第 20 回宿題

- 提出課題を解き kibaco に答を入力して下さい.
- 自習課題は提出する必要はありません．理解を深めるために自習しましょう.

提出課題 20.1

図 20.1 の囚人のジレンマについて，プレイヤー１が「協力－おうむ返し」戦略を選び，プレイヤー２が「「非協力－おうむ返し」戦略を選んだとする.

問 1 上記戦略でこのゲームを 5 回繰り返したときの履歴を以下の選択肢から選びなさい

(A)	$(C,C)(C,C)(C,C)(C,C)(C,C)$	(B)	$(D,D)(D,D)(D,D)(D,D)(D,D)$
(C)	$(C,D)(D,C)(C,D)(D,C)(C,D)$	(D)	$(D,C)(C,D)(D,C)(C,D)(D,C)$
(E)	$(C,C)(D,D)(D,D)(D,D)(D,D)$	(F)	$(C,C)(C,C)(C,C)(C,C)(D,D)$

利得を割引利得とし，割引因子を δ とする．上記の戦略で，このゲームを $2m$ 回繰り返したとき，以下の問いに答えなさい.

問 2 プレイヤー１の利得を，以下の選択肢の中から選びなさい.

問 3 プレイヤー２の利得を，以下の選択肢の中から選びなさい.

(A)	-2	(B)	6	(C)	$\frac{(-2+6\delta)(1-\delta^{2m})}{1-\delta}$	(D)	$\frac{(6-2\delta)(1-\delta^{2m})}{1-\delta}$
(E)	$\frac{(-2+6\delta)(1-\delta^{m})}{1-\delta^2}$	(F)	$\frac{(6-2\delta)(1-\delta^{m})}{1-\delta^2}$	(G)	$\frac{(-2+6\delta)(1-\delta^{2m})}{1-\delta^2}$	(H)	$\frac{(6-2\delta)(1-\delta^{2m})}{1-\delta^2}$

上記の戦略で，このゲームを無限回繰り返したとき，以下の問いに答えなさい.

問 4 プレイヤー１の利得を，以下の選択肢の中から選びなさい.

問 5 プレイヤー２の利得を，以下の選択肢の中から選びなさい.

(A)	-2	(B)	6	(C)	0	(D)	4
(E)	$\frac{-2+6\delta}{1-\delta}$	(F)	$\frac{6-2\delta}{1-\delta}$	(G)	$\frac{-2+6\delta}{1-\delta^2}$	(H)	$\frac{6-2\delta}{1-\delta^2}$

1 \ 2	C	D
C	(4 , 4)	(-2 , 6)
D	(6 , -2)	(0 , 0)

図 20.1: 成分ゲームとなる囚人のジレンマ

提出課題 20.2

無限に繰り返す囚人のジレンマにおいて，トリガー戦略を選び合うことが，割引率が１に近いときにはナッシュ均衡になることを一般的に証明し，その割引率の値を求めてみよう．以下の [a],[b],[c],[d] に当てはまる式を選択肢 (A)-(H) から選びなさい.

図 20.2 のような囚人のジレンマを考える.
ここで C は協力，D は非協力を示している．囚人のジレンマであることから，x, y, z, w は

$y > x > w > z$

1＼2	C	D
C	(x,x)	(z,y)
D	(y,z)	(w,w)

図 20.2: 囚人のジレンマ（一般形）

の関係にある．
お互いがトリガー戦略を選んでいるときには，すべての回で (C, C) が選ばれる．ナッシュ均衡であることを示すためには，各プレイヤーがトリガー戦略を選び合っている時に，そこから 1 人のプレイヤーが戦略を変えても利得が高くならないことを示せば良い．
そこでプレイヤー 2 がトリガー戦略を選んでいる時に，プレイヤー 1 が戦略を変化させたとする（プレイヤー 2 の方が戦略を変えた場合も同じ論理で示せる）．このとき，プレイヤー 1 の行動によって，次の 2 つの場合に分けられる．

すべての回で C を選ぶ場合

例えば，トリガー戦略から

- 相手が何をしてもすべての回で C を選ぶ
- 協力－オウム返し戦略にする
- 1 回目から 4 回目までオウム返し戦略にして，そのあとトリガー戦略にする

などに戦略を変化させても，やはりすべての回で (C, C) が選ばれる．このような場合は利得は同じであるので，利得は高くならない．

少なくとも 1 回は D を選ぶ場合

このとき最初に D が選ばれる回を t 回目としよう（$t \geq 1$）．このとき t 回目のプレイヤー 1 の（その時点での）利得は，トリガー戦略を選んでいたとき x であったが，(D, C) に変わったことで利得が [a] になる．このとき利得は [b] だけ増加する．これに対し，$t + 1$ 回目以降のプレイヤー 1 の（各時点での）利得は，トリガー戦略を選んでいるときは x であったが，プレイヤー 2 が D を選んでくるため，**最大で** [c] しか獲得できない．
トリガー戦略を選んでいたときの t 回目以降の割引利得の合計は（t 回目の割引が δ^{t-1} であることに注意し，無限級数の和を計算すれば）

$$x\delta^{t-1} + x\delta^{t} + x\delta^{t+1} + x\delta^{t+2} \cdots = \delta^{t-1}(x + x\delta + x\delta^2 + \delta^3 \cdots) = \delta^{t-1}\left(\frac{x}{1-\delta}\right)$$

である．これに対して，異なる戦略を選んだときの利得は最大値で

$$\mathrm{a}\delta^{t-1} + \mathrm{c}\delta^{t} + \mathrm{c}\delta^{t+1} + \mathrm{c}\delta^{t+2} \cdots = \delta^{t-1}(\mathrm{a} + \mathrm{c}\delta + \mathrm{c}\delta^2 + \mathrm{c}\delta^3 \cdots) = \delta^{t-1}\left(\mathrm{a} + \frac{\mathrm{c}}{1-\delta}\right)$$

である．もし

$$\delta^{t-1}\left(\frac{x}{1-\delta}\right) \geq \delta^{t-1}\left(\mathrm{a} + \frac{\mathrm{c}\delta}{1-\delta}\right)$$

であれば，他の戦略に変えても利得は大きくならない．これを解くと

$$\delta \geq \frac{\mathrm{a} - x}{\mathrm{a} - \mathrm{c}} = \mathrm{d}$$

となる．$\delta \geq \mathrm{d}$ であれば，トリガー戦略を選び合うことがナッシュ均衡になる．
別の見方をしてみよう，トリガー戦略から異なる戦略に変えることで t 回目の利得は b だけ増加するが，そのあと少なくとも $x-\mathrm{c}$ は減少する．よってトリガー戦略に変えた時に増加する利得は

$$\delta^{t-1}\left(\mathrm{b}-(x-\mathrm{c})\delta-(x-\mathrm{c})\delta^2+\cdots\right)=\delta^{t-1}\left(\mathrm{b}-\frac{(x-\mathrm{c})\delta}{1-\delta}\right)$$

である．よって $\mathrm{b}-\frac{(x-\mathrm{c})\delta}{1-\delta}\geq 0$ であれば良い．これを解いても同様に

$$\delta \geq \mathrm{d}$$

が得られる．

[a]-[d] の選択肢

(A)	x	(B)	y	(C)	z	(D)	w
(E)	$y-x$	(F)	$w-x$	(G)	$\frac{y-x}{y-w}$	(H)	$\frac{y-w}{y-z}$

提出課題 20.3
囚人のジレンマの繰り返しゲームにおいて，毎回，確率 p で次のプレイは継続するが，確率 $1-p$ で終了して，それ以上プレイが継続しないようなゲームを考える
囚人のジレンマの有限回の繰り返しでは，最終回に裏切ることが原因で，全ての回で協力しないことが唯一のナッシュ均衡であった．このゲームではいつが最終回か分からないため，協力し合うことがナッシュ均衡となる可能性がある．
図 20.3 の囚人のジレンマを考え，割引因子を δ とする．このとき確率 p がいくつ以上ならば，トリガー戦略を選び合うことがナッシュ均衡になるか考えてみよう．以下の a から f に当てはまる数値を答えよ．

1 \ 2	A	B
A	(4,4)	(0,5)
B	(5,0)	(1,1)

図 20.3: 囚人のジレンマ

ここで C は協力，D は非協力を示している．
ここでお互いがずっと協力した場合に得られるとき，どのくらいの利得が得られるかを考える．その割引利得和の期待値を v_C とする．このとき第 1 回目は必ずプレイされ，プレイヤーは第 1 回に利得 4 を獲得する．そして，確率 $1-p$ でゲームは終わり，それ以降の利得は 0 となり，確率 p でゲームは続く．ゲームが続いたときそれ以降に得られる利得は，第 1 回と同じ状況になっていると考えられるので v_C であることから

$$v_C = 4+\delta\{(1-p)\times 0+p\times v_C\}$$

という式が得られる．よって $v_C=\frac{a}{b-p\delta}$ である．
ここでお互いがずっと協力しない場合の割引利得和の期待値を v_D とする．同様に

$$v_D = 1+\delta\{(1-p)\times 0+p\times v_D\}$$

から，$v_D=\frac{c}{d-p\delta}$ となる．

トリガー戦略から，プレイヤー 1 が戦略を変更したとして，その戦略で，t 回目までは協力して，$t+1$ に協力しなかったとする．利得は $t+1$ 回目で 4 から 5 に上昇するが，そのあと $t+2$ 回目以降に得られる利得は最大で v_D となる．一方，利得を変更しないときは v_C が得られる．よって

$$\delta^t v_C \geq \delta^t \{5 + \delta((1-p) \times 0 + p \times v_D)\}$$

であれば，どんな戦略よりもトリガー戦略を選んだときの利得が高くなる．これより $p \geq \frac{1}{e\delta}$ であれば，トリガー戦略を選び合い，互いに協力を続けることがナッシュ均衡になる．
$\delta = 3/4$ のときは，$p \geq 1/f$ であれば協力が達成できる．