Jeux Stratégiques et Raisonnement Rationnel

M1 Miage 2015–2016 *Intelligence Artificielle*

Stéphane Airiau

LAMSADE

DAUPHINE
UNIVERSITÉ PARIS

But du cours

- On veut modéliser la décision d'agents rationnels
- La décision d'un agent peut avoir un effet sur d'autres agents

La Théorie des Jeux est un moyen de faire cette modélisation

Dans les exemples suivants, ***deux*** personnes doivent prendre une décision en ***même temps***, ***sans connaître*** la décision de l'autre.

On représentera la situation à l'aide d'une table :

- la table a autant de ligne que le premier joueur à d'actions
- la table a autant de colonne que le second joueur à d'actions
- dans une case, on indique l'utilité du premier joueur (celui qui choisit la ligne) et du second joueur (celui qui choisit la colonne)
- plus l'utilité est grande, plus le joueur aime cet état.

Dilemme du prisonnier

Deux complices Ligne (L) et Colonne (C) sont arrêtés par la police et sont interrogés dans des pièces séparées.
Du point de vue d'un des complices, disons L, quatre situations sont possibles :

- C coopère avec la police et L refuse, la police réduit donc la peine de C à une peine légère.
- C ne coopère pas, mais L si, et donc C reçoit une lourde peine.
- C et L ne coopèrent pas avec la police, qui ne peut donc prouver la culpabilité totale. C et L reçoivent une peine "moyenne".
- C et L coopèrent, chacun reçoit une peine assez lourde.

	C est fidèle	C trahit
L est fidèle	2,2	4,1
L trahit	1,4	3,3

Le jeu du poulet

Dans *Rebel Without a Cause*, Buzz met au défi le personnage joué par James Dean, appelé Jim : ils doivent faire une course avec des voitures volée en se dirigeant vers une falaise. Le premier qui freine ou qui saute de la voiture a perdu.

	Jim continue	Jim freine
Buzz continue	-10,-10	5,0
Buzz freine	0,5	1,1

Cap ou pas cap ?

Deux amis font un pari, par exemple de venir en cours le lendemain dans une tenue ridicule. Si seul un des deux tient sa promesse, il sera ridicule. S'ils la tiennent tous les deux, cela sera un succès. Si personne ne tient sa promesse, personne ne sera embarassé, mais ce serait mauvais pour leur amitié.

	pas cap	cap
pas cap	1,1	2,0
cap	0,2	3,3

Problème du rendez-vous

	G	D
H	2,2	4,3
B	3,4	1,1

- **Problème :** Match de foot ou Opéra avec son partenaire ?
- **Requirements :**
  - être ensemble !
  - profiter du spectacle qui vous convient le plus !

Questions

- Existe-t-il toujours un choix rationnel ?
- Si oui, comment le définir ?
- Si oui, comment le trouver ?

Représentation

Représentation générique du jeu est appelée jeu en forme normale.

Definition (Jeu en forme normale)

Un **jeu en forme normale (NFG)** est $(N, (S_i)_{i\in N}, (u)_{i\in N})$ où

- N est l'ensemble de n joueurs
- S_i est l'ensemble des stratégies/actions du joueur i.
- $u_i : S_1 \times \cdots \times S_n \to \mathbb{R}^n$ est l'utilité du joueur i : étant donné la stratégie de chacun des joueurs, la fonction retourne l'utilité du joueur i

Vocabulaire :

- un élément $s = \langle s_1, \ldots, s_n \rangle$ of $S_1 \times \cdots \times S_n$ est appelé profil des stratégies
- Soient $s \in S_1 \times \cdots \times S_n$ et $s'_i \in S_i$. On écrit (s'_i, s_{-i}) le profil des stratégie qui est le même que s sauf pour le joueur i qui joue la stratégie s'_i, c-a-d $(s'_i, s_{-i}) = \langle s_1, \ldots, s_{i-1}, s'_i, s_{i+1}, \ldots, s_n \rangle$

Que feriez-vous ?

- $N = \{\mathtt{L}, \mathtt{C}\}$
- $S_{Row} = S_{Column} = \{confiance,\ trahison\}$
- u_{Row} et u_{Column} sont définis par.

L \ C	trahison	confiance
trahison	2,2	4,1
confiance	1,4	3,3

Ici, on ne veut pas utiliser le nom des actions (et ce qui est rattaché à ces noms), on veut seulement tenir compte des utilités

Peut-on utiliser un principe général pour choisir une action ?

Stratégie Dominante

Definition (Dominance forte)

Une stratégie $x \in S_i$ pour le joueur i **domine (fortement)** un strategie $y \in S_i$ si le joueur i préfère (strictement) x à y indépendemment de la stratégie employée par les autres joueurs, i.e. $\forall s \in S_1 \times \cdots \times S_n$, $u_i(x, s_{-i}) > u_i(y, s_{-i})$

exemple : dilemme du prisonnier :

L \ C	trahison	confiance
trahison	2,2	4,1
confiance	1,4	3,3

Les deux joueurs ont des stratégies dominantes : trahir !

Du point de vue du joueur ligne L

- si C trahit, L a intérêt de trahir
- si C fait confiance, L a intérêt de trahir !

Stratégie Dominante

Problème du rendez-vous :

	G	D
H	2,2	4,3
B	3,4	1,1

Est-ce que les joueurs ont une stratégie dominante ?

Definition (Meilleurs réponse)

La stratégie s_i du joueur i est une **meilleure réponse** à un profil de stratégie s_{-i} ssi

$$\forall s_i' \in S_i,\ u_i(s_i, s_{-i}) \geqslant u_i(s_i', s_{-i}).$$

Definition (Equilibre de Nash)

Un profil de stratégie $s \in S_1 \times \cdots \times S_n$ est un **équilibre de Nash** si chaque s_i est une meilleure réponse à s_{-i}, c-a-d

$$(\forall i \in N) \left(\forall s_i' \in S_i\right) \; u_i(s_i, s_{-i}) \geqslant u_i(s_i', s_{-i})$$

Un équilibre de Nash est un profil de stratégie pour lequel aucun joueur ne peut améliorer son utilité en changeant seul sa stratégie.

Le problème du rendez-vous à deux équilibres de Nash : $\langle H, D \rangle$ and $\langle B, G \rangle$.

	G	D
H	2,2	4,3
B	3,4	1,1

Equilibre de Nash

$\texttt{L} \setminus \texttt{C}$	trahison	confiance
trahison	2,2	4,1
confiance	1,4	3,3

Il y a un unique équilibre de Nash : les deux joueurs trahissent !

Definition (Etat Pareto optimal)

Un profil de stratégies s est **Pareto optimal** s'il n'existe pas de profil de stratégies s'
$\forall i \in N \ u_i(s') \geqslant u_i(s)$ and $\exists i \in N u_i(s') > u_i(s)$

Un profil de stratégies est Pareto optimal quand on ne peut améliorer le bien-être d'un individu sans détériorer celui d'un autre

Avoir confiance (ne pas trahir) est Pareto optimal.

discussion : Il semble rationnel de trahir ! Cela semble contre-intuitif, car les deux joueurs ont intérêt à avoir confiance l'un en l'autre.

⇨Il y a donc un conflit entre une solution **stable** (l'équilibre de Nash dans lequel aucun joueur n'a intérêt à changer de stratégie) et une solution efficace (qui est Pareto optimale), ici, l'équilibre de Nash est dominée au sens de Pareto.

Existence d'un équilibre de Nash

L \ C	Pile	face
Pile	+1,-1	-1,+1
Face	-1,+1	+1,-1

Ce jeu n'admet pas d'équilibre de Nash.

Récapitulatif

- Quand il n'y a pas de stratégie dominante, un équilibre la meilleure chose qui reste.
- Un jeu n'a pas forcément un équilibre de Nash !
- Si un jeu possède un équilibre de Nash, il n'est pas forcément unique !
- Toute combinaison de stratégie dominante est un équilibre de Nash
- Un équilibre de Nash n'est pas forcément optimal au sens de Pareto
- Deux équilibres de Nash n'ont pas forcément les mêmes utilités

Definition (Stratégie Mixte)

Une stratégie mixte p_i pour un joueur i est une distribution de probabilité sur son espace de stratégie S_i.

exemple :$S_i = \{1,2,3\}$, le joueur décide de jouer la stratégie 1 avec une probabilité $\frac{1}{3}$, strategie 2 avec une probabilité de $\frac{1}{2}$ et la dernière action avec la probabilité $\frac{1}{6}$. Cette stratégie mixte sera notée $\left\langle \frac{1}{3}, \frac{1}{2}, \frac{1}{6} \right\rangle$.

Soit un profil de stratégies mixtes $p = \langle p_1, \ldots, p_n \rangle$, l'utilité espérée pour l'agent i est donc :

$$E_i(p) = \sum_{s \in S_1 \times \cdots \times S_n} \left(\left(\prod_{j \in N} p_j(s_j) \right) \times u_i(s) \right)$$

Problème du rendez vous

	G	D
H	2,2	4,3
B	3,4	1,1

L'utilité espérée du joueur ligne est donc :

$$xy \cdot 2 + x(1-y) \cdot 4 + (1-x)y \cdot 3 + (1-x)(1-y) \cdot 1 = -4xy + 3x + 2y + 1$$

Etant donnée un profil de stratégie mixte $p = \langle p_1, \ldots, p_n \rangle$, on écrit (p'_i, p_{-i}) le même profil de stratégie mixte que p sauf pour le joueur i qui joue la stratégie mixte p'_i, i.e., $(p'_i, p_{-i}) = \langle p_1, \ldots, p_{i-1}, p'_i, p_{i+1}, \ldots, p_n \rangle$.

Definition (Equilibre de Nash en stratégie mixte)

Un **équilibre de Nash en stratégie mixte** est un profil de stratégie mixte p tel que $E_i(p) \geqslant E_i(p'_i, p_i)$ pour chaque joueur i et toute autre stratégie mixte p'_i pour le joueur i.

Problème du rendez-vous

	G	D
H	2,2	4,3
B	3,4	1,1

Soit la stratégie mixte $\langle \frac{3}{4}, \frac{1}{4} \rangle$.
aucun joueur n'a intérêt de changer de stratégie

$$E_{row}(T) = \frac{3}{4} \cdot 2 + \frac{1}{4} \cdot 4 = \frac{5}{2} \qquad E_{row}(B) = \frac{3}{4} \cdot 3 + \frac{1}{4} \cdot 1 = \frac{5}{2}$$

(les joueurs sont indifférents)

Theorem (J. Nash, 1950)

Tout jeu fini en forme normale possède au moins un équilibre de Nash.

note : La démonstration est non-constructive (i.e. elle ne donne pas le moyen de calculer un équilibre de Nash), elle utilise un théorème de point fixe (celui de Brouwer ou celui de Kakutani selon la version) pour garantir l'existence.

J.F. Nash. Equilibrium points in n-person games. in *Proc. National Academy of Sciences of the United States of America*, 36 :48-49, 1950.

Calcul d'un équilibre de Nash

Complexité : C'est un problème difficile, ce problème appartient à une classe appellée PPAD.

Daskalakis, Goldberg, Papadimitriou : **The complexity of computing a Nash equilibrium**, in *Proc. 38th Ann. ACM Symp. Theory of Computing (STOC)*, 2006

Il existe des algorithmes pour calculer des solutions pour certaines classes de jeux.

Y. Shoham & K. Leyton-Brown : **Multiagent Systems**, Cambridge University Press, 2009. (Chapter 4)

Nisan, Roughgarden, Tardos & Vazirani : **Algorithmic Game Theory**, Cambridge University Press, 2007. (chapters 2, 3)

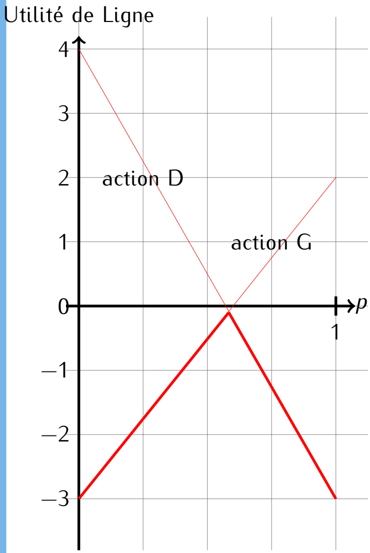
Jeu à somme nulle

La somme des utilités des agents pour une stratégie pure est égale à 0. On ne marque donc qu'un seul nombre (l'autre étant l'opposé !).

	G	M	D
H	30	-10	20
B	10	20	-20

Pour résoudre ce jeu, on peut calculer le MinMax en version probabiliste. On peut plus facilement travailler sur le jeu suivant

	G	D
H	2	-3
B	-3	4



	G	D
H	2	-3
B	-3	4

- Si `C` joue `G`
  `L` obtient $u(p) = 2p - 3 \times (1-p)$
- Si `C` joue `D`
  `L` obtient $u(p) = -3p + 4 \times (1-p)$

Pour maximiser son score, `C` cherche à minimiser l'utilité de `L`

- Il joue donc `G` avant que les droites se croisent
- il joue `D` après

Sachant cela, `L` va maximiser son utilité et jouer le point où les droites se croisent.