繰り返しゲーム

２０１９年１２月２日
ゲーム理論入門　第８回講義

荒木一法

繰り返しゲーム

- これまでに扱ったゲームの事例（値下げ競争ゲームなど）が繰り返しプレイされる状況全体を一つのゲームとして扱う。このとき繰り返される元のゲームは全体ゲームと区別して、ステージ・ゲーム(stage game)と呼ばれる。
  - 有限回繰り返しの場合：部分ゲーム完全均衡は一つに絞られるが、実際に観察される現象と一致しない。
  - 無限回繰り返しの場合：部分ゲーム完全均衡は数多く存在（フォーク定理） ⇒ 絞込みの失敗！

第7章　繰り返しゲーム

1. 繰り返し囚人のジレンマ
2. フォーク定理
3. 利己的行動と利他的行動
4. 不完全情報とシグナル

１　繰り返し囚人のジレンマ

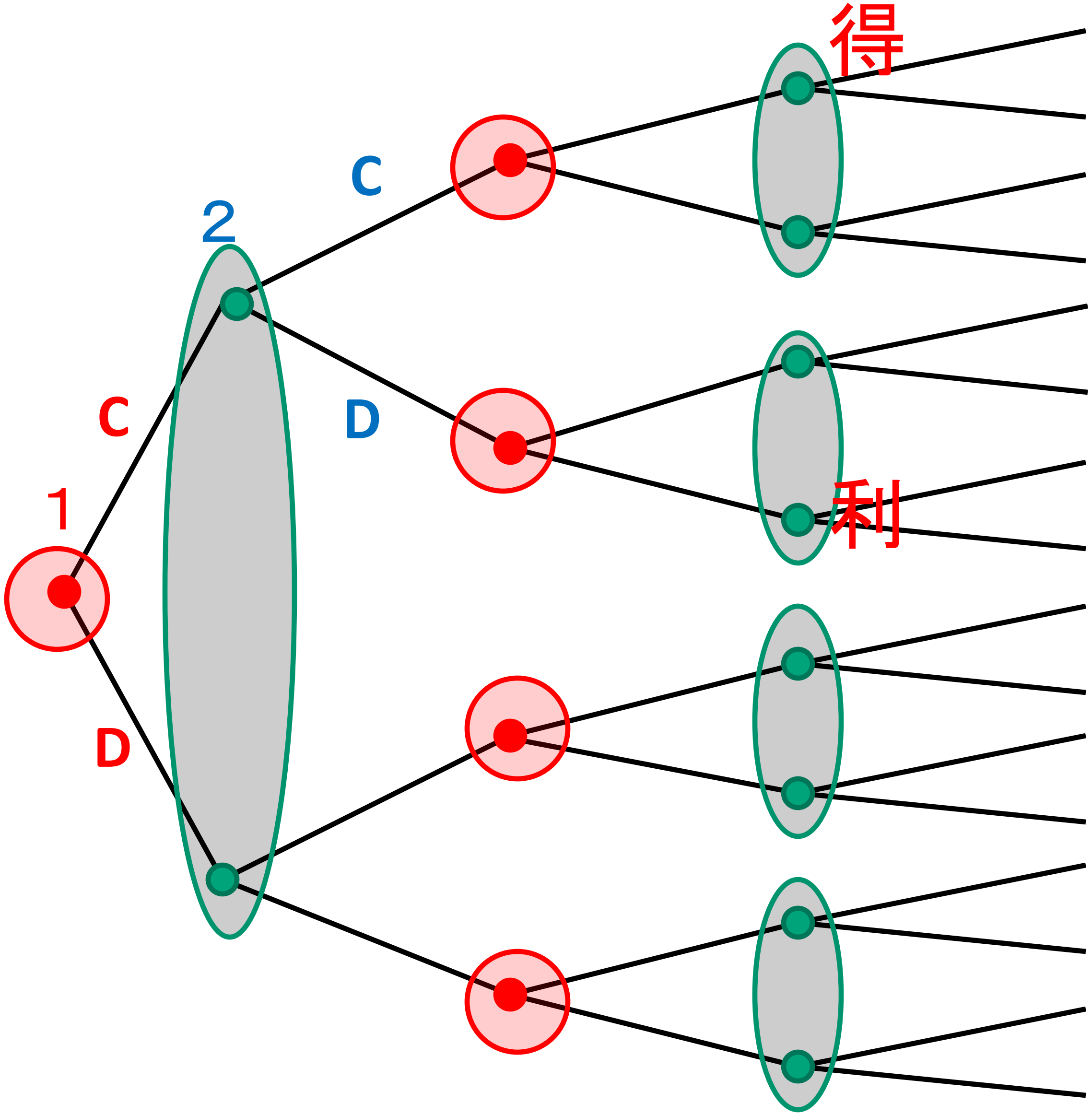
繰り返しゲームは、同じゲーム（例えば囚人のジレンマ）が同じプレーヤーについて複数回プレイされるゲーム。囚人のジレンマを２回繰り返すことが１つのゲームとなる。このとき、繰り返される対象となるゲームは、全体ゲームの一部となり、ステージ・ゲーム（Stage Game）と呼ばれる。

以下では、完全情報の繰り返しゲーム（プレーヤーが過去のゲーム展開を完全に把握している）を考察対象とする。

囚人のジレンマ

1 \ 2	協力（C）	裏切り（D）
協力（C）	5, 5	0, 7
裏切り（D）	7, 0	1, 1

囚人のジレンマが2回繰り返される場合



繰り返しゲームの戦略

囚人のジレンマが２回繰り返されるゲームの戦略の数は、それぞれのプレーヤーに$2^5 = 32$個ある。（意思決定をする情報集合の数がA，Bともに5つあるため。）列挙すると

(C,C,C,C,C),(C,C,C,C,D),(C,C,C,D,C), (C,C,D,C,C),(C,D,C,C,C)

(D,C,C,C,C),(C,C,C,D,D),(C,C,D,C,D),(C,D,C,C,D),(D,C,C,C,D)

(C,C,D,D,C),(C,D,C,D,C),(D,C,C,D,C),(C,D,D,C,C),(D,C,D,C,C)

(D,D,C,C,C),(C,C,D,D,D),(C,D,C,D,D),(C,D,D,C,D),(C,D,D,D,C)

(D,C,C,D,D),(D,C,D,C,D),(D,C,D,D,C),(D,D,C,C,D),(D,D,C,D,C)

(D,D,C,C,D),(D,D,C,D,C),(D,D,D,C,C),(D,C,D,D,D),(D,D,C,D,D)

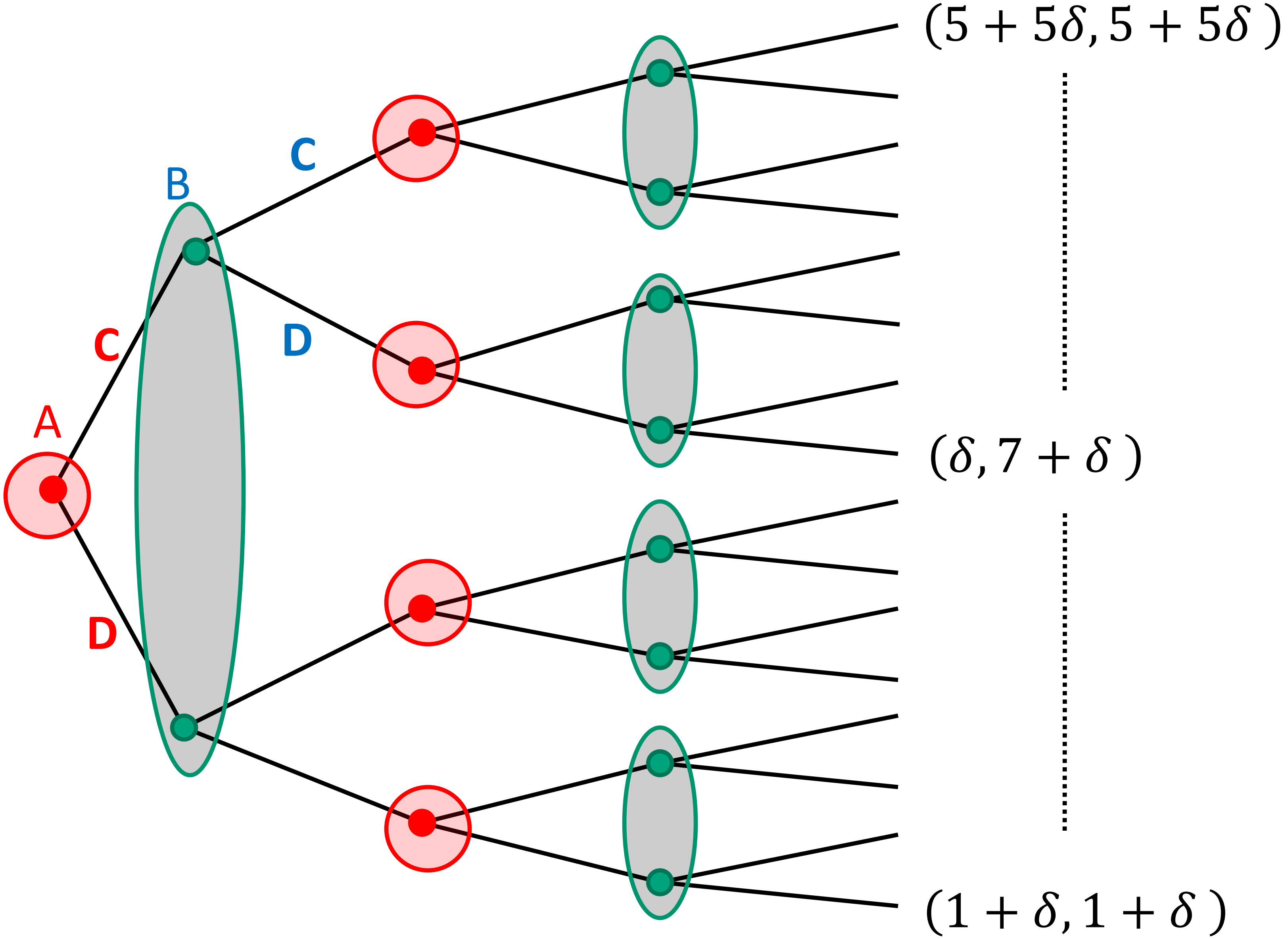
(D,D,D,C,D),(D,D,D,D,D)

繰り返しゲームの利得の定義

- 一番簡単なのは「単純和」。例えば1回目の利得が5で2回目が1なら5＋1＝6で6を繰り返しゲームの利得として定義することもできる。
- しかし、「単純和」による定義は無限回繰り返しゲームには拡張できない。1を無限回たしても、5を無限回たしても、ともに「無限大∞」で区別できない。
- ここでは、時間的におくれる利得には割引因子をかけた上で「単純和」をとる方法を採用

$$\delta : \text{割引因子} \quad (0 < \delta < 1)$$

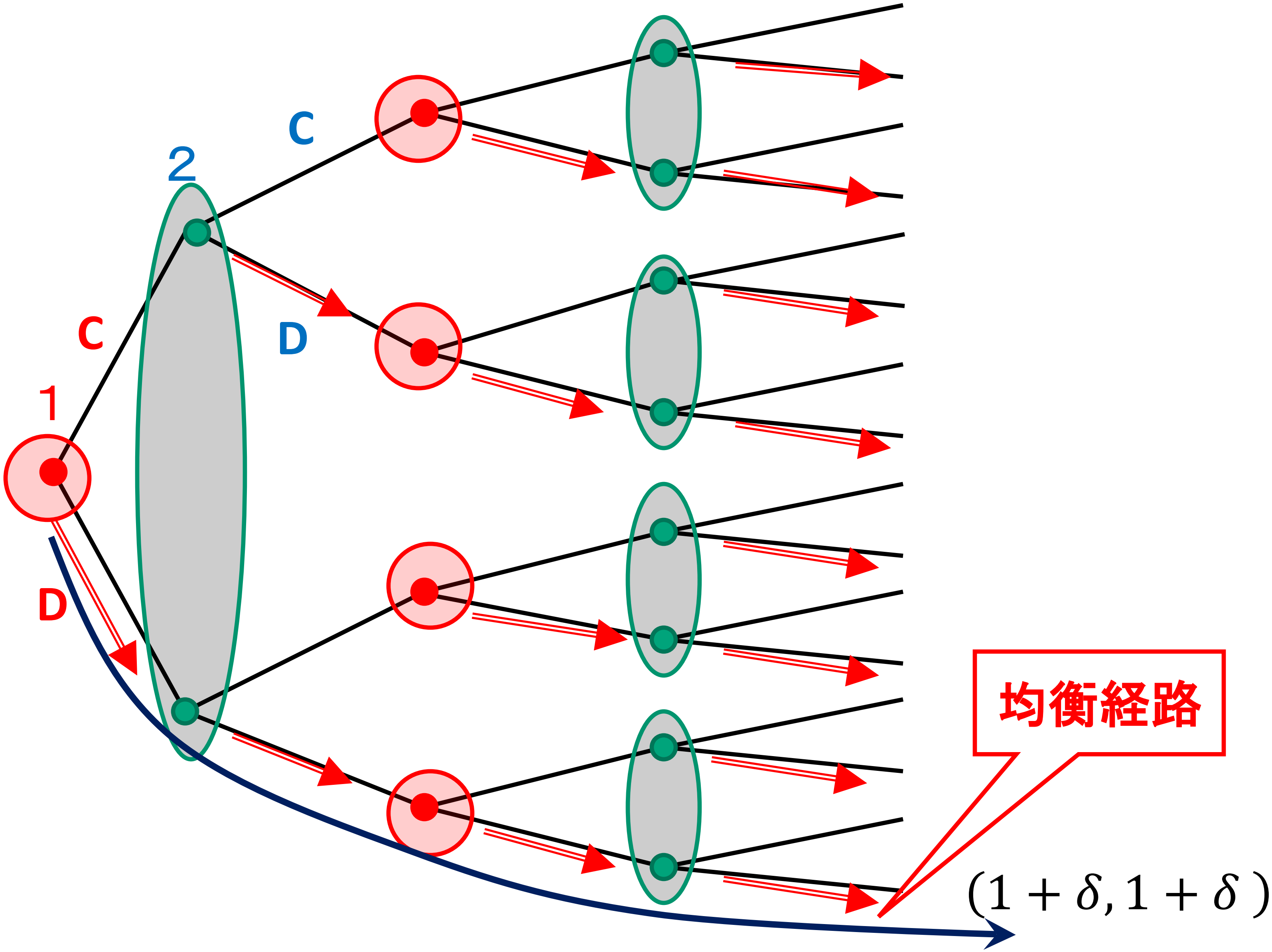
２回繰り返される場合の利得



部分ゲーム完全均衡

- ２回目の意思決定を行うプレーヤー2にとってはDが支配戦略
- したがって、２回目の意思決定を行うプレーヤー１はDを選ぶ
- ２回目でAがDを選ぶことを予想するとプレーヤー2は1回目でDを選ぶ
- プレーヤー２が１回目からDを選ぶことを予想するとプレーヤー１は１回目からDを選ぶ
- 繰り返し回数が有限である限り、何度繰り返しても同じ理由で、部分ゲーム完全均衡では、最初から両者ともにDを選択し続ける。

部分ゲーム完全均衡の均衡経路



戦略の数

完全情報の繰り返しゲームの場合、繰り返しの回数が増えると戦略の数は指数関数的に増加する。例えば、５回繰り返しゲームの戦略の数は、ゲームの展開（履歴）パターンは１回目は1、２回目は4、３回目は16、４回目は64、５回目は256通りとなるので

$$2^{(1+4+16+64+256)} = 2^{341} =?$$

無限回繰り返しゲームの戦略と利得

- 戦略：完全なゲームプランとして、すべての情報集合で何をするかを予め決めた計画。情報集合の数が無限大の場合は無限個の戦略が存在
- 利得：将来の利得を1期あたりδで割り引いた値の総和を基準化して表すこととする。

$$a_1 + a_2\delta + a_3\delta^2 + \cdots + a_n\delta^{n-1} + \cdots$$

$(1-\delta)$をかけて基準化

$$(1-\delta)\left(a_1 + a_2\delta + a_3\delta^2 + \cdots + a_n\delta^{n-1} + \cdots\right)$$

基準化の根拠

各回の利得がrで一定とした場合の、無限回繰り返しゲームの利得の現在価値の総和は

$$A = r + r\delta + r\delta^{2} + \cdots + r\delta^{n-1} + \cdots$$

両辺にδを乗じると

$$\delta A = r\delta + r\delta^{2} + r\delta^{3} + \cdots + r\delta^{n} + \cdots$$

$$A - \delta A = (1-\delta)A = r$$

$$A = \frac{r}{1-\delta}$$

Aに$1-\delta$を乗じると「平均的な」利得になる！

無限回繰り返しゲーム

例えば、5回繰り返しゲームの戦略の数は、ゲームの展開（履歴）パターンは1回目は1、2回目は4、3回目は16、4回目は64、5回目は256通りとなるので　$2^{(1+4+16+64+256)}=2^{341}=?$

以下では、膨大な戦略の中で比較的単純な次の4つの戦略に絞って分析する。 ①常にC　②常にD

③トリガー：1回目はCを選び、2回目以降は前回相手がCを選べばC，一度でも相手がDを選べば、その後は相手の行動にかかわらずDを続ける。

④しっぺ返し：1回目はCを選び、2回目以降は、前回の相手の行動と同じ行動を選択する。

繰り返し囚人のジレンマ（$\delta \geq \frac{1}{3}$の場合）

	All C	**All D**	**トリガー**	**しっぺ返し**
All C	5 5	0 7	5 5	5 5
All D	7 0	1 1	$7-6\delta$ δ	$7-6\delta$ δ
トリガー	5 5	δ $7-6\delta$	5 5	5 5
しっぺ返し	5 5	δ $7-6\delta$	5 5	5 5

（トリガー,トリガー）はナッシュ均衡

相手がトリガー戦略をとっているとき、自らもトリガー戦略をとり続ける場合の基準化された利得とある時点でトリガーから逸脱した場合の基準化された利得を比較する。トリガー戦略から得る利得が大きければ、（トリガー、トリガー）はナッシュ均衡！

利得の推移

5, 5, 5, ……..5, 5, 5, …….（トリガー戦略）

逸脱

5, 5, 5, ……..7, 1, 1, …….（逸脱）

逸脱時点でそれ以後の基準化された利得を比較

トリガー:　5　　逸脱: $7(1-\delta)+\delta$

$\delta \geq \frac{1}{3}$　であれば「トリガー」が有利（逸脱する誘因がない）

（しっぺ返し、しっぺ返し）はナッシュ均衡

利得表からは $\delta \geq \frac{1}{3}$ が満たされる場合には、（しっぺ返し、しっぺ返し）もナッシュ均衡となるように見える。しかし、この利得表は4種類の戦略のみが考慮されているため、他の戦略を含めて考えると、ナッシュ均衡となるとは限らない。トリガー戦略の分析と同様、「しっぺ返し」からの逸脱した場合としない場合の利得（Cが続くので基準化された利得は5）を比較し、（しっぺ返し、しっぺ返し）がナッシュ均衡になる条件を求める。

プレーヤー１、２がともに「しっぺ返し」を選択している状態から、ある期(t期)にプレーヤー２のみが一時的に逸脱し、次の期($t+1$期)には「しっぺ返し」に戻るとき両プレーヤーが各期にとる行動は次のようになる。

	$\boldsymbol{t}$	$\boldsymbol{t+1}$	$\boldsymbol{t+2}$	$\boldsymbol{t+3}$	$\boldsymbol{t+4}$
Player 1	C	D	C	D	C
Player 2	D	C	D	C	D

利得の列は（7,0,7,0,7…）となるのでt期以降のプレーヤー２の利得の現在価値は

$$(7+7\delta^2+7\delta^4+\cdots)(1-\delta)=\frac{7}{(1-\delta^2)}(1-\delta)=\frac{7}{1+\delta}$$

この値が逸脱しない場合の基準化された利得5を上回る条件は　$\delta \geq {}^2/_5$

プレーヤー1、2がともに「しっぺ返し」を選択している状態から、ある期(t期)にプレーヤー2のみが一時的に逸脱しDを選択、次の期($t+1$期)にもDを選択するとき両プレーヤーが各期にとる行動は次のようになる。

	$\boldsymbol{t}$	$\boldsymbol{t+1}$	$\boldsymbol{t+2}$	$\boldsymbol{t+3}$	$\boldsymbol{t+4}$
Player 1	C	D	D	D	D
Player 2	D	D	D	D	D

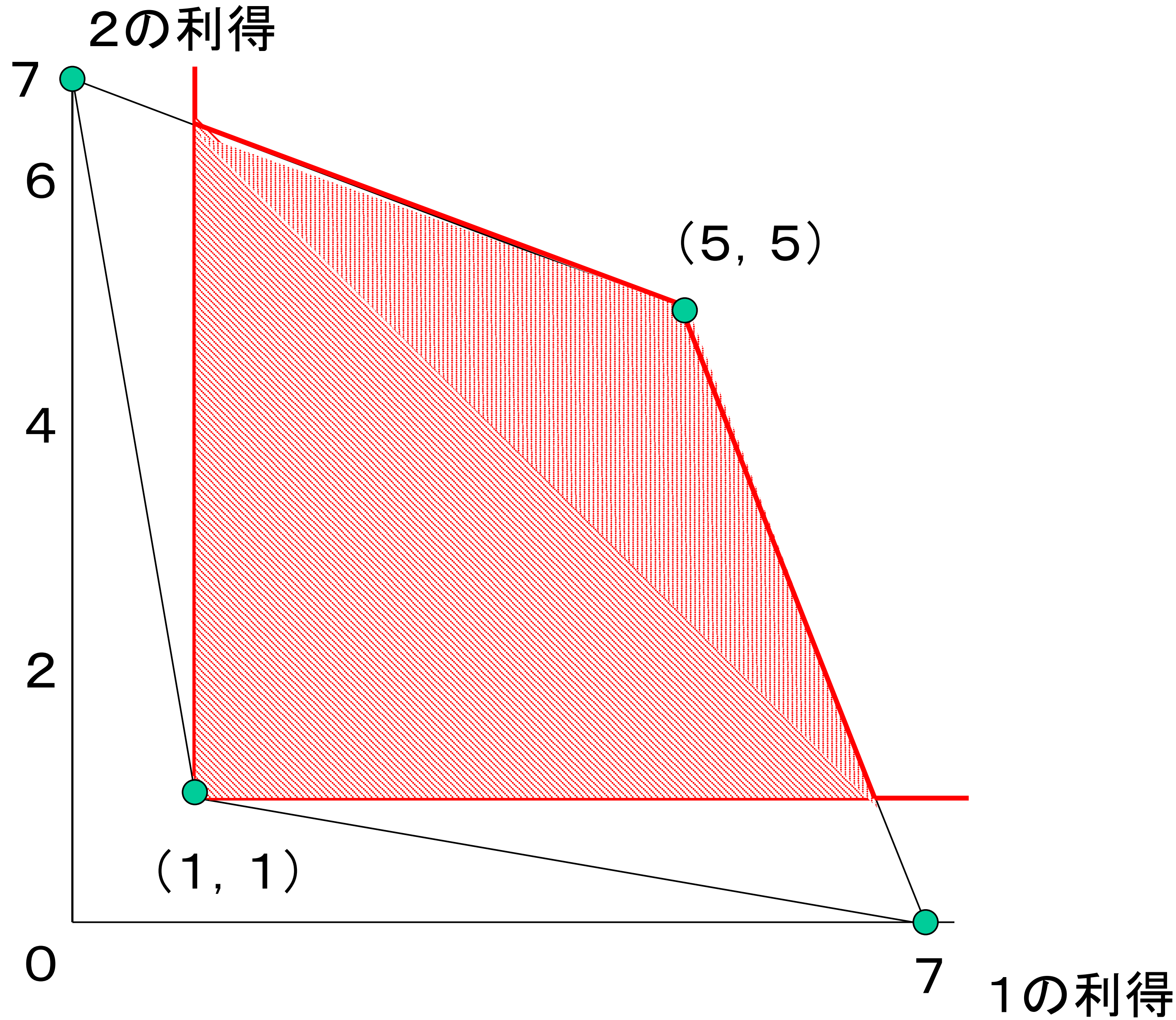
利得の列は(7,1,1,1,1…)となるのでt期以降のプレーヤー2の利得の現在価値は

$$(7+\delta+\delta^2+\delta^3+\cdots)(1-\delta)=7(1-\delta)+\delta$$

この値が逸脱しない場合の基準化された利得5を上回る条件は　$\delta \geq {}^{1}/_{3}$ となる。以上から$\delta \geq {}^{2}/_{5}$ であれば逸脱するインセンティブはない。

２ フォーク定理

- 両者が「トリガー」を取り続ける限りCの選択が続く。
- 他のパターン（例えばCとDを交互にえらぶパターン）からの逸脱にたいして「トリガー」をとる戦略を考えれば、やはり十分に大きなδについて（トリガー, トリガー）がナッシュ均衡になる。このとき基準化された利得は、(3, 3)　となる。
- 同様に様々なパターンからの逸脱に対するトリガーを想定し、δを十分に大きく（1に十分に近く）とれば、次の図の赤の斜線部分（個人合理的な利得）に「ほぼ」対応する基準化された利得の組み合わせが均衡として支持される　⇒　**「フォーク定理」**


2の利得
7
6
4
2
0
(5, 5)
(1, 1)
7
1の利得