IA04 – Systèmes Multi-Agents

Agents stratégiques - Théorie des jeux

Khaled Belahcene

khaled.belahcene@hds.utc.fr

Hénoïk Willot

henoik.willot@hds.utc.fr

Résumé des épisodes précédents

- Le cadre du *choix social* suppose que les agents *collaborent* pour former une société.
- Cependant, parmi les *règles de décision collective* que cette assemblée peut être amenée à adopter, *aucune* (ou presque) ne garantit *structurellement* que les agents ont intérêt à manifester *sincèrement* leurs convictions
- Aujourd'hui, nous *abandonnons* d'emblée l'hypothèse *collaborative*

Le dilemme du prisonnier

> *Deux suspects (Alice et Bob) sont arrêtés par la police. Mais les agents n'ont pas assez de preuves pour les inculper, donc ils les interrogent séparément en leur faisant la même offre :*
>
> - *"Si tu dénonces ton complice et qu'il ne te dénonce pas, tu seras remis en liberté et l'autre écopera de 10 ans de prison."*
> - *"Si tu le dénonces et lui aussi, vous écoperez tous les deux de 5 ans de prison."*
> - *"Si personne ne dénonce l'autre, vous aurez tous deux 6 mois de prison. »*

Représentation sous forme normale

	Bob se tait	Bob trahit
Alice se tait	Alice : -0.5, Bob : -0.5	Alice : -10, Bob : 0
Alice trahit	Alice : 0, Bob : -10	Alice : -5, Bob : -5

Un *jeu sous forme normale* est un triplet $(N, \mathbf{A}, \mathbf{u})$ où:

- N est un ensemble fini de *joueurs*
- Chaque joueur $i \in N$ dispose d'un nombre fini A_i d'actions possibles, et $\mathbf{A} = A_1 \times \ldots \times A_n$
- $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbf{A}$ est un *profil d'actions*
- *Intéraction stratégique :* Le *gain* $u_i : \mathbf{A} \rightarrow \mathbb{R}$ de chaque joueur dépend des actions choisies par l'ensemble des joueurs
- $u_i(a_1, \ldots, a_n)$ est le *gain* du joueur i étant donné le profil d'actions $(a_1, \ldots, a_n) \in \mathbf{A}$

Que va-t-il se passer ?

	Bob se tait	**Bob trahit**
Alice se tait	Alice : -0.5, Bob : -0.5	Alice : -10, Bob : 0
Alice trahit	Alice : 0, Bob : -10	Alice : -5, Bob : -5

Hypothèses

1. Tout est dans la matrice : pas de perte de réputation, ou de vengeance possible
2. Rationnalité des agents : chacun maximise ses propres gains

Le point de vue d'Alice

- si Bob se tait :

	Bob se tait
Alice se tait	Alice : -0.5
Alice trahit	*Alice : 0*

Alice a intérêt à trahir

- si Bob trahit :

	Bob trahit
Alice se tait	Alice : -10
Alice trahit	*Alice : -5*

Alice a intérêt à trahir

Alice décide de trahir : elle a intérêt à trahir dans tous les cas. On dit que la trahison est une *stratégie strictement dominante* pour Alice.

Premier bilan

- Bien entendu, Bob fait le même raisonnement.
- Il trahit Alice.
- Les anciens complices écopent tous deux de 5 ans de prison. --

> *Bien qu'ils aient collectivement intérêt à coopérer, le comportement individuel "rationnel" (sous une certaine hypothèse de rationnalité) et égoïste des agents peut les conduire dans une situation lamentable.*

Efficacité collective

Etant donné un jeu sous forme normale, on définit une relation binaire entre deux profils :

- $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbf{A}$ *domine faiblement au sens de Pareto* $\mathbf{a}' = (a'_1, \ldots, a'_n) \in \mathbf{A}$ si les gains de tous les agents sont au moins aussi bon pour $\mathbf{a}$ que pour $\mathbf{a}'$
- On dit qu'une solution est *efficace* ssi elle n'est pas strictement dominée au sens de Pareto : il n'existe pas de profil d'actions au moins aussi favorable pour tous les agents, et strictement meilleure pour au moins l'un d'eux

> *On vient de voir que le dilemme du prisonnier admettait un équilibre en solutions (individuellement) dominantes qui s'avère (Pareto-)inefficace*

Pertinence du modèle

Le dilemme du prisonnier capture la difficulté de se coordonner, due à l'impossibilité d'instaurer la confiance

Il peut servir de modèle à de nombreuses situations réelles :

- entreprises qui se livrent une guerre commerciale (par ex. sur les prix, ou sur la publicité)
- Etats qui décident de limiter leurs émissions de gaz à effet de serre
- Colocataires qui se partagent la bande passante du réseau
- Echappées du tour de France
- course aux armements

Cependant, c'est un modèle qui repose sur des hypothèses nombreuses et fortes : information parfaite, actions simultanées, gains déterministes, "rationnalité" égoïste

Existence d'un équilibre

On considère les jeux suivants. Admettent-ils un *équilibre en stratégies dominantes* ?

	L	R
T	2, 2	2, 1
B	1, 3	3, 2

	L	R
T	2, 2	2, 2
B	2, 2	2, 2

	L	R
T	1, 2	2, 1
B	2, 1	1, 2

On est amené à relâcher la définition d'un équilibre. On cherche à modéliser la notion de *stabilité*

Meilleure réponse

Que jouera un joueur "rationnel" qui anticipe l'action de ses adversaires ?

- **Notation :** étant donné un joueur $i \in N$,

$$(a'_i, \mathbf{a}_{-i}) := (a_1, \ldots, a_{i-1}, a'_i, a_{i+1}, \ldots, a_n)$$

  est le profil où tous les agents sauf i jouent a_i et i joue a'_i.

- **Définition :** On dit que $a_i^\star \in A_i$ est une *meilleure réponse* du joueur i au profil (partiel) d'actions $\mathbf{a}_{-i}$ si $u_i(a_i^\star, \mathbf{a}_{-i}) \geq u_i(a'_i, \mathbf{a}_{-i})$ pour toute $a'_i \in A_i$.

- **Définition :** Le profil d'actions $\mathbf{a}$ est un *équilibre de Nash* (en stratégies pures) lorsque chaque joueur joue la meilleur réponse aux actions des autres : pour tout joueur $i \in N$, a_i est une meilleure réponse à $\mathbf{a}_{-i}$

> *Aucun joueur n'a intérêt à dévier unilatéralement de sa stratégie.*

Existence d'un équilibre

On considère (encore) les jeux suivants. Admettent-ils un *équilibre de Nash* ?

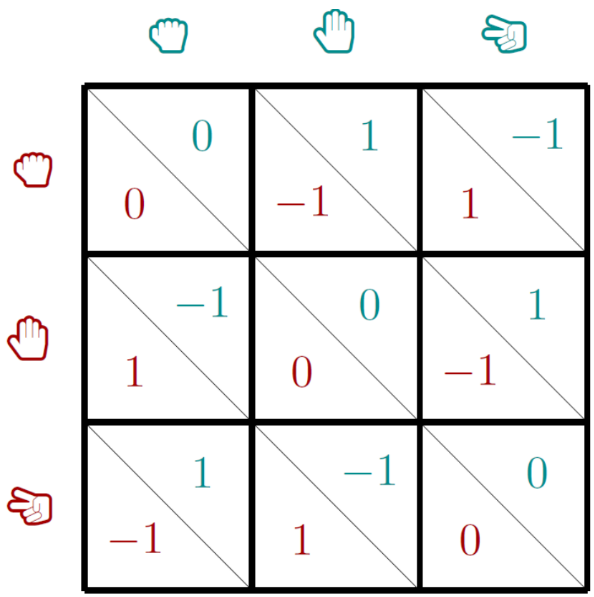
	L	R
T	2, 2	2, 1
B	1, 3	3, 2

	L	R
T	2, 2	2, 2
B	2, 2	2, 2

	L	R
T	1, 2	2, 1
B	2, 1	1, 2

On est (encore) amené à relâcher la définition d'un équilibre. On cherche (toujours) à modéliser la notion de *stabilité*

Intermède ludique



Ceci est un exemple de *jeu à somme nulle* : les gains des joueurs se neutralisent. On parle aussi de jeu *purement compétitif*

- Equilibre(s) de Nash ?
- Comment joueriez-vous ?

Stratégies mixtes

- Jusqu'à présent, les stratégies envisageables par le joueur i se limitaient à l'ensemble de ses actions (on parle de *stratégies pures*)
- On étend ce cadre en autorisant le joueur i à jouer des actions de A_i de manière *probabiliste*
- Pour tout ensemble fini X, on note $\mathcal{P}(X)$ l'ensemble de toutes les distributions de probabilité sur

$$\mathcal{P}(X) = \{p : X \to [0,1] | \sum_{x \in X} p(x) = 1\}$$

- Une *stratégie mixte* s_i une distribution de probabilité dans cet ensemble $S_i = \mathcal{P}(A_i)$
- Un profil de stratégies mixtes s $= (s_1, \ldots, s_n)$ est un élément de $S_1 \times \ldots \times S_n$
- l'*utilité espérée* du joueur i pour le profil de stratégies mixtes $\mathbf{s}$ est :

$$u_i(\mathbf{s}) = \sum_{a \in A} u_i(a) \cdot \prod_{j \in N} s_j(a_j)$$

Exemple : la bataille des sexes

Alice\Bob	Ballet	Foot
Ballet	2 \ 4	0 \ 0
Foot	0 \ 0	8 \ 3

Supposons qu'Alice aille voir le Ballet avec une probabilité de 75% et que Bob se rende à coup sûr au football.

- que valent s_{Alice} et s_{Bob} ?
- quelle est l'utilité de chacun des joueurs ?

Meilleure réponse et équilibre de Nash mixtes

Définitions : (l'adaptation stratégies pures $\rightarrow$ stratégies mixtes ne pose aucun problème)

> *Théorème : (Nash, 1951) tout jeu fini sous forme normale admet un équilibre de Nash en stratégies mixtes.*

Schéma de démonstration : preuve *non-constructive*

- on définit une fonction f *continue* correspondant à une mise à jour des stratégies des agents
- on montre que $\mathbf{a}$ est un équilibre de Nash ssi c'est un *point fixe* de f
- on utilise un *théorème du point fixe* (Brouwer) dans le simplexe (qui est convexe et compact) pour f

Calcul d'équilibres de Nash

Pour le jeu de la guerre des sexes, on note $s_{Alice} = (p, 1-p)$ et $s_{Bob} = (q, 1-q)$

Alice\Bob	**Ballet**	**Foot**
Ballet	2 \ 4	0 \ 0
Foot	0 \ 0	8 \ 3

- Pour q donné, quelle est la meilleure réponse d'Alice ?
- pour p donné, quelle est la meilleure réponse de Bob ?
- Représenter les *correspondances de meilleure réponse* dans le plan (p, q)
- Déterminer l'ensemble des équilibre de Nash de ce jeu.

Equilibre corrélé

Pour éviter de se retrouver dans une situation adversariale préjudiciable à tous, les agents peuvent décider avant de jouer de corréler leur stratégie à un *signal extérieur*

Exemples :

- la *guerre des sexes*

Alice\Bob	**Ballet**	**Foot**
Ballet	2 \ 1	0 \ 0
Foot	0 \ 0	1 \ 2

NE= { (A : Ballet, B : Ballet), (A : Foot, B : Foot), (A : (2/3,1/3), B : (1/3,2/3)) }

$\longrightarrow$ situation inégalitaire ou peu satisfaisante

Supposons qu'Alice et Bob décident de *jouer à pile ou face* leur soirée. Utilité espérée ?

Equilibre corrélé

Pour éviter de se retrouver dans une situation adversariale préjudiciable à tous, les agents peuvent décider avant de jouer de corréler leur stratégie à un *signal extérieur*

Exemples :

- le *bras de fer*

Ligne\Colonne	**Forcer**	**Céder**
Forcer	-2 \ -2	2 \ -1
Céder	-1 \ 2	1 \ 1

- voitures qui se font face (game of chicken)
- animaux défendant leur territoire (faucon vs colombe)
- attitude politique

NE = {(Forcer, Céder), (Céder, Forcer) , ((1/2, 1/2), (1/2, 1/2))}

$\rightarrow$ invention des feux de croisement

Autres questions

- calcul des équilibres de Nash dans le cas général

    > - *dans NP (vérification polynomiale)*
    > - *polynomial pour certaines classes de jeu*
    > - *inconnu dans le cas général*

- autres notions d'équilibre
- jeux séquentiels

    > - *représentation arborescente sous forme extensive*
    > - *stratégie (mixte) = une (distribution d') action(s) par noeud*
    > - *on peut toujours se ramener à une forme normale, avec une combinatoire d'actions*

- prix de l'anarchie
- équilibres corrélés
- ...