Agents et Théorie des jeux

Joël Quinqueton
LIRMM
Montpellier, France

Le besoin de communication

- Le dilemme du prisonnier
  - chaque prisonnier peut avouer ou non
  - si aucun n’avoue : 2 ans
  - si les 2 avouent: 4 ans
  - si un seul avoue: il est libre et l’autre a 5 ans
- Stratégies sur des longues suites:
  - populations de stratégies
  - génération proportionnelle au score de la génération précédente
  - stratégie “donnant donnant” dominante

Jeux à 2 joueurs

	J2: d1	J2: d2
J1: d1	x,x	u,v
J1: d2	v,u	y,y

- Jeux symétriques
- Jeu d'accord social ssi:
  - x ou y >0 et , si x et y >0, alors x=y
  - u ou v négatif

Dilemme des prisonniers

	J2: d1	J2: d2
J1: d1	-2,-2	-5,0
J1: d2	0,-5	-4,-4

- x = -2, y = -4
- u = -5, v = 0
- Ce n’est pas un jeu d’accord social

Jeu de coordination

	J2: d1	J2: d2
J1: d1	1,1	-1,-1
J1: d2	-1,-1	1,1

- $x = y = 1$
- $u = v = -1$
- C’est un jeu d’accord social

Jeu de coopération

	J2: d1	J2: d2
J1: d1	1,1	-3,3
J1: d2	3,-3	-2,-2

- $x = 1$, $y = -2$
- $u = -3$, $v = 3$
- C’est un jeu d’accord social

Equilibre de Nash

- Un jeu non coopératif décrit un cadre institutionnel dans lequel chaque joueur arrête seul ses choix sans consulter les autres joueurs.
- L'équilibre de Nash décrit une issue d'un jeu non coopératif dans lequel aucun joueur n'a intérêt à modifier sa stratégie, compte tenu des stratégies des autres joueurs.

Définition

- Soit un jeu non-coopératif à n joueurs, et $s^*= (s^*_1,...,s^*_n)$ une combinaison de choix stratégiques de ces n joueurs
  - s^*_i est le choix stratégique du joueur i
  - $s^*_i \in S_i$, l'ensemble des stratégies praticables par le joueur i.
  - $u_i (s^*_1,...,s^*_n)$ est le gain du joueur i lorsque s* est sélectionné.
- $\forall s_i \in S_i \; u_i (s^*_1,\ldots s^*_i,\ldots s^*_n) > u_i (s^*_1,\ldots s_i,\ldots s^*_n)$

Optimum de Pareto

- Préférence au sens de Pareto $>_P$:
  - Entre les états des agents d'un jeu
  - $(u_1, \ldots u_i, \ldots u_n) >_P (u'_1, \ldots u'_i, \ldots u'_n)$ ssi $\forall i \in [1,n]$, $u_i \geq u'_i$
- Optimum de Pareto: maximum de cette relation
- on ne peut augmenter le gain d'un agent sans diminuer celui d'un autre

Dilemme des prisonniers

	J2: d1	J2: d2
J1: d1	-2,-2	-5,0
J1: d2	0,-5	-4,-4

- équilibre de Nash: (d2,d2)
- optimum de Pareto: (d1,d1)
- Pour l'atteindre: suite de coups avec mémoire

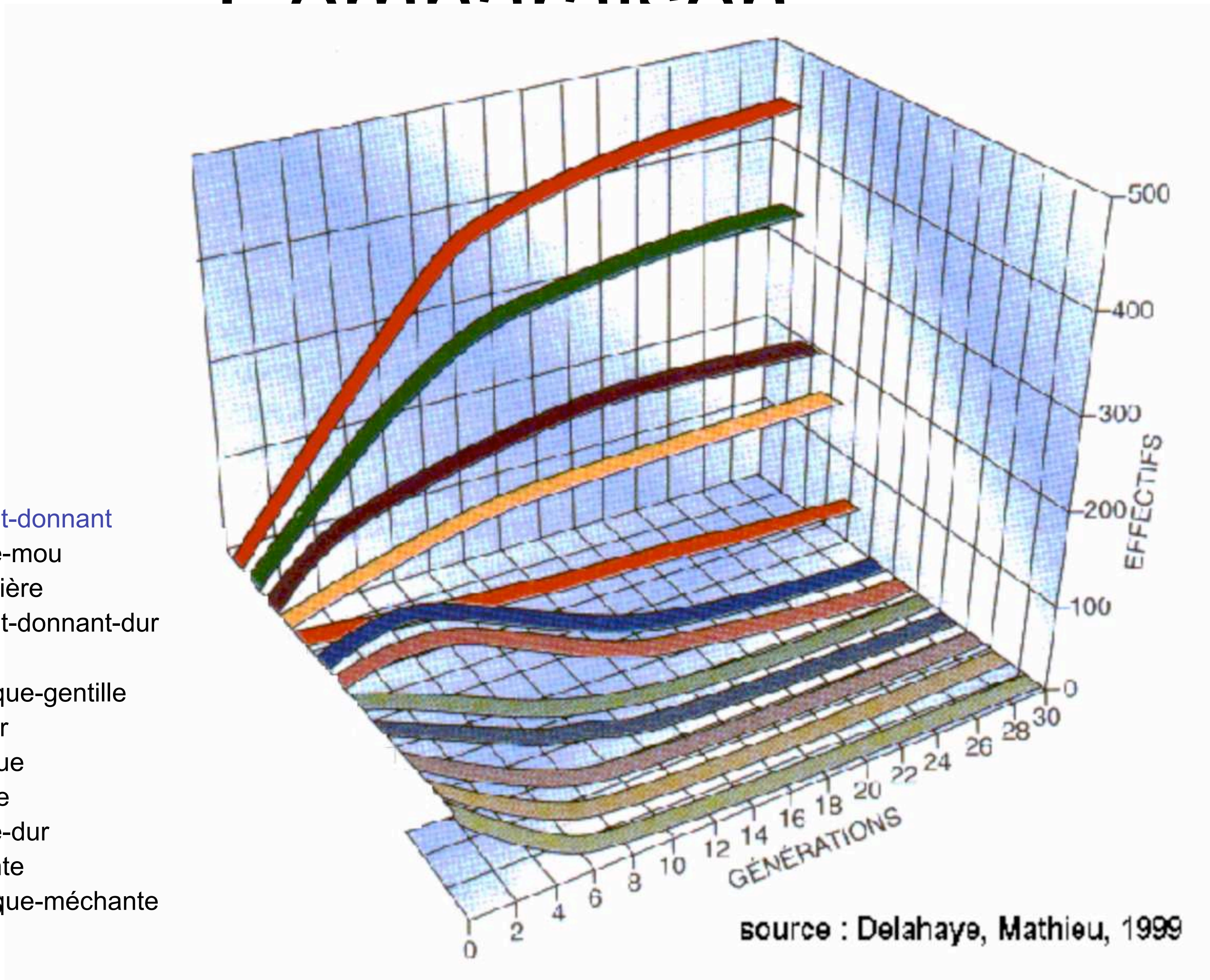
Stratégies sur des suites (1)

- **Donnant-donnant** : coopération au premier tour, puis stratégie précédente du partenaire.
- **Majorité mou** : choix majoritaire de l'adversaire, coopération si égalité et au premier tour.
- **Rancunière** : coopération, puis défection permanente si le partenaire fait une fois défection
- **Donnant-donnant dur** : coopération, sauf si le partenaire a trahi une des 2 fois précédentes
- **Gentille** : toujours coopérer
- **Périodique gentille** : séquence cyclique de deux coopération, puis une défection

Stratégies sur des suites (2)

- **Sondeur** : séquence trahir, coopérer, coopérer
- **Lunatique** : défection une fois sur deux en moyenne (séquences aléatoires)
- **Méfiante** : défection au premier tour, puis stratégie précédente du partenaire
- **Majorité dur:** choix majoritaire de l'adversaire, défection si égalité et au premier tour.
- **Méchante** : toujours faire défection
- **Périodique méchante** : séquence cyclique de deux défections, puis une coopération

Comparaison



- Donnant-donnant
- Majorité-mou
- Rancunière
- Donnant-donnant-dur
- Gentille
- Périodique-gentille
- Sondeur
- Lunatique
- Méfiante
- Majorité-dur
- Méchante
- Périodique-méchante

Tournoi du 50e anniversaire

- Chaque équipe pouvait soumettre plusieurs programmes
- Équipe Nick Jennings (Southampton)
- Stratégie = séquence de 10 coups pour se reconnaître
  - Si oui: stratégie maitre esclave
  - Si non: stratégie méchante

Combiner les préférences de plusieurs agents

D'après Toby Walsh et al

How do we combine preferences *fairly*?

Organiser un dîner

- Les agents ont des préférences individuelles
  - Alice & Bob préfèrent le poisson à la viande
  - Carol préfère l viande au poisson
- Le préférences peuvent être conditionnelles
  - S'il y a du poisson, Alice préfère le vin blanc au vin rouge
  - S'il y a de la viande, Alice préfère le vin rouge au vin blanc

Organiser un dîner

- Plusieurs notions d'optimalité
  - La viande est Pareto-optimale
    - La changer en poisson serait préjudiciable à Carol
  - Le poisson est majoritairement optimal
    - Une majorité d'invités préfère le poisson

Aggrégation de préférences

- Représenter les préférences de chaque agent
  - Réseaux CP partiels
  - Contraintes douces (floues)
  - …
- Chaque agent vote
  - Est-ce que A > B?
- Comment additionner les votes?
  - Organiser une élection!

La sémantique du vote

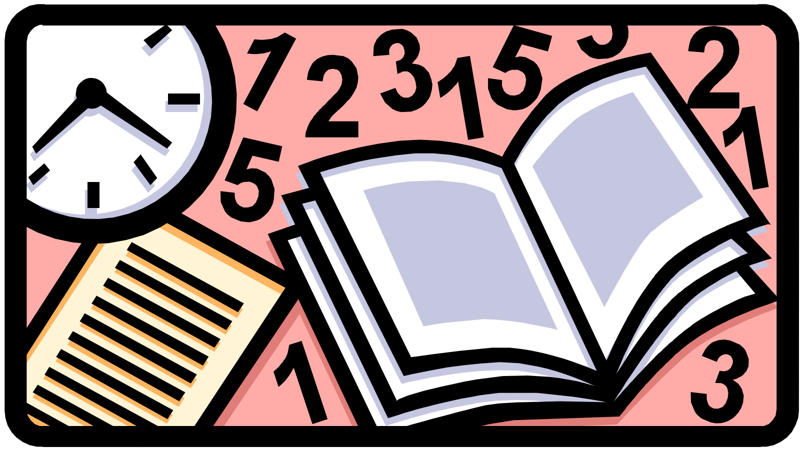
- Ordre de Pareto
  - A >p B ssi A>B or A et B indifférents pour tous les agents
- Ordre majoritaire
  - A >maj B ssi
    #meilleur > (#pire + #incomparable)
  - On peut aussi ignorer les agents qui sont indifférents
- Ordre max
  - A >max B ssi
    # meilleur > max(# pire, #incomparable)

La sémantique du vote

- Ordre lexicographique
  - A $>_{lex}$ B ssi
    - Pour l’agent 1, A>B
    - Ou l’agent 1 est indifférent entre A et B et pour l’agent 2, A > B
    - Ou …
- Ordre des rangs
  - A $>_r$ B ssi somme des rangs(A) < somme des rangs(B)
  - Rang = nombre minimum de permutations empirantes pour atteindre l’optimum

Propriétés de base

- Ordres
  - $>_p$ et $>_{lex}$ sont des ordres partiels stricts
  - transitifs, irreflexifs et antisymétriques
  - $>_{maj}$ et $>_{max}$ ne le sont pas
  - Uniquement irreflexifs et antisymétriques
  - $>_r$ est un ordre total



Propriétés de base

- Optimality

  A est >-optimal ssi il n’y a pas de B avec B > A

- Existence
  - Des résultats Pareto-optimal, majoritaire-optimal, max-optimal, lex-optimal, rang-optimal existent toujours

Equité

- Est-ce que nous combinons *équitablement* les préférences des agents?
  - Est-ce que le théorème d'Arrow s'applique?
  - Pas directement, les votes (et resultats) peuvent inclure l'incomparabilité

Cinq propriétés d'équité

- Libre
  - Tous les orders finaux sont possibles
- Transitive
- Indépendant d'alternatives non pertinentes
  - L'ordre final de 2 valeurs dépend seulement du vote des agents sur ces 2 valeurs
- Monotone
  - Un agent changeant B>A ou B et indifférents en A>B fait que A est davantage préféré
- Non-dictatorial
  - L'ordre final dépend de plus d'un agent

Le théorème d’Arrow

- Libre
- Transitive
- Indépendant d’alternatives non pertinentes
- Monotone
- Non-dictatorial

  – Le théorème d’Arrow dit essentiellement qu’aucune agrégation produisant un ordre *total* ne peut satisfaire les 5 propriétés d’équité

  – Kenneth Arrow (Prix Nobel d’Economie 1972) prouva ce théorème dans sa thèse
  – Cf livre de 1951 *Social Choice and Individual Values (Wiley and son).*
  – de "A Difficulty in the Concept of Social Welfare", The Journal of Political Economy, Volume 58, Issue 4 (August, 1950), pages 328-346

Généraliser le théorème d'Arrow

- Considérer un système de vote à partir d'un ensemble d'ordres *totaux* et produisant un ordre *partiel*
  - Type de vote plus spécifique
  - Type de résultat plus général

- On peut montrer qu'un tel système de vote ne peut être équitable
  - De là, tout système de vote avec un type de vote plus général et un type de résultat plus spécifique ne peut être équitable

Généraliser le théorème d'Arrow

- Le résultat est un peu surprenant
- On peut toujours utiliser l'incomparabilité pour résoudre le conflit
  - Si 50% des agents disent A>B, et les autres B>A alors on peut déclarer A et B incomparable s
- Mais il y a des limites à l'équité que l'on peut atteindre
  - Si 99% des agents disent A>B, et les autres B>A alors ne peut-on déclarer A>B et *équitablement* compensaer les 1% sur d'autres paires?

Dictatures

- Dictateur fort
  - agent dont les préférences sont toujours l'ordre final
- Dictateur
  - agent qui, s'il dit que A>B, alors A>B est dans l'ordre final
- Dictateur faible
  - agent qui, s'il dit que A>B, alors B>A n'est jamais dans l'ordre final

Dictatures

- Dictateur fort
  - Même un système inéquitable comme l'ordre lexicographique n'a pas de dictateur fort
  - Si le 1er agent dit que A et B sont incomparables, alors le 2ème agent peut les ordonner
- Dictateur
  - L'ordre de Pareto n'a pas de dictateur
- Dictateur faible
  - Contrairement aux dictateurs (forts), non nécessairement unique

L’agrégation de préférences ne peut être équitable

- Soit un système de vote depuis des ordres totaux vers un ordre partiel
- Il y a 2 votants et 3 choix
- Le système de vote est libre, monotone et insensible aux alternatives non pertinentes
- Alors il y a au moins un dictateur faible

D’où le théorème d’Arrow sur les ordres totaux

Some examples

- Ordre de Pareto
  - Tous les agents sont des dictateurs faibles
- Ordres majoritaire et maximum
  - Non transitifs
- Ordre lexicographique
  - Le premier agent est un dictateur faible
  - En fait, le premier agent est un dictateur
- Ordre des rangs
  - Non indépendant des alternatives non pertinentes

Une application

- Planification de réunion
  - Contraintes fortes
    - E.g. la réunion ne peut être un dimanche
  - Contraintes faibles
    - E.g. la réunion doit être dans un lieu proche
  - Préférences
    - E.g. dates et lieux que les utilisateurs préfèreraient

Conclusions

- Représentation des préférences
  - Méthodes par facteurs
  - Méthodes quantitatives comme les contraintes faibles
  - Méthodes qualitatives comme les réseaux CP
- Combiner préférences & contraintes
  - Compiler des preferences en contraintes fortes
- Combiner les préférences de plusieurs agents
  - On peut généraliser le théorème d'Arrow pour montrer que ce ne peut être équitable