

ENSEIRB

2ième année Informatique, 2015/2016

Option Jeux

DM du 13/04/16- Corrigé

Exercice 1 (10 pts) Jeu de Poker

1- Analysons les “comportements raisonnables” de A,B.

- si A_3/B_j :

soit σ_{pass} la stratégie où A choisit de passer et σ_m la stratégie où A monte : pour toute stratégie (mixte) τ de B,

$$r(\sigma_{pass}, \tau) = -1 \leq 1 \leq r(\sigma_m, \tau).$$

Donc A monte.

- si A_i/B_1 :

soit τ_{pass} la stratégie où B choisit de passer et τ_m la stratégie où B monte : dans le sous-jeu issu du sommet carré,

$$r(\tau_m) = 2 > 1 = r(\tau_{pass}).$$

Donc B passe.

- si A_i/B_3 : avec les mêmes notations

$$r(\sigma, \tau_m) = -2 < 1 = r(\sigma, \tau_{pass}).$$

Donc B monte

- si A_2/B_j : utilisons les notations ci-dessus. Les 6 donnes sont équiprobables, donc $\mathbb{P}(\{j = 1\}) = \mathbb{P}(\{j = 3\}) = \frac{1}{2}$.

La stratégie τ de B la meilleure est : si $j = 1$ B passe, sinon B monte (voir ci-dessus).

$$\mathbb{E}(r(\sigma_m, \tau)) = \frac{1}{2}(1 + (-2)) = \frac{-1}{2}$$

$$\mathbb{E}(r(\sigma_{pass}, \tau)) = \frac{1}{2}((-1) + (-1)) = -1$$

Donc la meilleure stratégie de A consiste à monter.

2- Calculons le gain moyen r de A pour chacune des 6 donnes :

$$\begin{aligned} A_1/B_2 : r &= (1-p) \cdot (-1) + p(1-q) \cdot 1 + pq \cdot (-2) \\ &= -3pq + 2p - 1 \end{aligned}$$

$$A_1/B_3 : r = (1-p) \cdot (-1) + p \cdot (-2) = -p - 1$$

$$A_3/B_2 : r = (1-q) \cdot 1 + q \cdot 2 = q + 1$$

$$A_2/B_1 : r = 1$$

$$A_2/B_3 : r = -2$$

$$A_3/B_1 : r = 1$$

D' où :

$$r = \frac{1}{6}(-3pq + p + q - 1) = -\frac{1}{2}\left(p - \frac{1}{3}\right)\left(q - \frac{1}{3}\right) - \frac{1}{9}.$$

3- On a vu en cours que, en posant :

$$\underline{v} := \max_{p \in [0,1]} \{ \min_{q \in [0,1]} \{ r(\sigma_p, \tau_q) \} \} \text{ et } \bar{v} := \min_{q \in [0,1]} \{ \max_{p \in [0,1]} \{ r(\sigma_p, \tau_q) \} \},$$

on a toujours $\underline{v} \leq \bar{v}$.

Ici, en choisissant $p = \frac{1}{3}$ on voit que :

$$\underline{v} \geq \min_{q \in [0,1]} \{ r(\sigma_{\frac{1}{3}}, \tau_q) \} = -\frac{1}{9}$$

et en choisissant $q = \frac{1}{3}$ on voit que :

$$\bar{v} \leq \max_{p \in [0,1]} \{ r(\sigma_p, \tau_{\frac{1}{3}}) \} = -\frac{1}{9}$$

Donc

$$\bar{v} \leq -\frac{1}{9} \leq \underline{v} \leq \bar{v},$$

ce qui montre que ces deux valeurs sont égales à $-\frac{1}{9}$ i.e.

$$\max_{p \in [0,1]} \{ \min_{q \in [0,1]} \{ r(\sigma_p, \tau_q) \} \} = \min_{q \in [0,1]} \{ \max_{p \in [0,1]} \{ r(\sigma_p, \tau_q) \} \} = -\frac{1}{9}.$$

Le jeu a donc la valeur $v = -\frac{1}{9}$.

Pour $(\bar{p} = \frac{1}{3} = \bar{q})$ on réalise :

$$v = \min_{q \in [0,1]} \{ r(\sigma_{\bar{p}}, \tau_q) \} = \max_{p \in [0,1]} \{ r(\sigma_p, \tau_{\bar{q}}) \}$$

i.e. \bar{p} (resp. \bar{q}) réalise la meilleure stratégie de A (resp. de B).

Par contre l'ensemble C_v des couples (p, q) tels que $r(\sigma_p, \tau_q) = v$ est donné par :

$$(p, q) \in C_v \Leftrightarrow (p - \frac{1}{3})(q - \frac{1}{3}) = 0 \Leftrightarrow p = \frac{1}{3} \text{ ou } q = \frac{1}{3}.$$

4- On a vu en cours que, pour un jeu à deux joueurs et à somme nulle, (σ, τ) est un équilibre de Nash ssi il réalise la valeur du jeu v .

5- Soit $p \neq \frac{1}{3}$.

Si $p > \frac{1}{3}$, B choisit $q := 1$ et obtient

$$r = [-\frac{1}{2} \cdot \frac{2}{3}(p - \frac{1}{3})] - \frac{1}{9} < -\frac{1}{9} = v$$

Si $p < \frac{1}{3}$, B choisit $q := 0$ et obtient

$$r = [\frac{1}{2} \cdot \frac{1}{3}(p - \frac{1}{3})] - \frac{1}{9} < -\frac{1}{9} = v.$$

6- Soit $q \neq \frac{1}{3}$.

Si $q > \frac{1}{3}$, A choisit $p := 0$ et obtient

$$r = -\frac{1}{6}(q - \frac{1}{3}) - \frac{1}{9} > -\frac{1}{9} = v$$

1. Les questions 4,5,6 du texte étaient *mal formulées* puisque l'ensemble C_v est infini; heureusement tout le monde a deviné que \bar{p}, \bar{q} désignaient, en fait, les paramètres des stratégies *optimales* de A et B

Si $q < \frac{1}{3}$, A choisit $p := 1$ et obtient

$$r = -\frac{1}{3}\left(q - \frac{1}{3}\right) - \frac{1}{9} > -\frac{1}{9} = v.$$

Exercice 2 (10 pts) **Équilibres de Nash**

1- Ce jeu a exactement un équilibre de Nash en stratégies pures : (T, T) .

On rappelle que le gain du joueur i ($i \in \{1, 2\}$) sur la partie h est défini par

$$R_i(h) := (1 - \delta) \cdot \sum_{t=1}^{\infty} \delta^{t-1} \cdot r_i(h_1^t, h_2^t) \quad (1)$$

i.e. le gain du joueur i est la moyenne de ses gains, dans le jeu Γ , aux tours $1, 2, \dots, t, \dots$ pondérés par les coefficients $\delta^0, \delta^1, \dots, \delta^{t-1}, \dots$

2- Calculons $R_2(h')$.

$$\begin{aligned} R_2(h') &= (1 - \delta)(3 + 2\delta + 3\delta^2 + \dots) \\ &= (1 - \delta)\left[\frac{3}{1 - \delta^2} + \frac{2\delta}{1 - \delta^2}\right] \\ &= (1 - \delta) \cdot \frac{3 + 2\delta}{1 - \delta^2} \\ &= \frac{3 + 2\delta}{1 + \delta}. \end{aligned}$$

Trahir Toujours

Notons TT la stratégie du joueur J1 (resp. J2) définie par : pour toute histoire $(h^1 \dots h^{t-1}) \in (S_1 \times S_2)^{t-1}$:

$$\text{TT}(h^1 \dots h^{t-1}) = T.$$

3- Montrons que $R_1(\text{TT}, \text{TT}) = 1$.

$$R_1(\text{TT}, \text{TT}) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} = 1$$

4- Soit σ_1 une stratégie du joueur J1.

$$\begin{aligned} R_1(\sigma_1, \text{TT}) &= (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} r_1(h^t(\sigma_1, \text{TT})) \\ &\leq (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \quad \text{car } r_1(*, \text{TT}) \leq 1 \\ &= 1 \end{aligned}$$

5- Par Q4, pour toute stratégie σ_1 de J1,

$$R_1(\sigma_1, \text{TT}) \leq R_1(\text{TT}, \text{TT}),$$

et par un raisonnement analogue, pour toute stratégie σ_2 de J2,

$$R_2(\text{TT}, \sigma_2) \leq R_2(\text{TT}, \text{TT}).$$

Donc, pour tout $\delta \in]0, 1[$, (TT, TT) est un équilibre de Nash du jeu Γ^δ .

Punir Éternellement

Notons PE la stratégie du joueur Ji définie par : pour toute histoire $h^1 \dots h^{t-1} \in (S_1 \times S_2)^{t-1}$:

$$\text{PE}(h^1 \dots h^{t-1}) = T \text{ si } \exists t_0 \in [1, t-1], h_{3-i}^{t_0} = T, \text{ PE}(h^1 \dots h^{t-1}) = C \text{ sinon}$$

i.e. Ji coopère tant que son partenaire coopère, et trahit éternellement, dès que son partenaire a trahi au moins une fois.

6- Remarquons que $h(\text{PE}, \text{PE}) = (C, C)^\omega$. Donc

$$R_1(\text{PE}, \text{PE}) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \cdot 2 = 2.$$

7- On détermine la partie jouée selon ces deux stratégies : $h = (C, C)^{t_0-1}(T, C)(T, T)^\omega$. On en déduit que

$$\begin{aligned} R_1(h) &= (1 - \delta) \left[\sum_{t=1}^{t_0-1} 2\delta^{t-1} + 3\delta^{t_0-1} + \sum_{t=t_0+1}^{\infty} \delta^{t-1} \right] \\ &= (1 - \delta) \left[\sum_{t=1}^{\infty} 2\delta^{t-1} + \delta^{t_0-1} - \sum_{t=t_0+1}^{\infty} \delta^{t-1} \right] \\ &= (1 - \delta) \left[\frac{2}{1 - \delta} + \delta^{t_0-1} - \delta^{t_0} \frac{1}{1 - \delta} \right] \\ &= 2 + \delta^{t_0-1}(1 - 2\delta). \end{aligned}$$

8- **Si** $\delta < \frac{1}{2}$:

alors $R_1(h) = 2 + \delta^{t_0-1}(1 - 2\delta) > 2 = R_1(\text{PE}, \text{PE})$ donc (PE, PE) n'est pas un équilibre de Nash.

Si $\delta \geq \frac{1}{2}$:

soit σ_1 une stratégie de J1, qui induit une partie contre PE différente de $h(\text{PE}, \text{PE})$: $\exists t_0 \in [1, \infty[, h^{t_0}(\sigma_1, \text{PE}) = (T, C)$.

Par définition de PE :

$\forall t \geq t_0 + 1, h^t(\sigma_1, \text{PE}) = (*, T)$ donc pour la partie $h := h(\sigma_1, \text{PE})$:

$$\forall t \in [1, t_0 - 1], r_1(h^t) = 2, \quad r_1(h^{t_0}) = 3, \quad \forall t \in [t_0 + 1, \infty[, r_1(h^t) \leq 1,$$

donc

$$\begin{aligned} R_1(h(\sigma_1, \text{PE})) &\leq (1 - \delta) \left[\sum_{t=1}^{t_0-1} 2\delta^{t-1} + 3\delta^{t_0-1} + \sum_{t=t_0+1}^{\infty} \delta^{t-1} \right] \\ &= 2 + \delta^{t_0-1}(1 - 2\delta) \\ &\leq 2. \end{aligned}$$

Dans ce cas, (PE, PE) est un équilibre de Nash.

On dit que (σ_1, σ_2) est un équilibre de Nash *uniforme* de la famille de jeux $(\Gamma^\delta)_{\delta \in]0, 1[}$ si,

$$\exists \delta_0 \in [0, 1[, \forall \delta \in]\delta_0, 1[, (\sigma_1, \sigma_2) \text{ est un équilibre de Nash de } \Gamma^\delta.$$

9- (PE, PE) vérifie la définition d'un équilibre de Nash *uniforme* de la famille $(\Gamma^\delta)_{\delta \in]0,1[}$: il suffit de choisir $\delta_0 := \frac{1}{2}$.

Punir et Pardonner

Notons PP_k (où k est un entier non nul), la stratégie du joueur J_i définie par :

$$\begin{aligned} PP_k(h^1 \dots h^{t-1}) &= T \text{ si } \exists t_0 \in [t-k, t-1], [(t_0 = 1 \text{ ou } h^{t_0-1} = (C, C)) \text{ et } h^{t_0} \neq (C, C)], \\ PP_k(h^1, \dots, h^{t-1}) &= C \text{ sinon} \end{aligned} \quad (2)$$

10- Soit σ_1 la stratégie de J_1 consistant à trahir aux tours 1, 2 puis coopérer éternellement.

On a alors :

$$h(\sigma_1, PP_1) = (T, C)(T, T)(C, C)^\omega$$

et le revenu de J_1 est

$$\begin{aligned} R_1((T, C)(T, T)(C, C)^\omega) &= (1 - \delta)[3 + \delta + 2\delta^2 + 2\delta^3 + \dots] \\ &= (1 - \delta)[1 - \delta + \sum_{t=1}^{\infty} 2\delta^{t-1}] \\ &= (1 - \delta)^2 + 2 \\ &> 2 \\ &= R_1(PP_1, PP_1). \end{aligned}$$

Donc, pour tout $\delta \in]0, 1[$, (PP_1, PP_1) n'est pas un équilibre de Nash de Γ^δ .

11- Soit h est une partie où J_2 joue la stratégie PP_k et J_1 joue la stratégie PP_k à tous les tours d'ordre $t \neq t_0$ mais trahit au coup $t_0 \in [1, \infty[$.

11.1 Vérifions que

$$h = (C, C)^{t_0-1} \cdot (T, C) \cdot (T, T)^k \cdot (C, C)^\omega$$

tours $t \leq t_0 - 1$:

il est clair que (C, C) est joué.

tour t_0 :

(T, C) est joué.

tour $t \in [t_0 + 1, t_0 + k]$:

Comme $t_0 + 1 \leq t \leq t_0 + k$ on a

$$t_0 \leq t - 1 \text{ et } t - k \leq t_0$$

or $h^{t_0-1} = (C, C)$ et $h^{t_0} \neq (C, C)$, donc les joueurs jouent (T, T) .

tour $t = t_0 + k + 1$:

pour tout $t' \in [t - k, t - 1] = [t_0 + 1, t_0 + k]$, on a

$$t' \neq 1 \text{ et } h^{t'-1} \neq (C, C)$$

donc PP_k prescrit au joueur 2 de jouer C et les joueurs jouent (C, C) .

tours $t \geq t_0 + k + 2$:

On démontre par récurrence sur t que

$$\forall t' \in [t - k, t - 1], (h^{t'-1} = (T, T) \neq (C, C)) \text{ ou } (h^{t'-1} = (C, C) \text{ et } h^{t'} = (C, C)).$$

Donc les joueurs jouent (C, C) .

Finalement h a bien la forme annoncée.

11.2

$$\begin{aligned}
R_1(h) &= (1 - \delta) \left[\sum_{t=1}^{t_0-1} \delta^{t-1} \cdot 2 + \delta^{t_0-1} \cdot 3 + \sum_{t=t_0+1}^{t_0+k} \delta^{t-1} \cdot 1 + \sum_{t=t_0+k+1}^{\infty} \delta^{t-1} \cdot 2 \right] \\
&= (1 - \delta) \left[\sum_{t=1}^{\infty} \delta^{t-1} \cdot 2 + \delta^{t_0-1} \cdot 1 - \sum_{t=t_0+1}^{t_0+k} \delta^{t-1} \right] \\
&= (1 - \delta) \left[\frac{2}{1 - \delta} + \delta^{t_0-1} (1 - \delta \dots - \delta^k) \right] \\
&= 2 + (1 - \delta) \delta^{t_0-1} (1 - \delta \dots - \delta^k).
\end{aligned}$$

On note φ la racine réelle positive de l'équation $P(x) = x^2 + x - 1 = 0$. Soit $\delta \in]\varphi, 1[$.

Remarquons que $P(1) = 1 > 0$ donc $1 > \varphi > \bar{\varphi}$.

12- Soit $\delta \in]\varphi, 1[$.

12.2 Soit h une partie où J2 joue la stratégie PP_2 (à tous les tours), et J1 joue la stratégie PP_2 à tous les tours d'ordre $t \leq t_0 - 1$ mais trahit au tour $t_0 \in [1, \infty[$. Le texte demandait de montrer que

$$R_1(h) \leq 2 + (1 - \delta) \delta^{t_0-1} (1 - \delta - \delta^2).$$

En fait les résultats à démontrer en Q12.1, Q12.2 sont faux (nous donnons plus loin une définition modifiée de PP_k qui conduit à des résultats corrects).

Traisons d'abord les questions Q12, Q13, Q14 sans modifier la définition (2).

Soit σ_1 la stratégie de J1 consistant à jouer $C^{t_0-1} \cdot T^\omega$. On a alors

$$h(\sigma_1, PP_k) = (C, C)^{t_0-1} \cdot (T, C) \cdot (T, T)^2 \cdot (T, C)^\omega$$

et

$$\begin{aligned}
R_1(\sigma_1, PP_2) &= (1 - \delta) \left[\sum_{t=1}^{t_0-1} \delta^{t-1} \cdot 2 + \delta^{t_0-1} \cdot 3 + \sum_{t=t_0+1}^{t_0+2} \delta^{t-1} \cdot 1 + \sum_{t=t_0+3}^{\infty} \delta^{t-1} \cdot 3 \right] \\
&= (1 - \delta) \left[\sum_{t=1}^{infy} \delta^{t-1} \cdot 2 + \delta^{t_0-1} - \delta^{t_0} - \delta^{t_0+1} + \sum_{t=t_0+3}^{\infty} \delta^{t-1} \right] \\
&= (1 - \delta) \left[\frac{2}{1 - \delta} + \delta^{t_0-1} (1 - \delta - \delta^2) + \delta^{t_0+2} \frac{1}{1 - \delta} \right] \\
&= 2 + (1 - \delta) \delta^{t_0-1} (1 - \delta - \delta^2) + \delta^{t_0+2}.
\end{aligned}$$

Comme la limite de cette expression est 3 lorsque $\delta \rightarrow 1$, il existe $\delta_0 \in]0, 1[$ tel que, $\forall \delta \in]\delta_0, 1[$

- l'inégalité de la question 12.1 est fautive

- (PP_2, PP_2) n'est pas un équilibre de Nash.

13- Le couple de stratégies (PP_2, PP_2) n'est donc pas un équilibre de Nash *uniforme* de la famille $(\Gamma^\delta)_{\delta \in]0, 1[}$.

14- Pour aucune valeur de $k!$ On a déjà vu aux questions 10 et 12 que pour $k \in \{1, 2\}$, (PP_k, PP_k) n'est pas un équilibre de Nash uniforme.

Pour $k \geq 3$, en reprenant la stratégie de la question 12 on construit une stratégie σ_1 telle que

$$h(\sigma_1, PP_k) = (C, C)^{t_0-1} \cdot (T, C) \cdot (T, T)^k \cdot (T, C)^\omega$$

et

$$\begin{aligned}
R_1(\sigma_1, PP_k) &= (1 - \delta) \left[\sum_{t=1}^{t_0-1} \delta^{t-1} \cdot 2 + \delta^{t_0-1} \cdot 3 + \sum_{t=t_0+1}^{t_0+k} \delta^{t-1} \cdot 1 + \sum_{t=t_0+k+1}^{\infty} \delta^{t-1} \cdot 3 \right] \\
&= (1 - \delta) \left[\sum_{t=1}^{infy} \delta^{t-1} \cdot 2 + \delta^{t_0-1} - \delta^{t_0} - \delta^{t_0+1} - \dots - \delta^{t_0+k-1} \sum_{t=t_0+k+1}^{\infty} \delta^{t-1} \right] \\
&= (1 - \delta) \left[\frac{2}{1 - \delta} + \delta^{t_0-1} (1 - \delta - \delta^2 - \dots - \delta^k) + \delta^{t_0+k} \frac{1}{1 - \delta} \right] \\
&= 2 + (1 - \delta) \delta^{t_0-1} (1 - \delta - \dots - \delta^k) + \delta^{t_0+k}.
\end{aligned}$$

De nouveau, il existe $\delta_0 \in]0, 1[$ tel que, $\forall \delta \in]\delta_0, 1[$

- l'inégalité de la question 12.1 est fausse

- (PP_k, PP_k) n'est pas un équilibre de Nash .

Modifions maintenant la définition de PP_k pour le joueur i :

$$\begin{aligned}
PP_k(h^1 \dots h^{t-1}) &= T \text{ si } \exists t_0 \in [t - k, t - 1], [h_i^{t_0} = C \text{ et } h_{3-i}^{t_0} = T] \\
PP_k(h^1, \dots, h^{t-1}) &= C \text{ sinon}
\end{aligned} \tag{3}$$

Cette modification n'influence pas, pour la stratégie σ_1 de la Q10, la partie $h(\sigma_1, PP_1)$. Donc pour tout $\delta \in]0, 1[$, (PP_1, PP_1) n'est pas un équilibre de Nash de Γ^δ .

12.1 Reprenons l'analyse de (PP_2, PP_2) . Soit h une partie où $J2$ joue la stratégie PP_2 (à tous les tours), et $J1$ joue la stratégie PP_2 à tous les tours d'ordre $t \leq t_0 - 1$ mais trahit au tour $t_0 \in [1, \infty[$. On définit une suite t_i par : t_{i+1} est le premier tour $\geq t_i + 3$ où $J1$ trahit.

On a alors

$$R_1(h) \leq (1 - \delta) \left[\sum_{t=1}^{t_0-1} \delta^{t-1} \cdot 2 + \sum_{i \in I} \delta^{t_i-1} (1 - \delta - \delta^2) \right]$$

(l'égalité est atteinte lorsque $J1$ joue T aux tours $t_i, t_i + 1, t_i + 2$ et C à tous les autres tours).

Donc, si $\delta \geq \varphi$

$$\begin{aligned}
R_1(h) &\leq (1 - \delta) \left[\sum_{t=1}^{t_0-1} \delta^{t-1} \cdot 2 + \sum_{i \in I} \delta^{t_i-1} (1 - \delta - \delta^2) \right] \\
&= (1 - \delta) \left[\frac{2}{1 - \delta} + \sum_{i \in I} \delta^{t_i-1} (1 - \delta - \delta^2) \right] \\
&= 2 + (1 - \delta) \sum_{i \in I} \delta^{t_i-1} (1 - \delta - \delta^2). \\
&\leq 2 + (1 - \delta) \delta^{t_0} (1 - \delta - \delta^2).
\end{aligned}$$

12.2 On en conclut que, pour $\delta \geq \varphi$, (PP_2, PP_2) est un équilibre de Nash. Donc (PP_2, PP_2) est un équilibre de Nash *uniforme* de la famille $(\Gamma^\delta)_{\delta \in]0, 1[}$.

14- Pour $k \geq 3$ on adapte le raisonnement de la question 12. Soit h une partie où $J2$ joue la stratégie PP_k (à tous les tours), et $J1$ joue la stratégie PP_k à tous les tours d'ordre $t \leq t_0 - 1$ mais trahit au tour $t_0 \in [1, \infty[$. On définit une suite t_i par : t_{i+1} est le premier tour $\geq t_i + k + 1$

où $J1$ trahit.

On a alors

$$R_1(h) \leq (1 - \delta) \left[\sum_{t=1}^{t_0-1} \delta^{t-1} \cdot 2 + \sum_{i \in I} \delta^{t_i-1} (1 - \delta - \dots - \delta^k) \right]$$

ce qui conduit, lorsque $\delta \geq \varphi$ à :

$$\begin{aligned} R_1(h) &\leq 2 + (1 - \delta) \delta^{t_0} (1 - \delta - \dots - \delta^k) \\ &\leq 2 + (1 - \delta) \delta^{t_0} (1 - \delta - \delta^2) \\ &< 2. \end{aligned}$$

Donc (PP_k, PP_k) est un équilibre de Nash *uniforme* de la famille $(\Gamma^\delta)_{\delta \in]0,1[}$ pour $k \geq 2$.