IA04 – Systèmes Multi-Agents

Partage de ressources

Khaled Belahcene

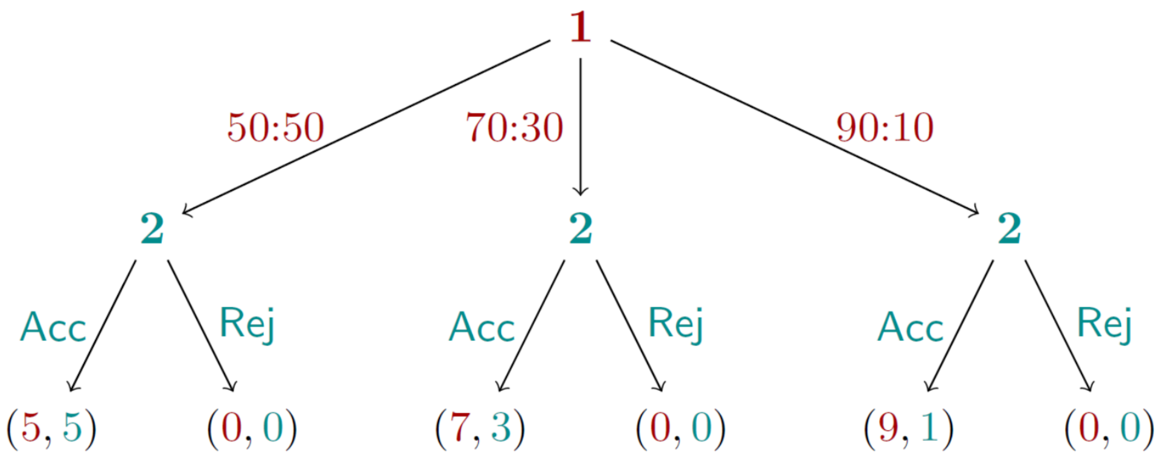
khaled.belahcene@hds.utc.fr

Hénoïk Willot

henoik.willot@hds.utc.fr

Jeux séquentiels

Exemple : Ultimatum !



- Chaque nœud précise le joueur dont c'est le tour
- Pour chaque joueur, une stratégie (pure) consiste à préciser l'action choisie à chaque nœud où c'est son tour de jouer
- On peut réécrire le jeu sous forme normale. Combien de lignes ?
- Quels équilibres de Nash en stratégies pures ?

Équilibre parfait

Définitions :

- Un *sous-jeu* est une partie de l'arbre contenant tous les successeurs de ses nœuds.
- un profil de stratégies $\mathbf{s}$ est un *équilibre parfait* du jeu séquentiel $\mathcal{G}$ lorsque c'est un équilibre pour tout sous-jeu de $\mathcal{G}$

Théorème :(Zermelo 1913, Shelten 1965)

Tout jeu séquentiel fini admet au moins un équilibre parfait en stratégies pures.

Preuve et calcul :

par récurrence, en partant des états terminaux (pour les jeux à 2 joueurs, voir les algorithmes minmax et alpha beta)

Interlude

X	*1*	O
O	O	X
2	*3*	X

Négociation bilatérale

- ressource divisible ?
- agents patients ?
- coopération ?
- rationalité limitée ?
- protocole de négociation

Aujourd'hui :

- ressource *divisible* (gains $(x, 1-x)$ avec $x \in [0,1]$)
- *pas de coopération*, joueurs *stratégiques*
- protocole de négociation : *offres alternées*
  - l'agent *A* commence par faire une offre $(x_0, 1-x_0)$
  - l'agent *B* peut accepter ou refuser et faire une contre-offre $(x_1, 1-x_1)$
  - etc.

Agents patients

- Si on ne fixe aucune limite de temps, les agents n'ont aucune raison de se mettre d'accord
- Si on fixe une limite, il faut prendre en considération une nouvelle action, le *conflit*, qui constitue la pire option pour les 2 joueurs
- Si la limite vaut 1…

> *C'est le jeu d'ultimatum : A propose de partir avec tout le gâteau, et B accepte*

- Si la limite vaut 2…

> *Quoique propose A, B refuse puis se retrouve en position d'ultimatum*

- Si la limite vaut N…

> *c'est le dernier qui joue qui rafle la mise !*

- S'il n'y a *pas de limite*

> *toute issue $(x, 1-x)$ est un équilibre parfait*
>
> *Preuve : considérer successivement deux joueurs constants. La MR est d'accepter immédiatement.*

Joueurs impatients

> *Le temps, c'est de l'argent*
>
> *(Benjamin Franklin, 1748 - Oncle Picsou, 1948)*

- On suppose maintenant que les joueurs préfèrent obtenir la ressource à l'instant t plutôt qu'à l'instant $t+1$
- 2 modèles fondamentaux
  - *coûts fixes* de marchandage, c_A et c_B : $u_i(x,t) = x - c_i t$

  (la patience d'un agent i diminue quand c_i augmente)
  - *taux d'actualisation* τ_A et τ_B : $u_i(x,t) = x\tau_i^t$

  (la patience d'un agent i diminue quand τ_i diminue)

Les équilibres obtenus dépendent de la *patience relative* des agents

Résultats concernant le modèle à coûts fixes

Les résultats dépendent de la *patience relative* des agents

Analyse du modèle à horizon fini

Modèle à taux d'actualisation

- On pose pour simplifier $\delta_A = \delta_B = \delta$
- horizon $n = 1$

> *un seul équilibre parfait :* $(1, 0)$

- horizon $n = 2$

> *un seul équilibre parfait :* $(1 - \delta, \delta)$

- horizon n

> *soit* u_n *la valeur obtenue par* A *au jeu à horizon* n
>
> *on à* $u_{n+1} = 1 - \delta u_n$
>
> *on peut vérifier* $u_n = \sum_{k=0}^{n-1} (-\delta)^k = \frac{1 + (-\delta)^n}{1 + \delta}$
>
> *et on a un unique équilibre parfait* $(u_n, 1 - u_n)$

Modèle à coûts constants

Vu en TD

Analyse du modèle à horizon infini

Résultats obtenus par Rubinstein (1982). Les états à l'équilibre dépendent de la patience relative des agents.

Modèle à taux d'actualisation

- un seul équilibre parfait $\left(\frac{1-\delta_B}{1-\delta_A\delta_B}, \frac{\delta_B(1-\delta_A)}{1-\delta_A\delta_B}\right)$

Modèle à coûts constants

- dans le cas $c_A < c_B$, le seul équilibre parfait est $(1, 0)$

> *B ne peut pas obtenir mieux que 0*

- dans le cas $c_B < c_A$, le seul équilibre parfait est $(c_B, 1 - c_B)$

> *B peut épuiser A mais ne peut pas obtenir mieux que* $1 - c_B$

- dans le cas $c_A = c_B = c$, tout partage $(x, 1 - x)$ où chaque joueur reçoit au moins c est un équilibre parfait

Autres approches

Heuristique : définir un *comportement* des agents susceptible d'en faire de "bons" négociateurs

Axiomatique : quel serait un "bon" partage ?

Enchères

- 1 vendeur, n acheteurs
- 1 objet à vendre
- Qui récupère l'objet ?
- A quel prix ?

Les agents

le vendeur

il n'a pas de préférence, hormis un *prix de réserve* en dessous duquel il ne préfère pas vendre

les acheteurs

- ils ont chacun leur *évaluation* de l'objet
  - évaluation *privée*, liée à la possession ou l'usage de l'objet
  - évaluation *commune*, si l'objet est destiné à être revendu plus tard
- leur utilité est *quasilinéaire* : pour l'acquéreur, gain = évaluation - prix payé

Le protocole

- *à l'anglaise* : le vendeur annonce son prix de réserve, puis les acheteurs peuvent surenchérir à la criée (par incrément $\geq \epsilon$). L'objet va au dernier enchérisseur
- *à l'hollandaise* : le vendeur annonce le prix d'achat en partant d'un prix très élevé, puis de manière descendante. Les acheteurs peuvent l'arrêter à tout moment et acquérir l'objet au dernier prix annoncé
- *enchère fermée* : les acheteurs font chacun une offre secrète, l'objet part au plus offrant

Quelle stratégie pour les acheteurs ?

> *ce n'est pas facile…*
>
> - *annoncer trop bas risque de vous faire perdre l'enchère, même si votre valuation est élevée*
> - *annoncer trop haut risque de conduire à un gain très faible, même si vous gagnez*

- enchères à l'anglaise : surenchérir tant que le prix est inférieur à son estimation v_i
- enchères à la hollandaise ou fermées :
  - essayer d'estimer l'enchère la plus haute de vos concurrents $v^\star$
  - enchérir $v^\star + \epsilon$, si c'est profitable, ou bien $v_i - \epsilon$, sinon (au cas où votre estimation serait erronée)

Un protocole sincère : l'enchère de Vickrey (1961)

> *enchère fermée - le vainqueur est celui qui annonce le prix le plus élevé mais il paie le deuxième prix le plus élevé*

- **Th :** *mécanisme sincère* - du point de vue de chaque acheteur, révéler sa valuation réelle est une stratégie dominante
- **Th :** *équivalence des revenus* - du point de vue du vendeur, les 4 mécanismes sont équivalents dès lors que les acheteurs sont neutres vis-à-vis du risque et que leurs valuations sont i.i.d.

Mais il reste des problèmes

- faible revenu pour le vendeur
- absence de monotonie
- collusion
- utilisation de fausses identités

Un cas particulier de mécanisme d'incitation

Quelle est la contribution du vainqueur de l'enchère à l'utilité collective hors paiement ?

> - *en sa présence : --* u_1
> - *en son absence : --* u_2
> - *sa contribution : --* $u_1 - u_2$
>
> $\rightarrow$ *c'est exactement le rabais qu'il obtient sur le paiement*
>
> *C'est le versement de ce rabais qui garantit la sincérité du mécanisme de l'enchère de Vickrey*

Mechanism design : ce procédé peut être appliqué à d'autres situations de marché, afin d'*inciter les agents* économiques *à adopter un comportement vertueux*

Partage équitable

Des agents *coopératifs* décident de partager un ensemble d'objets entre eux...

> *Prenant ses responsabilités dans la lutte contre le réchauffement climatique, l'UTC décide entre autres mesures de limiter les émissions de GES liées aux déplacements des étudiants.*
>
> *Un budget global de n tonnes équivalent CO_2 est alloué à l'ensemble des étudiants de filière pour se rendre sur leur lieu de stage.*
>
> *Comment le répartir entre les étudiants ?*

Comment faire ?

Approche normative : optimum social

- *périmètre du problème* : une situation économique se caractérise par la répartition des ressources

  > *on représente un partage des ressources sous la forme d'un profil σ, où σ_i représente les ressources allouées à l'agent i*

- *rationalité* : les individus sont les mieux placés pour juger leur propre bien-être

  > *on décrit l'état du monde comme le profil des utilités des agents* $\mathbf{u} = (u_1(\sigma_1), \ldots, u_n(\sigma_n))$

Critères d'optimalité

- maximiser le *bien-être utilitariste* (Bentham, J.S. Mill) : $\sum_i u_i(\sigma_i)$
- maximiser le *bien-être égalitaire* (Rawls) : $\min_i u_i(\sigma_i)$
- maximiser le *bien-être de Nash* : $\prod_i u_i(\sigma_i)$
- minimiser les inégalités (mesurées par exemple par l'*indice de Gini*)
- *efficacité* : il n'est pas possible d'améliorer le sort de tous les agents
- *absence d'envie* : $\forall i \forall j\ u_i(\sigma_i) \geq u_i(\sigma_j)$

Protocole centralisé

- Principe : une *autorité centrale* calcule l'allocation σ permettant d'atteindre l'optimum social

- Obstacles :

  - *élicitation des préférences* des agents
  - vulnérabilité à la *manipulation*
  - *calcul* de l'optimum (facile pour bien-être utilitariste et utilités additives, mais très souvent NP-difficile)

Protocole décentralisé

- Principe : en partant d'une allocation initiale arbitraire, les agents s'*échangent des lots* tant que ces échanges sont *individuellement profitables* pour chacun des agents

- Les échanges peuvent s'accompagner de *transferts monétaires*, qui somment à zéro sur l'ensemble des agents

Théorèmes : (Sandholm, 1998)

- *rationalité collective* : tout échange individuellement profitable augmente strictement le *bien-être utilitariste*

- *convergence* : toute suite d'allocation résultant d'une dynamique de tels échanges converge vers l'*optimum social* défini en termes *utilitaristes*

Limitations :

- cela peut nécessiter des échanges arbitrairement complexes (en termes de nombre d'agents ou de ressources)
- cela peut nécessiter un nombre arbitrairement long d'échanges

Cas particulier du marché immobilier

… ou, plus précisément, de tout marché où il y a autant de ressources que d'agents

Procédure décentralisée : *Top Trading Cycles*

Algorithme proposé par Gale et étudié/publié par Shapley et Scarf en 1974.

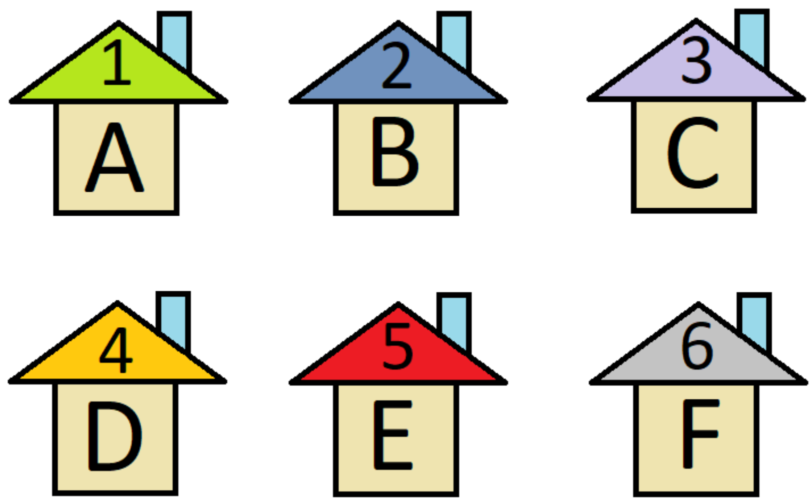
> - *chaque agent pointe vers l'agent possédant son objet préféré*
> - *identifier un cycle d'échange, i.e. une liste d'agents ,* $[a_1, \ldots a_k]$*, avec* $a_k = a_1$ *telle que* a_{i+1} *est pointé par l'agent pointé par* a_i
> - *apparier chaque agent* a_i *du cycle avec l'obet de l'agent qu'il pointe*
> - *retirer les agents (et leurs objets) qui se pointent eux même et recommencer*

Cas particulier du marché immobilier

Exemple

Soient les agents $\{A, B, C, D, E, F\}$ et les maisons $\{1, 2, 3, 4, 5, 6\}$, avec les préférences et l'affectation initiale suivantes :

$\geq_A$	$\geq_B$	$\geq_C$	$\geq_D$	$\geq_E$	$\geq_F$
3	3	3	2	1	2
2	5	1	5	3	4
4	6	6	6	2	5
1	2	5	4	4	6
6	4	2	1	6	3
5	1	4	3	5	1



Cas particulier du marché immobilier

... ou, plus précisément, de tout marché où il y a autant de ressources que d'agents

Procédure décentralisée : *Top Trading Cycles*

Algorithme proposé par Gale et étudié/publié par Shapley et Scarf en 1974.

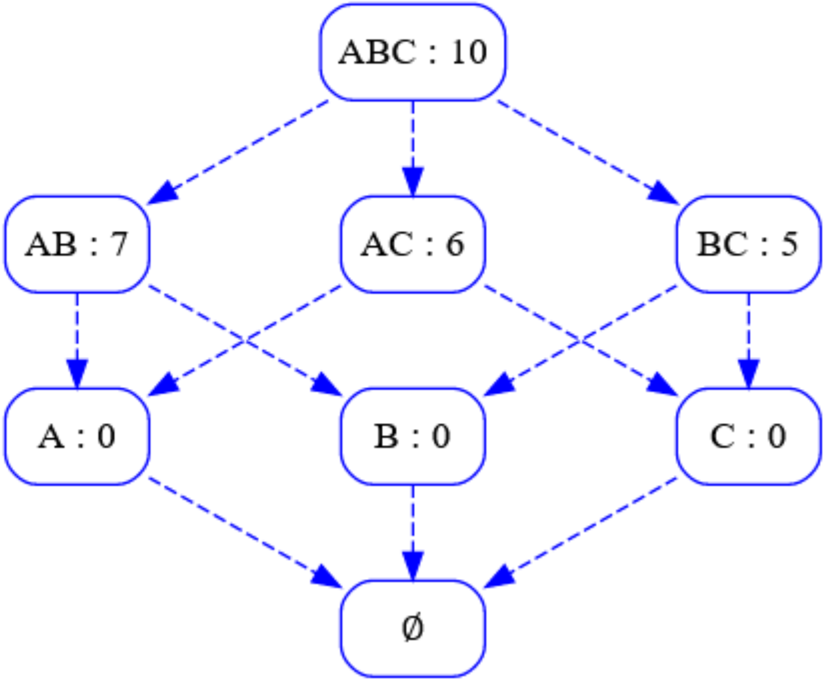
> - *chaque agent pointe vers l'agent possédant son objet préféré*
> - *identifier un cycle d'échange, i.e. une liste d'agents ,* $[a_1, \ldots a_k]$*, avec* $a_k = a_1$ *telle que* a_{i+1} *est pointé par l'agent pointé par* a_i
> - *apparier chaque agent* a_i *du cycle avec l'obet de l'agent qu'il pointe*
> - *retirer les agents (et leurs objets) qui se pointent eux même et recommencer*

- *TTC* est une procédure à la fois *individuellement rationnelle*, *efficace* et *sincère*
- ... et c'est *la seule*
- ... et elle peut nécessiter des cycles arbitrairement longs

Jeux coopératifs

Jeu coopératif (*transferable-utility coalition game*): les actions sont des coalitions d'agents

- Ex 1 : un jeu quelconque



Jeux coopératifs

- Ex 2 : le traité de Rome
  - 6 pays partenaires ayant chacun un nombre de votes : Allemagne (4), France (4), Italie (4), Belgique (2), Pays-Bas (2), Luxembourg (1)
  - une proposition est acceptée ssi elle obtient au moins 12 voix
- **Quelles coalitions se forment ?**

> *Question difficile, à laquelle on répond en introduisant de nouvelles notions de stabilité, par exemple le coeur du jeu, lié à l'existence d'un partage des gains tel qu'aucune sous-coalition n'ait intérêt à quitter la grande coalition N.*
>
> *Une condition nécessaire et suffisante pour la formation de la grande coalition consiste à redistribuer à chaque sous-coalition une valeur au moins égale à ce qu'elle obtiendrait seule. On obtient un programme linéaire (un système d'inéquations).*

- **Comment partager les gains ?**

> *le calcul du cœur étant généralement difficile, on va se focaliser sur une approche heuristique (consistant à proposer des façons de faire), et d'inspecter leurs propriétés*

Approche pragmatique : indices de pouvoir

- *contribution marginale* : lorsque l'agent i rejoint la coalition $C \subseteq N \setminus i$, le surplus augmente de $u(C \cup i) - u(C)$
- principe : donner à chaque agent la valeur "moyenne" de sa contribution
- *valeur de Banzhaf (1965)* : les sous-ensembles de $N \setminus \{i\}$ sont équiprobables

$$\beta_i(N, v) = \frac{1}{2^{|N|-1}} \cdot \sum_{C \subseteq N \setminus i} v(C \cup i) - v(C)$$

- *valeur de Shapley (1953)* : les ordres d'arrivée des joueurs sont équiprobables

$$\phi_i(N, v) = \frac{1}{|N|!} \cdot \sum_{C \subseteq N \setminus i} |C|! \cdot (|N| - |C| - 1)! \cdot (v(C \cup i) - v(C))$$

Propriétés de la valeur de Shapley

La valeur de Shapley vérifie les propriétés suivantes

- **Efficacité** : la totalité du surplus est distribuée, i.e. $\sum_{i \in N} \phi_i(N, v) = v(N)$
- **Symétrie** : des joueurs interchangeables reçoivent le même paiement
- **Joueur nul** : si le joueur i n'a aucune synergie avec les autres (i.e. $v(C \cup i) - v(C) = v(i)$ pour toute coalition $C \subseteq N\ i$) alors il reçoit sa valeur individuelle $\phi_i(N, v) = v(i)$
- **Additivité** : $\phi_i(N, v_1 + v_2) = \phi_i(N, v_1) + \phi_i(N, v_2)$

et c'est la seule à vérifier les 4 (Shapley 1953).

Par ailleurs, si le jeu est *convexe*, (i.e $v(C \cup C') \geq v(C) + v(C') - v(C \cap C')$ pour toutes coalitions C et C'), la valeur de Shapley se situe dans le cœur du jeu : elle est *stable*