2015年度　ゲームの理論 a 期末試験解答

グレーヴァ香子

1. (a) 最適反応に下線をつける。

P1 \ P2	L	R
A	1, 0	$\underline{5}$, $\underline{1}$
B	$\underline{2}$, $\underline{1}$	0, 0
C	0, $\underline{5}$	4, 4

ゆえに純戦略によるナッシュ均衡は (A, R) と (B, L)。

(b) ２回繰り返しにすると、純戦略の部分ゲーム完全均衡になるには、２期目は (A,R) または (B,L) がプレイされなければならない。どちらにせよ、P2 は同じ利得 1 を２期目にはもらうことになる。これがポイント。

１期目に G のナッシュ均衡以外の行動の組み合わせである (C, R) をして、２期目は (A,R) または (B,L) という形の、G^2 の <u>任意の</u> 純戦略の組み合わせを考える。P2 は R から逸脱して L をした方が１期目の利得は高く、２期目も利得は必ず 1 なので、R をすることは最適反応ではない。ゆえにこのような部分ゲーム完全均衡は存在しない。

（注：上にあるように、何かが存在しないことを証明するには「どうやってもできない」ことを書かなくてはならないが、「１期目に (C,R) が行われたら、２期目は (A,R) を、他の行動の組み合わせが行われたら (B,L) をする戦略が部分ゲーム完全均衡でない」を証明しただけの人が大多数であった。）

(c) 罰として使えるのは (C,R) より２人とも利得が低いナッシュ均衡である (B, L) である。従って、t 期の期初までの歴史を $h \in [\{A, B, C\} \times \{L, R\}]^{t-1}$ と書くと、グリム・トリガー戦略は

$$s_1(h) = \begin{cases} C & \text{if } h = \emptyset \text{ or } (C, R)^{t-1} \\ B & \text{otherwise} \end{cases}$$

$$s_2(h) = \begin{cases} R & \text{if } h = \emptyset \text{ or } (C, R)^{t-1} \\ L & \text{otherwise.} \end{cases}$$

このとき、部分ゲームとしては (i) $h = \emptyset$ または $(C, R)^{t-1}$ という歴史の後部分ゲーム群（経路上の部分ゲーム群）と (ii) そうでない歴史の後の部分ゲーム群に分けて考えればよい。(ii) のグループの部分ゲームについては G のナッシュ均衡をその部分ゲーム内では何があっても繰り返すだけなので、このグループの部分ゲームにおいては、<u>任意の $\delta \in (0, 1)$ について</u> ナッシュ均衡である。（ここも書いておくとベスト。）
(i) のグループの部分ゲームについては、相手プレイヤーが s_i に従っているとして、s_j が最適であることを示すために、s_j から one-step deviation をしても総利得が高まらないということを示す。
P2 は s_2 に従うとして、P1 が s_1 に従うと、今後ずっと (C,R) がおこるので総利得は

$$4 + \delta \cdot 4 + \delta^2 \cdot 4 + \cdots = \frac{4}{1 - \delta}.$$

一期だけ他の行動に変えて、その後 s_1 に従うとする。他の行動に変えてもっとも利得が高くなるのは A に変えたときである。その後は (B,L) がずっとプレイされるので、one-step deviation で得られる最大の総利得は

$$5 + \delta \cdot 2 + \delta^2 \cdot 2 + \cdots = 5 + \frac{2\delta}{1 - \delta}.$$

従って、P2 が s_2 に従っているときに、P1 にとって s_1 が最適である条件は

$$\frac{4}{1-\delta} \geqq 5 + \frac{2\delta}{1-\delta} \iff \delta \geqq \frac{1}{3}.$$

P1 が s_1 に従うとして、同様に P2 について s_2 の利得と one-step deviation の最大利得を比較する。s_2 からの利得は

$$4 + \delta \cdot 4 + \delta^2 \cdot 4 + \cdots = \frac{4}{1-\delta}.$$

One-step deviation としては L に変えると最大の 5 が１期間得られる。その後は (B,L) がずっとプレイされるので、one-step deviation で得られる最大の総利得は

$$5 + \delta \cdot 1 + \delta^2 \cdot 1 + \cdots = 5 + \frac{\delta}{1-\delta}.$$

ゆえに P1 が s_1 に従っているとき、P2 にとって s_2 が最適である条件は

$$\frac{4}{1-\delta} \geqq 5 + \frac{\delta}{1-\delta} \iff \delta \geqq \frac{1}{4}.$$

２人とも s_i が最適反応であるためには δ の下限は大きい方の数字、すなわち 1/3 である。

（注：例年にもまして、今年は「グリム・トリガー戦略の組み合わせを求めなさい」という問いに対して解答をしていないで、δ の下限だけを求めている人が多かった。このような不注意はもちろん減点である。別に試験のときだけでなく、「自分は何をすることを要求されているのか」を的確に理解することは今後の人生でずっと重要です。）

2. 解答その１：企業 X の行動戦略ということで解いてみる。企業 X には２つの情報集合が Nature の選択に応じて存在し、それぞれについて純行動 High または Low を選べるので、X の行動戦略は４つあり (G_h のときの行動、G_ℓ のときの行動）の順に書くとする。Y の戦略は情報集合がただ一つであるので High または Low のみである。G_h が選ばれる確率を p としたときの期待利得による行列表現は以下のようになる。

X \ Y	High	Low
(H, H)	$3p + (1-p)$, 3	$2p$, 0
(H, L)	3, $3p$	2, $(1-p)$
(L, H)	1, $3(1-p)$	0, p
(L, L)	$p + 3(1-p)$, 0	$2(1-p)$, 1

$p = 0.1, 0.5$ いずれにせよ、X の最適反応は相手がどちらの戦略でも (H, L) である。あとは、Y の最適反応が場合によって異なり、$p = 0.1$ のときは Low、$p = 0.5$ のときは High が (H, L) に対する最適反応である。従って

$p = 0.1$ のときのベイジアン・ナッシュ均衡は $(s_X, s_Y) = ((\text{H, L}), \text{Low})$

$p = 0.5$ のときのベイジアン・ナッシュ均衡は $(s_X, s_Y) = ((\text{H, L}), \text{High})$ である。

解答その２：タイプ別の３人ゲームとして解いてみる。

G_h タイプの X については Low 戦略が High 戦略に厳密に支配されているので、High 戦略をとる。G_ℓ タイプの X は、同様に Low 戦略を取るのが、Y が何を行っても最適である。

G_h タイプは High を、G_ℓ タイプは Low を行うときの、Y の２つの純戦略の期待利得を計算すると

$$Eu_Y((H, L), H) = 3p$$
$$Eu_Y((H, L), L) = 1 - p$$

ゆえに $p = 0.1$ のときは Low、$p = 0.5$ のときは High が (H, L) に対する最適反応である。あとは同じ。

3. (a) まず、完全ベイジアン均衡の逐次合理性から、Ask (Ask'), Y の後の W_g と W_b の１点からなる情報集合における意思決定は、最大化問題なので、W_g は NB を、W_b は B を選ぶのが最適である。

さて、W が一括戦略 Ask (Ask') を取るとして、将来の行動は上のようになっているとする。このとき、まず一括戦略の下での整合的な K の信念は g タイプが (0.6) の確率すなわち $r = 0.6$ とすることである。したがって、K が Y を行う時の期待利得は (0.6) の確率で裏切られず 10 を得、(0.4) の確率で裏切られて -10 を得ることになる。これは N を行うときの期待利得 0 より大きいので、K の最適戦略（K は情報集合が一つしかないので、行動が戦略である）は Y となる。

最後に、上の今後の各プレイヤーの行動を所与として、W の最初の２つの情報集合において最適行動をしているかをチェックする。W_g は Ask に従うと 10 の利得を得るが、Not にしてしまうと利得は 0 であるから Ask が最適である。W_b も Ask' に従うと 10 を得るが、Not' にすると 1 しか得られないので、Ask' が最適である。

以上のことより、W の戦略は (W_g の最初の情報集合での行動、W_b の最初の情報集合での行動、W_g の２つ目の情報集合での行動、W_b の２つ目の情報集合での行動) の順に書くとして、

((Ask, Ask', NB, B), Y, $r = 0.6$)

というものである。

(b) （本試験の問題では最後の情報集合２つにおいて W の添字を入れるのを忘れていて、すみません！しかし添字があってもなくても、問題を正しく理解していれば分析にはまったく影響がないことがわかるはず。）

解答その１：新しいゲームでは、bad タイプの W は Ask' をしても、$-1, 0, 10 - C < 0$ のどれかしか得られないので、Not' が最適である。従って両方のタイプの W が Ask(Ask') をする一括均衡は存在しない。

解答その２： (a) の分析と同様にまず２つ目の W_g, W_b の情報集合における最適行動を求める。今度は上の情報集合でも、下の情報集合でも NB をするのが最適になる。

再び、W が一括戦略 Ask (Ask') を取るとして、将来の行動は (NB, NB) とする。このとき、一括戦略の下での整合的な K の信念は、またもや $r = 0.6$ であり、今度は Y を選んだあと、どちらの W も NB をしてくれるので利得はかならず 10 になり、これは N を選んだときの 0 より大きいので、Y が最適戦略である。

最初に戻って、W_g の１つ目の情報集合を考える。Ask をすると Y, NB と続き、利得は 10、Not をすると 0 であるから Ask が最適である。W_b の一つ目の情報集合を考えると、Ask' をすると Y, NB と続き利得は 0、しかし Not' をすると利得は 1 であるから、Aks' は W_b にとって最適ではない。

ゆえに、両方のタイプの W が Ask(Ask') をする一括均衡は存在しない。