Série 3 : Dynamiques multi-agents

Exercice 1: Dynamique de meilleures réponses

On considère une famille composée d'un père (p), d'une mère (m), d'une fille (f) et d'un garçon (g). Chacun doit prendre une décision binaire, aller au cinéma (oui ou non). Leur décision dépend de celles des autres. Les stratégies des agents sont les suivantes :

— le père préfère toujours prendre la même décision que la mère ;
— le fils s'oppose toujours au père, quelque soit sa décision ;
— la mère préfère choisir comme son fils et sa fille lorsque ceux-ci sont d'accord ; et sinon elle choisit comme le père.
— la fille est complètement indépendante.

1. Indiquez pourquoi les dynamiques de réponses améliorantes et de meilleures réponses sont équivalentes dans ce problème

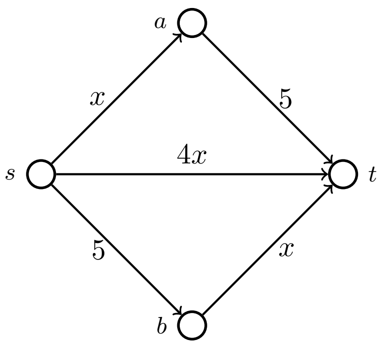
2. Donnez les tables de meilleures réponses des différents agents.

3. Vous considèrerez les deux cas : selon que la fille choisit d'aller ou cinéma ou non. Dans chacun des cas, indiquez si il existe une situation d'équilibre. Indiquez également si une dynamique de meilleure réponse permet, dans chacun des cas, de converger vers une situation d'équilibre.

Exercice 2: Jeu de congestion

On considère un jeu dit de "congestion", dans lequel 6 agents doivent se déplacer sur un réseau pour aller d'un nœud s à un nœud t. Dans ce problème :

— les stratégies pour les agents sont les chemins possibles pour aller de s à t. Une situation de jeu est la collection $S = \langle s_1, \ldots s_6 \rangle$ des stratégies des agents.
— le coût à passer sur un arc est, pour un agent, donné par une fonction linéaire $cout(e, x) = ax+b$ où x est le nombre d'agents empruntant cet arc
— le coût pour un agent i d'emprunter un chemin $\pi = [e_1, \ldots, e_k]$ dans une situation S, est donné par $\sum_{e_j \in \pi} cout(e_j, S(e_j))$, où $S(e_j)$ est le nombre d'agents empruntant l'arc dans S.
— une réponse améliorante consiste à changer son chemin pour un chemin de moindre coût



1. En partant de la situation initiale où tous les agents empruntent le chemin $\pi = [(s, a), (a, t)]$, indiquez la séquence d'actions observées dans une dynamique de meilleures réponses.

2. Exhibez à présent une situation d'équilibre qui peut être atteinte par une dynamique de réponses améliorantes, mais pas pas une dynamique de meilleure réponse.

3. Pour une situation S, considérez la fonction $\phi(s) = \sum_{e \in G} \sum_{i=1}^{S(e)} cout(e, i)$, où rappelons-le $S(e)$ est le nombre d'agents empruntant l'arc e. Calculez sa valeur pour la séquence de meilleures réponses de la question 1. Montrez que la valeur de cette fonction décroît nécessairement après une réponse améliorante. Que pouvez-vous en déduire sur les séquences de réponses améliorantes dans ce problème ?

4. Un concepteur du réseau vous suggère de remplacer l'arc (s, t) par un arc (a, b) de coût fixe nul. Que pouvez-vous dire de la situation où tous les agents empruntent le chemin $[(s, a), (a, b), (b, t)]$?

Exercice 3: Dynamiques avec apprentissage : Fictitious Play et Regret Matching

On considère la situation de jeu donnée par la matrice suivante :

$R\backslash B$	a	b
a	$(3,2)$	$(0,0)$
b	$(0,0)$	$(2,3)$

Il s'agit d'un jeu de coordination : le joueur B a une préférence pour la coordination sur l'action a, le joueur R a une préférence sur l'action b, mais les deux joueurs préfèrent être coordonnés sur la même action que non-coordonnés.

1. Calculer sur 4 itérations les coups joués si les deux agents suivent une stratégie de fictitious play, en partant de l'information a priori suivante : pour R : $(a:3,b:1)$ et pour B : $(a:1,b:2)$. Autrement dit, R dispose d'un échantillon initial où il a observé que B avait joué 3 fois a et 1 fois b, tandis que R dispose d'un échantillon où B a joué 1 fois a et 2 fois b. Si deux actions sont également préférées, on pose par convention que les agents choisissent l'action par ordre alphabétique.

2. Donnez un autre exemple d'information a priori pour lequel la convergence s'effectuera sur un équilibre de Nash différent.

3. Si l'agent R joue 10 fois l'action b, puis constamment la même action a et que l'autre agent B joue en fictitious play, est-il vrai que ce joueur R "borné" aura à la limite (c'est-à-dire, pour un nombre d'itérations suffisamment grand) un gain cumulé plus grand que celui de B ?

4. En supposant que l'agent B joue la séquence $[b,b,a,a]$, indiquez les coups qui seraient joués par l'agent R en suivant une dynamique de *regret matching*. On fera ici l'hypothèse que R débute avec l'action a, et qu'il joue toujours l'action qui a la plus forte probabilité selon la règle vue en cours.