ゲーム論Ｉ 第十回

上條　良夫

テキスト紹介

- 渡辺隆裕　2004『図解雑学ゲーム理論』ナツメ社
- 佐藤嘉倫　2008『ゲーム理論:人間と社会の複雑な関係を解く』新曜社
- 武藤滋夫　2001『経済学入門シリーズ　ゲーム理論入門』日本経済新聞社
- 佐々木宏夫　2003『入門ゲーム理論:戦略的思考の科学』日本評論社
- 船木由喜彦　2004『演習ゲーム理論』新世社

講義のキーワード

- 有限回繰り返しゲーム
- 無限回繰り返しゲーム
  - 割引利得の計算
- フォーク定理
- アクセルロッドの実験
  - Tit for tat 戦略

繰り返しゲーム

- 標準形ゲームでは、プレイヤーが同時に意思決定を行い、それぞれが利得を獲得し、それでゲームが終了するような状況を想定した。
- しかしながら、私達が日常的に行うようなゲーム的状況では、意思決定を一度行ったらそれで終わりというときはまれであり、似たようなゲーム的状況に繰り返し直面することのほうがむしろ多いだろう。
- このような同一のゲーム的状況に繰り返し直面する、というような状況も、展開形ゲームを用いることにより分析することが可能である。

- このように、同一の標準形ゲームを繰り返し行うようなゲームを、繰り返しゲームとよぶ。

- 繰り返しゲームは二種類に分類できる。
  - 有限回繰り返しゲーム
  - 無限回繰り返しゲーム

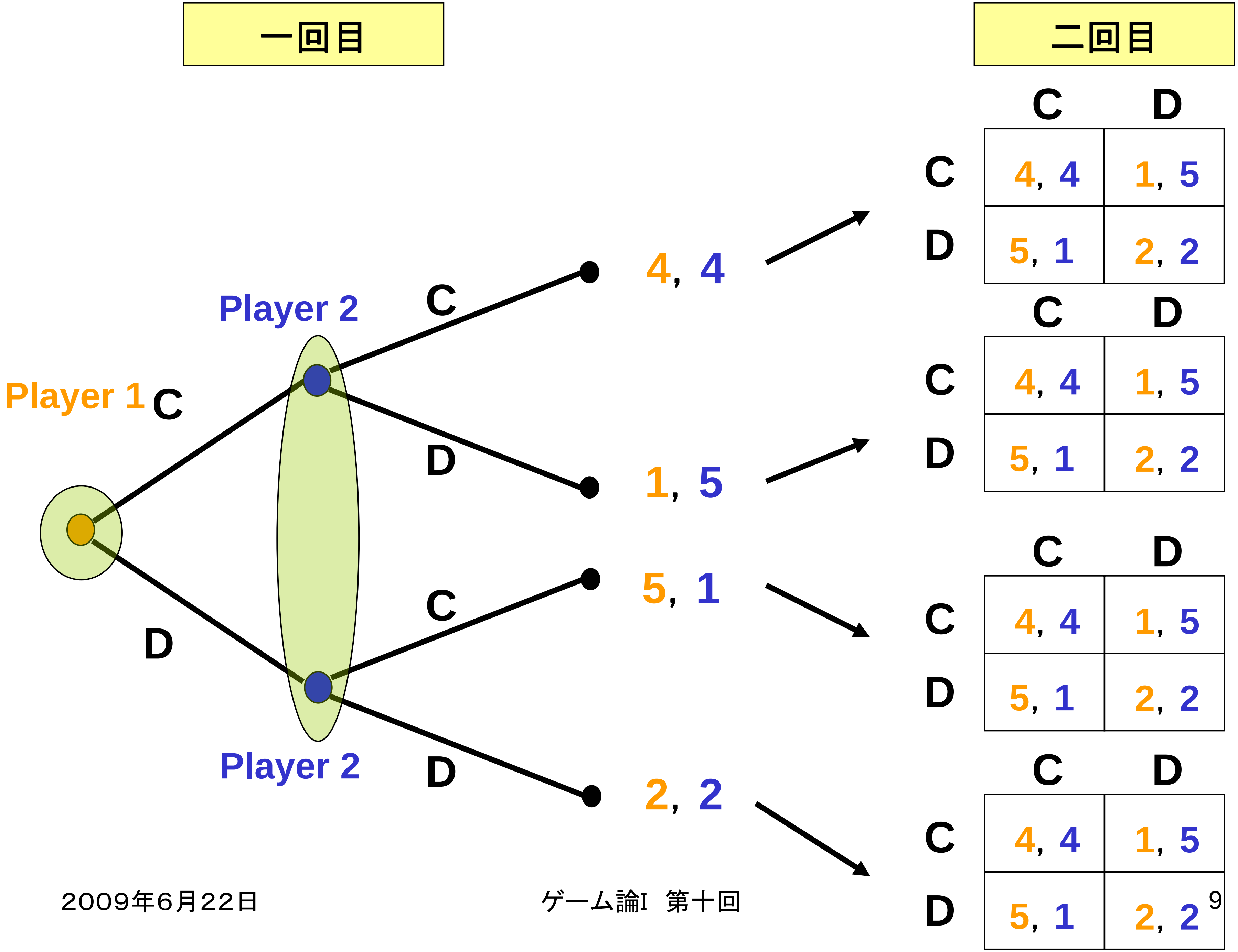
- 繰り返しゲームの分析により我々が知りたいことは、
  - 同一の標準形ゲームを繰り返すことが、いかにプレイヤーのおかれている戦略的状況を変化させ、各回の標準形ゲームでの行動に影響を与えるのか。
  - 囚人のジレンマゲームなどで観察された、残念で少々やっかいな（そして、場合によっては直感に反するかもしれない）結果は、繰り返しゲームを行うという（より日常的な）状況にするとどのように変わるのだろうか。

有限回繰り返しゲーム

- 囚人のジレンマゲームを複数回繰り返し行う状況を考えよう。
- 繰り返し回数には実は結果は依存しないので、2回繰り返しケースを考える。
- 各回の囚人のジレンマを行うときには、プレイヤーは同時かつ独立に意思決定を行うが、過去に相手が取った行動(と、当然自分の行動も)は観察することができる。

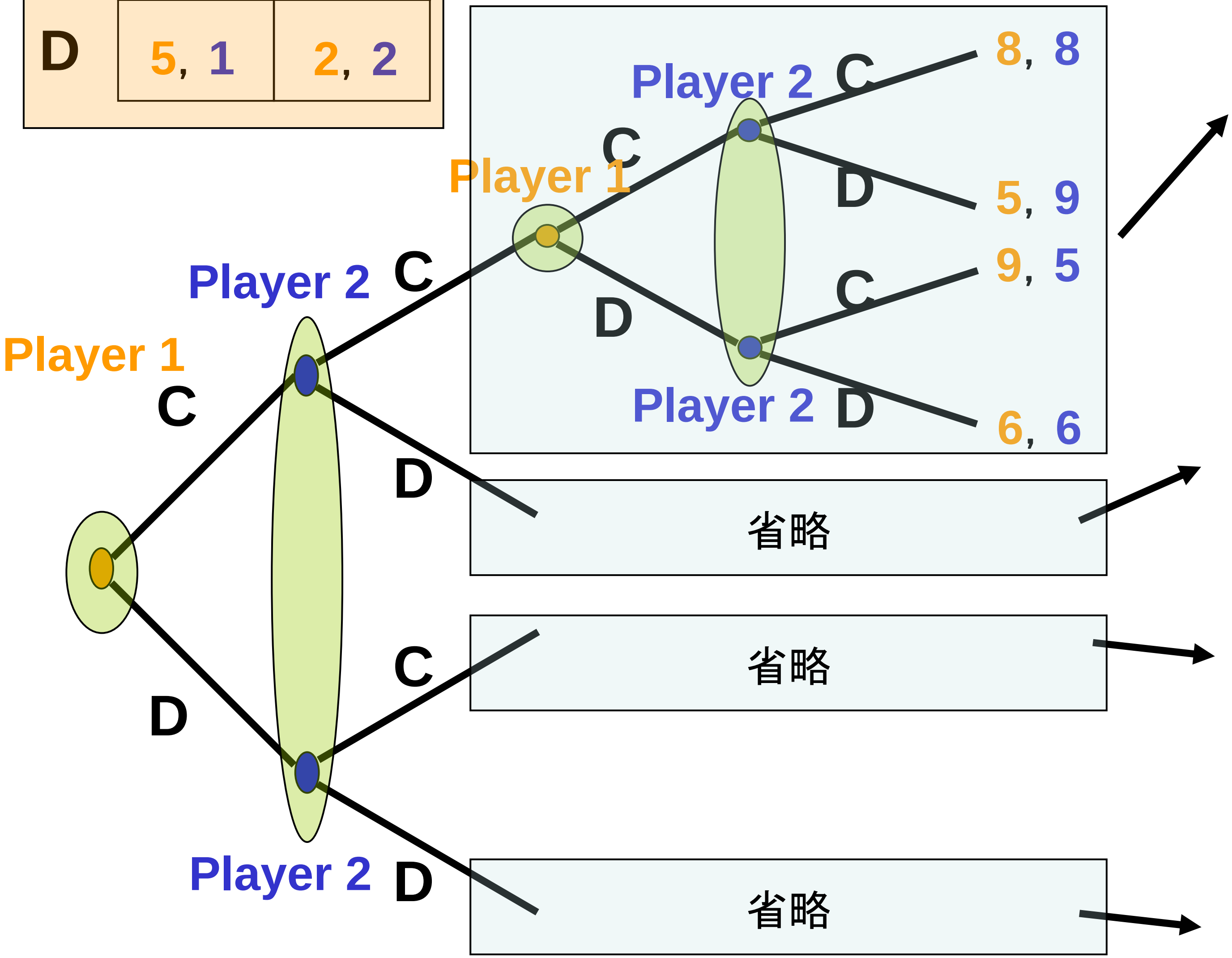
- 以下のような囚人のジレンマゲームを2回繰り返す。
- C (cooperationo) : 協力
- D (defection): 裏切り

	C	D
C	4, 4	1, 5
D	5, 1	2, 2



一回目
二回目
Player 1
Player 2
Player 2
C
D
C
D
C
D
4, 4
1, 5
5, 1
2, 2
C D
C 4, 4 1, 5
D 5, 1 2, 2
C D
C 4, 4 1, 5
D 5, 1 2, 2
C D
C 4, 4 1, 5
D 5, 1 2, 2
C D
C 4, 4 1, 5
D 5, 1 2, 2

	C	D
C	4, 4	1, 5
D	5, 1	2, 2



	C	D
C	8, 8	5, 9
D	9, 5	6, 6

	C	D
C	5, 9	2, 10
D	6, 6	3, 7

	C	D
C	9, 5	6, 6
D	10, 2	7, 3

	C	D
C	6, 6	3, 7
D	7, 3	4, 4

4つの部分ゲームがあるが
すべて Nash 均衡は
(D, D)

つまり、第一ステージの
結果に関わりなく、二回目の
囚人のジレンマゲーム
の結果は (D, D) である

これは、標準形ゲームでは、
プレイヤーの利得に定数を加える
ことにより新しいゲームを作っても
Nash 均衡は変化しない
という性質のため。

(C,C) のあと
の部分ゲーム

	C	D
C	8, 8	5, 9
D	9, 5	$\boxed{6, 6}$

(C,D) のあと
の部分ゲーム

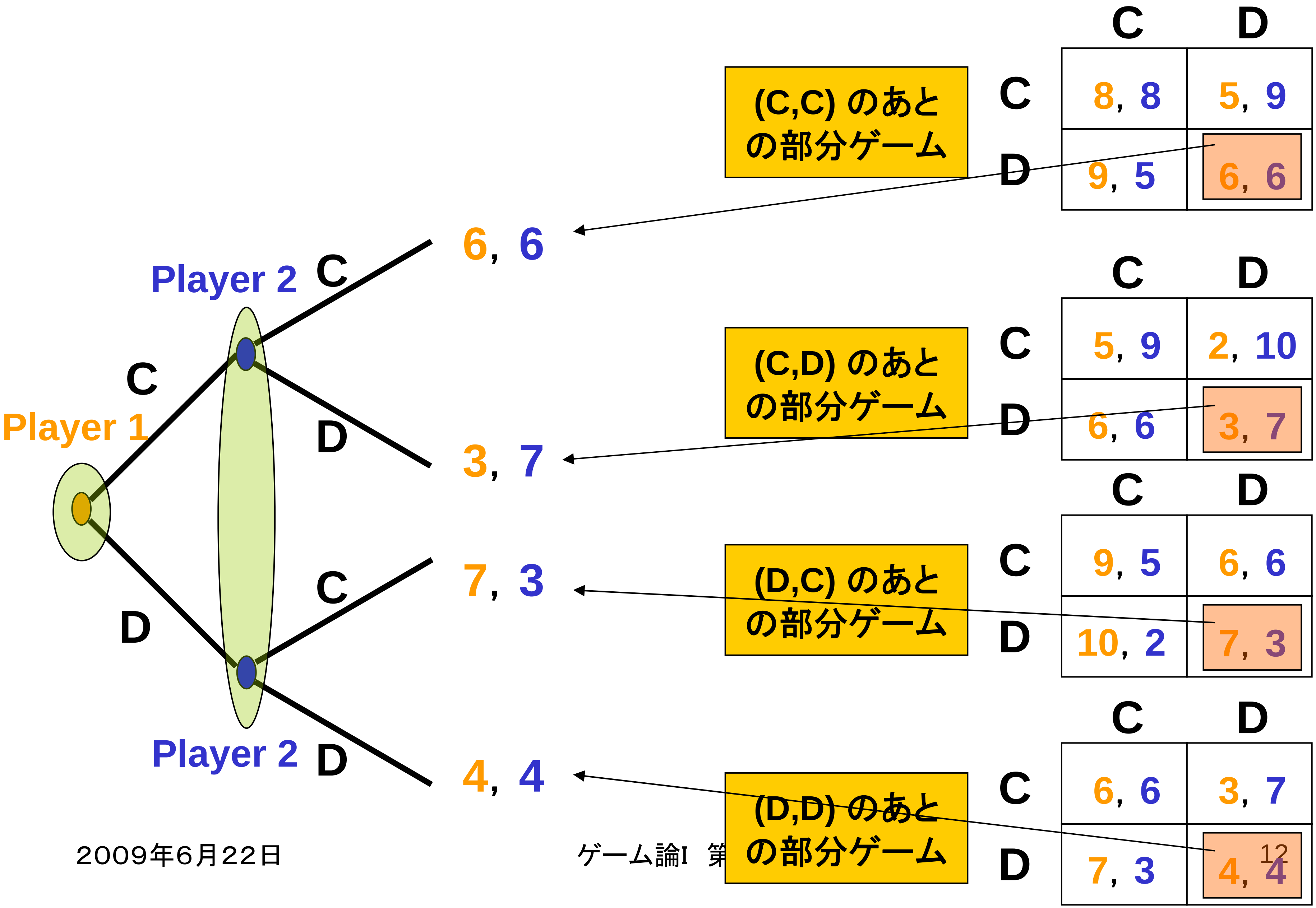
	C	D
C	5, 9	2, 10
D	6, 6	$\boxed{3, 7}$

(D,C) のあと
の部分ゲーム

	C	D
C	9, 5	6, 6
D	10, 2	$\boxed{7, 3}$

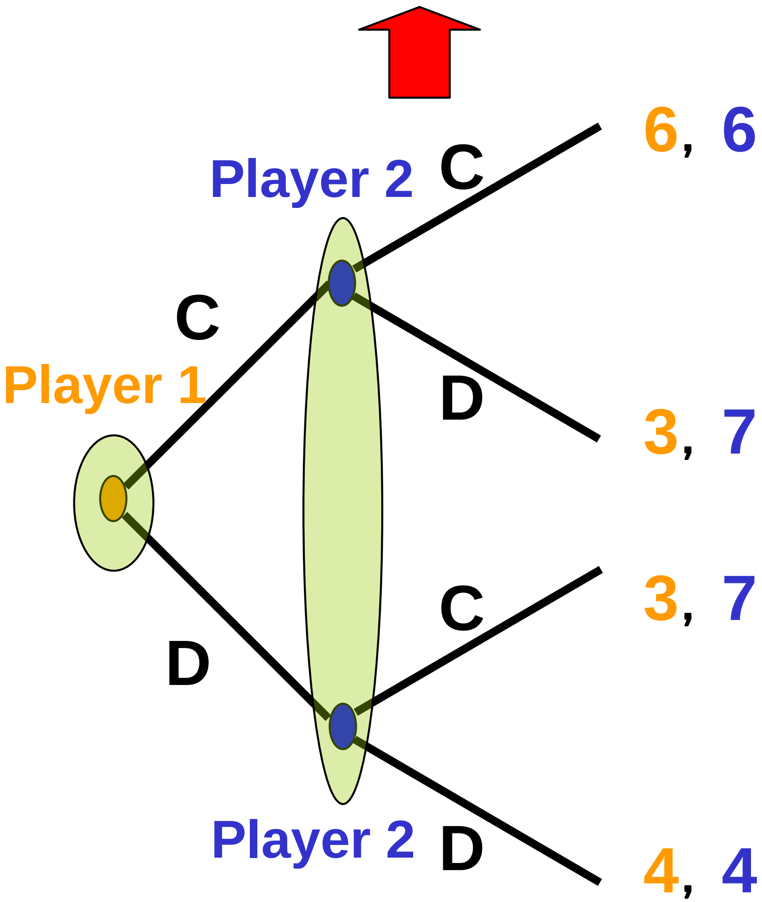
(D,D) のあと
の部分ゲーム

	C	D
C	6, 6	3, 7
D	7, 3	$\boxed{4, 4}$


Player 1
Player 2
Player 2
C
D
C
D
C
D
6, 6
3, 7
7, 3
4, 4
(C,C) のあとの部分ゲーム
(C,D) のあとの部分ゲーム
(D,C) のあとの部分ゲーム
(D,D) のあとの部分ゲーム
C D
C 8, 8 | 5, 9
D 9, 5 | 6, 6
C D
C 5, 9 | 2, 10
D 6, 6 | 3, 7
C D
C 9, 5 | 6, 6
D 10, 2 | 7, 3
C D
C 6, 6 | 3, 7
D 7, 3 | 4, 4

	C	D
C	6, 6	3, 7
D	7, 3	4, 4

つまり、第一ステージの
Nash 均衡も
(D, D)



部分ゲーム完全均衡は、

Player 1 は第一ステージでは D を選択
Plauer 1 は第二ステージでは
  第一ステージの結果が(C, C) のときは D を選ぶ
  第一ステージの結果が(C, D) のときは D を選ぶ
  第一ステージの結果が(D, C) のときは D を選ぶ
  第一ステージの結果が(D, D) のときは D を選ぶ

Player 2 も同じ

部分ゲーム完全均衡は、

(D/D/D/D/D, D/D/D/D/D)

である。

- 囚人のジレンマゲームの二回繰り返しゲームの分析からわかったこと
  - 結局、毎回 (D, D) が達成されるだけ

- なぜだろうか？
  - 二回目（最終回）の囚人のジレンマゲームを考えると、一回目の囚人のジレンマゲームの結果が何であれ（これまで実現した利得がなんであれ）、それが囚人のジレンマゲームと同じ構造を持つことは変わらない。
  - これにより、結局、二回目（最終回）の囚人のジレンマゲームでは、(D, D) だけが Nash 均衡である。
  - このことを踏まえると、結局一回目（前の回）の囚人のジレンマゲームの結果は、二回目（次の回）の囚人のジレンマゲームの結果には影響を与えない。
  - よって、一回目（その前の回）の囚人のジレンマゲームでも（D,D）が達成されるのである。

- 実は、このような結論は、最終回が存在する限りいつでも成立してしまう。

- 言い換えれば、囚人のジレンマゲームを10回、100回、1000回と繰り返しても、毎回 (D,D) が実現し続ける、ということになるのである。

- 例えば、1000回繰り返す場合を考えると、100回目のゲーム、999回目のゲーム、998回目のゲーム、．．．．、と考えていき、結局 (D, D) が実現し続けることになる。

- さて、このような思考方法は、本当に我々が長期的な関係にある相手（配偶者、家族、友人、長期的な契約相手、ビジネスパートナー）に対して意思決定をする際に用いている思考方法を近似しているといえるのだろうか？

- 無限回繰り返しゲーム（最終回の存在しないゲーム）のほうが適切である

無限回繰り返しゲーム

- 囚人のジレンマゲームを繰り返し行う状況を考えよう。
- 今度は、無限に囚人のジレンマゲームを繰り返すとする。
- 無限回繰り返しゲームで難しい点は、いかにして無限に生じる利得を評価するのか、という点。
- 例えば、囚人のジレンマゲームで無限に (C,C) が実現されるとすると、そのときの利得は
  - 4 + 4 + 4 + … + 4 + … = ∞
- 無限に大小関係はつけられない。

- 利得の割引現在価値
  - 同じ利得でも、今日の利得は一年後の利得よりうれしい。（今日の1万円は一年後にもらえる1万円よりも、今日の時点で評価するとうれしい）
- 利得の割引因子を　δ　（0＜δ＜1）とする。
  - 今日の利得 1 を今日評価すると 1
  - 時期（明日、一年後）の利得1を今日評価すると

$$1 \times \delta = \delta$$

  - さらに時期（明後日、二年後）の利得1を今日評価すると

$$1 \times \delta \times \delta = \delta^2$$

- 今日利得1もらえ、時期にも利得1もらえ、さらに時期にも利得1もらえる状況を、今日評価すると

$$1 + \delta + \delta^2$$

- となる。

- 利得1を毎期もらえる状況を、今日評価すると

  1＋δ＋δ²＋δ³＋δ⁴＋. . .

- となる。

- これは、初項1、公比δの無限等比級数の和なので

$$\boxed{1+\delta+\delta^{2}+\delta^{3}+\ldots=\frac{1}{1-\delta}}$$

$$a+a\delta+a\delta^{2}+a\delta^{3}+\ldots=\frac{a}{1-\delta}$$

- 無限回繰り返しゲームの利得を考える際には、各回の標準形ゲームから得られる利得の割引現在価値により評価して、その和を考えることにする。

- 例えば、囚人のジレンマゲームの無限回繰り返しゲームでずっと (C,C) が実現すれば、

$$4 + 4\delta + 4\delta^2 + 4\delta^3 + \ldots = \frac{4}{1-\delta}$$

	C	D
C	4, 4	1, 5
D	5, 1	2, 2

- さて、以下では囚人のジレンマゲームを無限回繰り返す状況の部分ゲーム完全均衡を求める。
- しかし、有限回のときと違って、無限回繰り返しゲームの部分ゲーム完全均衡をすべて導出することは普通は出来ない。
- なので、これから、ある部分ゲーム完全均衡を求めることにする。
- 面白いことに、この部分ゲーム完全均衡では、有限回のときとはまったく異なる行動が生じることになる。
- 部分ゲーム完全均衡を求めるといったが、実際にこれからみせる方法は、ある意味のある戦略の組を考えて、これが実際に部分ゲーム完全均衡を構成することを示すことになる。

- では、無限回繰り返しゲームの戦略とはなんだろうか？
  – それは過去の出来事（過去の標準形ゲームの結果）を観察した上で、今日の行動を決定するような、完全な計画である。

- 例
  – 常に C をとる。（All C）
  – 常に D をとる。(All D)
  – 奇数回にはC をとり、偶数回には D をとる。
  – 過去に相手が裏切ったことがある場合には D をとり、それ以外では C をとる。(永久懲罰、Trigger)

- 以下では、次のような Trigger 戦略の亜種を考えて、δ がある範囲内にあるときには、両者がこの戦略をとっている状態が部分ゲーム完全均衡であり、さらに均衡パスでは常に (C,C) が達成されていることを示す。

- Trigger’ 戦略
  - 第一回目には C をとる。
  - 第二回目以降では、過去に自分か相手かが D をとったことがある場合には D を選び、それ以外では C を選ぶ。

- Trigger’戦略を互いに取り合っている状態が部分ゲーム完全均衡であることを示す。

- ケース1：　過去に自分か相手かのどちらかが D をとっているような部分ゲーム。

  – このとき、相手は必ず今後ずっと D を取り続けることになる。
  – よって、自分も D を取り続けることが最適反応である。
  – つまり、Trigger’ 戦略に従うことが、自分にとって最適である。

ケース2：過去に自分も相手もDをとったことがないような部分ゲーム。

– このとき、自分が Trigger’ 戦略に従えば、（今期からみた）利得の割引現在価値の和は

$$4+4\delta+4\delta^2+4\delta^3+\ldots=\frac{4}{1-\delta}$$

– では、**今期の行動を D にするとして**、利得の割引現在価値の和を最も高くするような戦略はどんなものか

– それは、今期からずっと D を取り続けることである。というのも、今期 D をとってしまうと、来期以降は相手は必ずDをとってくるので。

– 今期以降ずっと D を取り続けたときの、利得の割引現在価値の和は

$$5+2\delta+2\delta^2+2\delta^3+\ldots=5+\frac{2\delta}{1-\delta}$$

- さて、以上より、戦略を Trigger' から常に D に切り替えるインセンティブが存在しないためには、

$$\frac{4}{1-\delta}\geq 5+\frac{2\delta}{1-\delta}\Leftrightarrow 4\geq 5(1-\delta)+2\delta\Leftrightarrow \boxed{\delta\geq\frac{1}{3}}$$

- さて、つまり、δ >=1/3 であれば、過去に自分も相手もDをとったことがないような部分ゲームにおいて、戦略をTrigger' から今期 D を取るようないかなる戦略にも変更する誘引は無い。

- 実は、この点だけチェックすれば、過去に自分も相手もDをとったことがないような部分ゲームにおいて、Trigger' からいかなる戦略にも変更する誘引が存在しないことを示したことになるのである。

- というのも、当該部分ゲームにおいて、Trigger' からある戦略に変更する誘引が存在するのならば、それは今期か次期か、あるいはどこかで D を選択することを意味するはず。それを仮に t 期とするのならば、この t 期から始まる部分ゲームは、過去に自分も相手もDをとったことがないような部分ゲームに相当している。しかし、すでに示したとおり、t 期に D を選択するような戦略で、Trigger' よりも高い利得を与えるものは存在しないのである。

- よって、$\delta \geq 1/3$ であれば、互いに Trigger’ を取り合っているような状態は、部分ゲーム完全均衡である。
- さらに、Trigger’ の定義より、均衡パスでは (C,C), (C,C), … を無限に繰り返すことになる。
- このように、無限回繰り返しゲームでは、有限回のときとは異なる部分ゲーム完全均衡が存在している。
- この結果は、囚人のジレンマ状態にあっても、それを繰り返し状況として認識することにより、ジレンマにおかれた当人間の力で協力状態を実現可能であることを示唆している。

- 以上の結果は、実は、囚人のジレンマゲームに限らず、他のすべての標準形ゲームの無限回繰り返しゲームで成立している結果の一例である。

- 大雑把に言えば、標準形ゲームを無限回繰り返すことにより、もとの標準形ゲームのナッシュ均衡利得以上の利得を平均して実現することが可能であり、さらに割引因子が１に十分近ければ、いかなるパレート最適かつ個人合理的な利得を達成することが可能である。

- このような結果は、ゲーム理論の世界では、古くから民間伝承のように伝わってたものなので、フォーク(Folk)定理とよばれている。

アクセルロッドの実験

- 長期的関係なの中でどのような戦略をとることが有効であるのかを調べるため、アクセルロッドはコンピュータプログラムを用いたシミュレーション実験を行った。
- 囚人のジレンマゲームを200回繰り返す。
- プログラムを募集して、14のプログラムが参加
- 各プログラムは総当りする
- 合計で一番高い利得を得たプログラムが優勝

- さて、優勝したのはどんなプログラムだろうか？

- 優勝したのは、非常に単純な tit for tat (しっぺ返し戦略、オウム返し戦略、お返し戦略) とよばれるプログラムだった。
  - 初回は必ず協力（C）をとる
  - 二回目以降は、直前の相手の行動をとる。（つまり、前の回に相手が C ならば今回自分はC,前の回に相手がDだったら今回自分はD）

- 複雑なプログラムも存在する中で、なぜ、tit for tat が優勝したのだろうか。

- この点を理解するため、仮に以下の四種類のプログラムだけがアクセルロッドの実験に参加した場合を考えてみよう。
  - All C
  - All D
  - Tit for tat
  - Trigger
- 計算を簡単にするために、繰り返し回数は100回とする。

	C	D
C	4, 4	1, 5
D	5, 1	2, 2

- 対戦成績がどうなるのか考えてみよう。

	All C	All D	TfT	Tri
All C		100, 500	400, 400	400, 400
All D			203, 199	203, 199
TfT				400, 400
Tri				

All C	900
All D	906
TfT	999
Tri	999

	C	D
C	4, 4	1, 5
D	5, 1	2, 2

- All C
  - 常に相手に勝てない戦略。協力する気のある人が周りに多ければそれなりに強い。ただし、だまそうとする戦略がいると、その戦略から一方的に搾取される。
- All D
  - 常に相手に勝つか引き分ける戦略。周りにかもがいると強い。
- Tit for tat
  - 常に相手に勝てない戦略。協力できそうな相手とは協力し、そうでない相手とは協力しない。ただし、Trigger とは異なり、相手の裏切りに対して寛容である。
- Trigger
  - 常に相手に勝てない戦略。協力できそうな相手とは協力し、そうでない相手とは協力しない。ただし、一度でも裏切った相手とは二度と協力しない。

- Tit for tat と Trigger に差がでるケース
  - All D
  - Tit for tat
  - Trigger
  - Tit for tat' (最初はDで後は前回の相手の手をまねる)

	C	D
C	4, 4	1, 5
D	5, 1	2, 2

- 対戦成績がどうなるのか考えてみよう。

	Tft’	All D	TfT	Tri
Tft’		200, 200	300, 300	202, 202
All D			203, 199	203, 199
TfT				400, 400
Tri				

Tft’	702
All D	606
TfT	899
Tri	801

	C	D
C	4, 4	1, 5
D	5, 1	2, 2

- アクセルロッドの実験からわかること
  - 相手の裏をかこうとするような戦略は、個々の対戦では強いかもしれないが、相手との協力する機会を失うことにより、総合するとそれほど高い利得をあげるわけではない。
  - 相手の善意を無条件で信じるような戦略は、相手が同じようなタイプか可能な限り協力することを目的とするような戦略であればそれなりに強いが、裏をかこうとする戦略にあたると大敗する。
  - 相手との協力関係を模索しつつ、それでいて相手からの搾取行動を許さず、また相手の裏切りに対して寛容であるような戦略が、様々な戦略が存在する環境では一番強い。

次回講義

- これまでの範囲でやりのこした話をやる
- or
- 不完備情報ゲーム