試行目まで TFT
 - ここで過去の履歴や 51 回目の裏切りに対する反応から、対戦相手を自分自身 (twin), TFT, RANDOM, それ以外の方略を見分ける
 - 4-1 対戦相手を twin 又は TFT と認識した場合：残りの試行全て協調
 - 4-2 対戦相手を RANDOM と認識した場合：残りの試行全て裏切り
 - 4-3 それ以外の方略と認識した場合：TFT の改変版, 15 回試行で 5 回裏切りを選択する
- RANDOM
 - RANDOM は kingmaker の役割を果たさなかった
 - * 高得点群と kingmaker の DOWNING と対戦した時, 得点の開きは 443 点. また, kingmaker の GRAASKAMP と対戦した時, 得点の開きは 290 点
 - * それに比べ, RANDOM は 157 点
 - 逆に, 高得点群において, RANDOM と対戦した時の得点が高いほど, 順位が低くなる傾向がある
 - * RANDOM との対戦で高得点を得るためには, なるべく早く「対戦相手は RANDOM である」と判断し, そこから全て裏切りを選択する必要がある
 - * しかし, 対戦相手を RANDOM と早く判断しようとする, その分リスクを伴う
 - FRIEDMAN: 最初は協調. 一度相手に裏切られると, そこから全て裏切る方略
 - * nice ルールであるが, 相手の裏切りに非常に厳しく対応する
 - * 対戦相手が RANDOM の時, 高い得点 (598 点). 優勝した TFT は 441 点
 - * しかし, nice ルールでない方略との対戦では成績が悪い
 - * その結果, 総合順位は低くなってしまった

3.3 Forgiveness

FRIEDMAN は相手に裏切り行動に対して極めて厳しく対応した. その結果, nice ルールでありながら, 低い順位に終わってしまった. nice ルールであるのと同時に, 相手の裏切り行動に対して「寛容 (Forgiveness)」に振舞うことが鍵となる.

- Forgiveness とは?
 - 相手の裏切りに対する寛容さ, 相手の裏切りを許すこと
- 「寛容さ」の比較
 - FRIEDMAN < SHUBIK < TFT
 - * FRIEDMAN: 相手が一回裏切った後は全て裏切り続ける.
 - * SHUBIK: 基本は TFT. しかし, 相手が裏切る度に裏切る回数が増えていく
 - * TFT: 裏切られたら裏切るが, 単独の裏切りに対しては寛容
- nice ルールでない方略の順位が低かった理由
 - 提出された方略のほとんどが「寛容さ」に欠ける方略だったから
 - * 自分が裏切るより先に裏切られた場合, すぐに裏切りで返す
 - * その裏切りが裏切りを呼び, 結局「お互い裏切り」状態が続いてしまった
 - 例)TFT 対 JOSS³の行動パターン (Table3 参照)
 1. 6 試行目で JOSS が裏切る
 - * 相手が協調した後に裏切るという 10%に引っ掛かったため

³JOSS: TFT の改良版. 相手が協調した場合, 90%の確率で協調する方略

2. 7 試行目から 24 試行目まで「片方協調, 片方裏切り」を交互に繰り返す
 - * ここで得られる得点は平均 2.5 点であり、「お互い協調」の 3 点より低い
3. 25 試行目に協調に対して JOSS がまた裏切ったため、それ以降「お互い裏切り」

TABLE 3
Illustrative Game Between TIT FOR TAT and JOSS

moves	1- 20	11111	23232	32323	23232
moves	21- 40	32324	44444	44444	44444
moves	41- 60	44444	44444	44444	44444
moves	61- 80	44444	44444	44444	44444
moves	81-100	44444	44444	44444	44444
moves	101-120	44444	44444	44444	44444
moves	121-140	44444	44444	44444	44444
moves	141-160	44444	44444	44444	44444
moves	161-180	44444	44444	44444	44444
moves	181-200	44444	44444	44444	44444

Score in this game: TIT FOR TAT 236, JOSS 241

Legend: 1 both cooperated
2 TIT FOR TAT only cooperated
3 JOSS only cooperated
4 neither cooperated

- TFT の「寛容さ」
 - TFT は nice ルールであり、かつ他の方略に比べ「寛容さ」を持っていた
 - 前回の試行のみに着目し、それ以前の裏切りについては過去のものとして割り切っている
 - よって、2 つの kingmaker に対しても上手く振舞っていた
- TIDEMAN AND CHIERUZZI の「寛容さ」
 - SHUBIK⁴と似ているが、ある条件を満たしたら 2 回協調し、裏切る回数もリセットする方略。また最後の 2 回は自動的に裏切る。
 - リセット条件：以下の 5 つの条件を全て満たした時
 1. 自分と相手との差が 10 点以上
 2. リセットしてから 20 試行以上経過
 3. 相手が連続して裏切ってこなかった
 4. 残り試行数が 10 試行以上
 5. 今までの相手の裏切り確率が 1/2 付近でない (RANDOM を識別する目的)
 - 条件を満たした時に見せる「寛容さ」によって、「お互い裏切り」状態から回避することが出来た
- NYDEGGER の「寛容さ」
 - 直近の 3 試行における行動履歴をもとに、行動を決定する方略
 - 注目すべき特徴：直近の 3 試行全てが「お互い裏切り」だった場合、協調を選択する
 - * TFT に似た方略 (FELD や JOSS) との対戦において、「お互い裏切り」状態を回避した
 - * しかし、「寛容さ」が欠けている方略 (FRIEDMAN や DAVIS) との対戦では、逆に搾取されてしまった

⁴SHUBIK：基本は TFT。しかし、相手が裏切る度に裏切る回数が増えていく

4 The Surprises and Their Implications

今回のコンピュータートーナメントの驚きべき結果と明らかになったことを述べる

1. TFT の優勝

- 提出された方略の中で最もシンプルな方略が優勝した
- 互惠行動は倫理的な行動であるだけでなく、行動主体にとっても極めて実用的な行動だということが示唆された

2. TFT に勝る方略が無かったこと

- 参加者は TFT に勝る方略を考えたはずだが、それにも関わらず TFT が優勝してしまった
- 原因は「裏切りの反響 (echo)」を見過ごしていた点にあると考えられる
 - － 裏切りが裏切りを呼び、結果的に「お互い裏切り」状態に陥った
- 洗練された方略を考え出すためには、裏切りが持つ効果を 3 つのレベルで考えなくてはならない
 1. 裏切りを選択する直接的な効果
 - － 協調するよりも裏切った方がもらえる点数が高いこと
 2. 裏切りを選択する間接的な効果
 - － 相手がこちらの裏切りに対して報復してくるか否か
 3. 裏切りを選択する長期的な効果
 - － 自分の裏切りがその後どのような結果に結びつくのか
- 絶好の機会に単独の裏切りをすることは 1,2 に観点で言えば有効である
- しかし、3 を考慮に入れたとすれば、その裏切りから「お互い裏切り」状態を招く危険性がある
 - － 提出された方略は 3 についての考慮を欠いていた

3. きれいな結果

- nice ルールである方略と nice ルールでない方略との順位の差が明確になった
 - － nice ルールの方略が上位 8 位、そうでない方略は全て 9 位以下
- nice ルールというたった一つの特性が成績の良し悪しを決めた

4. 「寛容さ」

- nice ルールが上位と下位を明確に分けた。
- 「寛容さ」は上位内での順位を決める要因となった
 - － 「寛容さ」を持っていればいるほど、パフォーマンスが良い
 - － TFT は前回の裏切りには厳しいが、過去の行動に対しては寛容な態度を取る
- 「寛容さ」は相手の裏切りから始まる「お互い裏切り」状態を避ける効果がある

5. 2 つの kingmaker と RANDOM

- 上位の順位を決めたのはたった 2 つの kingmaker (GRAASKAMP と DOWNING) であった
- RANDOM は kingmaker の役目を果たせなかった
 - － 提出された方略は対戦相手が RANDOM かどうかを迅速に判断出来たから
 - － しかし、見切りの早さが原因で RANDOM でない相手に対して誤った行動をしてしまう方略も見受けられた
- 不可解な行動を取る対戦相手に対して、正確に RANDOM かどうかを見極めることが必要

6. TFT に勝る方略

- 今回の大会では TFT が優勝したが、TFT に勝る方略は比較的簡単に見つけ出すことが出来る
- 次の 3 つの方略はいずれも出場していれば優勝していた方略である

- 半返し方略 (TF2T)
 - * TFT の改変版．相手に二回連続裏切られたら裏切る方略
 - * TFT よりも大きい「寛容さ」を持つ
- 探索木方略 (LOOK AHEAD)
 - * 人工知能の分野で頻繁に使われる技法である探索木を使用した方略
 - * 探索木はチェスの大会でも使われるアルゴリズムである
- 改良型 DOWNING (REVISED DOWNING)
 - * kingmaker となった DOWNING の改良型方略
 - * DOWNING は相手の手に関わらず，最初に 2 試行で裏切りを選択する方略であった．
 - ・ そのため，「お互い裏切り」状態へ
 - ・ この欠点を変更し，最初に協調から始めるようにしている
- TFT の代わりに，これら 3 つの方略が出場していた場合に獲得する得点
 - 半返し方略：532 点
 - 探索木方略：520 点
 - 改良型 DOWNING：542 点
 - (参考)TFT：504 点
- この結果から分かることは以下の 3 つ
 1. 今大会に出場した方略の多くは早い試行で裏切りを選択している
 2. TFT に勝る 3 つの方略は出場した方略に比べ，「寛容さ」を持つ
 3. DOWNING は相手の行動を悲観的に考えたため，それが原因で優勝を逃した
 - 相手が自分の協調に対して反応してくれない，協調で返してくれないという悲観的な想定⁵をしたために，最初の 2 回を裏切りにしていた

7. たくさんのサプライズがあったこと

- たくさんのサプライズがあったこと自体が驚きである
- これはこの大会の結果が協調行動について多くの洞察を与えてくれたことを意味している
- 様々な専門分野から出場してもらったが，全ての参加者は...
 - 自分の利得に目がくらみ，
 - 十分な「寛容さ」を持たず，
 - 相手に対して悲観的な想定をしてしまった．
- 方略の有効性というのは自分の特性だけではなく，対戦相手の性質に依存する
 - そのため，今回の結果だけで「効果的な方略」を断定することはできない
 - しかし，有効な選択をするために必要とされる概念について新たな洞察を提供した
- 有効な選択を行うために必要とされるものは以下の 3 つであることが分かった
 - nice
 - Forgiveness(寛容さ)
 - optimistic(楽観的)

⁵DOWNING は自分が協調した後に相手が協調する確率と自分が裏切った後に相手が協調する確率の初期値をともに 1/2 に設定していた．これは自分が協調しても相手が協調を返すかどうか分からないという悲観的な考えによるものである．

TABLE 2
Tournament Scores

Player	Other Players														Average Score	
	TIT FOR TAT	TIDEMAN AND CHIERUZZI	NYDEGGER	GROFMAN	SHUBIK	STEIN AND RAPOPORT	FRIEDMAN	DAVIS	GRAASKAMP	DOWNING	FELD	JOSS	TULLOCK	(Name Withheld)		RANDOM
1. TIT FOR TAT (Anatol Rapoport)	600	595	600	600	600	595	600	600	597	597	280	225	279	359	441	504
2. TIDEMAN AND CHIERUZZI	600	596	600	601	600	596	600	600	310	601	271	213	291	455	573	500
3. NYDEGGER	600	595	600	600	600	595	600	600	433	158	354	374	347	368	464	486
4. GROFMAN	600	595	600	600	600	594	600	600	376	309	280	236	305	426	507	482
5. SHUBIK	600	595	600	600	600	595	600	600	348	271	274	272	265	448	543	481
6. STEIN AND RAPOPORT	600	596	600	602	600	596	600	600	319	200	252	249	280	480	592	478
7. FRIEDMAN	600	595	600	600	600	595	600	600	307	207	235	213	263	489	598	473
8. DAVIS	600	595	600	600	600	595	600	600	307	194	238	247	253	450	598	472
9. GRAASKAMP	597	305	462	375	348	314	302	302	588	625	268	238	274	466	548	401
10. DOWNING	597	591	398	289	261	215	202	239	555	202	436	540	243	487	604	391
11. FELD	285	272	426	286	297	255	235	239	274	704	246	236	272	420	467	328
12. JOSS	230	214	409	237	286	254	213	252	244	634	236	224	273	390	469	304
13. TULLOCK	284	287	415	293	318	271	243	229	278	193	271	260	273	416	478	301
14. (Name Withheld)	362	231	397	273	230	149	133	173	187	133	317	366	345	413	526	282
15. RANDOM	442	142	407	313	219	141	108	137	189	102	360	416	419	300	450	276