コミットメントを含む繰り返しゲーム

柴　直樹

虚偽のコミットメントの効用と，それを防ぐメカニズムを探るための基本的な枠組として，繰り返し囚人のジレンマゲームを拡張したモデルを提案する．さらに，そのモデルを用いた計算機実験により，コミットメントの履行に着目した三つの典型的な戦略の優位性について比較を行う．3戦略にノイズを加えた場合のコンピュータシミュレーションと，多様な戦略プールの中から淘汰によって優位な戦略が次世代に残るという進化ゲームの手法を用いた計算機実験を通して，コミットメントを破棄する戦略に比べて，コミットメントを履行する戦略の方が有利であるという結果を示す．

キーワード：繰り返し囚人のジレンマ，コミットメント，進化ゲーム

1. はじめに

意思決定主体が，将来の行動をあらかじめ表明することがある．経済学では，このような行為を指して「コミットメント」といい，通常，表明した行動を確実に実行することの約束を意味する．しかし，約束が守れないことがわかっているにもかかわらず，意思決定主体が将来の行動を表明することがしばしば行われる．

マイクロソフト社によるWindows 95の数回にわたる販売延期には，当初から実現不可能なリリース時期をアナウンスすることにより，販売戦略を有利に進めたという見方がある[1]．このように，企業が虚偽のコミットメントを意図的に行い，戦略的に「嘘をつく」ことにより，自身を有利な立場に導くことができる場合がある．本稿では，このような虚偽のコミットメントの効用と，それを防ぐメカニズムを探るための基本的な枠組として，繰り返し囚人のジレンマゲームを拡張したモデルを提案する．さらに，そのモデルを用いた計算機実験により，コミットメントの履行に着目した三つの典型的な戦略の比較を行う．

2. 繰り返し囚人のジレンマ

戦略形ゲームとしての本来の囚人のジレンマは，例えば表1にあるような利得行列を持つ2人ゲームである．それぞれのプレイヤは，協調する (C)，裏切る (D) という二つの戦略を持っており，例えば，両者のプレイヤが互いに協調し合えば，ともに3という利得を得るが，一方が協調しもう一方が裏切った場合には，それぞれ0,5という利得を得ることになる．両者とも裏切った場合には，ともに1という利得しか得られない．

表1　囚人のジレンマゲームの利得行列

	C	D
C	3,3	0,5
D	5,0	1,1

ゲーム理論においては，プレイヤに何らかの合理性を仮定し，その仮定のもとで起こりうるゲームの結果を「解」と呼んでいる．解概念の代表は，相手の戦略に対して自己の利得を最大化する戦略の選択が互いに整合し合うような戦略の組を指す「ナッシュ均衡」である．表1のゲームでは，Dの組が唯一のナッシュ均衡となる．結局，個々のプレイヤが自己の利得を最大化しようとして行動した場合には，互いに相手を裏切るということになり，その結果，ともに1という利得しか得られない．

一方で，Cの組合せは，Dの組合せよりも2人のプレイヤ双方の利得を増加させることができる「パレート優位」な戦略の組となっている．このように，囚人のジレンマは，個人合理性（ナッシュ均衡）と，全体合理性（パレート優位性）とが矛盾するゲームであり，両者間の調和をいかにして計るかというのは，ゲーム理論の課題の一つである．その解決のためのアプローチの一つが，本稿の主題である「繰り返しゲーム」という分析枠組である．

我々が行う主体間の取引の多くは，長期的な継続を前提に行われる．商取引はもちろん，交友関係，婚姻といった私的なものに至るまで，最初から期限つきの

しば　なおき
千葉工業大学 社会システム科学部
〒275-0016 習志野市津田沼2-17-1

関係が設定されることは少ない．つまり，意思決定主体間の相互作用が，戦略形ゲームが想定するように1回だけで終わるのではなく，長期にわたり継続的，反復的に行われることは多い．

繰り返しゲームは，上記のような1回きりのゲーム（戦略形ゲーム）が反復して行われるゲームである．繰り返しの基本単位となる戦略形ゲームを，成分ゲームという．各回の成分ゲームでは，利得行列に基づいて利得が決定され，ゲームの反復に伴い，個々のプレイヤの利得が累積され，それを個々のプレイヤのゲーム全体での利得と考える．利得の総和を求めるに当たっては，割引因子を設定し，将来の利得の現在価値を割り引いて考えることが多い．

繰り返しゲームでは，時間の経過にそってゲームは進行する．このような状況を取り扱うゲーム理論での一般的なモデルは，展開形ゲームと呼ばれる．展開形ゲームの解概念としては，戦略形ゲームにおけるナッシュ均衡の概念を拡張した，部分ゲーム完全均衡（以下，単に均衡と呼ぶ）と呼ばれるものがある．繰り返し囚人のジレンマにおいては，ゲームの反復回数が有限の場合には，ゲームの終了時からさかのぼって均衡を求めることができる（後向き帰納法という）．この場合，やはり，常に裏切るという戦略の組が唯一の均衡となり，結果として，2人のプレイヤは互いに裏切りあい続けることになる．結局，有限回の繰り返しでは，個人合理性と全体合理性の調和を計ることはできない．

繰り返しゲームでは，成分ゲームの反復回数として，有限回のみならず，無限回を考える場合がある．現実の状況で，ゲームが無限回繰り返されることはない．しかし，ゲームの反復がいつ終了するかわからない状況は我々のまわりに多く存在する．無限回の繰り返しゲームは，このような状況のモデルと考えてよい．無限回の反復では，有限回の場合とは全く異なるゲームの解が現れる場合がある．繰り返し囚人のジレンマでは，割引因子が十分に大きい値をとるならば，ナッシュ均衡として，上述の「常に裏切る」という戦略以外の組合せが存在する[2]．その結果，互いに協調しあうというプレイが実現できる．つまり，長期的な関係を前提とした場合，互いに自己の利益を最大化するという目的で行動して（個人合理性），しかも互いに協調しあう（全体合理性）という結果を引き出すことができるようになるのである．

繰り返しゲームにおいては，過去のゲーム履歴に基づいて，毎回の成分ゲームでの選択を行うルールが戦略となるため，戦略空間は非常に多様なものとなる．Axelrodは繰り返し囚人のジレンマの様々な戦略同士をコンピュータ上で戦わせる選手権を行い，tit-for-tat戦略（以下，tft戦略とする）が，多様な戦略の中で総じて高い利得を稼ぐ「強い戦略」であると結論づけた[3]．tft戦略とは，最初は協調（C）し，2回目以降は前回に対戦相手の選んだ選択肢と同じ選択肢を選ぶ戦略である．相手の裏切り（D）に対して，ただちにDで報復する，つまり「しっぺ返し（tit-for-tat）」というわけである．

繰り返し囚人のジレンマには，支配戦略は存在しない[3]．つまり，相手の戦略によらず，常に最適な（つまり最も高い利得を稼げる）戦略というものは存在しない．したがって，tft戦略も必勝戦略ではない．しかし，多様な戦略の中で，平均的に高い利得を稼ぐことができ，結局，それらの利得を合計したとき，他の多くの戦略に比べて高い成績を残すことができる．Axelrodの開催した2回の選手権には，それぞれ，15と63通りの戦略同士でリーグ戦を行いその合計得点を競わせているが，そのいずれにおいてもtft戦略が優勝している[3]．tft戦略は，相手の1手前の手のみを参照する戦略であるため，その決定メカニズムは非常に単純である．本稿でも，拡張された繰り返し囚人のジレンマゲームで，tft戦略をベースにしたいくつかの戦略をコンピュータ上で競わせる．

3. コミットメントを含む繰り返しゲーム

上記の繰り返しゲームを，コミットメントを含むゲームのモデルに拡張する．ここでの議論は，一般にn人のゲームとして定義可能であるが，簡単化のためここでは二人による囚人のジレンマを基本としたゲームに説明を限定する．

3.1 成分ゲーム

繰り返しゲームの成分ゲームとなる2人ゲームは，次のようなものである．通常の囚人のジレンマと同様，2人戦略形ゲームの利得行列が与えられている．

ステップ1：それぞれのプレイヤは，CとDのいずれかを「宣言」として選択し，表明する．その際，相手の宣言が何であるかを知らずに，どちらを選択するかを決める．

ステップ2：次に，相手と自分の宣言が何であるかを知ったうえで，CとDのいずれかを「行動」として選択する．その際，相手の行動が何であるかを知ら

ずに，どちらを選択するかを決める．

ステップ3：ステップ2で2人のプレイヤによって選択された行動から，与えられた利得行列にしたがって各プレイヤの利得が決定される．

自分の宣言と行動との間にどのような関係を持たせるかは，プレイヤの自由である．宣言した通りに行動する「正直なプレイヤ」も，宣言と行動が一致しない「嘘つきプレイヤ」も，ゲームのルールとしては許される．また，相手の宣言によって得られる情報を，自分の行動の選択にどのように使用するかもプレイヤの自由である．この2人ゲームを成分ゲームとして反復するのが，今回提示するモデルである．

宣言で何を選択するかは，獲得する利得には直接的には影響しない．この点については注釈が必要であろう．本稿では，コミットメントとは，法的な強制力を伴う契約などとは異なり，その不履行によって，直接当人が懲罰を受けるものとは考えない．ただし，コミットメントの不履行は，将来，他のプレイヤから協調を引き出すことを困難にする．例えば電話でのレストランの予約は，多くの場合予約金を払うことなく行われる．もしコミットメントを履行しなくても（つまり，連絡なく予約をすっぽかしても），キャンセル料を払わずに済ませることは可能かもしれないが，同じレストランでの予約を再度行うことは困難になるだろう．このように，宣言と行動との間の関係性は，直接その回での成分ゲームの利得には反映されないが，将来，他のプレイヤとの関係に影響を与えることで，結果的にそのプレイヤの利得を左右するものとなる．

3.2 繰り返しゲームの戦略プール

上述の2人ゲームを成分ゲームとして反復して行うのが，今回の繰り返しゲームである．成分ゲームのステップ1からステップ3までを1回のゲームとし，反復して，毎回得られる利得に割引因子をかけて総計した値を，繰り返しゲームの各プレイヤの利得とする．コンピュータ上でのシミュレーションでは，成分ゲームの反復回数は有限回で打ち切られるため，成分ゲームでの利得の単純和を各プレイヤの利得とし，簡単化のため割引因子は考えないものとする（あるいは割引因子を1とおいたと考えてもよい）．

本稿では，戦略空間を通常のtft戦略をベースにした戦略に限定して考える．すなわち，各プレイヤは1回前の相手の選択にのみ基づいて，自分の選択を行うものとする．成分ゲームにおいて，宣言と行動に関する二つの意思決定が必要であるため，繰り返しゲームにおける各プレイヤの戦略は，次の2種類の戦略の組みとして表される．

1. 宣言戦略

   相手の1回前の宣言と行動の両方の選択に基づき，毎回の宣言として，CとDのどちらを選ぶかを定めたルールである．CとDからなる集合をSと表す，すなわち，$S=\{C, D\}$とすると，宣言戦略は，関数
   $$p: S\times S\rightarrow S$$
   として表される．ここで，$p(x, y)$は，1回前の成分ゲームにおいて，相手が宣言としてxを選択し，行動としてyを選択した場合の今回の自己の宣言として選ばれる選択肢である．次では，$p(x, y)$をp_{xy}と表す．例えば，p_{CD}は，1回前の成分ゲームにおいて，相手が宣言Cと行動Dを選択した場合の今回の宣言を表す．

2. 行動戦略

   相手の1回前の宣言と行動の両方の選択と，さらに今回の相手の宣言に基づき，今回の行動として，CとDのどちらを選ぶかを定めたルールである．すなわち，行動戦略は，関数
   $$a: S\times S\times S\rightarrow S$$
   として表される．ここで，$a(x, y, z)$は，1回前の成分ゲームにおいて，相手が宣言と行動として，それぞれxをyを選択し，さらに今回の成分ゲームにおいて相手が宣言としてzを選択した場合の今回の自己の行動として選ばれる選択肢である．以下では，$a(x\ y, z)$をa_{xyz}と表す．例えば，a_{CDC}は，1回前の成分ゲームにおいて，相手が宣言C，行動Dを選択し，さらに今回の成分ゲームにおいて相手がCと宣言した場合の今回の行動を表す．

ゲーム開始時の初回は，宣言と行動のいずれについても「協調（C)」を選択するものとし，これはすべての戦略について共通であるとする．

すべての宣言戦略からなる集合は，$S^{S\times S}$と表されるので，その個数は$2^4=16$とおりである．同様に，行動戦略については，$2^8=256$とおり存在する．これら二つの戦略は独立であるので，すべての戦略は$2^{12}=4096$とおり存在することになる．

次では，戦略を次のようなCかDの12文字のビット列によるコードとして表現する．
$$b_0b_1b_2b_3b_4b_5b_6b_7b_8b_9b_{10}b_{11}$$
表2に，各文字b_iの意味を示す．表からわかるとお

表2 戦略ビット列のコーディング規則

b_0	p_{CC}
b_1	p_{CD}
b_2	p_{DC}
b_3	p_{DD}

b_4	a_{CCC}
b_5	a_{CCD}
b_6	a_{CDC}
b_7	a_{CDD}
b_8	a_{DCC}
b_9	a_{DCD}
b_{10}	a_{DDC}
b_{11}	a_{DDD}

り，b_0 から b_3 までの最初の4ビットは宣言戦略を表し，b_4 から右の8ビットは行動戦略を表す．

本稿では，4096とおりの戦略の中で，コミットメントの履行に焦点を当てるため，特に次の三つの戦略に着目することにする．

1. Honest 戦略（以下，H 戦略とする）

CDCDCCDDCCDD

行動戦略は，1回前の相手の行動と同一の選択肢を取る戦略である．宣言も，相手の1回前の行動における選択肢をそのまま用いている．したがって，宣言と行動が必ず一致する．つまり，相手の行動のみに着目し通常の tft 戦略と同様に決定した行動をそのまま宣言する．言い替えると，この戦略を用いるプレイヤは，毎回コミットメントを確実に履行する，「正直な」プレイヤである．

2. Liar 戦略（以下，L 戦略とする）

CCCCCCDDCCDD

最初の4ビットがすべてCになっていることから，必ず「協調（C）」すると宣言するが，実際の行動においては，1回前の相手の行動と同一の選択肢を取る戦略である．つまり，宣言については常に協調すると見せかけて，行動に関しては tft 戦略と同様，1回前の相手の行動をそのまま真似るため，宣言通りには行動せずに時には「裏切る（D）」こともある．つまり，コミットメントを必ずしも履行しない，「嘘をつく」戦略である．

3. Clever 戦略（以下，C 戦略とする）

CDCDCDDDDDCD

宣言戦略は，Honest 戦略と同じである．行動戦略は，1回前の対戦において対戦相手の宣言と行動が一致している場合（a_{CCC}, a_{CCD}, a_{DDC}, a_{DDD}）には，相手の今回の宣言で提示された選択肢を自分の行動として選択するが，一致していない場合は「裏切る（D）」つまりこのプレイヤは，前回の対戦結果から，宣言と行動の一致/不一致に基づいて相手の宣言の信頼性を判断し，信頼できる場合（一致している場合）には相手の宣言に合わせた行動を選択するが，信頼できない場合（一致しない場合）には「裏切る（D）」．

C 戦略を取るプレイヤは，宣言の段階では，H 戦略と同様，相手の前回の行動に対する tft ルール（相手の行動の手をそのまま真似る）に従って行動するというコミットメントを行うが，今回の相手の宣言に基づいて，自分の行動を変えてくる．必ずしも宣言通りの行動を取らないという意味では，L 戦略と同様，コミットメントを履行しないプレイヤである．しかし，前回のゲームで相手プレイヤが「裏切る（D）」という行動をコミットメント通りに行っている場合には，今回のゲームにおいて相手が「協調（C）」するというコミットメントを信用し，「裏切る」というコミットメントを履行せず協調で応じる．相手が信頼できると見るや，宣言Dを取り下げて行動Cで応じて互いに協調するという点で，L 戦略のコミットメントの不履行とはその意味合いが異なる．また，コミットメントを破棄した（宣言と行動の一致しない）プレイヤに対しては，宣言Cを表明している場合でも行動Dで報復しコミットメントを破棄する．しかし，自分から先にコミットメントを破棄することはやらない．また，H 戦略と L 戦略は，いずれも今回の対戦における相手の宣言を意思決定の材料にに用いない戦略であるのに対して，C 戦略は相手の宣言を情報として利用する「巧みな」プレイヤである．

以降では，この三つの戦略を中心に行った，二つの計算機実験の結果について報告する．

4. ノイズを伴う繰り返しゲーム

容易にわかるように，これらの三つの戦略に従うプレイヤを対戦させると，いかなる組合せについても(同一の戦略同士の場合も含めて)，宣言と行動の双方において常に協調し続けるプレイが得られる．つまり，どのプレイヤも，宣言と行動にCの列を取り続ける．

通常の繰り返しゲームにおいて，プレイヤが本来の戦略で定まる選択肢とは別な選択肢を確率的に取る，いわゆるノイズの入った繰り返しゲームを考える場合がある．実際のゲームにおいても，まれに発生する判断のエラーや，自己の選択が相手に正しく伝わらない事態が発生することをモデル化したものである．ノイ

ズが加わった場合に，互いに多くの協調を引き出せる戦略であるかどうかというのは，繰り返しゲームの戦略評価の際の観点の一つとして用いられる．本稿でも，上述した三つの戦略について，確率的に本来の戦略とは違った手を取る戦略を設定することでプレイにノイズを加え，コンピュータシミュレーションを行った．

毎回の成分ゲームにおいて，各戦略のプレイヤには確率 $p(0 \leq p \leq 1)$ でそれぞれの本来の戦略に基づく手とは異なる手を取らせる．ただし，このようなノイズを加える場合でも，初回は，宣言と行動の両者で互いに協調しあうことや，H戦略はコミットメントを守る（宣言と行動が一致する）といったそれぞれの戦略の特徴的な関係は保ったうえで，ノイズを加えるものとする．これは，各戦略を特徴づける重要な部分に関しては，判断ミスや伝達ミスが発生しないように注意してプレイするということを反映したものと解釈できる．例えば，正直なプレイヤはコミットメントの遵守に厳格に望むというわけである．具体的には次のように各戦略のプレイにノイズを加える．

1. H戦略

   常に宣言と行動が一致しているものとする．2回目以降の成分ゲームにおいて，確率 p で宣言と行動の両方をH戦略ルールによる選択肢から反転させる．

2. L戦略

   宣言は常に C を取り続けるものとする．つまり，ノイズは行動にのみ加わる．2回目以降の成分ゲームにおいて，行動のみが確率 p でL戦略のルールに従わないものとする．

3. C戦略

   宣言は上述したL戦略と同様，ノイズの影響を受けないものとする．さらに，前回の対戦で相手の言動が一致しない場合には，必ず報復(D)に出る．ノイズが加わるのは，2回目以降の成分ゲームにおいて，前回の対戦で相手の宣言と行動が一致した場合に限る．そのとき，行動のみが確率 p でC戦略ルールに従わないものとする．

シミュレーションは，三つの戦略のプレイヤの総当たり戦（同じ戦略同士の対戦も含む）を行い，その総成績を比較することとする．その際，各戦略のルールから逸脱する確率 p は，3戦略ともに共通とする．1回の対戦では，100回の繰り返しを1セットとし，それを10セット行いその合計利得を求めた．1回の囚人のジレンマの利得行列は，表1にあるものを用いた．

表3　総当たり戦の結果（$p=0.2$）

自分＼相手	H	L	C	総利得
H	3295	3335	3154	9784
L	3255	3295	2771	9321
C	3114	3156	2555	8825

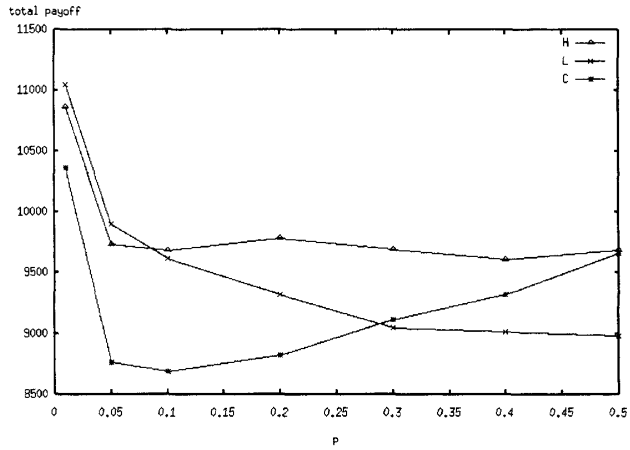

図1　総利得の変化

4.1　シミュレーションの結果

表3に示すのは，$p=0.2$ とした場合の総当たり戦の結果である．表には，二つの戦略を対戦させた場合のそれぞれの戦略のプレイヤが得た100回の繰り返し10セット分の利得の合計と，それらの総利得を示してある．

また，本来の戦略ルールから逸脱する確率 p を，0.01～0.5の範囲で変化させた場合の三つの戦略の総利得の変化を，図1に示す．

4.2　ディスカッション

図1が示す通り，H戦略は，p の値にかかわらず概していい成績を示しているが，L戦略とC戦略は，p の変化によってその優位性が変化している．また，$p=0.2$ の場合を見ると，表3からわかるように，H戦略は対戦する相手の戦略にかかわらず，平均して高い利得を得ている．H戦略とL戦略は，ともに相手の宣言を意思決定に利用しない戦略である．したがって，これら二つの戦略の中で対戦する限りにおいては，宣言と行動との間の関連性はゲームの結果に影響を与えないため，宣言での選択は意味を持たない．つまり，これらの2戦略内で対戦する限り，両者ともコミットメントを含まない通常の繰り返し囚人のジレンマのtft戦略と本質的に同じである．よって，H戦略は自分と対戦した場合も，L戦略と対戦した場合も同程度の利得を獲得できる．

表4　総当たり戦の結果（$p=0.5$）

自分＼相手	H	L	C	総利得
H	3263	3133	3288	9684
L	3393	3263	2323	8979
C	3233	3708	2717	9658

また，H戦略は常に宣言と行動とが一致する戦略であるので，C戦略のプレイヤから多くの協調を引き出すことができる．確率pでのノイズは，宣言と行動の一致に関しては影響を与えない設定になっているため，この関係は，pの値によらない．結果として，対戦相手や不確実性の変化（ノイズの程度）によらず，平均的に高い利得を獲得することに成功している．

L戦略は，pが小さい環境においてはいい成績を示すが，pの増加に伴い総利得は減少する．表4に，$p=0.5$のときの総当たり戦の結果を示す．表からわかる通り，pが大きい環境においては，L戦略は宣言と行動とが乖離する傾向が強まり，C戦略のプレイヤからの協調が得られない．逆に，戦略Lから提示される協調（C）の手に対して，C戦略は裏切り（D）で報復することが多くなるため，C戦略はL戦略から多くの利得を獲得することに成功している．ただし，C戦略自身でも宣言と行動とが乖離する傾向が強まるため，C戦略同士での対戦では高い利得を稼ぐことはできなくなる．

結局，コミットメントを確実に履行する戦略（H）が総じて有利であり，意図的に嘘をつきコミットメントを破棄する戦略（L）が有利ではないことが示されている．また，相手がコミットメントを履行したか否かに基づき，相手のコミットメントを利用して選択を行う戦略（C）も，総じて有利な戦略というわけではない．ただし，コミットメントを履行しないプレイヤに対し，確実に報復するC戦略プレイヤは，意図的に「嘘をつくプレイヤ」が有利になることを防ぐ役割を果たしている．また，コミットメントを確実に履行するプレイヤは，環境の不確実性の大小にかかわらず，コミットメントの信頼性に答えるC戦略プレイヤの存在によって，必ずしも情報を利用して巧みに行動することなく，利益を上げることができる．

5．進化ゲームアプローチによる3戦略の評価

現実の状況では，上記のシミュレーションのように小数の戦略に閉じた環境で競争することは少ない．ここでは，多種多様の戦略がひしめきあう雑多な集団の中で，H/L/Cの3戦略が生き残ることができるかどうか（Axelrodのいう「たくましさ」[3]）に着目してみよう．これらの三つの戦略を含む4096通りすべての戦略プールの中から，自戦略との対戦も含めた総当たり戦を行い，総得点の優れた戦略をある規則に従って選択して残していくという淘汰メカニズムを組み込んだプログラムを作成した．プログラムの処理の流れは，次の通りである．

ステップ1：4096個の戦略（個体）を初期の戦略プールとする．

ステップ2：戦略プール中のすべて個体について，自分との対戦も含めて総当たり戦で，繰り返しゲームの対戦を行う．繰り返しゲームにおける繰り返し回数Nは，パラメータとしてあらかじめ与えておく．

ステップ3：総当たり戦の各対戦で得られた利得を合計した総利得を各個体について求め，個体間に差がなければ（つまりすべての個体について総利得が同じならば），シミュレーションを終了する．総利得に個体差がある場合は，ステップ4へ進む．

ステップ4：総利得をもとに，後述の淘汰メカニズムに従って戦略（個体）の選択を行い，残った戦略集団を新たな戦略プールとして設定し，ステップ2へ戻る．

淘汰メカニズムについては，様々な方法が考えられるが，今回は次のような単純なものを用いた．現在の戦略プールに当たる戦略集合を，

$$S=\{s_1, s_2, \cdots, s_n\}$$

と表したとする．また，戦略s_iの総当たり戦による総利得を$t(s_i)$と表すとする．$t(s_i)(i=1, 2, \cdots, n)$の最大値と最小値をそれぞれ，$t_{\max}, t_{\min}$とする．次期の戦略プールとなる戦略集合$S'$は，次のように定められる．

$$S'=\{s_i | t(s_i) \geq p \cdot t_{\max}+(1-p) \cdot t_{\min}\}$$

ここで，$p(0 \leq p \leq 1)$は，淘汰圧に相当するパラメータであり，区間$[t_{\min}, t_{\max}]$を$p:(1-p)$に内分する点を求め，その値以上の総利得を稼いだ戦略は，次期の戦略プールに残す．pが大きいと，次期戦略プールに残す戦略の割合が減るため，pが大きければ大きいほど淘汰圧が高いこと，つまり次期戦略プールに残ることのできる戦略の割合が減ることを表している．

繰り返しゲームの反復回数Nの変化は，結果に大きな影響は与えなかった．次節では，$N=100$とした場合の結果を示す．

表5　残存した戦略数とそのビットパターン

p	総淘汰回数	残存戦略数	共通ビットパターン
0.9	4	1	DDDCDDDDDDDD
0.8	4	1	DCCDDCDDCDDD
0.7	4	20	C***C*D**D*D
0.6	4	65	C***C*D*****
0.5	6	121	C***C*******
0.4	7	270	C***C*******
0.3	11	361	C***C*******
0.2	21	464	C***C*******
0.1	42	397	C***C*D*****

表6　三つの戦略が戦略プールから退出した淘汰回数

p	L戦略	H戦略	C戦略
0.9	1	1	1
0.8	1	1	1
0.7	1	1	2
0.6	1	-	-
0.5	2	-	-
0.4	2	-	-
0.3	-	-	-
0.2	-	-	-
0.1	-	-	-

6. 実験結果

表5に，プログラムの実行結果を示す．表は，淘汰圧パラメータpの設定それぞれについて，最終的に総利得の優劣がなくなり残存した戦略数，それまでに要した総淘汰回数（上述の淘汰メカニズムのステップ2〜ステップ4が何回実行されたか），残存した戦略に共通に見られるビットパターンを示している．例えば，p=0.9のときには，淘汰回数4回で最終的に一つの戦略が生き残り，その戦略のビットパターンが，DDDCDDDDDDDDであったことを示している．また，p=0.7のときには，残存する戦略数が20となり，それらに見られる共通のビットパターンは，C＊＊＊C＊D＊＊D＊Dである．“＊”は，残存した戦略の該当するビットに，CとDの両方が現れていることを意味する．

表6には，既述のH/L/Cの三つの戦略が最終的に戦略プールに残存したか否か，また残存しなかった場合に何回目の淘汰で戦略プールから退出したかを示す．表中の“−”は，該当する戦略が最後まで残存したことを表す．例えば，p=0.9においては，H/L/Cの3戦略とも，第1回目の淘汰において戦略プールから退出していることを，また，p=0.6では，L戦略は第1回目の淘汰において戦略プールから退出しているが，他の二つの戦略は最終的に総利得の優劣がなくなるまで戦略プールに残ったこと，つまり4回（表5参照）の淘汰で排除されずに最後まで戦略プールに残ったことを示している．

7. ディスカッション

表5からわかる通り，淘汰圧pが大きい環境では，残存する戦略バリエーションは少なく，裏切る（D）という手を多用する戦略が生き残る．しかし，pが小さくなるに従って，残存する戦略のバリエーションは大きくなり，$p \leq 0.7$では，H/L/Cと共通な次のビットパターンが見られるようになる．

C＊＊＊C＊＊＊＊＊＊＊

この二つのCビットは，

$$p_{CC}=C, \quad a_{CCC}=C$$

であり，それぞれ，

1. 相手が前回の対戦で，宣言Cと行動Cを選択した場合は，今回の自己の宣言はCを選ぶこと
2. 相手が前回の対戦で，宣言Cと行動Cを選択し，さらに今回の宣言でCを選択した場合は，今回の自己の行動はCを選ぶこと

を表している．前回の対戦で相手が「協調」のコミットメントを履行した場合には，自分は「協調」のコミットメントを表明し，しかも，今回の対戦で相手が「協調」のコミットメントを表明している場合には，自分も今回の「協調」のコミットメントを履行し「協調」という行動をとるという戦略を表している．「協調」のコミットメントの履行に対して，同じく「協調」のコミットメントを確実に履行する戦略が，戦略プールに共通に生き残っている．しかし，「裏切り」コミットメント（つまり，「裏切り」の約束を守る）に関しては，共通なパターンは見られない．つまり，生き残りのためには，相手の「協調」のコミットメントの履行は情報として重要だが，「裏切り」についてはそれほど重要ではないということである．しかも，これらの戦略は「協調」のコミットメントの履行に対しては「協調」で応じてくれる．その結果，「協調」のコミットメントを確実に履行することで，多くの戦略から協調を引き出すことができるような環境が作られていることがわかる．「協調」のコミットメントを意図的に破棄することは，有利な戦略ではない．

表6からわかる通り，淘汰圧の大きさを表すパラメータpが十分小ければ，H/L/Cの3戦略とも戦略プールに残存できていることがわかる．ただし，Lに比べ，HとCはよりpの大きい，つまりより競争の厳しい環境下でも生き残ることのできる，より「たくましい」戦略であるということが結論づけられる．淘汰圧との関係でも，意図的にコミットメントを破棄する戦略Lは得策ではないことがわかる．

また，淘汰圧と「たくましさ」との関係では，HとCとの間に大きな差は見られない．相手の宣言から得られる情報を利用する「巧みな」戦略は，巧みさゆえに有利というわけではなさそうである．

8. おわりに

拡張された繰り返し囚人のジレンマにおいて，相手の1回前の選択に基づいて意思決定を行う戦略について，三つの戦略を中心にそれらの優位性を検討した．その結果，コミットメントを破棄する戦略に比べて，コミットメントを履行する戦略の方が有利であるという結果が得られた．説苑（ぜいえん）に「巧詐は拙誠に如かず」という訓話がある．巧みに人をあざむくよりも，拙くとも誠意をもって行動すべしというわけである．

今回は，戦略プールが淘汰後に安定した状態に到達した後に，突然変異による新たな戦略が侵入した場合のこれらの戦略安定性については検討していない．また，ランダムに生成した初期個体に対して，突然変位や戦略の交差といった操作を行った場合に，どのような戦略に進化するのかについては，現時点ではまだシャープな結果は得られていない．これらについては，今後の課題としたい．

参考文献

[1] R. A. Spinello: “Case Studies in Information Technology Ethics”, Prentice Hall, 2002.

[2] 岡田章:「ゲーム理論」, 有斐閣, 1996.

[3] アクセルロッド著: 松田裕之訳,「つきあい方の科学」, ミネルヴァ書房, 1998.