COURS DE THEORIE DES JEUX

Shmuel ZAMIR

CNRS, EUREQua Paris 1 et LEI/CREST

En collaboration avec

Rida LARAKI

CNRS, Laboratoire d'Econométrie de l'Ecole Polytechnique

Les auteurs tiennent à remercier Thomas Boulogne et Maximilien Laye, doctorants au Laboratoire d'Econométrie de l'Ecole Polytechnique, pour leurs corrections et remarques.

Contents

Chapter 1

Introduction

La théorie des jeux est un champ des mathématiques qui a pour objet d'établir et d'étudier les principes et les règles mathématiques pouvant intervenir dans l'analyse des différents types de comportement et des issues possibles lors d'une interaction stratégique entre plusieurs preneurs de décisions (appelés agents en économie et joueurs en théorie des jeux).
En effet, dans le vie de tous les jours, des preneurs de décisions (hommes politiques, consommateurs, producteurs, comités d'entreprise, traders, citoyens) ont à faire un choix parmi plusieurs actions possibles. Dans un grand nombre de problèmes décisionnels, au moins les deux premiers des aspects suivants sont présents :

- il y a au moins deux preneurs de décision;
- il y a une interaction entre les décisions dans le sens où l'issue finale pour un des preneurs de décision dépend non seulement de l'action qu'il a choisie mais aussi des actions choisies par d'autres décideurs;
- il y a un ou plusieurs éléments d'incertitude.

Un des buts de la théorie des jeux est d'abord de créer des modèles mathématiques de base. Ces modèles essaient de synthétiser tous les éléments essentiels pour décrire l'interaction, puis d'introduire des concepts de solution pour décrire les issues possibles d'un jeu, et enfin, d'appliquer ces outils pour mieux comprendre les phénomènes sociaux mais aussi pour prédire les conséquences d'une interaction stratégique.

Il est facile de citer plusieurs exemples en sciences sociales caractérisés par les aspects présentés ci-dessus et utilisant naturellement la théorie des jeux comme outil de modélisation et d'analyse.

Economie

- dans un marché : le prix d'un bien dépend des décisions de tous les intervenants (pouvoirs public, producteurs, consommateurs) ;
- en situation d'oligopole : à une demande fixée des consommateurs, ce sont les choix de tous les producteurs en compétition qui vont déterminer le prix d'un bien ;
- le marchandage entre acheteur et vendeur ;
- dans une enchère publique ou privée : l'issue (i.e. l'agent qui a obtenu l'objet et le prix qu'il a payé pour l'acquérir) est déterminée par les actions de tous les enchérisseurs, et aussi par le type d'enchère utilisée par l'organisateur.

Science Politique et Relations Internationales

- la compétition électorale entre les candidats pour les différents postes;
- La compétition entre les partis politiques pour gagner les éléctions;
- le vote législatif, le vote pour une politique extérieure ou une politique économique;
- la négociation entre pays afin d'établir des accords internationaux: cessation des hostilités, traités de désarmement, accords économiques et commerciaux...

Droit

- La compétition entre le défendeur et la partie adverse: comment construire son dossier, à quels moments faut-il révéler un argument, quels sont les faits qu'il faut cacher et ceux qu'il faut dévoiler, comment induire l'autre partie en erreur et comment utiliser l'argument de l'autre contre lui mais aussi faire attention à ne pas utiliser un argument qui risque de se retourner contre soi...
- Quand une loi doit être adoptée, il faut analyser toutes les possibilités d'utilisation de cette loi et les différentes conséquences de chaque utilisation (par exemple la loi sur la présomption d'innocence, le principe de précaution, la mise en garde à vue...)

Politique et relations internationales, biologie, ...

Chaque situation décrite plus haut fait appel à un modèle, un type de comportement et une solution adéquats. Voici quelques exemple d'applications réelles de quelques concepts introduits dans ce cours :

L'équilibre de Nash

C'est un concept de solution (ou plus précisément une propriété d'équilibre) pour les jeux non coopératifs.

Si aucun joueur n'a intérêt à changer individuellement sa stratégie en connaissant les stratégies choisies par les autres joueurs alors le vecteur (le profil) des stratégies est appelé un *équilibre de Nash*.

Cette notion a permis d'étudier et d'analyser :

- la compétition oligopolistique en économie;
- les conflits entre états, partis politiques, personnes;
- le rapport entre la longueur de la trompe d'une abeille et la longueur du pistil d'une fleur.

Les modèles de jeux répétés

Ils permettent de mettre en évidence (et de comprendre) l'apparition de certains comportements et phénomènes lors d'une interaction prolongée dans le temps (une situation qui se répète) :

- les menaces, les promesses, les punitions et les phénomènes de réputation;
- la corrélation entre les comportements des acteurs dans les marchés financiers;
- l'émergence des normes sociales (telle que le comportement coopératif entre agents individualistes);

La valeur de Shapley

C'est un concept de solution pour les jeux coopératifs.

Imaginons une situation où des individus peuvent former des coalitions (**ex.** des partis politiques). Chaque coalition peut engendrer un certain gain (**ex.** le nombre de sièges obtenus aux législatives). Un agent (**ex.** un parti politique) ne peut participer qu'à une et une seule coalition.

La question théorique qui se pose est de définir un partage de la valeur (**ex.** le nombre de ministres pour chaque parti dans la coalition majoritaire). Ce partage du gain devrait respecter l'apport de chaque agent dans la coalition.

La valeur de Shapley est la première notion répondant à cette problématique. Elle est aussi l'une des plus appliquées ; on l'a par exemple utilisée pour calculer:

- l'indice de pouvoir dans un comité comme par exemple le conseil de sécurité des Etats Unis d'Amérique;
- l'impact d'un citoyen et d'un état dans l'élection présidentielle américaine;
- l'allocation des coûts dans divers domaines publics tels que la télévision, les aéroports... ; le prix des appels téléphoniques dans un système de communication.

Part I

Préférences

Chapter 2

Représentation ordinal, fonction d'utilité

La théorie de la décision modélise le comportement d'un agent face à des situations de choix entre plusieurs alternatives. Par exemple vous avez le choix entre plusieurs menus repas composé chacun d'une entrée, un plat et un dessert. Nous verrons sous quelles conditions les préférences peuvent être représentées par une fonction à valeur réelle dite *fonction d'utilité.* Parfois, vous avez aussi à choisir entre plusieurs possibilités où le résultat peut être aléatoire (par exemple, jouer au Lotto, jouer à la roulette dans un Casino où ne pas dépenser votre argent). Nous établissons aussi les conditions nécessaires et suffisantes sur les préférences entre les lotteries sous lesquelles un agent se comporte exactement comme s'il maximisait l'espérance mathématique d'une fonction d'utilité.

2.1 Préférences rationnelles

Nous partons d'un ensemble fini X d'alternatives. Nous allons définir une *relation de préférence faible* $\succeq$ sur X, ainsi que des propriétés simples qui font de $\succeq$ une *relation de préférence rationnelle.*

Définition 1 *Une relation de préférence faible est une relation binaire $\succeq$ sur X. C'est-à-dire la donnée d'un ensemble de paires d'éléments (x, y) de $X \times X$ pour lesquels nous écrivons $x \succeq y$.*

$x \succeq y$ est interprétée par : "l'agent préfère faiblement x à y". On écrira $x \not\succeq y$ pour dire que x n'est pas faiblement préféré à y (ce qui n'implique pas que y est faiblement préféré à x).

A partir d'une relation de préférence faible, on peut définir deux relations $\succ$ et $\approx$:

- On dira que $x \approx y$ si $x \succeq y$ et $y \succeq x$.
- On dira que $x \succ y$ si $x \succeq y$ et $y \not\succeq x$.

$x \approx y$ se lit, l'agent est indifférent entre x et y. $x \succ y$ est interprétée : l'egent préfère strictement x à y.

Une relation de préférence faible doit vérifier des bons axiomes pour être appelée rationnelle (à vous de voir si vous trouvez que les préférences d'un agent rationnel devraient les vérifier).

- **Complétude :** Pour tout x et y, au moins une des deux propositions suivantes est vraie : $x \succeq y$ ou $y \succeq x$.
- **Transitivité :** Pour tout x, y et z, si $x \succeq y$ et $y \succeq z$ alors $x \succeq z$.

La complétude exige que pour tout couple d'alternatives (x, y) l'agent est toujours capable de dire, je préfère strictement x à y, ou bien je préfère strictement y à x ou sinon, je suis indifférent entre x et y. On exclut donc les cas où l'agent est incapable de comparer entre deux alternatives. La transitivité exige un peu plus. Cela semble être logique à première vue mais détrompez vous. En effet, supposer que vous avez le choix, pour vous rendre à votre travail, entre trois alternatives : la voiture, le bus et le métro. Si la différence du temps de trajet entre le deux moyens de transport est inférieure à 5mn, vous privilégiez le moyen de transport le plus confortable, sinon vous penchez plutôt pour le moyen le plus rapide. Supposons que la voiture est plus confortable que le bus qui est plus confortable que le métro. Supposons aussi que la durée du trajet en voiture est $t_v = 18mn$, en bus la durée est $t_b = 14mn$ et en métro elle est de $t_m = 10mn$. Vous préférez donc la voiture au bus car la voiture est plus confortable et que $t_v - t_b = 4mn < 5mn$. De même vous préférez le bus au métro pour les mêmes raisons. Cependant, vous préférez le métro à la voiture car $t_v - t_m = 8mn > 5mn$! vos préférences dans ce cas ne sont pas transitives.

Définition 2 *Une relation de préférence faible est dite rationnelle si elle est complète et transitive.*

2.2 Représentation ordinale des préférences rationnelles

Supposons que pour comparer les alternatives sur X, notre agent dispose d'une fonction à valeur réelle $u(\cdot)$ sur X telle que $x \succeq_u y$ si et seulement si $u(x) \geq u(y)$. On pourrait interpréter $u(x)$ comme étant l'utilité, ou le plaisir, que l'agent peut retirer de l'alternative x. C'est pour cette raison que u est appelée *fonction d'utilité* et notre agent est appelé : un utilitariste.

Proposition 1 *Soit $u : X \to R$. Alors $\succeq_u$ est une relation de préférence rationnelle.*

Preuve. Car l'ordre faible $\geq$ sur R est complet et transitif. ■

En fait, la réciproque de ce théorème est vraie quand X est fini ou dénombrable.

Théorème 1 *Supposons que X soit fini ou dénombrable, et soit $\succeq$ une relation de préférence faible sur X. Alors, il existe une fonction d'utilité u telle que $\succeq_u$coïncide avec $\succeq$ si et seulement si $\succeq$ est une relation de préférence rationnelle.*

Preuve. Puisque X est dénombrable, il est possible d'indexer ses éléments par l'ensemble des entiers : $X = \{x_1, x_2, ..., x_n, ...\}$. Notons $X_n = \{x_1, ..., x_n\}$. Remarquez que puisque $\succeq$ est rationnelle sur X, elle l'est aussi sur X_n et ce pour tout n.

On va construire u par récurrence - étape par étape-. On va donc supposer qu'on a déjà construit les valeurs de u pour tout $x \in X_n$ de telle sorte que $\succeq = \succeq_u$ sur X_n puis on utilisera cela pour construire une valeur pour x_{n+1} de telle sorte que $\succeq = \succeq_u$ sur X_{n+1}.

Posons $u(x_1) = 0$. Il est évident (puisque X_1 contient un seul élément) que $\succeq$ et $\succeq_u$ induisent le même ordre sur X_1.

Maintenant, supposons que $\succeq = \succeq_u$ sur X_n. Introduisons alors deux sous ensembles de X_n :

$$\begin{aligned} \overline{X_n} &= \{x \in X_n : x \succeq x_{n+1}\} \\ \underline{X_n} &= \{x \in X_n : x_{n+1} \succeq x\} \end{aligned}$$

Puisque $\succeq$ est complète on a : $\overline{X_n} \cup \underline{X_n} = X_n$ et par transitivité on a, $\forall x \in \overline{X_n}$ et quelque soit $y \in \underline{X_n}$, $x \succeq y$ (ou d'une manière équivalente, par l'hypothèse de récurrence, $u(x) \geq u(y)$).

- **Cas 1 :** $\overline{X_n} = \emptyset$. Ceci veut dire que pour tout $x \in X_n$, $x_{n+1} \succ x$. On pose alors $u(x_{n+1}) = [\max_{x \in X_n} u(x)] + 1$.
- **Cas 2 :** $\underline{X_n} = \emptyset$. Ceci veut dire que pour tout $x \in X_n$, $x \succ x_{n+1}$. On pose alors $u(x_{n+1}) = [\min_{x \in X_n} u(x)] - 1$.

- **Cas 3 :** $\overline{X_n} \neq \emptyset$ et $\underline{X_n} \neq \emptyset$. Il existe alors $\overline{x_n} \in \overline{X_n}$ et $\underline{x_n} \in \underline{X_n}$ tels que

$$\begin{aligned} u(\overline{x_n}) &= \min_{x \in \overline{X_n}} u(x) \\ u(\underline{x_n}) &= \max_{x \in \underline{X_n}} u(x) \end{aligned}$$

  Par définition on a $\overline{x_n} \succeq x \succeq \underline{x_n}$. On a trois sous cas (par complétude) :

  - **Cas 3-1 :** $\overline{x_n} \approx x$, dans ce cas on pose $u(x_{n+1}) = u(\overline{x_n})$
  - **Cas 3-2 :** $\underline{x_n} \approx x$, dans ce cas on pose $u(x_{n+1}) = u(\underline{x_n})$
  - **Cas 3-3 :** $\overline{x_n} \succ x \succ \underline{x_n}$, dans ce cas on pose $u(x_{n+1}) = \frac{u(\overline{x_n})+u(\underline{x_n})}{2}$

Il reste maintenant à vérifier que, dans chacun des cas possibles, notre extension de u à X_{n+1} est telle que $\succeq = \succeq_u$ sur X_{n+1}. Ceci est laissé en exercice au lecteur. ■

Nous pouvons nous poser la question de l'unicité de la représentation d'une relation de préférences par une fonction d'utilité. Il est facile de voir que la composition de u par une fonction strictement croissante ne change par l'ordre induit par $\succeq_u$. La proposition suivante nous donne la réciproque (la preuve est laissée en exercice).

Proposition 2 *Soient u et u' deux fonctions d'utilité sur X. Alors $\succeq_u$ coïncide avec $\succeq_{u'}$ si et seulement si il existe une fonction f : $\{u(x), x \in X\} \to R$, strictement croissante, telle que $u' = f \circ u$.*

2.3 Préférences sur choix incertains

Jusque là on n'a pas regardé l'ensemble des choix (ou d'alternatives X). $x \in X$ peut aussi bien représenter un panier-repas qu'un portefeuille d'actions. Cependant, il est clair que comparer deux portefeuilles d'actions est d'une nature beaucoup plus complexe que comparer deux paniers repas. La grande difficulté réside dans le fait que le prix d'une action dépend de plusieurs événements incertains. Pour modéliser un problème de choix très complexe on a donc besoin de plus d'abstraction et plus d'axiomatisation.

Sans même être un grand investisseur, vous et moi sommes confrontés souvent à choisir entre des événements à conséquences incertaines. Par exemple choisir entre le métro (risque de grève) et la voiture (risque d'embouteillage, pollution) ; entre acheter un Lotto, un millionnaire ou miser sur un cheval ?

Soit $Z = \{z_1, ..., z_n\}$ un ensemble fini de prix. Soit $P = \Delta(Z)$ l'ensemble des lois de probabilités sur Z.

$$P = \{(p_1, ..., p_n) \in \left(R^+\right)^n : \sum_{k=1}^{n} p_k = 1\}$$

Pour une raison ou une autre, notre agent doit exprimer ses préférences entre les lotteries (les éléments de P).

P est donc son ensemble d'alternatives. Il joue le rôle de X dans la section précédente.

2.4 Les axiomes de von-Neumann et Morgenstern

En utilisant la structure particulière de (un ensemble convexe) P et l'interprétation des éléments P (comme des loteries) il est possible de définir d'autres axiomes sur les préférences de notre agent.

Soit donc $\succeq$ une relation de préférence faible sur P.

- **Axiome de rationalité :** $\succeq$ est une relation de préférence rationnelle.

- **Axiome d'indépendance :** Pour tout $p, q, r \in P$, et $\lambda \in]0, 1[$, $p \succ q$ implique $\lambda p + (1-\lambda)r \succ \lambda q + (1-\lambda)r$.
- **Axiome de continuité :** Pour tout $p \succ q \succ r$, Il existe $\varepsilon \in]0, 1[$ (très petit) tel que $(1-\varepsilon)p + \varepsilon r \succ q \succ \varepsilon p + (1-\varepsilon)r$

L'axiome d'indépendance se comprend bien si on interprète correctement l'alternative $\lambda p + (1-\lambda)r$. En effet, celle-ci veut dire qu'avec la probabilité λ l'alternative p est sélectionnée, et avec la probabilité $(1-\lambda)$, c'est r qui est sélectionnée. Si notre agent préfère p à q, il devrait aussi préférer $\lambda p + (1-\lambda)r$ à $\lambda q + (1-\lambda)r$. L'axiome de continuité est interprétable aussi. Si l'agent préfère strictement p à q et si p_ε est très proche de p, alors l'agent doit encore préférer strictement est p_ε à q. Cet axiome doit aussi être utilisé avec précaution. En effet, si je vous propose le choix entre $q =$ gagner 10 euros d'une manière certaine et $p =$ gagner 1000 euros d'un manière certaine, vous allez choisir sans aucun doute 1000 euros. Si maintenant je vous donne le choix entre $q = 10$ euros d'une manière certaine et $p_\varepsilon =$ avec probabilité $\frac{999.999}{1.000.000}$ vous gagnez 1000 euros et avec probabilité $\frac{1}{1.000.000}$ vous mourrez, la plus part d'entre vous vont peut être préférer ne pas risquez leur vie pour seulement 1000 euros : le prix de la vie est peut-être infini pour certain... Notre modèle ne tient pas compte donc de ce genre phénomène. Quoique, si on se trouve en ce moment à Palaiseau et que je vous je propose de vous donner tout de suite 10 euros ou sinon, vous prenez votre voiture, vous m'enmener jusqu'à la défense puis je vous donne 1000 euros, plusieurs d'entre vous vont peut être préférer m'enmener jusqu'à la défense pour gagner 1000 euros... mais en voiture de Palaiseau à la défense, la probabilité d'un accident mortel est non nul ! L'axiome de continuité n'est donc pas si idiot que ça, on est tous obligé d'accepter certains risques car sinon la vie n'a plus aucun intérêt.

2.5 Utilité espérée

Supposons que notre agent associe l'utilité $u(z)$ à chacun des prix $z \in Z$, où u est une fonction de Z vers R. L'agent peut donc associer à toute loterie $p \in P$, une valeur $\widetilde{u}(p)$, appelée utilité espérée, avec $\widetilde{u}(p) = \sum_{z \in Z} p(z)u(z)$. $\widetilde{u}(p)$ est tout simplement l'espérance de la variable aléatoire $u(\widetilde{z})$ sous la probabilité p.

Soit alors $\succeq_{\tilde{u}}$ la relation de préférence correspondante sur P.

Proposition 3 *$\succeq_{\tilde{u}}$ satisfait les axiomes de rationalité, d'indépendance et de continuité.*

Preuve. L'axiome de continuité est du au fait de la continuité de $p \to \widetilde{u}(p)$. L'axiome de l'indépendance est de une conséquence de la linéarité de $\widetilde{u}(\cdot)$. Enfin, la rationalité est une conséquence du fait que la relation d'ordre $\geq$ sur R est complète. ■

Il peut sembler que cette forme de préférence est très restrictive. Elle reste néanmoins très commode pour le modélisateur (vous, nous), car la donnée de $cardinal(Z)$ paramètres (les utilités associées aux évènements certains) suffit à elle seule pour définir les préférences sur un ensemble beaucoup plus complexe et non dénombrable : les loteries sur Z.

Un des résultats les plus importants en économie est le suivant. En effet, von Neumann et Morgenstern ont montré l'équivalence entre les axiomes (simples et intuitives) que nous avons définis sur $\succeq$ et la possibilité d'une représentation de la forme $\succeq_{\tilde{u}}$.

Théorème 2 *(von Neumann et Morgenstern)*
Soit $\succeq$ une relation de préférence faible sur $P = \Delta(Z)$. Il existe une fonction d'utilité $u(\cdot) : Z \to R$ telle que $\succeq$ coïncide avec $\succeq_{\tilde{u}}$ si et seulement si $\succeq$ vérifie les axiomes de rationalité, d'indépendance et de continuité.

La partie "si" a déjà été montrée. Montrer la partie "seulement si" est un peu plus complexe. Soit donc $Z = \{z_1, ..., z_n\}$ et soit $\succeq$ une relation de préférence rationnelle sur P qui vérifie les axiomes d'indépendance et de continuité. On va commencer par montrer quelques lemmes utiles.

Lemme 1 *Il existe $\overline{z}$ et $\underline{z}$ dans Z tel que, pour tout $z \in Z$, $\delta_{\overline{z}} \succeq \delta_z \succeq \delta_{\underline{z}}$, où δ_z est la masse de Dirac au point z (recevoir le prix z avec probabilité 1).*

Preuve. Puisque $\succeq$ est rationnel sur P, elle est sur $H = \{\delta_{z_1}, ..., \delta_{z_n}\}$. Par la section précédente, il existe une fonction v telle que $\succeq = \succeq_v$ sur H. Soit alors $\overline{z}$ et $\underline{z}$ dans Z tels que

$$\begin{aligned} v(\overline{z}) &= \max_{z\in Z} v(\delta_z) \\ v(\underline{z}) &= \min_{z\in Z} v(\delta_z) \end{aligned}$$

Preuve. Puisque pour tout $z \in Z$, $v(\delta_{\overline{z}}) \geq v(\delta_z) \subset v(\delta_{\underline{z}})$, on déduit que $\delta_{\overline{z}} \succeq \delta_z \succeq \delta_{\underline{z}}$. ■ ■

Lemme 2 *Pour tout $p \in P$, $\delta_{\overline{z}} \succeq p \succeq \delta_{\underline{z}}$.*

Preuve. Soit $p = \sum_{k=1}^m p_k \delta_{z_k}$. On va prouver le lemme par récurrence sur m. Pour $m = 1$, c'est fait dans le lemme précédent. Supposons que cela est vrai pour toute combinaison convexe de m élément dans $H = \{\delta_{z_1}, ..., \delta_{z_n}\}$ et montrons que cela est vrai aussi pour une combinaison convexe de $m+1$ éléments. Si $p = \sum_{k=1}^m p_k \delta_{z_k} + p_{m+1}\delta_{z_{m+1}}$ avec $p_{m+1} > 0$ (sinon c'est fini). Dans ce cas, on peut écrire $p = (1-p_{m+1})q + p_{m+1}\delta_{z_{m+1}}$ avec $q = \frac{\sum_{k=1}^m p_k \delta_{z_k}}{\sum_{k=1}^m p_k}$.

L'hypothèse de récurrence implique que

$$\delta_{\overline{z}} \succeq q \succeq \delta_{\underline{z}}$$

mais aussi que

$$\delta_{\overline{z}} \succeq \delta_{z_{m+1}} \succeq \delta_{\underline{z}}$$

En utilisant l'axiome d'indépendance une fois on déduit que

$$\delta_{\overline{z}} \succeq (1-p_{m+1})q + p_{m+1}\delta_{\overline{z}}$$

En utilisant l'axiome d'indépendance une deuxième fois on déduit donc que

$$(1-p_{m+1})q + p_{m+1}\delta_{\overline{z}} \succeq (1-p_{m+1})q + p_{m+1}\delta_{z_{m+1}}$$

Enfin, on conclut en utilisant la transitivité. ■

Lemme 3 *Pour tout $1 > \alpha > \beta > 0$, pour tout $p \succ q$, $\alpha p + (1-\alpha)q \succ \beta p + (1-\beta)q$.*

Preuve. Par l'axiome d'indépendance, on déduit que

$$p \succ \frac{\beta}{\alpha}p + (1 - \frac{\beta}{\alpha})q$$

En appliquant encore l'axiome d'indépendance, on déduit que

$$\alpha p + (1-\alpha)q \succ \alpha \left[\frac{\beta}{\alpha}p + (1 - \frac{\beta}{\alpha})q\right] + (1-\alpha)q$$

On conclut en remarquant que $\alpha \left[\frac{\beta}{\alpha}p + (1 - \frac{\beta}{\alpha})q\right] + (1-\alpha)q = \beta p + (1-\beta)q$. ■

Lemme 4 *Pour tout p tel que $\delta_{\overline{z}} \succ p \succ \delta_{\underline{z}}$, il existe un unique $\alpha^* \in]0,1[$ tel que $p \approx \alpha^*\delta_{\overline{z}} + (1-\alpha^*)\delta_{\underline{z}}$.*

Preuve. L'unicité (en cas d'existence) est une conséquence du lemme précédent. Définissons $\alpha^* = \sup\{\alpha \in [0,1] : p \succeq \alpha\delta_{\overline{z}} + (1-\alpha)\delta_{\underline{z}}\}$. L'axiome de continuité implique que $\alpha^* > 0$. Puisque α^* est le sup, on déduit que $\forall 0 < \beta < \alpha^*$, il existe $\alpha \in]\beta, \alpha^*[$ tel que $p \succeq \alpha\delta_{\overline{z}} + (1-\alpha)\delta_{\underline{z}}$. Le lemme précédent implique que $\alpha\delta_{\overline{z}} + (1-\alpha)\delta_{\underline{z}} \succ \beta\delta_{\overline{z}} + (1-\beta)\delta_{\underline{z}}$. Par transitivité de la relation de préférence on donc que $\forall 0 < \beta < \alpha^*$, $p \succeq \beta\delta_{\overline{z}} + (1-\beta)\delta_{\underline{z}}$.

Maintenant, il y a trois possibilités : ■

- **Preuve. Cas 1 :** $p \succ \alpha^*\delta_{\overline{z}} + (1-\alpha^*)\delta_{\underline{z}}$. Ceci implique en particulier que $\alpha^* < 1$. Par l'axiome de continuité, il existe alors $\varepsilon > 0$ tel que

$$\begin{aligned} p &\succ (1-\varepsilon)\left[\alpha^*\delta_{\overline{z}} + (1-\alpha^*)\delta_{\underline{z}}\right] + \varepsilon\delta_{\overline{z}} \\ &= \beta\delta_{\overline{z}} + (1-\beta)\delta_{\underline{z}} \end{aligned}$$

avec $\beta = (1-\varepsilon)\alpha^* + \varepsilon > (1-\varepsilon)\alpha^* + \varepsilon\alpha^* = \alpha^*$. En contradiction avec le fait que α^* est le sup.

Cas 2 : $\alpha^*\delta_{\overline{z}} + (1-\alpha^*)\delta_{\underline{z}} \succ p$. Encore par l'axiome de continuité, on déduit qu'il existe $\varepsilon > 0$ tel que

$$(1-\varepsilon)\left[\alpha^*\delta_{\overline{z}} + (1-\alpha^*)\delta_{\underline{z}}\right] + \varepsilon\delta_{\underline{z}} \succ p$$

En posons alors $\beta = (1-\varepsilon)\alpha^* < \alpha^*$ on déduit que

$$\beta\delta_{\overline{z}} + (1-\beta)\delta_{\underline{z}} \succ p$$

En contradiction avec le fait que $\forall 0 < \beta < \alpha^*$, $p \succeq \beta\delta_{\overline{z}} + (1-\beta)\delta_{\underline{z}}$.

Donc seulement le cas qui suit reste possible.

Cas 3 : $\alpha^*\delta_{\overline{z}} + (1-\alpha^*)\delta_{\underline{z}} \approx p$. C'est le résultat souhaité, le lemme est prouvé. ∎

Maintenant on peut prouver le théorème de vNM.

Preuve. Si $\delta_{\overline{z}} \approx \delta_{\underline{z}}$, alors pour tout $p \in P$, $\delta_{\overline{z}} \approx p \approx \delta_{\underline{z}}$. On prend alors n'importe quelle fonction constante sur P.

Maintenant, si $\delta_{\overline{z}} \succ \delta_{\underline{z}}$. Soit alors $p \in P$. Il existe alors un unique $\alpha^*(p) \in [0,1]$ tel que

$$p \approx \alpha^*(p)\delta_{\overline{z}} + (1-\alpha^*(p))\delta_{\underline{z}}$$

Posons alors $f(p) = \alpha^*(p)$. Montrons d'abord que $f(p) > f(q)$ si et seulement si $p \succ q$. En effet, si $p \succ q$ alors on a (par transitivité) $\alpha^*(p)\delta_{\overline{z}} + (1-\alpha^*(p))\delta_{\underline{z}} \succ \alpha^*(q)\delta_{\overline{z}} + (1-\alpha^*(q))\delta_{\underline{z}}$ ce qui n'est possible (d'après un des lemmes ci-dessus) que si $\alpha^*(p) > \alpha^*(q)$. La réciproque se montre de la même manière. Donc on a bien $\succeq = \succeq_f$.

Il nous suffit donc de montrer que $f(p)$ est linéaire. En effet, par l'axiome de l'indépendance, on déduit que :

$$\lambda p + (1-\lambda)q \approx \lambda\left[\alpha^*(p)\delta_{\overline{z}} + (1-\alpha^*(p))\delta_{\underline{z}}\right] + (1-\lambda)\left[\alpha^*(q)\delta_{\overline{z}} + (1-\alpha^*(q))\delta_{\underline{z}}\right]$$

Et par unicité de $\alpha^*(\lambda p + (1-\lambda)q)$ on déduit que

$$\alpha^*(\lambda p + (1-\lambda)q) = \lambda\alpha^*(p) + (1-\lambda)\alpha^*(q)$$

La linéarité de $f(p)$ implique donc que

$$f(p) = \sum_{k=1}^{n} p_k f(\delta_{z_k})$$

En posant $u(z_k) = f(\delta_{z_k})$ le théorème de vNM est démontré. ∎

Nous pouvons nous poser la question de l'unicité de la représentation. Il est facile de voir pour tous nombres réels $a > 0$ et b et pour toute fonction d'utilité u, on a $\succeq_{\widetilde{u}} = \succeq_{\widetilde{au+b}}$. La proposition suivante nous donne la réciproque (la preuve est laissée en exercice).

Proposition 4 *Soient u et u' deux fonctions d'utilité sur X. Alors $\succeq_{\widetilde{u}}$ coïncide avec $\succeq_{\widetilde{u'}}$ si et seulement s'il existe $a > 0$ et b tel que $u' = a \times u + b$.*

Chapter 3

Aggrégation, choix social

L'issue d'un vote est un sujet très important qu'on étudiera plusieurs fois le long de ce cours. Ici, nous avons envie de vous présenter un résultat très célèbre d'Arrow qui ne rentre pas dans l'optique des notions développées plus tard. C'est une approche normative (axiomatique) de l'issue d'un problème de choix social.

3.1 Le problème d'agrégation des préférences

Soit A un ensemble fini d'alternatives (**ex.** des candidats pour un poste, plusieurs portefeuilles d'investissement, différentes politiques monétaires). N est un ensemble fini de décideurs appelés ici votants (**ex.** la société, un comité, ... etc.)

Chaque membre $i \in N$ a un ordre de préférence sur les alternatives dans A exprimé comme une relation binaire P_i. Ainsi: xP_iy signifie le votant i préfère x à y (P_i va jouer le rôle de $\succ$ dans le chapitre précédent).

Un profil de préférences : $P = P_1, P_2, \ldots, P_n$ consiste en une suite d'ordres de préférences P_i, $i = 1, \ldots n$. Dans le cas de trois alternatives $A = \{x, y, z\}$, P_i peut être représenté de la manière suivante:

$$P_i \quad = \quad \begin{array}{c} x \\ y \\ z \end{array} \quad \text{signifiant: } xP_iy \text{ et } yP_iz.$$

Un des problèmes qui se pose est de choisir un vainqueur entre les alternatives x, y et z, étant donné les préférences individuelles (**ex.** une société veut désigner un président ou un comité veut choisir un candidat pour un prix ou un travail...etc).

Une possibilité pour répondre à cette question est de considérer le vote majoritaire simple (a gagne devant b si une majorité de personne préfère a à b). Le vote majoritaire fonctionne bien dans le cas de deux alternatives. S'il y a plusieurs alternatives, on pourrait essayer de répondre à la question d'agrégation en considérant le vote majoritaire entre deux alternatives puis deux autres... etc. Une solution raisonnable est de considérer *le vainqueur de Condorcet* défini comme l'alternative $c \in A$ telle que: quelque soit $a \neq c$, une majorité de votants dans N préfère c à a.

Le problème est qu'il n'y a pas toujours un vainqueur de Condorcet, comme le montre l'exemple suivant: si $A = \{x, y, z\}$ et $N = \{1, 2, 3\}$ avec le profil de préférences suivant

1	2	3
x	y	z
y	z	x
z	x	y

. En effet, x gagne (suivant le vote majoritaire) devant y qui gagne devant z qui gagne devant x: il n'y a donc pas de vainqueur de Condorcet.

Une possibilité pour résoudre ce problème est d'organiser plusieurs tours (un peu comme un tournoi) pendant lesquels deux candidats se confrontent et seulement le gagnant des deux candidats passe au tour suivant. Cependant, cette solution cache un défaut majeur. Dans l'exemple précédent, en jouant sur l'agenda éléctoral, on pourrait élire n'importe quelle alternative. En effet, supposons que l'on adopte des élections en deux tours telles que: au premier tour, l'alternative a se présente contre b puis au second tour la gagnante entre a et b se présente contre c. Il y a donc trois possibilités: $(a, b) = (x, y)$, $(a, b) = (x, z)$ et $(a, b) = (y, z)$. Le problème est que n'importe quelle alternative peut sortir gagnante :

$$\begin{array}{ccccc} x & & x & & \\ & \rightarrow & & \Longrightarrow & z \\ y & & z & & \end{array}$$

$$\begin{array}{ccccc} x & & z & & \\ & \rightarrow & & \Longrightarrow & y \\ z & & y & & \end{array}$$

$$\begin{array}{ccccc} z & & y & & \\ & \rightarrow & & \Longrightarrow & x \\ y & & x & & \end{array}$$

Pour traiter ce problème d'agrégation des préférences, Arrow a suggéré une approche normative. Il remarque que le problème consiste en la recherche d'une fonctionnelle F (une sorte de boite noire) qui à un profil de préférences $P_1, \ldots, P_N$ de la société, associe un classement sur A (un ordre de préférence sur A) ou encore un vainqueur. Maintenant il suffit de mettre les propriétés que l'on désire obtenir puis de voir s'il existe une fonction qui satisfait à ces propriétés.

3.2 Théorème d'impossibilité d'Arrow

Soit A un ensemble d'alternatives et soit P un ordre de préférence sur A (une relation binaire sur A). On rappel que

- Si $x, y \in A$ alors xPy signifie x est préféré à y.
- P est complet si, pour tout x et y, on a soit xPy soit yPx. (i.e. chaque paire d'alternatives est comparable.)
- P est *transitif* si pour tout x, y et z dans A:

$$\left.\begin{array}{c} xPy \\ yPz \end{array}\right\} \Rightarrow xPz$$

- P est *antisymétrique* si pour tout x et y dans A:

$$\left.\begin{array}{c} xPy \\ yPx \end{array}\right\} \Rightarrow x = y$$

Définition 3 *Une relation biniare est dite linéaire si elle est rationnelle (complète et transitive) et antisymétrique (il n'y a pas d'indifférence entre deux alternatives différentes).*

On dénote par L l'ensemble des préférences linéaires sur A. N est la société (votants, joueurs, etc.), $S \subset N$ est une coalition. A est l'ensemble des alternatives (candidats, décisions, ...)

L^s dénote l'ensemble des profils de préférences (linéaires) des membres de S:

$$L^s = \{P^s = (P_i)_{i \in S}, P_i \in L\}$$

Définition 4 *F est une fonction de bien être social si F est une fonction de L^N vers L.*

Les propriétés (ou axiomes) désirées pour F:

- **Unanimité** (U):
  si $xP_iy, \forall i \in N$ **alors** $xF(P^N)y$ (U est appelée la Pareto optimalité);
- **Indépendance vis-à-vis des Alternatives Non Pertinentes** ($IANP$):
  $\forall x, y \in N$, **si** P^N et Q^N sont deux profils tels que $\forall i \in N, xP_iy \Leftrightarrow xQ_iy$ (la relation de préférence entre x et y est la même pour les deux profils) **alors**: $xF(P^N)y \Leftrightarrow xF(Q^N)y$;

L'unanimité veut dire que, si tous les votants préfèrent x à y, la fonction de bien être social devrait aussi préférer x à y. L'indépendance veut dire que la préférence sociale finale entre x et y devrait seulement dépendre des préférences des agents entre x et y, indépendamment de leurs préférences pour les autres alternatives.

Définition 5 *Une fonction de bien être social F est **Dictatoriale** s'il existe un joueur j dans N: $\forall P^N \in L^N$, $F(P^N) = P_j$. j est appelé un dictateur.*

Il est facile de vérifier qu'une fonction Dictatoriale satisfait U et $IANP$. Le résultat surprenant est le suivant: dès qu'il y a au moins trois alternatives, la fonction dictatoriale est la seule satisfaisant à la fois U et $IANP$!!!

Théorème 3 ***(Arrow)***
S'il y a au moins trois alternatives dans A, la seule fonction de bien être social F qui satisfait à la fois les propriétés U et $IANP$ est la fonction dictatoriale.

Preuve. Soient x et y deux alternatives différentes dans A ($x \neq y, x, y \in A$) et soit S une coalition. S est dite *décisive* pour x contre y si, pour tout $P^N \in L^N$, si $xP_iy, \forall i \in S$ et que $yP_jx, \forall j \in N \setminus S$, alors $xF(P^N)y$.

Puisque F vérifie U, N est décisive pour n'importe quelle paire de deux alternatives distinctes. Etant donné que F est IANP, S est décisive pour x contre y si et seulement s'il existe P^N tel que xP_iy, $\forall i \in S$, yP_jx, $\forall j \in N \setminus S$ et $xF(P^N)y$.

Soit D la famille des coalitions qui sont décisives au moins pour une paire (ordonnée) (x, y). Soit V une coalition dans D qui est minimale en termes de nombre de votants: $|V| \leq |S|, \forall S \in D$ (On appellera V: une coalition décisive minimale). Il est clair que $D \neq \emptyset$ (car par la propriété U, N est dans D) et que $V \neq \emptyset$ (car l'ensemble $\emptyset$ n'est pas décisif par la propriété U).

On va commencer par montrer que V est réduite à un seul élément, $j \in N$. Ensuite, on montrera que $\{j\}$ est décisive pour toute paire d'alternatives. Enfin, nous montrons que j est un dictateur (ce qui fait en tout trois lemmes).

Lemme 5 $|V| = 1$.

Preuve du lemme:

Soit $j \in V$. On définit $T = V - \{j\}$, $W = N \setminus V$. Il existe x et y tels que V est décisive pour x contre y. Soit $z \in A \setminus \{x, y\}$. (Il existe un tel z car $|A| \geq 3$.)

Soit le profil suivant :

$$P^N : \begin{array}{ccc} \{j\} & T & W \\ x & z & y \\ y & x & z \\ z & y & x \end{array}$$

Puisque V est décisive pour x contre y, on déduit que $xF(P^N)y$. Puisque $T \notin D$ (car V est minimale), on déduit que $yF(P^N)z$ (sinon T serait décisive pour z contre y.) Ainsi, par transitivité (car $F(P^N) \in L$), on déduit que $xF(P^N)z$.

D'où: $\{j\}$ est décisive pour x contre z et donc: $V = \{j\}$ (car sinon V ne serait pas minimale) et donc $\{j\}$ est décisive pour x contre tout $t \in A \setminus \{x\}$.

Lemme 6 $\{j\}$ *est décisive contre toute paire d'alternatives:*

Preuve du lemme:
Soit $z, w \in A \setminus \{x\} : z \neq w$

$$P^N : \begin{array}{cc} \{j\} & W \\ w & z \\ x & w \\ z & x \end{array}$$

$xF(P^N)z$, puisque $\{j\}$ est décisive pour x contre z.
$wF(P^N)x$, puisque F est U.

Donc, $wF(P^N)z$. Ceci veut dire que $\{j\}$ est décisive pour w contre z. En d'autres termes $\{j\}$ est décisive pour toute alternative dans $A \setminus \{x\}$ contre toute alternative dans $A \setminus \{x\}$. En plus, on a vu que $\{j\}$ est décisive pour x contre toute alternative. Ainsi, pour compléter la preuve que $\{j\}$ est décisive pour toutes paires d'alternatives, on doit montrer qu'elle est décisive pour toute alternative dans $A \setminus \{x\}$ (disons, w) contre x. Soit $z \neq w$ deux alternatives dans $A \setminus \{x\}$.

$$P^N : \begin{array}{cc} \{j\} & W \\ w & z \\ z & x \\ x & w \end{array}$$

$wF(P^N)z$, puisque $\{j\}$ est décisive pour w contre z. $zF(P^N)x$, puisque F est U.

Donc, $wF(P^N)x$. En d'autres termes $\{j\}$ est décisive pour w contre x. D'où, $\{j\}$ est décisive pour n'importe quelle paire d'alternatives.

Lemme 7 $\{j\}$ *est un dictateur.*

Preuve du lemme:
Soit P^N un profil de préférences, et soient $x, y \in A, x \neq y$. Supposons que xP_jy. On veut montrer que $xF(P^N)y$. Soit $z \in A \setminus \{x, y\}$. En utilisant l'IANP, supposons que xP_jzP_jy et que zP_iy et que zP_ix pour tout $i \in A \setminus \{j\}$.

$$P^N : \begin{array}{cc} \{\text{j}\} & \text{W} \\ \hline \vdots & \vdots \\ \text{x} & \text{z} \\ \vdots & \vdots \\ \text{z} & \text{x} \ldots \text{y} \\ \vdots & \vdots \\ \text{y} & \end{array}$$

$xF(P^N)z$, puisque $\{j\}$ est décisive pour x contre z. $zF(P^N)y$, car F est U. D'où, $xF(P^N)y$ ce qui termine la preuve du théorème d'Arrow. ■

Ce résultat montre l'impossibilité de trouver une bonne manière d'agréger les préférences individuelles. Certaines personnes diraient que IANP demande trop. Certes, mais Robert Aumann, un chercheur éminent en théorie des jeux, vous expliquera IANP de la manière suivante. Imaginez qu'un comité veuille attribuer un poste (ou un prix) à un candidat. Après de longues discussions, il décide enfin de donner le poste au candidat x. A ce moment là, on apprend que le candidat $y \neq x$ se désiste. IANP veut dire que la décision ne devrait pas changer après cette annonce (ce qui apparaît donc raisonnable). Le théorème d'impossibilité montre que ces deux propriétés ne peuvent pas toujours être respecté simultanément en démocratie. Un exemple connu (par exemple lors des recrutements d'un maître de conférence par une faculté) est le suivant. Il peut arriver qu'entre trois candidats, la moitié des votants (dans le conseil scientifique) classent le premier candidat A en première position et le second candidat B en deuxième position et que l'autre moitié des votants fait l'inverse : elle classe B en première position et classe A en seconde position. Le troisième candidat C est classé par tout le monde en troisième position. Il est tout-à-fait possible (et c'est déjà arrivé) qu'à la fin des délibérations, le candidat C se trouve classé premier, pour des raisons de compromis ou de lutte de pouvoir. On voit donc que le candidat le moins préféré peut se retrouver le mieux classé (U est violée). Aussi, il suffit qu'une personne change d'avis dans le camp adverse pour faire émerger une majorité et donc changer le classement du final en faveur de A ou de B, et C se retrouvent donc dernier (IANP est violée aussi).

En fait, ce théorème apparaît sous des formes variées dans la littérature. En voici deux exemples.

Définition 6 *f est une fonction de choix social si $F : L^N \longrightarrow A$.*

f représente l'alternative élue par la société étant données les préférences individuelles (ceci est notre situation de départ concernant l'élection d'un unique candidat).

Définition 7 *f est strictement monotone si pour tout P^N, **si** Q^N augmente la position relative de y relativement à P^N (i.e. si $\forall z \in A \backslash \{y\}$, $\forall i \in N$, $yP_iz \Rightarrow yQ_iz$) et si $f(P^N) = x$, **alors** $f(Q^N) \in \{x, y\}$.*

La stricte monotonie veut dire que: si seulement la position de y a augmenté alors soit le choix social n'a pas changé, soit il devient y.

Théorème 4 *Si f est strictement monotone et son rang A^* ($A^* := f(L^N)$) contient au moins trois alternatives, alors f est dictatoriale.*

(Pour prouver ce résultat, on utilise le théorème d'Arrow).

Définition 8 *Une fonction de choix social est manipulable s'il y a une situation (P^N) pour laquelle, si chacun présente ses vraies préférences, alors il y a un joueur qui, en mentant sur ses vraies préférences, peut modifier le choix social vers un nouveau résultat meilleur pour lui.*

Par exemple, si le système f choisit le candidat qui a le plus grand nombre de votants le mettant en première position (et qu'en cas d'égalité entre plusieurs candidats il choisit le candidat a) alors dans la situation suivante:

$$\begin{array}{ccccccc} & & & & & & i \\ a & a & a & b & b & b & c \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & b \\ & & & & & & a \end{array}$$

Si le joueur i donne ses vraies préférences a serait élu. Cependant, s'il ment sur sa préférence et fait comme si son classement était le suivante:

$$\begin{array}{c} b \\ c \\ a \end{array}$$

ceci va changer le résultat de l'élection en b au lieu de a.

Théorème 5 *Si une fonction de choix social f n'est pas manipulable et si son rang ($A^* = f(L^N)$) contient au moins trois alternatives alors elle est dictatoriale.*

<u>**Quelles sont les conclusions de cette analyse:**</u> il est évident qu'on est plutôt intéressé par l'analyse des fonctions de choix social qui ne sont pas dictatoriales. Par le théorème précédent, on sait qu'elles seront nécessairement manipulables. La question est donc : comment minimiser le mieux possible la distorsion causée par cette manipulation ?

Plus précisément, supposons que nous soyons intéressé par l'application d'une certaine fonction f qui apparaît raisonnable si tout le monde vote en respectant ses vraies préférences. Peut-on construire une fonction $\widetilde{f}$ qui prend la manipulation en compte et neutralise son effet ? C'est-à-dire, supposons que les vraies préférences soient données par P^N, et que l'on souhaite atteindre le résultat $f(P^N)$. Si la fonction est $\widetilde{f}$, et les joueurs jouent $\tilde{P}^N$ alors $\widetilde{f}(\tilde{P}^N)$ va être choisie. Peut-on déterminer $\widetilde{f}$ de telle sorte que $\widetilde{f}(\tilde{P}^N) = f(P^N)$?

C'est pour résoudre de telles questions que la théorie des jeux intervient.

Chapter 4

Affectation, le problème de mariage

Le problème de Mariage, introduit par Gale et Shapley en 1962, se présente de la manière suivante. On dispose de n hommes $X = \{x_1, x_2, ..., x_n\}$ et n femmes $Y = \{y_1, y_2, ..., y_n\}$. Chaque homme a un ordre de préférence linéaire sur les femmes (i.e une relation binaire complète, transitive et antisymétrique, voir la section précédente) et vice-versa. Par exemple, l'homme x_i classe la femme y_2 en première position (c'est elle qu'il préfère le plus) puis la femme y_1 en deuxième position (il choisirait y_1 si y_2 ne veut pas de lui), etc. On notera alors $y_2 \succ_{x_i} y_1$ pour exprimer que x_i préfère y_2 à y_1.

On pourrait imaginer dans le rôle des hommes par exemple des universités (resp. hôpitaux) et dans le rôle des femmes les candidats (resp. infirmières).

Exemple: Deux hommes $\{A, B\}$ et deux femmes $\{a, b\}$. Les deux hommes préfèrent a à b. La femme a préfère A à B et b préfère B à A. Ceci peut se présenter sous la forme matricielle suivante:

$$\begin{array}{cc} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} A \\ \\ B \end{array} & \begin{bmatrix} (1,1) & (2,2) \\ & \\ (1,2) & (2,1) \end{bmatrix} \end{array}$$

Les lignes représentent les hommes et les colonnes les femmes. $(1,1)$ indique que l'homme A classe la femme a en première position et que la femme a classe l'homme A en première position.

On souhaite affecter à chaque homme exactement une femme et vice-versa.

Définition 9 *$\mu \subset X \times Y$ est une affectation si $\forall x \in X$, $\exists! y \in Y$ tel que $(x, y) \in \mu$ et $\forall y \in Y$, $\exists! x \in X$ tel que $(x, y) \in \mu$.*

L'ensemble des affectations peut donc être identifié à l'ensemble des permutations de $\{1, ..., n\}$.

Notation: Si μ est une affectation, on notera μ^x l'unique femme affectée à x par μ et on notera par μ_y l'unique homme affecté à y par μ.

Par exemple $\{(A, b), (B, a)\}$ est une affectation dans l'exemple précédent. Cependant, celle-ci n'est pas raisonnable car l'homme A et la femme a se préfèrent mutuellement et rien ne les empêchent de se mettre ensemble. On dit que cette affectation est instable.

Définition 10 *Une affectation μ est dite stable s'il n'existe aucun couple $(x, y) \in X \times Y$ tel que x préfère y à μ^x et que y préfère x à μ_y.*

Dans l'exemple précédent, $\{(A, a), (B, b)\}$ est une affectation stable. Se pose ainsi le problème de l'existence d'une affectation stable pour n'importe quel problème de mariage.

Théorème 6 *(Gale & Shapley)*
Pour tout problème de Mariage, il existe une affectation stable.

Preuve.
La preuve est constructive et se présente sous forme d'un algorithme. Elle est illustrée à l'aide de l'exemple 3×3 suivant (les hommes sont notés A, B et C et les femmes a, b et c).

$$\begin{array}{c} \\ A \\ \\ B \\ \\ C \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} (1,2) & (2,1) & (3,3) \\ \\ (1,3) & (3,3) & (2,1) \\ \\ (3,1) & (2,2) & (1,2) \end{bmatrix} \end{array}$$

- **Premier jour :** chaque homme convoite la femme qu'il préfère. Si une femme est convoitée par plus d'un homme, elle choisit celui qu'elle préfère le plus, les autres hommes sont rejetés.

  Dans l'exemple : A et B convoite a, et C convoite c. a préfère A à B. B est donc rejeté.

- **Deuxième jour :** les hommes qui ont été rejetés le tour précédent vont convoiter leur second choix et ceux qui n'ont pas été rejetés restent en attente. Puis, chaque femme va garder celui qu'elle préfère le plus parmi les nouveaux arrivants et l'ancien convoitant et rejette les autres.

  Dans l'exemple : B, l'unique rejeté, va convoiter c. c qui avait déjà C comme prétendant va changer pour B et rejette C.

- L'algorithme continue jusqu'à ce que toute femme soit convoitée au moins une fois. Ceci constitue alors une affectation (chaque homme convoite une et une seule femme et chaque femme est convoitée par au moins un homme).

  Dans l'exemple : le troisième jour, C convoite b. A est en attente chez a et B continue à convoiter c. La procédure s'arrête.

L'algorithme s'arrête toujours en un nombre fini d'étapes (il y en a au plus n^2-2n+2). En effet, puisqu'un homme ne peut pas faire une deuxième proposition à une femme qui l'a déjà rejeté une fois dans le passé, toute femme reçoit à un moment où à un autre au moins une proposition.

Soit μ l'affectation obtenue suivant la procédure précédente. Alors, μ est stable. En effet, si ce n'était pas le cas, alors il existerait un couple (x, y) tel que x préfère y à μ^x et que y préfère x à μ_y. Mais alors il existe un jour pendant lequel x avait été rejeté par y (car sinon il n'aurait jamais convoité μ^x). Puisqu'une femme ne rejette que pour obtenir mieux, et nécessairement, par transitivité de la relation de préférence, y préfère μ_y à x. Une contradiction.

Dans l'exemple : $\{(A, a), (B, c), (C, b)\}$ est une affectation stable . ■

Remarque : On aurait pu appliquer l'algorithme en inversant les rôles. Les femmes convoitent et les hommes choisissent. Ainsi, dans l'exemple 3×3, le premier jour, a convoite C, b convoite A et c convoite B. L'algorithme converge donc en une seule étape, et on obtient une autre affectation stable $\{(A, b), (B, c), (C, a)\}$. Ceci montre donc qu'on n'a pas toujours l'unicité de l'affectation stable.

On peut montrer que l'algorithme suivant lequel les hommes convoitent et les femmes éliminent donne l'affectation stable que les hommes préfèrent le plus parmi les affectations stables. De même, l'algorithme suivant lequel les femmes convoitent et les hommes éliminent donne l'affectation stable que les femmes préfèrent le plus parmi toutes les allocations stables. Entre ces deux extrêmes, il peut exister plusieurs autres affectations stables. En fait on a un résultat encore plus robuste : l'affectation stable que les hommes préfèrent le plus parmi toutes les allocations stables est Pareto-optimale dans l'ensemble des affectations (stables ou pas). C'est-à-dire qu'il n'existe aucune affectation (stable ou pas) qui donne, au moins autant à tous les hommes et strictement plus pour au moins un homme, que l'affectation stable la plus préférée par les hommes.

Ainsi dans l'exemple 3×3 précédent, entre $\mu_1 = \{(A,a),(B,c),(C,b)\}$ et $\mu_2 = \{(A,b),(B,c),(C,a)\}$, tous les hommes préfèrent μ_1 : A préfère a à b, B est indifférent entre μ_1 et μ_2 car dans les deux cas il obtient c, et C préfère b à a. Dans cet exemple, toutes les femmes préfèrent μ_2 à ce qu'elles obtiendraient dans toute autre affectation, car elles ont toutes eu l'homme qu'elles préfèrent le plus.

Une autre propriété très jolie concerne la structure de l'ensemble des affectations stables. En effet, soit deux affectations stable μ_1 et μ_2. Définissons le maximum de point de vue des hommes comme étant

$$\mu_1 \vee_h \mu_2 = \{(x, \mu_1^x \vee_h \mu_2^x), x \in X\}$$

où

$$\mu_1^x \vee_h \mu_2^x = \begin{cases} \mu_1^x & si \quad \mu_1^x \succ_x \mu_2^x \\ \mu_1^x & si \quad \mu_1^x = \mu_2^x \\ \mu_2^x & si \quad \mu_2^x \succ_x \mu_1^x \end{cases}$$

De même nous définissons $\mu_1 \wedge_h \mu_2$, $\mu_1 \vee_f \mu_2$ et $\mu_1 \wedge_f \mu_2$, où les deux derniers concernent les femmes.

Théorème 7 *Sit μ_1 et μ_2 deux affectations stables alors :*

- *$\mu_1 \vee_h \mu_2$, $\mu_1 \wedge_h \mu_2$, $\mu_1 \vee_f \mu_2$ et $\mu_1 \wedge_f \mu_2$ sont toutes des affectations stables ;*
- *$\mu_1 \vee_h \mu_2 = \mu_1 \wedge_f \mu_2$ et $\mu_1 \wedge_h \mu_2 = \mu_1 \vee_f \mu_2$.*

La preuve de ce résultat est laissée en exercice. Le premiers résultat affirme que l'ensemble des affectations stables a *une structure de treillis* pour les relations $\vee_h$ et $\wedge_h$: l'affectation maximum et l'affectation minimum de deux affectations stables sont aussi des affectations stables. Le second montre qu'en se restreignant aux affectations stables, les intérêts entre les hommes et les femmes sont antagonistes : si on augmente dans les préférences des hommes, on diminue dans celle des femmes et vice-versa. En fait on peut montrer un résultat beaucoup plus robuste : pour n'importe quelles affectations stables μ_1 et μ_2 et pout tout couple $(x, y) \in \mu_1$ si dans l'affectation μ_2, x se voit affecter une femme qu'il préfère strictement plus à y, alors la femme y se voit affecter dans μ_2 un homme qu'elle préfère strictement moins à x, et vice-versa.

Part II

Jeux non coopératifs

Chapter 5

Exemples introductifs

Les jeux non coopératifs modélisent les interactions où les agents sont libres de choisir leurs actions et où un agent rationnel cherche à maximiser son propre bien-être (si un agent se rend compte qu'il a une stratégie admissible b lui permettant d'obtenir une meilleure utilité que celle obtenue avec la stratégie a alors il ne devrait pas jouer a). Pour incorporer dans un tel modèle de jeu dit non coopératif des considérations sociales telles que la justice, la jalousie, l'amour du partage, etc, on doit interpréter la fonction de paiement pour un joueur comme étant son utilité (de von Neumann et Morgenstern en présence d'incertitude).

Les exemples qui suivent (jeux d'échecs et jeu de Gale) permettent de comprendre les deux manières de modéliser une interaction: c'est-à-dire les jeux sous forme extensive et celle des jeux sous forme normale (ou stratégique). Enfin, on introduira plusieurs concepts de solution.

5.1 Le théorème de Von Neumann

5.1.1 Le jeu d'échec

On considère le jeu d'échec dans lequel: quand une partie repasse trois fois par une même configuration alors on dira qu'il y match nul (égalité entre les joueurs). Le Blanc est le joueur qui commence et on considère qu'il n'y a pas de contrainte de temps. Une stratégie d'un joueur est un plan d'actions (une fonction) qui lui dicte ce qu'il doit faire dans toutes les étapes possibles du jeu.

Théorème 8 *Une et une seule des trois assertions suivantes est vraie pour le jeu d'échec:*

(i) le joueur Blanc a une stratégie gagnante;

(ii) le joueur Noir a une stratégie gagnante;

(iii) chacun des joueurs a une stratégie qui lui garantit au moins l'égalité.

Une stratégie gagnante étant une stratégie qui permet de gagner, et ce quoi que fasse l'autre joueur.

Preuve. (En forme d'exercice)

Soit $(F_n)_{n\geq 1}$ la famille des jeux G à deux joueurs (un joueur Blanc et un joueur Noir) tels que (1) G commence à une position arbitraire sur un échiquier, (2) les joueurs dans G jouent séquentiellement suivant les règles du jeu d'échec et enfin (3) le jeu G s'arrête sur une égalité entre les joueurs dès lors que le nombre de coups effectué sans que l'un des joueurs gagne la partie dépasse n.

1) Montrer (par récurrence) que les jeux dans F_n satisfont le théorème précédent.
2) En remarquant que la longueur maximale pour le jeu d'échec est finie, en déduire le résultat. ■

Remarque:

Dans la preuve précédente, nous n'avons pas utilisé spécifiquement les règles du jeu d'échec, mais seulement utilisé les propriétés suivantes:

(P1) le jeu est à deux joueurs (I et II) qui jouent séquentiellement;

(P2) le jeu a une longueur maximale finie;

(P3) il y a seulement trois résultats possibles: I gagne, II gagne, Egalité entre I et II.

On en déduit donc le théorème suivant de Von Neumann (attribué aussi à Zermelo).

Théorème 9 *Pour tout jeu satisfaisant (P1), (P2) et (P3), une et une seule des trois assertions suivantes est vraie:*

(i) le joueur I a une stratégie gagnante;

(ii) le joueur II a une stratégie gagnante;

(iii) chacun des joueurs a une stratégie qui lui garantit au moins l'égalité.

5.1.2 Le jeu de Gale

Une autre exemple d'application du théorème de Von Neumann est le jeu suivant introduit par David Gale. On considère un échiquier (une tablette de chocolat) $n \times m$. Il y a deux joueurs I et II. Ils choisissent alternativement une case dans l'échiquier. Lorsqu'une case (i, j) est choisie, toutes les cases $(i\prime, j\prime)$ situées au Nord-Est ($i\prime \geq i$ et $j\prime \geq j$) sont éliminées (ne peuvent plus être choisies pas la suite).

Le joueur qui choisit la case $(1, 1)$ est déclaré perdant et l'autre joueur est déclaré gagnant.

Le joueur I est celui qui commence le jeu.

Exercice:
On suppose que n et m sont finis pour les questions de 1) à 4).
1) Qu'est ce qu'une stratégie pour un joueur ?
2) Construire explicitement une stratégie gagnante pour le joueur I dans le cas $n = m$.
3) En remarquant que ce jeu satisfait (P1), (P2) et P(3) en déduire qu'il y a seulement deux possibilités: (i) le joueur I a une stratégie gagnante ou bien (ii) le joueur II a une stratégie gagnante. (En d'autre termes, montrer que l'égalité est impossible dans ce jeu).
4) Montrer que si II a une stratégie gagnante alors I aussi a une stratégie gagnante. En déduire que dans ce jeu, le joueur I a une stratégie gagnante.
5) Etudier le jeu quand n (et/ou m) est infini (commencer par le cas $n = 2$ et $m = \infty$).

Remarque:
La construction explicite d'une stratégie gagnante pour le joueur I dans le cas fini ($n \neq m$) reste encore un problème ouvert pour lequel un prix est offert.

Chapter 6

Modélisation des jeux sous forme extensive

Un jeu est une modélisation d'une situation d'interaction décisionnelle. Les ingrédients minimaux nécessaires pour décrire un jeu sont :

- Les joueurs (preneurs de décisions, appelés aussi agents);
- les actions possibles pour chaque joueur;
- les règles qui déterminent l'ordre dans lequel chaque joueur doit choisir ses actions et aussi qui spécifient quand le jeu se termine;
- le résultat du jeu pour chaque fin possible.

Les jeux sous forme extensive constituent un des modèles les plus simples où tous ces aspects sont présents (pensez au jeu d'échecs). Il s'agit d'un modèle où les joueurs choisissent séquentiellement leurs actions (suivant certaines règles), jusqu'à un moment où le jeu est déclaré fini.

6.1 Le modèle standard

Les jeux sous forme extensive ont naturellement une structure arborescente, et peuvent donc être décrits par l'utilisation d'un graphe (ou un arbre) orienté, appelé l'arbre de décision ou encore l'arbre du jeu.

Définition 11 *On appellera un arbre de jeu tout triplet ordonné (X, x_0, f) vérifiant:*

- *X est un ensemble (au plus dénombrable);*
- *$x_0 \in X$;*
- *f est une application de $X \backslash \{x_0\}$ vers X telle que: quel que soit $x \in X \setminus \{x_0\}$, il existe un unique entier $n \geq 1$ tel que $f^{(n)}(x) = x_0$*
  où, $f^{(n)} = f \circ f ... \circ f$, n fois.

X désigne l'ensemble des noeuds, x_0 est la racine de l'arbre et $f(x)$ est l'unique *prédécesseur immédiat* du noeud x. L'ensemble T des noeuds terminaux est l'ensemble des noeuds qui ne sont les prédécesseurs d'aucun autre noeud: $T = \{x \in X | f(y) \neq x, \forall y \in X\}$.

S'il existe n tel que $f^n(x) = y$ alors y est appelé tout simplement un *prédécesseur* de x (pour $n = 1$ c'est un prédécesseur immédiat).
x est appelé un *successeur immédiat* de $f(x)$. L'ensemble des successeurs immédiats d'un noeud x est identifié à l'ensemble des actions possibles du joueur qui va intervenir au noeud x.

Définition 12 *Un jeu sous forme extensive standard* Γ *est un sextuple ordonné:*

$$\Gamma = \left(N, X, x_0, f, \left(X^i\right)_{i\in N}, r\right)$$

où:
N est l'ensemble des joueurs (au plus dénombrable).
(X, x_0, f) est un arbre orienté.
$\left(X^i\right)_{i\in N}$ est une partition de $X \setminus T$.
$r : T \to O$ est la fonction résultat, où O est l'ensemble des résultats possibles (en anglais: set of outcomes).

X^i est l'ensemble des noeuds où le joueur i doit jouer.
Si X et fini, alors on dit que le jeu est fini.

Définition 13 ***Le Déroulement du jeu sous forme extensive :***
Le jeu sous forme extensive standard : $\Gamma = \left(N, X, x_0, f, \left(X^i\right)_{i\in N}, r\right)$ *se déroule de la manière suivante:*

- *le joueur $J(x_0)$ (l'unique joueur i tel que $x_0 \in X^i$) commence par choisir une action possible $x_1 \in A(x_0)$, où $A(x_0)$ est l'ensemble de tout les successeurs immédiats de x_0 (c'est-à-dire l'ensemble des actions admissibles à x_0 où encore: $f^{-1}(x_0)$).*
- *puis, le joueur $J(x_1)$ choisit une action x_2, où x_2 est un successeur immédiat de x_1;*
- *etc...;*
- *jusqu'à ce que le jeu atteigne un noeud terminal disons $x_k = t \in T$. Dans ce cas, le jeu est dit terminé.*

La séquence $x_0, x_1, ..., x_k$ est appelée une partie (ou réalisation du jeu). Le résultat du jeu (associé à cette partie) est donné par $r(t) \in O$. Enfin, le jeu et sa description sont supposés, en général, être une connaissance commune des joueurs.

Un dessin sous la forme d'un arbre orienté vous permet de représenter un jeu sous forme extensive. Vous commencez par la racine, vous tracez des arcs dirigés pour chacun des successeurs immédiats x_1 de x_0. Vous continuez le même procédé à partir de chaque nouveau noeud x_1 et ce jusqu'à l'arrivée à un noeud terminal.

Chaque arc dirigé de la forme $a \to b$ où b est un successeur immédiat de a correspond en réalité à une action possible du joueur intervenant au noeud a (il a une signification réelle telle que: acheter, vendre, voter pour le candidat Chirac, etc...)

Dire que la description est une connaissance commune des joueurs veut dire que chacun connaît la description, chacun sait que les autres joueurs connaissent la description, chacun sait que les autres savent qu'il connaît la description, etc... et ce jusqu'à l'infini.

Le cas où des joueurs ont seulement des informations partielles sera modélisé plus tard.

Exemple 1:

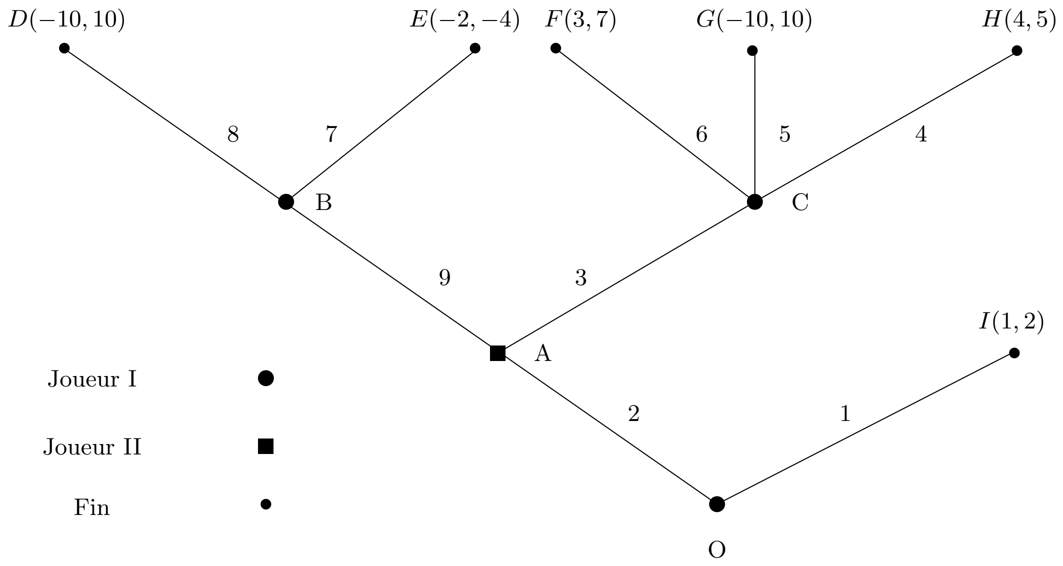

Figure 6.1:

Dans le jeu décrit dans la figure 6.1, la racine du jeu est le noeud O, qui est un noeud d'action du joueur I, celui-ci peut choisir l'action 1 ou 2. L'action 1 va mener à la fin du jeu avec le résultat (1,2): cela veut dire que le joueur 1 reçoit 1 unité d'utilité et le joueur 2 reçoit 2 unités. Le résultat (3,7) est atteint si par exemple le joueur 1 choisit l'action 2, puis le joueur 2 l'action 3 puis le joueur 1 l'action 6. La séquence d'actions (2,3,6) correspond à la séquence de noeuds (0,A,C,F) décrivant donc une partie spécifique du jeu.

Définition 14 *Une partie p du jeu est une séquence de noeuds (d'arcs dirigés adjacents) reliant l'origine de l'arbre orienté du jeu à un noeud terminal:* $p = (x_0, x_1, \ldots, x_k)$ *t.q.* $x_k \in T$ *et* $x_m = f(x_{m+1})$ *pour* $m = 0, \ldots, k-1$. *On notera* $\mathcal{P}$ *l'ensemble de toutes les parties.*

Il résulte donc qu'il y a une correspondance unique entre les points terminaux et l'ensemble des parties du jeu. Par exemple, dans l'exemple 1, le noeud terminal E correspond à la partie $(2, 9, 7)$. Cependant, il se peut que plusieurs parties aient le même résultat (mais des points terminaux différents). Par exemple, dans l'exemple 1, les parties $(2, 3, 5)$ et $(2, 9, 8)$ induisent le même résultat $(-10, 10)$ mais des noeuds terminaux différents: G et D.

Exemple 2 : le jeu d'échecs

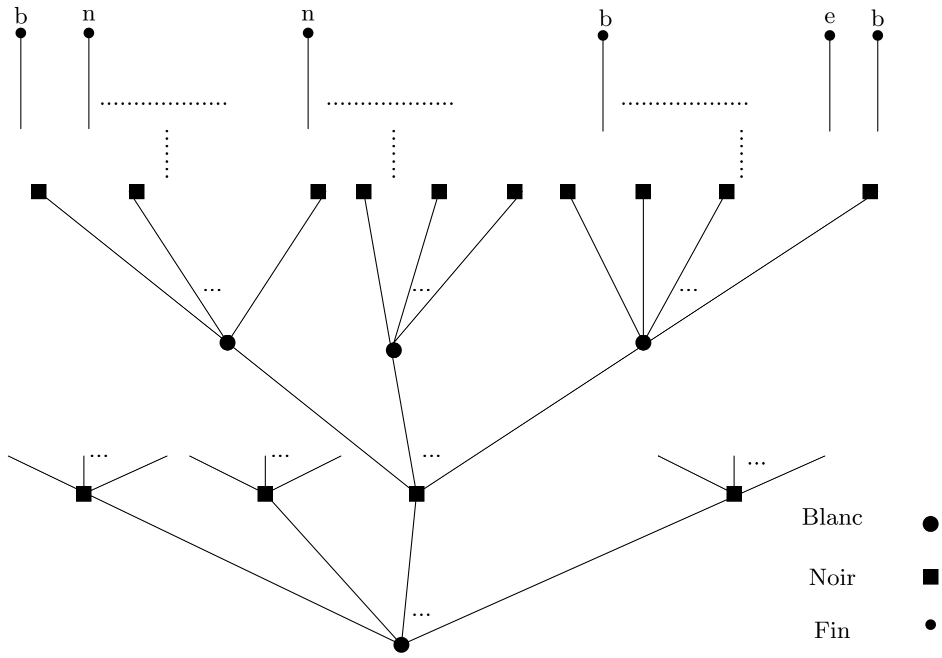

Figure 6.2:

C'est un exemple typique d'un jeu réel qui peut être mis sous la forme d'un jeu sous forme extensive (voir la figure 6.2). Pour cela, il nous faudra énumérer toutes les positions du jeu (correspondant donc aux noeuds) ainsi que toutes les actions possibles à chaque position. La fonction résultat prendra ses valeurs dans $O = \{E, B, N\}$. (E pour égalité, B si le joueur Blanc gagne et N si c'est le joueur Noir qui gagne). Une partie est donc une succession d'actions pour chaque joueur et ce jusqu'à ce que le jeu se termine: géométriquement c'est un chemin partant de la racine x_0 et se terminant à un point terminal t_0.

Cependant, dans la réalité, le nombre de parties possibles pour le jeu d'échecs est tellement grand (plus que le nombre d'atomes dans l'univers) que sa résolution, théoriquement possible, est très difficile en pratique (même avec l'aide d'un ordinateur très puissant).

Remarque:

Etant donné un jeu sous forme extensive standard $\Gamma = (N, X, x_0, f, (X^i)_{i \in N}, r)$ et un noeud $z \in X$, la position z peut être vue comme la racine d'un sous-jeu commençant à z. (Imaginer que vous interveniez au milieu d'une partie d'échecs, ceci peut être vu comme un nouveau jeu, remarquez aussi que l'on a considéré (implicitement) des sous-jeux dans la preuve du théorème de von Neumann). Bien sûr, il y a autant de sous-jeux que de noeuds X.

Définition 15 *Soit $\Gamma = (N, X, x_0, f, (X^i)_{i \in N}, r)$ un jeu sous forme extensive standard et soit $z \in X$. Le sous-jeu Γ_z commençant à z est le jeu sous forme extensive standard suivant:*

$$\Gamma_z = (N, X_z, z, f, (X_z^i)_{i \in N}, r)$$

avec: $X_z = \left\{x \in X | \exists n, \;\; s.t. \;\; f^{(n)}(x) = z\right\}$ et $X_z^i = X^i \cap X_z$.

Ainsi, certains joueurs d'échecs savent comment forcer la partie vers le gain à partir d'une certaine configuration. Chaque bon joueur essaie de diriger la partie vers une configuration (un sous-jeu) plus favorable pour lui (à partir de laquelle il est sûr de gagner). Naturellement un champion comme Kasparov doit connaître beaucoup de stratégies gagnantes dans plusieurs sous-jeux.

Quelques cas spéciaux:

- $\Gamma_{x_0} = \Gamma$;

- Pour tout point terminal $t \in T$, $\Gamma_t = \left(N, t, t, f, (\emptyset)_{i \in N}, r\right)$ est tout simplement le jeu trivial où aucun des joueurs n'a à choisir une action et où le résultat final est $r(t)$.

6.2 Le joueur hasard

Jusque là, on a modélisé seulement les situations où, à chaque noeud non terminal, on a associé un preneur de décision, libre de choisir n'importe quelle action.

Donc, la transition d'un noeud x_1 à son successeur immédiat x_2 est l'oeuvre d'un choix d'un décideur qui aurait aussi bien pu prendre une autre décision.

Cependant, dans plusieurs situations (par exemple les jeux de cartes comme le Bridge ou le Poker) certaines transitions peuvent être la conséquence d'un processus aléatoire exogène suivant une certaine loi de probabilité p (par exemple la répartition des cartes entre les joueurs). On a donc envie de modéliser les situations où les deux successeurs immédiats d'un noeud correspondent à : ¨(1) il va faire beau demain” et ¨(2) il ne va pas faire beau demain”, où encore: ¨(1) le CAC 40 sera plus élevé dans un mois” et ¨(2) le CAC 40 sera plus bas dans un mois”. La réalisation de tels évènements correspond souvent à un choix aléatoire. La probabilité p peut être calculée par des statistiques sur le passé. Cependant, les vraies probabilités (et donc le bon modèle) peuvent parfois être très difficiles d'accès.

En fait, il est facile d'étendre notre modèle de base en rajoutant un nouveau joueur: le joueur 0 (qui va jouer le rôle du hasard). On doit simplement spécifier les noeuds où le joueur hasard va intervenir, et avec quelles probabilités les transitions se font vers les successeurs immédiats. Puis, on ajoute l'hypothèse de connaissance commune du jeu ainsi enrichi.

Sur un dessin, en plus des Noms pour chaque action du joueur Hasard, on rajoute la probabilité avec laquelle il va choisir cette action.

Exemple 3

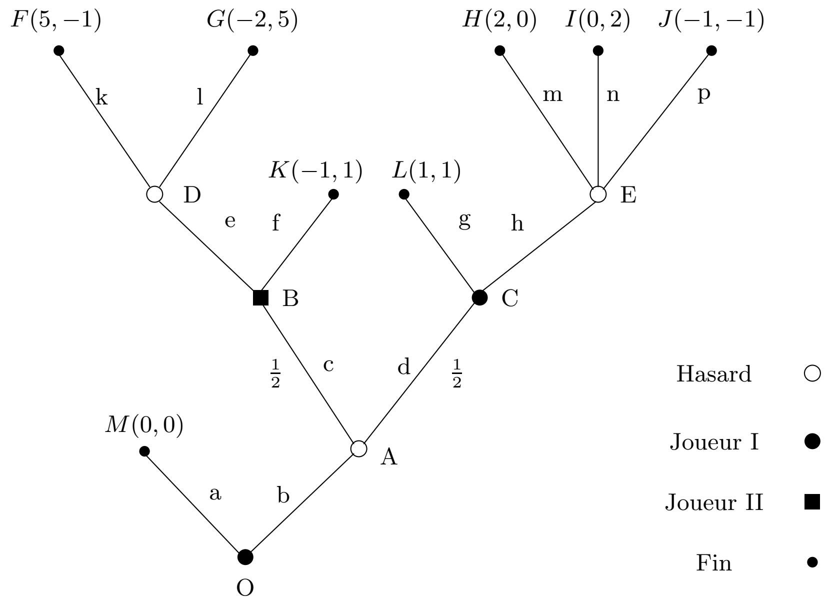

Figure 6.3:

Si par exemple (voir figure 6.3) le joueur I choisit l'action b pour son premier mouvement au noeud O alors, le hasard (la nature) choisit l'action suivante (c ou d) suivant la distribution de probabilité $(\frac{1}{2}, \frac{1}{2})$. Ainsi, avec probabilité $\frac{1}{2}$ le jeu arrive au noeud C où c'est encore le tour du joueur I de jouer et avec la probabilité $\frac{1}{2}$ le jeu arrive au noeud B où c'est au joueur II de jouer. La partie (O, A, B, D, G) décrit une réalisation possible de ce jeu dans laquelle le joueur I choisit b, le hasard sélectionne c, le joueur II choisit l'action e puis le hasard choisit au noeud D l'action l. Le résultat de cette partie est donc $(-2, 5)$ (c'est-à-dire que le joueur I perd 2 unités et le joueur II gagne 5 unités).

6.3 Information imparfaite et ensembles d'information

Jusqu'à présent, on a modélisé seulement les situations dites à information parfaite: explicitement les jeux où les joueurs connaissent avec exactitude le noeud où ils interviennent (le noeud où la partie se trouve). Leur connaissance du jeu leur permet alors de savoir tout ce qui s'est passé depuis le début du jeu. Il leur suffit pour cela de remonter l'arbre de jeu (par exemple en appliquant successivement la fonction prédécesseur f et ce jusqu'à l'arrivée à la racine de l'arbre).

Cependant, dans plusieurs situations, quand un joueur est amené à prendre une décision, il ne connaît pas avec exactitude la position du jeu. La raison pourrait être par exemple que les décisions se font simultanément (ex. le vote secret, R&D, le choix de l'événement qui va faire la une du journal, etc).

Exemple 4

Un jeu très simple où les joueurs jouent simultanément est le jeu dit de ¨Pile ou Face". Dans ce jeu, deux joueurs (I et II) choisissent simultanément (soumettent chacun une enveloppe fermée) une des deux faces d'une pièce de monnaie : (Pile ou Face) noté (P et F). Si les deux choisissent la même face (P, P) ou (F, F) alors le joueur II gagne (1 unité) et le joueur I perd (1 unité). S'ils choisissent des faces différentes $((P, F)$ ou $(F, P))$ alors le joueur II perd (1 unité) et le joueur I gagne (1 unité). On utilise la figure suivante pour décrire ce jeu:

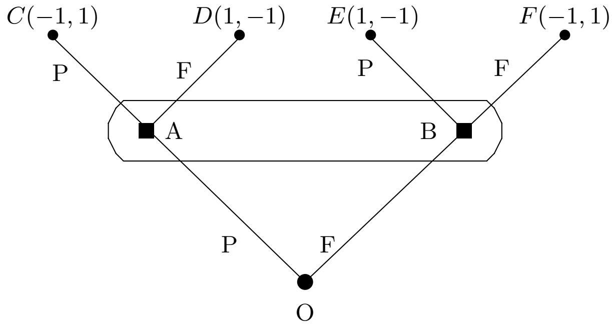

Figure 6.4:

Le jeu décrit dans la figure 6.4 contient un aspect important. En effet, dans le jeu de Pile ou Face, les noeuds A et B ne sont pas deux points de décisions différents pour le joueur II mais un unique point de décision car le joueur II ne peut pas distinguer entre ces deux positions. Pour exprimer cela, on encercle A et B comme indiqué dans la figure 6.4. On dit que $\{A, B\}$ est un ensemble d'information pour le joueur II.

En plus des jeux simultanés, les ensembles d'information peuvent modéliser des situations où un joueur ne sait pas le choix effectué par le joueur hasard. Par exemple un investisseur ne sait pas forcément, au moment même où il décide d'acheter une action, si la bourse de Hong Kong est en hausse ou en baisse. Les ensembles d'information peuvent aussi servir pour les situations où certains joueurs oublient des informations qu'ils ont acquis dans le passé (on peut imaginer une entreprise dans laquelle les décisions se prennent par des personnes différentes à chaque point de décision).

On doit donc préciser en plus dans la description d'un jeu, pour chaque joueur i, que i ne distingue pas entre les noeuds $(a, b, ..., c)$.
Cependant, pour que le joueur i ne puisse réellement pas faire une distinction entre les deux noeuds a et b, il est nécessaire que a et b aient le même nombre de successeurs (c'est-à-dire que le joueur en question ait le même nombre d'actions dans les deux cas, sinon a et b ne peuvent pas être considérés comme indistinguables). Par exemple, dans le jeu présenté dans la figure 6.5, A et B ne peuvent pas être dans le même ensemble d'information:

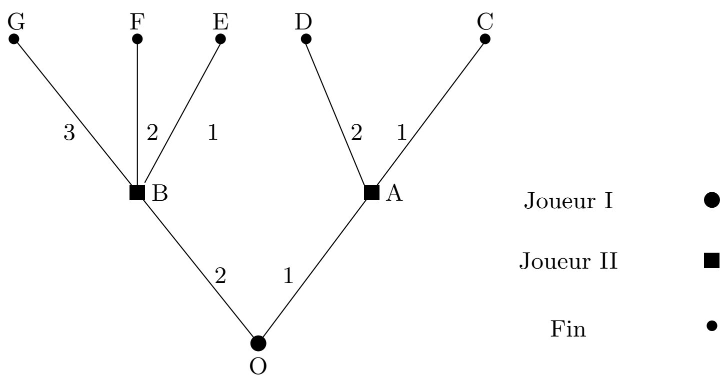

Figure 6.5:

Cette dernière condition est en fait nécessaire mais non suffisante : il faut en plus que les actions du problème modélisé, auxquelles les successeurs immédiats font références, soient les mêmes. En effet, si dans le noeud a les deux actions du joueur sont (Vendre et Acheter) et dans le noeud b ils sont (Investir et Epargner) alors il est absurde de qualifier a et b d'indistinguables.

Ainsi, dans la description du jeu, un nouvel élément appelé *structure d'information* fait son apparition dans la modélisation.
Une structure d'information consiste en une famille d'ensembles: $\left(U_j^i\right)_{\substack{i\in N\\ j=1,\dots,k_i}}$, où $\left\{U_1^i,\dots U_{k_i}^i\right\}$ est une partition de X_i. Chaque U_j^i constitue l'ensemble des noeuds que le joueur i ne distingue pas (appelé un ensemble d'information ou encore: un point de décision).

Sur un dessin, pour désigner l'ensemble d'information contenant les noeud $a, b..., c$ on les entourent tous ensemble à l'aide d'une courbe fermée (un cercle déformé).

Exemple 5:

Soit le jeu suivant en trois étapes: le joueur I choisit un nombre x dans $\{1,2\}$ qu'il ne révèle pas au joueur II. Dans la deuxième étape, un nombre y dans $\{1,2\}$ est choisi par le joueur hasard suivant la distribution de probabilité $(1/5, 4/5)$ et dans la troisième étape, le joueur II qui connaît y mais pas x choisit un nombre z dans $\{1,2\}$. Après le choix du triplet (x, y, z) le jeu se termine. La fonction résultat est la suivante:

$$\begin{array}{lcl} r(1,1,1)=(-5,5) & , & r(2,1,1)=(9,-9) \\ r(1,1,2)=(-3,3) & , & r(2,1,2)=(3,-3) \\ r(1,2,1)=(5,-5) & , & r(2,2,1)=(3,-3) \\ r(1,2,2)=(-9,9) & , & r(2,2,2)=(11,-11) \end{array}$$

Ce jeu peut être représenté par la figure suivante :

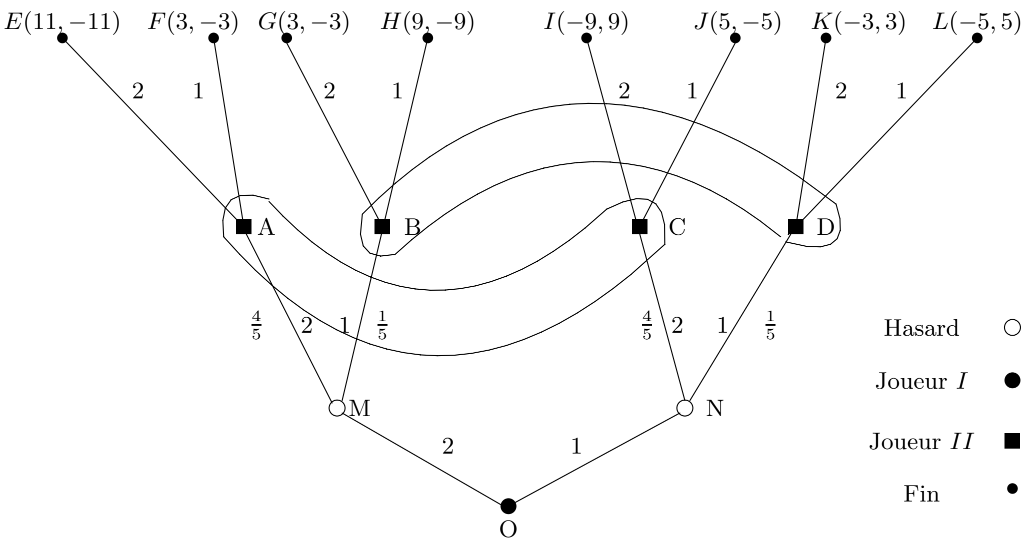

Figure 6.6:

Les ensembles d'information du joueur I sont $\{A, C\}$ et $\{B, D\}$. Les noeuds B et D sont dans le même ensemble d'information parce que le joueur II ne connaît pas la valeur de x (le choix du joueur I). Cependant, puisqu'il connaît le choix du joueur hasard, A et B sont donc distinguables par le joueur II (ils ne sont donc pas dans le même ensemble d'information).

On peut considérer une autre variante de ce jeu dans laquelle le joueur II ne connaît pas non plus la réalisation du joueur hasard. Ainsi, au moment où il doit choisir z, il ne connaît ni x ni y. Dans ce cas, le jeu peut être schématisé comme suit:

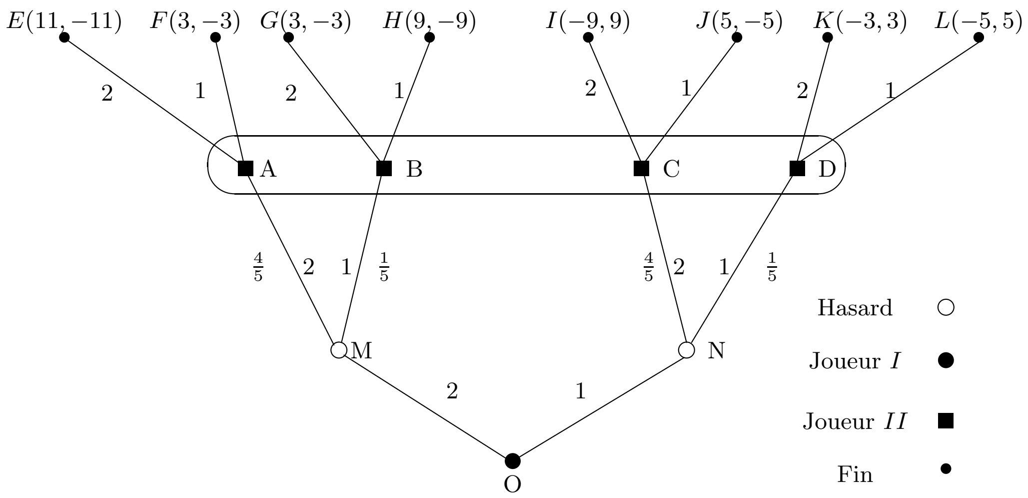

Figure 6.7:

6.4 Le modèle général

Maintenant, nous pouvons définir le modèle général de jeu sous forme extensive:

Définition 16 *Un jeu sous forme extensive, noté* Γ *est donné par:*

$$\Gamma = (N, X, x_0, f, X^0, \left(p_j^0\right)_{j \in X^0}, (X^i)_{i \in N}, \left(U_j^i\right)_{\substack{i \in N \\ j=1,\ldots,k_i}}, r)$$

où:

- $(N, X, x_0, f, (X^i)_{i \in N \cup \{0\}}, r)$ *est un jeu sous forme extensive standard;*
- *L'indice* 0 *(en puissance) désigne le joueur Hasard;*
- *Pour chaque* $j \in X^0$, p_j^0 *est la probabilité de transition utilisée par le joueur 0 pour choisir un successeur à* j*;*
- $\left(U_j^i\right)_{\substack{i \in N \\ j=1,\ldots,k_i}}$ *constitue la structure d'information des joueurs autres que le joueur Hasard. Pour chaque* $i \in N$, $\left\{U_1^i, \ldots U_{k_i}^i\right\}$ *est une partition de* X_i*;*
- $\forall U_j^i$, *si* $x_1 \in U_j^i$ *et* $x_2 \in U_j^i$ *alors* $A(x_1) = A(x_2)$ *qu'on peut aussi noter par* $A(U_j^i)$.

- ***Si**, pour tout i et j : U_j^i est réduit à un singleton **alors** le jeu est dit à information parfaite. Sinon, il est dit à **information imparfaite.***

Le Déroulement du jeu sous forme extensive :

Le jeu $\Gamma = (N, X, x_0, f, X^0, \left(p_j^0\right)_{j\in X^0}, (X^i)_{i\in N}, \left(U_j^i\right)_{\substack{i\in N \\ j=1,\dots,k_i}}, r)$ ce déroule exactement de la même manière qu'un jeu sous forme extensive standard, à quelques petites différences près: si le jeu arrive au noeud x et si le joueur $J(x)$ (l'unique joueur i tel que $x \in X^i$) est le joueur Hasard, alors l'action suivante est sélectionnée en utilisant la loi de probabilité p_x^0. Si le jeu est au noeud x, que le joueur $J(x)$ n'est pas le joueur Hasard, et que $x \in U_j^{J(x)}$ alors le joueur $J(x)$ choisit une action dans $A(U_j^{J(x)})$. Ceci va mener le jeu (le pointeur) vers un successeur immédiat de x et ce jusqu'à l'arrivée à un noeud terminal.

Les sous-jeux :

Un sous-jeu d'un jeu sous forme extensive et à information imparfaite est un sous-arbre dans lequel la structure d'information est préservée (plus précisément un sous-arbre qui inclut tout ensemble d'information qu'il croise). Par exemple, le jeu de la figure 6.7, il y a un seul sous-jeu qui est le jeu lui même car dans tout autre noeud le sous-arbre induit contient un ensemble d'information non inclut strictement dans le sous-arbre. En effet, dans cet exemple, il est absurde de considérer qu'à partir du noeud N, on est en présence d'un jeu à part entière. Sinon, arrivé à C où à D, le joueur II peut penser par exemple qu'il est au noeud A ou encore au noeud B. Le joueur II, doit prendre donc en considération des éléments qui ne sont pas décrit seulement par les successeurs de N. Le sous-arbre commençant au noeud N ne peut constituer un jeu à part entière.

6.5 La notion de stratégie

La notion de stratégie constitue un concept central en théorie des jeux, à ne pas confondre avec la notion d'action. Pour les jeux sous forme extensive, c'est un plan détaillé d'actions sur tout ce qu'un joueur a l'intention de faire dans toutes les positions possibles, en respectant l'information dont il dispose.

Définition 17 *Une stratégie pour un joueur dans un jeu sous forme extensive est une fonction qui associe à chaque ensemble d'information, une action dans l'ensemble des actions possibles à cet ensemble d'information.*
*Plus précisément: **si** $\left\{U_1^i, U_2^i, \dots\right\}$ est la partition d'information du joueur i, **si** $A\left(U_j^i\right)$ est l'ensemble des actions possibles à U_j^i **et si** $A^i = \cup_{j=1,2,..}\left[A\left(U_j^i\right)\right]$ **alors** une stratégie s_i du joueur i est une fonction: $s_i : \left\{U_1^i, U_2^i, \dots\right\} \to A^i$ telle que $\forall j$:*

$$s_i\left(U_j^i\right) \in A\left(U_j^i\right).$$

On notera S_i l'ensemble de toutes les stratégies possibles du joueur i.

Exemple 6:

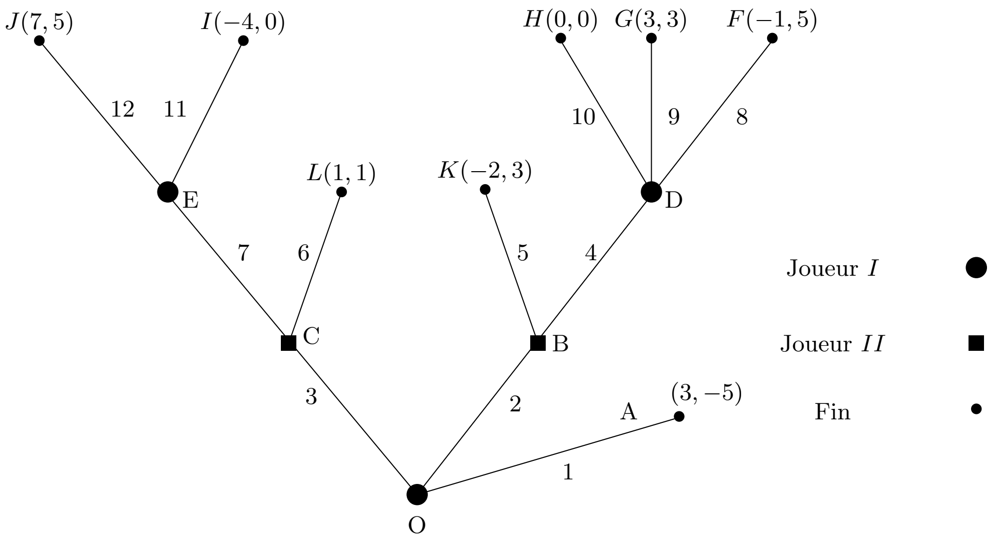

Figure 6.8:

Soit le jeu de la figure 6.8. Les ensembles d'information pour les joueurs sont:

$$\begin{aligned} U^I &= \{\{O\}, \{D\}, \{E\}\} \\ U^{II} &= \{\{B\}, \{C\}\} \end{aligned}$$

Une stratégie possible S^I pour le joueur I est la suivante:

$$S^I(\{O\}) = 3, \qquad S^I(\{D\}) = 8, \qquad S^I(\{E\}) = 12$$

Une stratégie possible S^{II} pour le joueur II est la suivante

$$S^{II}(\{B\}) = 4, \qquad S^{II}(\{C\}) = 6$$

Ainsi, en suivant la stratégie (le plan d'action) S^I, le joueur I choisit l'action 3 quand il est à l'ensemble d'information $\{O\}$ (le début du jeu), l'action 8 à l'ensemble d'information $\{D\}$ et l'action 12 à l'ensemble d'information $\{E\}$.

Il est important de noter que les instructions d'une stratégie n'ont pas un impact sur toutes les parties. En effet, si dans l'exemple précédent les joueurs jouent suivant les stratégies S^I et S^{II} alors, puisque à $\{C\}$ le joueur II choisit l'action 6, l'ensemble d'information $\{E\}$ du joueur I n'est jamais atteint et donc $S^I(\{E\})$ n'a aucun impact sur ce morceau du jeu. L'ensemble d'information $\{D\}$ n'est jamais atteint (puisque $S^I(\{O\}) = 3$) et ce quel que soit ce qui se passe après. Donc la partie de la stratégie $S^I(\{D\})$ peut être omise sans créer aucune ambiguïté sur la suite du jeu. Cependant, pour des raisons méthodologiques, S^I doit être défini pour tous les ensembles d'information du joueur I. Par exemple, une stratégie pourrait cacher une menace. Le joueur II pourrait menacer le joueur I en disant: tu fais ce que tu veux, mais je te fais remarquer que tu n'as pas du tout intérêt à aller à droite car voila ce que je ferais dans cette éventualité. Ainsi, même si le jeu ne va pas à droite, la décision de ne pas aller à droite est une conséquence de la stratégie du joueur II.

Dans cet exemple 6, la paire de stratégies (S^I, S^{II}) détermine une partie unique (suivant la trajectoire 3 puis 6) qui se termine par le résultat $(1, 1)$. Il est facile de voir que ceci est toujours le cas s'il n'y a pas de joueur hasard.

Proposition 5 *Dans tout jeu sous forme extensive et sans joueur hasard, tout vecteur de stratégies* $S = \left(S^i\right)_{i \in N}$ *détermine une unique partie (et donc un unique résultat).*

6.6 Le théorème de von Neumann (revisité)

Avec nos notations formelles, on peut maintenant reformuler le théorème de Von Neumann:

Théorème 10 *(Von Neumann)*
Dans tout jeu fini à deux joueurs (I et II), sous forme extensive et à information parfaite, ***si*** *l'ensemble des résultats: O contient seulement trois éléments: (I gagne), (II gagne) et (Egalité), une et une seule des propositions suivantes est satisfaite:*

- *(i) le joueur I a une stratégie gagnante.*
- *(ii) le joueur II a une stratégie gagnante.*
- *(iii) les deux joueurs ont des stratégies qui garantissent au moins l'égalité.*

Dans le cas où il n'y a pas de joueur hasard, la preuve est exactement la même que pour le jeu d'échec (par récurrence sur la longueur maximale d'une partie allant du noeud x_0 et ce jusqu'à un noeud terminal). Cependant, l'étape de récurrence peut être faite de deux manières différentes : **A)** une induction en amont (backward) et **B)** une induction en aval (forward).

La preuve plus haut pour le jeu d'échec utilise une induction en aval (forward). L'idée capturée par l'induction en amont est très importante en théorie des jeux et mérite d'être citée, car entre autres, elle permet la construction explicite des stratégies optimales des joueurs. Le cas où il y a un joueur hasard nécessite certaines modifications (on reviendra sur ce point un peu plus tard dans le cours dans le chapitre sur le raffinement).

Schéma des preuves en amont et en aval: soit G_n l'ensemble de tous les jeux sous forme extensive, à deux joueurs, à information parfaite et avec seulement trois résultats possibles: I gagne, II gagne et égalité. Notre hypothèse de récurrence est que le théorème de von Neumann est vrai pour tout $k \leq n$. On veut prouver que le théorème est vrai pour $k = n + 1$. L'induction en aval (voir le jeu d'échec plus haut) consiste à partir de la racine de l'arbre, à descendre d'un pas puis d'utiliser l'hypothèse de récurrence. L'induction en amont est plus constructive : les points de départ sont les noeuds juste avant les points avant terminaux (i.e. les points qui ont comme successeurs uniquement des noeuds terminaux). Soit $g_{n+1} \in G_{n+1}$. Comme en programmation dynamique, nous réduisons d'en bas la longueur de g_{n+1} et puis nous utilisons l'hypothèse de récurrence. Nous définissons donc un nouveau $g_n \in G_n$ coïncidant avec g_{n+1} de la racine jusqu'aux noeuds avant-terminaux de g_{n+1}. Nous définissons, le résultat dans g_n d'un noeud terminal a_t dans g_{n+1} comme étant le meilleur résultat que pourrait obtenir le joueur $J(a_t)$ dans g_{n+1} en partant du noeud a_t. En effet, partant d'un noeud avant-terminal, on est face à un (sous) jeu en une étape et il est facile de déterminer la meilleur décision qu'un joueur rationnel doit prendre à ce noeud : le joueur en question gagne s'il existe une action terminale qui lui permet de gagner, il fait égalité s'il n'existe pas d'action gagnante et perd autrement. Puisque g_n a une longueur maximale égale à n, par hypothèse de récurrence, nous déduisons que g_n satisfait le théorème de Von Neumann. Il est maintenant facile de vérifier que si (i) est vrai pour g_n alors (i) est vrai pour g_{n+1} etc.

Exercices:

1) pourquoi ce théorème ne peut pas s'appliquer au jeu de Bridge ?

2) montrer que dans le cas d'information imparfaite, ce théorème n'est pas toujours vrai (utiliser pour cela le jeu de Pile et Face décrit plus haut).

3) donner un contre exemple dans le cas d'un nombre infini dénombrable de noeuds (par exemple aux étapes paires, le joueur I choisit un élément dans $\{0, 1\}$ et dans les étapes impaires II choisit un élément dans $\{0, 1\}$. Une partie correspond a une suite infinie de 0 et de 1 et peut donc être identifiée à un élément de l'intervalle $[0, 1]$)...

Chapter 7

Construction des jeux sous forme normale (ou stratégique)

Dans le jeu de Pile et Face, chaque joueur a un seul ensemble d'information et deux actions P et F. Ainsi, chaque joueur a deux stratégies qu'on notera aussi par convenance: P et F. Si le joueur I joue P et le joueur 2 joue F alors le résultat est (1,-1),.... Ceci peut être décrit par l'utilisation de la matrice suivante :

Joueur II

		P	F
Joueur I	P	1,-1	-1,1
	F	-1,1	1,-1

Figure 7.1:

Cette description est appelée la forme stratégique du jeu (aussi appelée la forme normale). Elle consiste en la donnée des ensembles de stratégies des joueurs et du résultat correspondant à chaque vecteur (profil) de stratégies.

7.1 La forme normale d'un jeu sous forme extensive

La mise en forme normale peut se faire d'une manière similaire au jeu de Pile ou Face dès qu'il n'y a pas de joueur hasard. Le point clé est que tout vecteur de stratégies $s = (s_i)_{i \in N}$ définit une trajectoire (partie) unique et donc un résultat unique $r(s)$. Ainsi, tant que nous sommes seulement intéressés par le résultat du jeu et non par le processus qui a permis d'atteindre ce résultat, on peut résumer le jeu en énumérant pour chaque joueur toutes ses stratégies et le résultat associé à chaque combinaison des stratégies.

Définition 18 *Un jeu sous forme stratégique (ou sous forme normale) est un triplet ordonné:*

$$\Gamma = \left(N, (S_i)_{i \in N}, (r_i)_{i \in N}\right)$$

où $N = \{1, \ldots, n\}$ est l'ensemble des joueurs, S_i est l'ensemble des stratégies du joueur i et r_i est une fonction réelle, définie sur le produit cartésien $S = S_1 \times S_2 .... \times S_n$.
La fonction r_i est appelée le paiement du joueur i.
Plus précisément: $r_i(s_1, \ldots, s_n)$ est le paiement du joueur i si le joueur 1 choisit la stratégie s_1, le joueur 2 choisit la stratégie s_2, etc.

Le déroulement d'un jeu sous forme stratégique est le suivant: simultanément, chaque joueur i choisit une stratégie s_i. Si $s = (s_i)_{i \in N}$ est le profil de stratégies choisit, alors chaque joueur i reçoit $r_i(s)$. Le jeu et sa description sont une information commune.

Quand S est fini, on dit que le jeu est fini. Dans ce cas, le jeu peut être représenté par une matrice n-dimensionnelle où chaque case représente un profil de stratégies $s \in S$ et contient un élément de R^n représentant le vecteur des paiements (des utilités de vNM): $r(s) = (r_i(s))_{i \in N}$. Pour cette raison, les jeux finis sont souvent appelés aussi les jeux matriciels.

Quand la forme extensive contient un joueur hasard, on peut continuer à avoir la même définition pour un jeu sous forme stratégique mais la fonction r doit être interprétée différemment: quand les vrais joueurs utilisent les stratégies $(s_i)_{i \in N}$, le résultat n'est pas déterministe: c'est une distribution de probabilité sur les points terminaux (voir l'exemple 9 ci-dessous). Pour permettre une analyse mathématique manipulable on doit supposer que les préférences des joueurs sur l'ensemble des résultats sont exprimées par une fonction d'utilité de von Neumann et Morgenstern. Plus précisément, si $u_i : O \to R$ est la fonction d'utilité du joueur i, alors l'utilité du vecteur aléatoire (o_1 avec probabilité p_1,...,o_k avec probabilité p_k) est tout simplement l'espérance: $\sum_{j=1}^{k} p_j u_i(o_j)$. Ainsi, si les paiements aux points terminaux d'un arbre sont interprétés comme étant les utilités de von Neumann et Morgenstern des joueurs associées à ces points terminaux, alors dans la forme stratégique correspondante, le paiement $r_i(s)$ est l'espérance d'utilité du joueur i si le vecteur de stratégies s est joué.
Cette construction est la transformation standard qui permet de mettre un jeu sous forme extensive, sous forme stratégique (voir ci-dessous la définition générale d'un jeu sous forme stratégique): elle associe à chaque jeu sous forme extensive un et un unique jeu sous forme stratégique. L'inverse est faux, il se peut qu'un jeu sous forme stratégique soit la représentation de plusieurs jeux sous forme extensive (voir ci-dessous).

Exemple 7

On considère le jeu en forme extensive suivant:

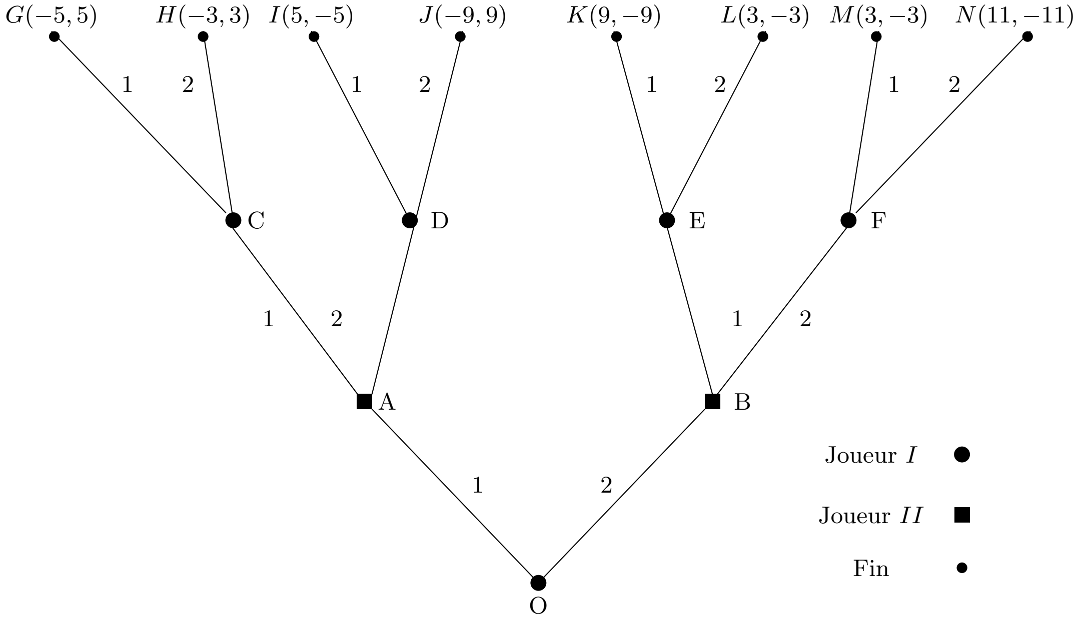

Figure 7.2:

Le joueur I a 5 ensembles d'information. Pour chaque ensemble d'information, le joueur I a deux actions possibles. Ainsi, une stratégie pour le joueur I est un quintuple d'actions $(x_O, x_C, x_D, x_E, x_F)$ qui liste une action pour chaque ensemble d'information. Le nombre de stratégies est donc: $2^5 = 32$. Le joueur II a seulement deux ensembles d'information avec chacun deux actions possibles. Ainsi, il a 2^2 stratégies, chacune consiste en une paire (y_A, y_B).

Chaque paire de stratégies détermine une partie et donc un paiement. Par exemple, si le joueur I utilise la stratégie $(2, 1, 1, 1, 2)$ et le joueur II $(2, 1)$, alors le résultat est le paiement $(9, -9)$. Au total, la forme stratégique du jeu va être donnée par la matrice 32×4 suivante:

Joueur II / Joueur I	(1,1)	(1,2)	(2,1)	(2,2)
(1,1,1,1,1)	(-5,5)	(-5,5)	(5,-5)	(5,-5)
(1,1,1,1,2)	(-5,5)	(-5,5)	(5,-5)	(5,-5)
(1,1,1,2,1)	(-5,5)	(-5,5)	(5,-5)	(5,-5)
(1,1,1,2,2)	(-5,5)	(-5,5)	(5,-5)	(5,-5)
(1,1,2,1,1)	(-5,5)	(-5,5)	(-9,9)	(-9,9)
(1,1,2,1,2)	(-5,5)	(-5,5)	(-9,9)	(-9,9)
(1,1,2,2,1)	(-5,5)	(-5,5)	(-9,9)	(-9,9)
(1,1,2,2,2)	(-5,5)	(-5,5)	(-9,9)	(-9,9)
(1,2,1,1,1)	(-3,3)	(-3,3)	(5,-5)	(5,-5)
(1,2,1,1,2)	(-3,3)	(-3,3)	(5,-5)	(5,-5)
(1,2,1,2,1)	(-3,3)	(-3,3)	(5,-5)	(5,-5)
(1,2,1,2,2)	(-3,3)	(-3,3)	(5,-5)	(5,-5)
(1,2,2,1,1)	(-3,3)	(-3,3)	(-9,9)	(-9,9)
(1,2,2,1,2)	(-3,3)	(-3,3)	(-9,9)	(-9,9)
(1,2,2,2,1)	(-3,3)	(-3,3)	(-9,9)	(-9,9)
(1,2,2,2,2)	(-3,3)	(-3,3)	(-9,9)	(-9,9)
(2,1,1,1,1)	(9,-9)	(3,-3)	(9,-9)	(3,-3)
(2,1,1,1,2)	(9,-9)	(11,-11)	(9,-9)	(11,-11)
(2,1,1,2,1)	(3,-3)	(3,-3)	(3,-3)	(3,-3)
(2,1,1,2,2)	(3,-3)	(11,-11)	(3,-3)	(11,-11)
(2,1,2,1,1)	(9,-9)	(3,-3)	(9,-9)	(3,-3)
(2,1,2,1,2)	(9,-9)	(11,-11)	(9,-9)	(11,-11)
(2,1,2,2,1)	(3,-3)	(3,-3)	(3,-3)	(3,-3)
(2,1,2,2,2)	(3,-3)	(11,-11)	(3,-3)	(11,-11)
(2,2,1,1,1)	(9,-9)	(3,-3)	(9,-9)	(3,-3)
(2,2,1,1,2)	(9,-9)	(11,-11)	(9,-9)	(11,-11)
(2,2,1,2,1)	(3,-3)	(3,-3)	(3,-3)	(3,-3)
(2,2,1,2,2)	(3,-3)	(11,-11)	(3,-3)	(11,-11)
(2,2,2,1,1)	(9,-9)	(3,-3)	(9,-9)	(3,-3)
(2,2,2,1,2)	(9,-9)	(11,-11)	(9,-9)	(11,-11)
(2,2,2,2,1)	(3,-3)	(3,-3)	(3,-3)	(3,-3)
(2,2,2,2,2)	(3,-3)	(11,-11)	(3,-3)	(11,-11)

Il est important de remarquer qu'il peut y avoir beaucoup d'informations inutiles dans la description sous forme stratégique. Par exemple, les stratégies $(2,1,1,1,2)$ et $(2,2,2,1,2)$ sont équivalentes dans un sens très robuste: quelque soit la stratégie du joueur II, la partie (et donc le résultat associé) est exactement le même pour les deux stratégies. Intuitivement, quand le joueur I joue 2 dans le noeud O, alors ce qu'il compte jouer dans les noeuds C et D n'a aucun impact sur le déroulement de la partie: ces deux stratégies sont représentées par les mêmes lignes dans la matrice. Si on identifie de telles stratégies, alors on peut représenter l'ensemble des stratégies du joueurs I par un triplet (x, z, z_2) où x est l'action que le joueur I compte choisir à O, et où z et z_2 sont les actions qu'il compte choisir après (dépendant de ce qu'il a choisi à O). Par exemple, les quatre stratégies $(2,1,1,1,2)$, $(2,1,2,1,2)$, $(2,2,1,1,2)$ et $(2,2,2,1,2)$ vont toutes être représentées par le triplet $(2,1,2)$. En groupant ensemble toutes les classes d'équivalence de la sorte, on obtient la forme stratégique suivante (qu'on appellera aussi, pour convenance, la forme stratégique du jeu):

(x, z_1, z_2) \ (y_A, y_B)	(1,1)	(1,2)	(2,1)	(2,2)
(1,1,1)	(-5,5)	(-5,5)	(5,-5)	(5,-5)
(1,1,2)	(-5,5)	(-5,5)	(-9,9)	(-9,9)
(1,2,1)	(-3,3)	(-3,3)	(5,-5)	(5,-5)
(1,2,2)	(-3,3)	(-3,3)	(-9,9)	(-9,9)
(2,1,1)	(9,-9)	(3,-3)	(9,-9)	(3,-3)
(2,1,2)	(9,-9)	(11,-11)	(9,-9)	(11,-11)
(2,2,1)	(3,-3)	(3,-3)	(3,-3)	(3,-3)
(2,2,2)	(3,-3)	(11,-11)	(3,-3)	(11,-11)

Exemple 8:

Dans le célèbre jeu d'enfant: (pierre, papier, ciseaux), chacun des deux joueurs doit choisir une action: pierre, papier ou ciseaux. Il y a un cycle entre ces trois symboles: la pierre casse les ciseaux, les ciseaux coupent le papier et le papier recouvre la pierre. Si les deux joueurs choisissent le même symbole, alors ils font match nul et aucun paiement n'est effectué. Sinon, le joueur ayant le symbole le plus fort gagne une unité et l'autre en perd une. Ainsi, la matrice suivante décrit ce jeu sous sa forme stratégique:

$$
\begin{array}{cc}
 & \textbf{Joueur } II \\
 & \begin{array}{ccc} \text{Pierre} & \text{Papier} & \text{Ciseaux} \end{array} \\
\textbf{Joueur } I \quad \begin{array}{c} \text{Pierre} \\ \text{Papier} \\ \text{Ciseaux} \end{array} & \begin{pmatrix} (0,0) & (-1,1) & (1,-1) \\ (1,-1) & (0,0) & (-1,1) \\ (-1,1) & (1,-1) & (0,0) \end{pmatrix}
\end{array}
$$

Figure 7.3:

Exemple 9

Considérons le jeu sous forme extensive dans la figure 7.3. Dans ce jeu, le résultat d'une partie n'est pas seulement déterminé par les choix des joueurs mais aussi par le choix du joueur Hasard. En fait, étant données les stratégies des joueurs, le résultat est une variable aléatoire pouvant prendre différentes valeurs. Pour bien voir cela, considérons la paire de stratégies (s, t) :

- s pour le joueur I: choisir b dans O et g dans C.
- t pour le joueur II: choisir e en B.

En supposant que les joueurs utilisent les stratégies s et t, la partie réalisée du jeu peut être une des trois possibilités:

- $b \to d \to g$, avec probabilité $\frac{1}{2}$, et le résultat final $(1,1)$.
- $b \to c \to e \to k$, avec probabilité $\frac{1}{6}$, et le résultat final $(5,-1)$.
- $b \to c \to e \to l$, avec probabilité $\frac{1}{3}$, et le résultat final $(-2,5)$.

Dans la matrice décrivant la forme stratégique, dans la case (s,t), il devrait y avoir le vecteur aléatoire suivant:

(x,y)	$(1,1)$	$(5,-1)$	$(-2,5)$
$P(x,y)$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$

Puisqu'on suppose que les paiements correspondent aux utilités de vNM l'utilité de ce vecteur aléatoire est donnée par l'espérance des utilités des points extrêmes:

$$\frac{1}{2}(1,1) + \frac{1}{6}(5,-1) + \frac{1}{3}(-2,5) = (\frac{2}{3}, 2)$$

Ainsi, dans cette exemple, la forme stratégique peut se décrire de la manière suivante:

$$\begin{array}{ccc} & & \text{II} \\ & & \begin{array}{cc} \text{f} & \text{e} \end{array} \\ \text{I} & \begin{array}{c} \text{a} \\ (b,h) \\ (b,g) \end{array} & \begin{pmatrix} (0,0) & (0,0) \\ (-\frac{3}{8}, \frac{7}{8}) & (\frac{7}{24}, \frac{15}{8}) \\ (0,1) & (\frac{2}{3}, 2) \end{pmatrix} \end{array}$$

Figure 7.4:

7.2 La correspondance entre les deux formes

La discussion précédente permet d'associer pour chaque jeu sous forme extensive, un unique jeu sous forme stratégique (à une identification près des classes de stratégies équivalentes). Le contraire est-il vrai?

Le première observation est que le jeu sous forme normale $\Gamma = \left(N, (S_i)_{i\in N}, (r_i)_{i\in N}\right)$ peut toujours se mettre sous la forme extensive suivante:

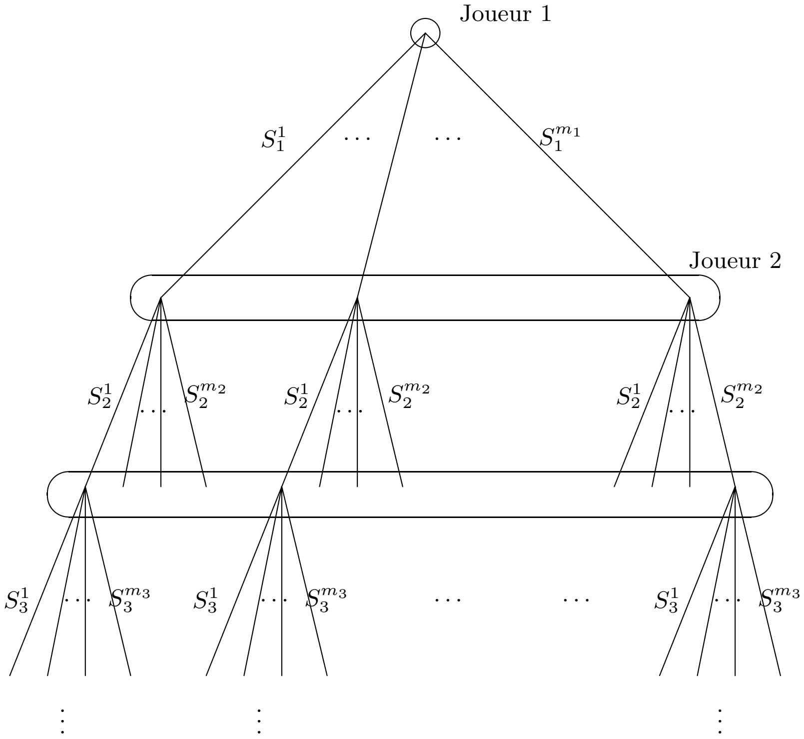

Figure 7.5:

Chaque joueur a un unique ensemble d'information (un unique point de décision). Cette dernière représentation est appelée la représentation extensive canonique. Celle-ci est déterminée d'une manière unique à une permutation près de l'ordre suivant lequel on liste les joueurs. Par exemple, le jeu de Pile ou Face :

Joueur II

		P	F
Joueur I	P	1,-1	-1,1
	F	-1,1	1,-1

Figure 7.6:

admet les deux formes canoniques suivantes :

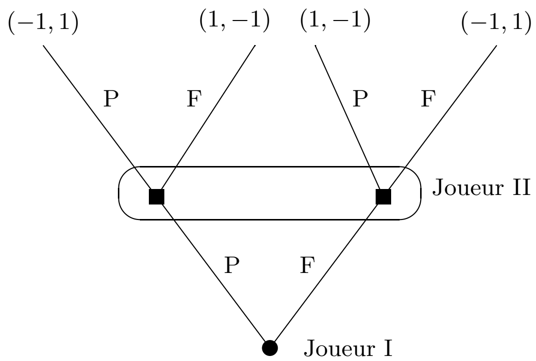

Figure 7.7:

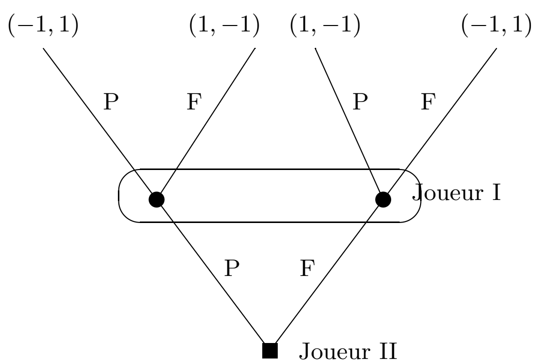

Figure 7.8:

Il est clair que ces deux formes sont essentiellement les mêmes.

Cependant, deux formes extensives ayant la même forme normale peuvent être fondamentalement différentes. Pour voir cela, considérons les trois jeux sous forme extensive suivants :

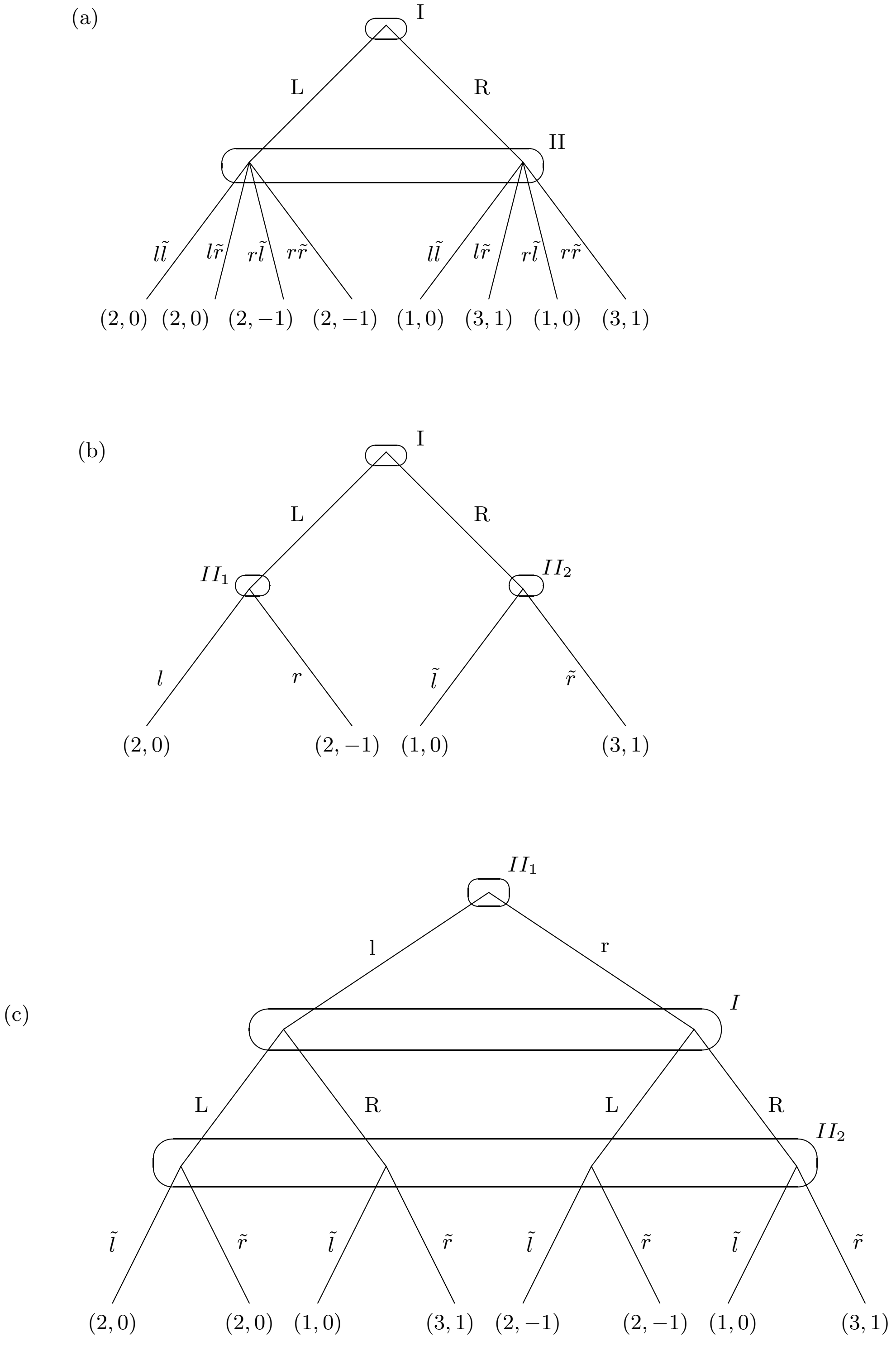

Figure 7.9:

Il est facile de vérifier que ces trois jeux ont tous la forme suivante:

$$\begin{array}{ccc}
 & & \text{Joueur II} \\
 & & \begin{array}{cccc} l\tilde{l} & l\tilde{r} & r\tilde{l} & r\tilde{r} \end{array} \\
\text{Joueur I} & \begin{array}{c} \text{L} \\ \\ \text{R} \end{array} & \begin{pmatrix} 2,0 & 2,0 & 2,-1 & 2,-1 \\ & & & \\ 1,0 & 3,1 & 1,0 & 3,1 \end{pmatrix}
\end{array}$$

Figure 7.10:

Si le jeu (a) est la représentation canonique de ce jeu matriciel, les deux autres jeux diffèrent par l'ordre des actions et la dynamique des parties, des aspects qui ne sont donc pas capturés par la forme stratégique. Ceci montre que la forme extensive est plus riche et plus détaillée que la forme stratégique.

Chapter 8

Quelques concepts de solution

Jusque là, on a exprimé (de deux manières différentes) un jeu. Cependant, on a pas encore développé les outils qui permettent d'analyser l'issue (où les issues) d'une telle interaction. En quelque sorte, on veut répondre à la question suivante: étant donné un jeu, que se passe-t-il ? La détermination des concepts de solution constitue le coeur de la théorie des jeux.

On peut définir un concept de solution comme un ensemble de lois (aboutissant à des équations mathématiques) qui permettent de sélectionner parmi toutes les issues possibles, un sous ensemble d'issues satisfaisant certaines propriétés, jugées désirables si les agents possèdent certaines facultés de raisonnement ou de comportement (rationalité, prudence, connaissance, etc...)

8.1 Elimination répétée des stratégies dominées, application aux enchères

Exemple 10

Considérons le jeu suivant (on va appeler le joueur I le joueur ligne et le joueur II le joueur colonne).

$$
\begin{array}{cc}
 & \text{Colonne} \\
 & \begin{array}{ccc} \text{L} & \text{M} & \text{R} \end{array} \\
\text{Ligne} \quad \begin{array}{c} \text{T} \\ \\ \text{B} \end{array} & \begin{pmatrix} 1,0 & 1,2 & 0,1 \\ & & \\ 0,3 & 0,1 & 2,0 \end{pmatrix}
\end{array}
$$

Figure 8.1:

Comparons les stratégies M et R pour le joueur colonne. On observe que:

- Si le joueur ligne joue T, la stratégie M donne au joueur colonne 2 alors que R donne seulement 1.
- Si le joueur ligne joue B alors, la stratégie M donne au joueur colonne 1 alors que la stratégie R donne seulement 0.

Et donc, indépendamment de ce que fait le joueur ligne, la stratégie M donne strictement plus au joueur colonne que la stratégie R.

Définition 19 *Une stratégie s_i du joueur i est dite strictement dominée s'il existe une stratégie t_i de i telle que:*
$\forall s_{-i} \in S_{-i}$, $r_i(s_i, s_{-i}) < r_i(t_i, s_{-i})$.
On dira dans ce cas que s_i est strictement dominée par t_i et que t_i domine strictement s_i.

Si s_i est strictement dominée par t_i alors, face à n'importe quelle stratégie des autres joueurs, en jouant t_i le joueur i gagne strictement plus que ce qu'il aurait gagné en jouant s_i.

Il parait donc naturel et logique de supposer qu'un joueur rationnel (cherchant à maximiser son utilité), ne va jamais jouer une stratégie strictement dominée.

Hypothèse 1: un joueur rationnel ne joue jamais une stratégie strictement dominée.

Hypothèse 2: tout les joueurs sont rationnels.

Maintenant, si les autres joueurs connaissent les stratégies et les paiements du joueur i et s'ils savent que le joueur i est un joueur rationnel, alors, ils peuvent tous anticiper que le joueur i ne va jamais jouer une stratégie strictement dominée.
Donc, **si** on est dans une situation où la rationalité de chacun des joueurs est une connaissance commune (chacun est rationnel, chacun sait que les autres sont rationnels, chacun sait que les autres savent qu'il est rationnel...etc) **alors** une stratégie strictement dominée peut être éliminée (dans le sens où tout le monde se comporte comme si cette stratégie n'existait pas).

La connaissance commune de la rationalité est importante car sinon l'élimination ne peut pas se faire. En effet, si par exemple le joueur 1 (irrationnel) compte utiliser la stratégie strictement dominée s_1 alors les autres joueurs sont obligés de prendre cela en considération (il se peut que la meilleur réponse face à s_1 soit différente de la meilleur réponse à une autre stratégie t_1 du joueur 1). Maintenant, si le joueur 1 est rationnel mais que le joueur 2 ne sait pas que le joueur 1 est rationnel, alors -par le même argument- le joueur 2 ne peut pas éliminer les stratégies strictement dominées du joueur 1. De même, si le joueur 1 ne sait pas que le joueur 2 sait que le joueur 1 est rationnel, alors, le joueur 1 ne peut pas considérer que le joueur 2 va anticiper que le joueur 1 ne va pas utiliser sa stratégie strictement dominée...

Remarquons aussi que l'élimination globale ne peut pas se faire si par exemple le joueur 1 ne connaît pas les paiements du joueur 2 car dans ce cas, il ne peut pas savoir quelles sont les stratégies strictement dominées du joueur 2.

Ainsi, sous l'hypothèse de connaissance commune de rationalité et du jeu, un bon concept de solution doit exclure toutes les stratégies strictement dominées: la dite solution du jeu est la même si on élimine de telle stratégies car elles ne sont pas jouées et tout le monde sait qu'elles ne le seront pas).

Hypothèse 3: la rationalité et le jeu sont une connaissance commune entre les joueurs.

Ainsi, dans notre l'exemple 11, la stratégie R peut être éliminée. Les joueurs considèrent donc qu'ils jouent le jeu suivant:

$$\begin{array}{cc} & \begin{array}{cc} L & M \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} (1,0) & (1,2) \\ (0,3) & (0,1) \end{pmatrix} \end{array}$$

Figure 8.2:

Mais, dans ce nouveau jeu, la stratégie B est strictement dominée par T. Ainsi, B peut être éliminée. On obtient donc le nouveau jeu réduit suivant:

$$\begin{array}{cc} & \begin{array}{cc} L & M \end{array} \\ T & \begin{pmatrix} 1,0 & 1,2 \end{pmatrix} \end{array}$$

Figure 8.3:

Dans ce jeu, L peut être éliminée. Ceci nous amène à un jeu où chaque joueur possède une unique stratégie: pour le joueur ligne T et pour le joueur colonne M avec $(1,2)$ comme paiement.

Ce procédé est appelé procédé d'élimination des stratégies strictement dominées. Quand ce procédé converge vers un unique résultat, on qualifie ce résultat comme étant la solution du jeu et on dit que le jeu est *résoluble par dominance.*

Une classe particulière de jeux résoluble par dominance est la classe de jeu où chaque joueur possède une stratégie qui domine strictement toutes les autres stratégies. Dans ce cas, on dit que le jeu possède une solution en stratégie strictement dominante. Remarquons aussi que dans ce cas on a pas besoin de l'hypothèse 3 pour déterminer le résultat du jeu car si tout les joueurs sont rationnels (et connaissent chacun son propre paiement) alors chacun va jouer sa (seule) stratégie strictement dominante...

Exemple 11 (le dilemme du prisonnier)

Un jeu très simple et en même temps très intriguant est le dilemme du prisonnier. Celui-ci a été très étudié dans la littérature et il est souvent introduit par l'histoire suivante :
Deux criminels ayant commis ensemble un crime sont arrêtés. Cependant, par manque de preuve, le procureur ne peut les condamner que si l'un des deux (ou les deux) avoue le crime. La police les sépare et fait une même offre à chacun. L'offre est la suivante:

(1) si vous avouez et que votre partenaire n'avoue pas alors on vous garantie l'immunité et vous repartez libre (2) si vous n'avouez pas alors que votre partenaire avoue alors vous aurez une peine maximale de 10 ans de prison (3) si aucun des deux n'avoue, alors on a suffisamment d'éléments pour vous faire condamner pour 1 an de prison et enfin (4) si vous avouez tous les deux alors la cour prendra cela en considération et vous condamnera à une peine réduite de 5 ans de prison. Pensez à cela et dites moi votre décision.

Ceci définit un jeu en forme stratégique. Chaque joueur a deux stratégies: A (Avouer) et NA (Ne pas Avouer). Si le résultat est exprimé en termes d'années d'emprisonnement, alors le jeu peut être décrit à l'aide de cette matrice :

$$\begin{array}{cc} & \text{Prisonnier 2} \\ & \begin{array}{cc} \quad A \quad & \quad NA \quad \end{array} \\ \text{Prisonnier 1} & \begin{array}{c} A \\ NA \end{array} \begin{pmatrix} (5,5) & (0,10) \\ (10,0) & (1,1) \end{pmatrix} \end{array}$$

Figure 8.4:

Pour enlever l'ambiguïté, récrivant le jeu en terme d'utilité, où l'utilité de sortir libre (acquittement) vaut 5 unités, l'utilité d'un an de prison vaut 4 et celle de 5 ans de prison vaut 1 et enfin l'utilité de 10 ans de prison vaut 0. On obtient alors le jeu suivant:

$$\begin{array}{cc} & \text{Prisonnier 2} \\ & \begin{array}{cc} \quad A \quad & \quad NA \quad \end{array} \\ \text{Prisonnier 1} & \begin{array}{c} A \\ NA \end{array} \begin{pmatrix} (1,1) & (5,0) \\ (0,5) & (4,4) \end{pmatrix} \end{array}$$

Figure 8.5:

Dans ce jeu NA est strictement dominé par A pour les deux joueurs. Par élimination des stratégies strictement dominées, on obtient la paire de stratégies (A, A) (les deux avouent) comme l'unique résultat avec un paiement de $(1, 1)$ (c'est-à-dire 5 ans d'emprisonnement chacun). L'aspect qui rend le dilemme du prisonnier intriguant et célèbre est que la paire de stratégie (NA, NA) est plus favorable aux deux (seulement 1 an d'emprisonnement chacun). Cependant, (NA, NA) est vulnérable aux déviations (chaque joueur est tenté de joueur A plutôt que NA et ce quelque soit sa croyance sur ce que va jouer son adversaire).

Exercice: l'ordre des éliminations ne compte pas:
Montrer que le jeu qui résulte d'une élimination répétée des stratégies strictement dominées est toujours le même: il est indépendant de l'ordre des éliminations.

Ainsi, si l'élimination des stratégies strictement dominées converge (pour un ordre d'élimination donné) vers un unique résultat s^* alors tout ordre d'élimination converge vers l'unique résultat s^*.

Cependant, l'élimination répétée des stratégies strictement dominées ne procurent **pas toujours un**

résultat unique comme le montre les exemples dans la figure qui suit où il n'y a aucune stratégie strictement dominée:

(a)

$$\begin{array}{ccc} & L & R \\ T & \multicolumn{2}{c}{\multirow{2}{*}{}} \end{array}$$

$$\begin{array}{c} \\ T \\ B \end{array} \begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} (1,2) & (2,3) \\ (2,2) & (2,0) \end{pmatrix} \end{array}$$

(b)

$$\begin{array}{c} \\ T \\ M \\ B \end{array} \begin{array}{c} \begin{array}{ccc} L & C & R \end{array} \\ \begin{pmatrix} (0,6) & (6,0) & (4,3) \\ (6,0) & (0,6) & (4,3) \\ (3,3) & (3,3) & (5,5) \end{pmatrix} \end{array}$$

Figure 8.6:

Cependant, dans le jeu (a), la stratégie B est un peu particulière car elle donne toujours un paiement au moins égal à celui de T et parfois strictement plus... On dit que B domine faiblement T.

Définition 20 *Une stratégie s_i du joueur i est dite faiblement dominée s'il existe une stratégie t_i de i telle que:*
(i) $\forall s_{-i} \in S_{-i}$, $r_i(s_i, s_{-i}) \leq r_i(t_i, s_{-i})$.
(ii) $\exists t_{-i} \in S_{-i}$, $r_i(s_i, t_{-i}) < r_i(t_i, t_{-i})$.
On dira dans ce cas que s_i est faiblement dominée par t_i et que t_i domine faiblement s_i.

On aurait donc envie encore de conseiller: une stratégie faiblement ne doit pas être jouée car il existe une autre stratégie qui donne autant sinon strictement plus. Ainsi, dans l'exemple (a) ci-dessus, ceci élimine T (faible dominance) puis R (stricte dominance) convergeant vers l'unique résultat (B, L) avec le paiement $(2, 2)$.

Hypothèse 1bis: un joueur rationnel ne joue jamais une stratégie faiblement dominé.

La meilleur justification de cette hypothèse est celle de la main tremblante de Selten. En effet, supposons que les joueurs ont une probabilité $\varepsilon > 0$ (très petite) de se tromper et de joueur n'importe quoi (on dit qu'ils tremblent). Désormais, toute stratégie qui était seulement faiblement dominée avant devient strictement dominée après la perturbation. L'hypothèse 1bis apparaît donc comme une conséquence de l'hypothèse 1 et de l'hypothèse selon laquelle chacun pense que les autres tremblent. Cependant, il est intéressant de remarquer qu'il y a une contradiction entre le désir d'éliminer une stratégie (considérer qu'elle ne va jamais être jouée) et dire que toutes les stratégies peuvent être jouées avec une probabilité strictement positive...

Exemple 14 (l'enchère de Vickrey)

En fait, l'élimination des stratégies faiblement dominées peut parfois se révéler très puissant comme le montre cet exemple d'application économique.

Le modèle d'enchère suivant a été proposé par Vickrey. Un objet indivisible (par exemple un tableau) est vendu suivant la procédure suivante: chaque acheteur potentiel soumet sous enveloppe une proposition b. L'acheteur qui soumet la plus grande offre gagne l'objet et paye pour l'acquérir le second meilleur prix offert. On suppose que chaque acheteur potentiel a une évaluation v pour l'objet qui reflète toute valeur objective ou subjective pour lui.

Remarquons d'abord que cette situation est très complexe qui, jusqu'à présent, ne décrit pas une jeu, si on ne rajoute pas d'autres hypothèses concernant les croyances de chaque acheteur sur les évaluations des autres acheteurs....etc. Remarquons aussi que l'ensemble des stratégies pour chaque joueur n'est pas fini mais un continuum.

Cependant, sans aucun autre ingrédient, on peut résoudre ce jeu en remarquant tout simplement que: proposer sa propre évaluation est une stratégie faiblement dominante pour chaque acheteur. Ainsi, une solution en stratégie faiblement dominée pour ce jeu est la suivante:

- chaque acheteur i propose son évaluation ($b_i = v_i$);

- l'objet est donné à l'acheteur ayant la plus haute évaluation au prix de la second plus haute évaluation.

  Expliquons les raisons pour lesquelles vous devez choisir $b = v$ (si vous étiez vous même un participant à l'enchère):

- d'abord, vous faites une bonne affaire dès lors que vous payez un prix inférieur à v et vous faites une mauvaise affaire dès lors que vous payez plus que v;

- conditionnelement au fait que vous gagnez l'objet, le prix que vous payez est indépendant de votre offre b. Celui-ci dépend seulement des autres offres;

- en offrant $b < v$ vous risquez une opportunité de gagner si jamais la plus haute offre des autres acheteurs est strictement entre b et v.

- En offrant $b > v$ vous risquez, en gagnant l'objet, de payer plus que votre évaluation (si jamais la deuxième meilleure offre est strictement entre v et b).

Il est facile de voir que l'effet des stratégies des autres joueurs sur le paiement de i est résumé dans y: la plus haute offre des joueurs autres que i ($y = \max_{j \neq i} b_j$). Pour cette raison, il nous suffit donc de tracer le profit d'un acheteur i ayant une évaluation v_i comme une fonction de la variable y et ce pour les trois types de stratégies possibles: $b_i < v_i, b_i = v_i$ $b_i > v_i$:

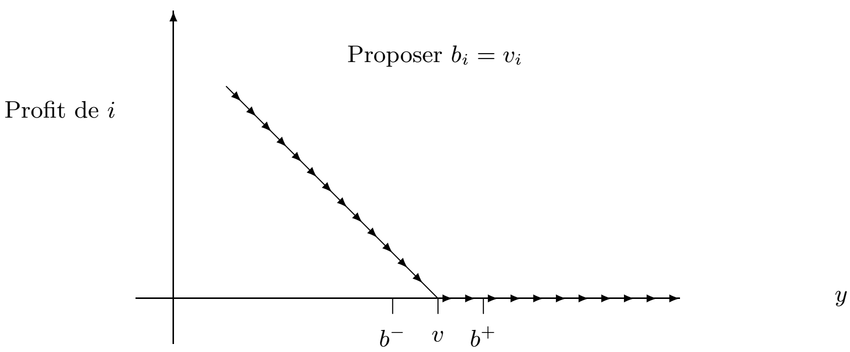

Figure 8.7:

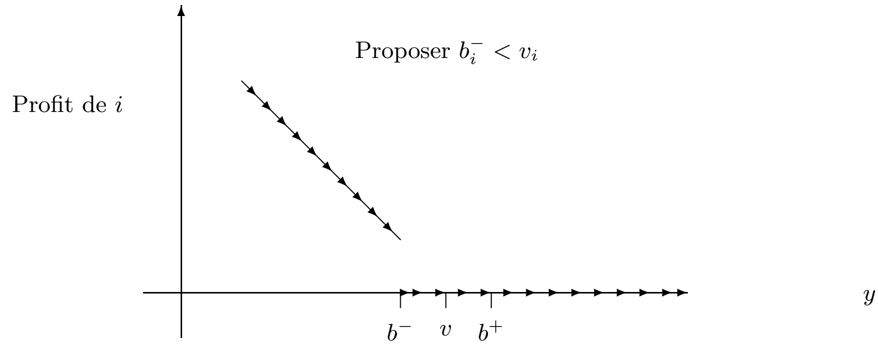

Figure 8.8:

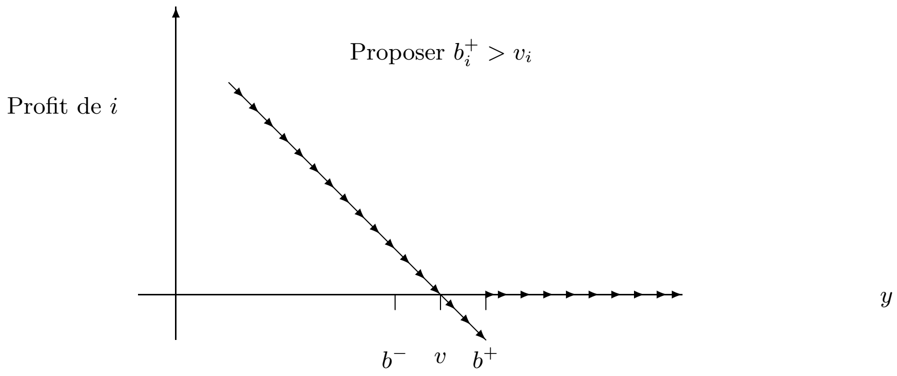

Figure 8.9:

Cela montre clairement que proposer $b = v$ donne un profit au moins égal à une autre stratégies et parfois même strictement plus. En d'autre terme, la stratégie $b = v$ est une stratégie faiblement dominante (c'est même l'unique sous l'hypothèse que toutes les valeurs de y sont possibles).

Dépendance par rapport à l'ordre:
Contrairement aux cas des stratégies strictement dominées, le jeu résultant (et donc le résultat) de l'élimination des stratégies qui sont seulement faiblement dominées peut dépendre de l'ordre suivant lequel les éliminations se font. Ceci est illustré dans l'exemple suivant:

$$\begin{array}{c c} & \begin{array}{ccc} L & M & R \end{array} \\ \begin{array}{c} T \\ M \\ B \end{array} & \begin{pmatrix} 1,2 & 2,3 & 0,3 \\ 2,2 & 2,1 & 3,2 \\ 2,1 & 0,0 & 1,0 \end{pmatrix} \end{array}$$

ordre d' élimination	résultat	paiement final
T,R,B,C	ML	$2,2$
B,L,C,T	MR	$3,2$

Figure 8.10:

Ceci rend contestable la prédiction qui résulte d'une élimination des stratégies faiblement dominantes. En effet, l'exemple précédent montre clairement que les premières éliminations peuvent être déterminantes sur le résultat final. Puisque nos hypothèses (1, 1bis, 2 et 3) ne spécifient pas l'ordre par rapport auquel les éliminations doivent être faites, il n'y a donc aucune raison (sans communiquer) que les joueurs s'accordent sur la stratégie à éliminer en premier ?

8.2 Equilibre de Nash

La notion d'élimination des stratégies dominées reste limitée: en comparant la performance de deux stratégies, les situations intéressantes, dans certain cas l'une est meilleure que l'autre et dans d'autres cas c'est le contraire a lieu. Le jeu (b) de la figure 7.9 illustre bien cette situation.

Pour le joueur ligne, M est préférable à T si le joueur colonne joue L, cependant M est moins bien que T si le joueur colonne choisit C. Ainsi, il n'y a pas de relation de dominance entre les deux stratégies.

Cependant, il y a un candidat potentiel pour la solution du jeu: la paire de stratégies (B, R) donnant le paiement $(5, 5)$. En quel sens? En fait, s'il est suggéré (par quelqu'un, par un raisonnement commun aux joueurs, après une négociation..) que (B, R) doit être joué, alors, aucun joueur n'a intérêt à changer tout seul en jouant une autre stratégie. En effet, si par exemple le joueur 1 envisage de jouer M à la place de B, il obtient seulement 4 unités (alors que son utilité en jouant B est de 5 unités). En fait, on vérifie facilement qu'aucun joueur n'a intérêt à changer seul de stratégie si (B, R) est la paire de stratégies prévue.

La notion d'équilibre caché derrière ce raisonnement est en faite l'une des plus importantes en théorie des jeux.

Supposons l'existence d'une sorte de norme sociale (**ex.** rouler à droite pour un conducteur de voiture) qui soit connaissance commune, et qui pour chaque interaction stratégique (où pour la situation en question), fait qu'un joueur donné sait ce que vont faire les autres joueurs, **alors** pour être effectivement stable cette norme doit au moins satisfaire à ce critère: aucun joueur n'a intérêt à dévier individuellement du comportement dicté par la norme - sinon la norme pourrait être violée puisqu'il existe au moins un individu qui, anticipant que tout le reste de la société va suivre la norme, lui a intérêt à la violer- (**ex.** aucun conducteur de voiture n'a intérêt à rouler à gauche si tout le monde roule à droite).

L'équilibre de Nash est simplement la traduction mathématique de cette notion de stabilité.

Définition 21 *(Nash)*
Un profil de stratégies $s^* = (s_i^*)_{i \in N} \in S$ *est un équilibre de Nash du jeu sous forme stratégique* Γ *si et seulement si,* $\forall i \in N, \forall s_i \in S_i : r_i(s_i^*, s_{-i}^*) \geq r_i(s_i, s_{-i}^*)$.

Ainsi, si les joueurs devaient se mettre d'accord sur un profil à jouer (resp. s'il y avait une théorie élaborée par de grands sages qui soit une connaissance commune et qui dicte -ou anticipe- le profile qui doit -ou va- être joué), alors ce profil est nécessairement un équilibre de Nash: sinon, au moins un joueur aura intérêt à dévier et donc à ne pas respecter l'accord établit avec les autres joueurs (resp. le conseil -la prédiction- des sages)

Ou encore, si dans une société, s'établit une certaine régularité dans le comportement volontaire (et non imposé) des agents dans une situation d'interaction et que cette régularité est observable par tous, alors, si les agents sont rationnels, ils sont nécessairement en train de jouer un équilibre de Nash (car sinon, un des agents à intérêt à changer de comportement et donc à violer cette régularité observée).
Cela explique pourquoi cette notion d'équilibre a un rôle important en sciences sociales.

Un des problèmes importants de la recherche actuelle est de trouver par quel procédé intellectuel ou autre (comme par exemple l'apprentissage) et sous quelles conditions, les comportements des agents convergent vers un équilibre de Nash et sinon, vers quoi il y a convergence.

Exemples

Dans le jeu suivant:

$$\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ M \\ B \end{array} & \begin{pmatrix} 3,7 & 0,0 \\ 4,1 & 5,3 \\ 6,5 & 4,7 \end{pmatrix} \end{array}$$

Figure 8.11:

(M, R) est l'unique équilibre de Nash.

Dans **le jeu de coordination** suivant:

$$\begin{array}{c c} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \\ \text{B} \end{array} & \begin{pmatrix} (1,1) & (0,0) \\ \\ (0,0) & (3,3) \end{pmatrix} \end{array}$$

Figure 8.12:

(T, L) et (B, R) sont tous les deux des équilibres de Nash.

Et dans le jeu **(la bataille des sexes)** suivant:

$$\begin{array}{c c} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \\ \text{B} \end{array} & \begin{pmatrix} (1,3) & (0,0) \\ \\ (0,0) & (3,1) \end{pmatrix} \end{array}$$

Figure 8.13:

(T, L) et (B, R) sont des équilibres de Nash.

Les deux derniers exemples montrent que cette notion à elle seule ne suffit pas pour déterminer une unique solution. Elle doit être considérée comme un méta concept: elle capture une propriété de stabilité qui devrait être satisfaite par tout concept de prédiction raisonnable.

Une question nécessaire pour la cohérence de notre raisonnement est la suivante : est-ce-que la solution obtenue par élimination répétée de stratégies dominées satisfait la condition de stabilité de Nash ? c'est ce que montre les propositions suivantes.

Proposition 6 *Si le jeu est soluble par élimination répétée de stratégies faiblement dominées, alors le vecteur de stratégies qui en résulte est un équilibre de Nash du jeu initial.*

Preuve. Il suffit de remarquer qu'après l'élimination d'une stratégie faiblement dominante, un équilibre du jeu réduit est aussi un équilibre du jeu initial. On conclut alors par récurrence. ■

Proposition 7 *Si le jeu est soluble par élimination répétée des stratégies **strictement** dominées, alors le vecteur de stratégies qui en résulte est l'unique équilibre de Nash du jeu initial.*

Preuve. Il suffit de remarquer qu'une stratégie strictement dominée ne peut jamais être jouée dans un équilibre de Nash du jeu initial. D'où après son élimination, le jeu de départ et le jeu réduit ont exactement les mêmes équilibres de Nash. ■

Ainsi, dans l'exemple 11, (T, M) est l'unique équilibre de Nash car c'est l'unique résultat d'une élimination répétée de stratégies strictement dominées. La même observation peut être faite pour le dilemme du prisonnier.

Le résultat de l'élimination de stratégies faiblement dominées est aussi un équilibre de Nash mais n'est pas l'unique équilibre de Nash. En effet, il se peut qu'un équilibre de Nash soit éliminé. Dans le jeu (a) figure 7.9, l'unique résultat des éliminations répétées des stratégies faiblement dominées est (B, L) (avec paiement $(2, 2)$), cependant (T, R) avec le paiement $(2, 3)$ est aussi un équilibre de Nash. Aussi, même si le résultat d'une élimination répétée de stratégies faiblement dominées est unique, il se peut qu'il qu'un équilibre de Nash soit éliminé, comme l'illustre l'exemple suivant :

$$\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} A \\ \\ B \end{array} & \begin{bmatrix} 0,0 & 2,1 \\ & \\ 1,2 & 1,2 \end{bmatrix} \end{array}$$

Ce jeu possède deux équilibres de Nash: (A, R) et (B, L). Or, R domine faiblement L, d'où, l'unique équilibre qui subsite après élimination des stratégies faiblement dominées est (A, R).

Enfin, il existe des jeux (le jeu de Pile ou Face par exemple) qui n'admettent aucun équilibre de Nash.

8.3 Le comportement prudent (maxmin)

Un autre aspect, non capturé par l'équilibre de Nash, et la notion de risque. En effet, considérant le jeu suivant:

$$\begin{array}{cc} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \\ \text{M} \\ \\ \text{B} \end{array} & \begin{pmatrix} (3,1) & (2,2) \\ & \\ (0,8) & (0,-1) \\ & \\ (-100,2) & (3,3) \end{pmatrix} \end{array}$$

Figure 8.14:

Il est facile de voir que (B, R) avec le paiement $(3, 3)$ est l'unique équilibre de Nash. Cependant, même si le joueur ligne croit que le joueur colonne va jouer R, va-t-il vraiment jouer B ? est-ce-que jouer B n'est pas trop dangereux ?, va-t-il vraiment prendre le risque de perdre 100 unités si pour une raison ou une autre (irrationalité, stupidité, vengeance, erreur) le joueur colonne joue L ? Ceci montre un autre aspect pour le concept de solution qui n'est pas présent dans le raisonnement sous-jacent à l'équilibre de Nash: les notions de risque et de sécurité. L'équilibre de Nash (B, R) est très dangereux pour le joueur ligne car il risque de perdre beaucoup si jamais le joueur colonne ne joue pas la stratégie d'équilibre alors qu'en jouant T, le joueur ligne se garantit au moins un paiement de 2 et ce quel que soit ce que fait le joueur colonne. Une règle de comportement qui prend compte le risque est la suivante.

Le joueur i est prudent (ou pessimiste) s'il croit (ou considère) qu'en utilisant la stratégie s_i, les autres joueurs vont choisir le vecteur de stratégies s_{-i} qu'il déteste le plus: $r_i(s_i, s_{-i}) = \min_{t_{-i} \in S_{-i}} r_i(s_i, t_{-i})$. Dans ce cas, s'il est rationnel, alors il va jouer la stratégie s_i^* qui va maximiser son paiement étant donné sa croyance :

$$r_i(s_i^*, s_{-i}^*) = \min_{t_{-i}} r_i(s_i^*, t_{-i}) = \max_{s_i} \min_{t_{-i}} r_i(s_i, t_{-i}) := \underline{v_i}$$

Donc, si le joueur i joue la stratégie s_i^*, alors il est sûr qu'au pire des cas d'obtenir le paiement: $\underline{v_i}$.

Définition 22 *s_i^* est appelée la stratégie maxmin (où la stratégie prudente), $\underline{v_i}$ est appelée: la valeur maxmin du joueur i. Un vecteur de stratégies où tous les joueurs utilisent une stratégie maxmin est appelé: un vecteur de stratégies maxmin.*

Pour trouver le valeur maxmin et les stratégies maxmin du joueur ligne, on rajoute une colonne dans laquelle on inscrit dans chaque ligne l, le paiement minimal du joueur ligne quand il joue cette ligne l.

	L	R	min
T	$(3,1)$	$(2,2)$	2
M	$(0,8)$	$(0,-1)$	0
B	$(-100,2)$	$(3,3)$	-100

Figure 8.15:

La valeur maximale de cette colonne étant 2, on déduit donc que la valeur maxmin du joueur ligne est $\underline{v_1} = 2$ et que sa stratégie maxmin est T.

D'une manière similaire, pour trouver le maxmin du joueur colonne, on rajoute une ligne où on inscrit le paiement minimal de joueur 2.

	L	R
T	$(3,1)$	$(2,2)$
M	$(0,8)$	$(0,-1)$
B	$(-100,2)$	$(3,3)$
min	1	-1

Figure 8.16:

On déduit donc que $\underline{v_2} = 1$ et que L est la stratégie maxmin du joueur colonne.

Par définition des stratégies maxmin, on déduit que le vecteur de paiements $r = (r_i)_{i\in N}$ associé à un vecteur de stratégies maxmin satisfait toujours: pour chaque joueur i, $r_i \geq \underline{v_i}$. En fait, cette inégalité peut être stricte. En effet, dans l'exemple précédent, la stratégie maxmin pour le joueur ligne est de jouer T et pour le joueur colonne est de jouer L. Le paiement associé au vecteur de stratégies maxmin est donc $(3,1)$ alors que la valeur maxmin du joueur ligne est 2.

Remarquons enfin que le vecteur de paiement associé au vecteur de stratégies maxmin n'est pas toujours bien défini car un joueur peut avoir plusieurs stratégies maxmin, comme le montre l'exemple suivant:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \\ \mathrm{B} \end{array} & \begin{pmatrix} (3,1) & (0,4) \\ & \\ (2,3) & (1,1) \end{pmatrix} \end{array}$$

Figure 8.17:

Dans ce jeu, si les deux joueurs jouent une stratégie maxmin, le paiement final peut être $(3, 2)$ ou $(1, 1)$.

L'équilibre de Nash et le maxmin sont deux concepts basés sur deux idées différentes: l'un capture la stabilité et l'autre la sécurité. Typiquement, ces deux concepts donnent des prédictions différentes. Cependant, il y a des jeux pour lesquels ces deux concepts coïncident: dans le dilemme du prisonnier (s_1, s_1) est l'unique prédiction pour les deux concepts. Ceci est en fait un cas particulier de la classe suivante:

Proposition 8 *Une stratégie faiblement dominante est aussi une stratégie maxmin.*

Corollaire 1 *Dans un jeu où chaque joueur a une stratégie faiblement dominante, le vecteur des stratégies dominantes est un équilibre de Nash et aussi un vecteur de stratégies maxmin.*

Corollaire 2 *Dans un jeu où chaque joueur a une stratégie strictement dominante, le vecteur des stratégies strictement dominantes est l'unique équilibre de Nash et aussi l'unique vecteur de stratégies maxmin.*

Chapter 9

Jeux à deux joueurs et à somme nulle

9.1 Motivation

Dans cette section, nous allons étudier une classe de jeux très importante pour plusieurs raisons. D'abord c'est la classe qui a donné naissance à la théorie des jeux comme discipline à part entière, mais aussi c'est une classe pour laquelle les deux idées à savoir celle de l'équilibre mais aussi celle de la prudence coïncident.

Soit un jeu à deux joueurs et sous forme normale $\Gamma = (\{1,2\}, S_1, S_2, h_1, h_2)$ où les joueurs sont notés par 1 et 2, les ensembles de stratégies sont S_1 et S_2 et les fonctions de paiements: sont r_1 et r_2.

Définition 23 *Un jeu à deux joueurs est dit strictement compétitif si pour toute paire de résultats x et y, si le joueur 1 préfère x à y alors le joueur 2 préfère y à x, autrement dit, pour toute paire de vecteurs de stratégies $s = (s_1, s_2)$ et $t = (t_1, t_2)$:*

$$r_1(s) \leq r_1(t) \Leftrightarrow r_2(s) \geq r_2(t)$$

Un exemple de jeux strictement compétitif est le jeu d'échec. En effet, les préférences du joueur Blanc sur les résultats sont $B \succ_b E \succ_b N$ alors que les préférences du joueur Noir sont: $B \prec_n E \prec_n N$. Une fonction de paiement simple permet de représenter les préférences du joueur Blanc:

$$r_b(s_b, s_n) = \begin{cases} 1 & si \quad B \\ 0 & si \quad E \\ -1 & si \quad N \end{cases}$$

Et pour le joueur 2 on pourrait prendre :

$$r_n(s_b, s_n) = \begin{cases} -1 & si \quad B \\ 0 & si \quad E \\ 1 & si \quad N \end{cases}$$

Clairement, il y a plusieurs fonctions de paiement pouvant décrire ces préférences, par exemple, ou aurait pu choisir la fonction suivante pour le joueur 2:

$$\widehat{r}_n(s_b, s_n) = \begin{cases} 0 & si \quad B \\ 1 & si \quad E \\ 3 & si \quad N \end{cases}$$

Cependant, la première représentation a une propriété mathématique avantageuse: toute issue du jeu donne lieu à une paire de paiement (r_b, r_n) avec $r_b + r_n = 0$.

Définition 24 *Un jeu à deux joueurs est dit à somme nulle si pour tout vecteur de stratégies* $s = (s_1, s_2)$ *on a* $r_1(s) + r_2(s) = 0$.

Ainsi, dans un jeu à somme nulle, pour toute configuration, le paiement total reçu par les joueurs est égal à zero. Généralement, on dit que le joueur 2 paye au joueur 1 le montant $r_1(s)$. Dans se cas, $r_1(s)$ est ce que gagne le joueur 1 et ce que perd le joueur 2, qui lui gagne donc $-r_1(s)$. Ainsi, la première simplification technique due à cette propriété est qu'il suffit de spécifier la fonction de paiement du joueur 1. La convention est de définir la fonction de paiement $r(\cdot)$ pour le joueur 1 et d'appeler le joueur 1 le maximiseur (puisqu'il est intéressé par les valeurs élevées de r et qu'il veut donc maximiser son paiement r). Le joueur 2 a donc la fonction de paiement $-r$ et il est appelé le minimiseur (car il veut maximiser $-r$, donc minimiser r).

9.2 Equilibre et valeur

Quand ce jeu est présenté sous une forme matricielle, la convention est de considérer que le maximiseur (le joueur 1) est le joueur ligne et que le minimiseur est le joueur colonne. Sur la matrice, on représente seulement le paiement du maximiseur (le joueur ligne). Ainsi, le jeu Pile ou Face, qui est un jeu à somme nulle, se présente sous la forme:

$$\begin{array}{ccc} & & \text{Joueur 2} \\ & & \begin{array}{cc} \text{P} & \text{F} \end{array} \\ \text{Joueur 1} & \begin{array}{c} \text{P} \\ \text{F} \end{array} & \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{array}$$

Figure 9.1:

La valeur maxmin pour le maximiseur dans un jeu à somme nulle est:

$$\underline{v} = \max_{s_1 \in S_1} \min_{s_2 \in S_2} r(s_1, s_2)$$

Ceci est le niveau de sécurité du joueur 1; il peut garantir (en utilisant une stratégie maxmin) un paiement au moins égal à $\underline{v}$.

Quand au joueur 2 (le minimiseur) il peut se garantir un paiement au moins égal à:

$$\max_{s_2 \in S_2} \min_{s_1 \in S_1} (-r(s_1, s_2)) = - \min_{s_2 \in S_2} \max_{s_1 \in S_1} r(s_1, s_2) = -\overline{v}$$

avec:

$$\overline{v} = \min_{s_2 \in S_2} \max_{s_1 \in S_1} r(s_1, s_2)$$

$\overline{v}$ est appelé la valeur minmax du jeu: le joueur 2 peut se garantir de ne pas payer plus que $\overline{v}$ au joueur 1 (donc de ne pas perdre plus que $\overline{v}$). En fait, $\overline{v}$ est le montant minimal ayant cette dernière propriété. Aussi, $\underline{v}$ est appelé la valeur maxmin du jeu.

Proposition 9 *Pour tout jeu à somme nulle on a:*

$$(\overline{v} =) \min_{s_2 \in S_2} \max_{s_1 \in S_1} r(s_1, s_2) \geq \max_{s_1 \in S_1} \min_{s_2 \in S_2} r(s_1, s_2) \, (= \underline{v})$$

Preuve. Il est facile de voir que:

$$\max_{s_1} h(s_1, s_2) \geq h(s_1, s_2) : \forall s_1, s_2$$

Donc

$$\overline{v} = \min_{s_2} \max_{s_1} h(s_1, s_2) \geq \min_{s_2} h(s_1, s_2), : \forall s_1$$

et finalement:

$$\overline{v} \geq \max_{s_1} \min_{s_2} h(s_1, s_2) = \underline{v}$$

■

En fait, l'inégalité $\overline{v} \geq \underline{v}$ peut être stricte. Par exemple, dans le jeu de Pile ou Face:

$$\begin{array}{cccccc} & & \text{Joueur 2} & & & \\ & & \text{P} & \text{F} & \text{min} & \\ \text{Joueur 1} & \begin{array}{c} \text{P} \\ \text{F} \end{array} & \multicolumn{2}{c}{\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}} & \begin{array}{c} -1 \\ -1 \end{array} & \text{maxmin} \\ & & & & -1 & \\ & \text{max} & 1 & 1 & 1 & \\ & & \text{minmax} & & & \end{array}$$

Figure 9.2:

Ainsi, dans ce jeu $\overline{v} = \min_{s_1} \max_{s_2} = 1$ alors que $\underline{v} = \max_{s_1} \min_{s_2} = -1$. Remarquons aussi que ce jeu n'admet pas non plus d'équilibre de Nash (et en particulier le vecteur des stratégies maxmin n'est pas un équilibre de Nash).

Cependant, dans le jeu suivant, les valeurs maxmin et minmax coïncident:

$$\begin{array}{ccc} & \begin{array}{ccc} \text{L} & \text{M} & \text{R} \end{array} & \text{min} \\ \begin{array}{c} \text{T} \\ \text{M} \\ \text{B} \end{array} & \begin{pmatrix} 3 & -5 & -2 \\ 1 & 4 & 1 \\ 6 & -3 & -5 \end{pmatrix} & \begin{array}{c} -5 \\ 1 \\ -5 \end{array} \\ \text{max} & \begin{array}{ccc} 6 & 4 & 1 \end{array} & \end{array}$$

Figure 9.3:

d'où: $\min_{s_2}\max_{s_1} h(s_1,s_2) = 1 = \max_{s_1}\min_{s_2} h(s_1,s_2)$. Autrement dit, dans ce jeu, le joueur 1 garantit de recevoir (du joueur 2) au moins égale à $v=1$ et le joueur 2, quant à lui, il garantit de ne pas payer au joueur 1 plus que la $v=1$. Ce montant semble être l'issue logique du jeu, et il est appelé la valeur du jeu.

Définition 25 *Un jeu à somme nulle a une valeur v si $\overline{v} = \underline{v}$ $(= v)$. Quand un jeu a une valeur, les stratégies maxmin et minmax sont aussi appelées des stratégies optimales.*

On a vu que dans le jeu de Pile ou Face, la valeur n'existe pas et que en même temps, l'équilibre de Nash n'existe pas. Dans le dernier exemple, la valeur existe, et comme par hasard, on vérifie qu'un équilibre de Nash existe aussi. On a même plus: l'unique équilibre de Nash de ce jeu est (M,R) qui correspond tout simplement au vecteur de stratégies maxmin et le paiement de l'équilibre de Nash n'est autre que la valeur. En fait, ceci n'est pas du hasard, mais un résultat général:

Proposition 10 *Dans un jeu à deux joueurs et à somme nulle, si la valeur existe et si s_1^*, s_2^* sont des stratégies optimales alors: $s^* = (s_1^*, s_2^*)$ est un équilibre de Nash avec comme paiement $(v,-v)$.*

Preuve. Remarquons tout d'abord que la condition d'équilibre de Nash, pour un jeu à somme nulle, est tout simplement la condition de point selle pour une fonction à deux variables. En effet, si $\left(s_1^0, s_2^0\right)$ est un équilibre de Nash si et seulement si :

$$\begin{array}{rcl} r(s_1^0,s_2^0) & \geq & r(s_1,s_2^0), : \forall s_1 \\ -r(s_1^0,s_2^0) & \geq & -r(s_1^0,s_2), : \forall s_2 \end{array}$$

soit :

$$r(s_1,s_2^0) \leq r(s_1^0,s_2^0) \leq r(s_1^0,s_2) \qquad \forall s_1, s_2$$

Maintenant, puisque s_1^* est une stratégie maxmin et que $\underline{v} = v$ on en déduit que $r(s_1^*, s_2) \geq v$, $\forall s_2$ (a). D'une manière similaire, puisque s_2^* est une stratégie minmax et que $\overline{v} = v$ on déduit que $r(s_1, s_2^*) \leq v$, $\forall s_1$ (b). En prenant $s_1 = s_1^*$ et $s_2 = s_2^*$ dans (a) et (b) on obtient $v \leq r(s_1^*, s_2^*) \leq v$, d'où $v = r(s_1^*, s_2^*)$. En remplaçant v par sa valeur dans (a) et (b) en conclut que :

$$r(s_1,s_2^*) \leq r(s_1^*,s_2^*) \leq r(s_1^*,s_2) \qquad \forall s_1, s_2$$

D'où (s_1^*, s_2^*) est un équilibre de Nash et $(v,-v)$ est le paiement associé. ■

Voici la réciproque de ce dernier résultat:

Proposition 11 *Dans un jeu à deux joueurs et à somme nulle, si (s_1^*, s_2^*) est un équilibre de Nash alors le jeu a une valeur v, $v = r(s_1^*, s_2^*)$ et (s_1^*, s_2^*) sont des stratégies optimales.*

Preuve. La condition d'équilibre de Nash implique:

$$r(s_1,s_2^*) \leq r(s_1^*,s_2^*) \leq r(s_1^*,s_2) \qquad \forall s_1, s_2$$

On déduit donc que

$$\max_{s_1 \in S_1} r(s_1,s_2^*) \leq r(s_1^*,s_2^*) \leq \min_{s_2 \in S_2} r(s_1^*,s_2)$$

D'où:

$$
\begin{aligned}
\overline{v} &= \min_{s_2}\max_{s_1} r(s_1, s_2) \\
&\leq \max_{s_1\in S_1} r(s_1, s_2^*) \\
&\leq r(s_1^*, s_2^*) \\
&\leq \min_{s_2\in S_2} r(s_1^*, s_2) \\
&\leq \max_{s_1}\min_{s_2} h(s_1, s_2) \\
&= \underline{v}
\end{aligned}
$$

Or, on a toujours $\underline{v} \leq \overline{v}$, d'où toutes les inégalités ci-dessus sont en fait des égalités. ■

Corollaire 3 *Si un jeu à deux joueurs a une valeur alors celle-ci est bien définie (i.e. si v est une valeur du jeu et que v' est aussi une valeur du jeu alors $v = v'$).*

Corollaire 4 *Dans un jeu à somme nulle, si (s_1^*, s_2^*) et (s_1^{**}, s_2^{**}) sont des équilibres de Nash alors:*

(i) $r\left(s_1^*, s_2^*\right) = r\left(s_1^{**}, s_2^{**}\right)$

(ii) (s_1^*, s_2^{**}) *et* (s_1^{**}, s_2^*) *sont aussi des équilibres de Nash.*

En fait, aucune de ces deux dernières propriétés n'est vraie pour un jeu à somme non nulle. En effet, pour le jeu de coordination (figure 8.12), (T, L) et (B, R) sont tous les deux des équilibres de Nash, cependant, ils induisent des paiements différents et ni (T, R) ni (B, L) n'est un équilibre de Nash.

9.3 Interprétation du théorème de von Neumann

Représentons le jeu d'échec comme un jeu à somme nulle avec comme fonction de paiement:

$$
r(s) = \begin{cases} 1 & si \quad B \\ 0 & si \quad E \\ -1 & si \quad N \end{cases}
$$

Rappelons ce que dit le théorème de Von Neumann: il y a seulement trois possibilités:

(i) le joueur blanc a une stratégie gagnante

(ii) le joueur noire a une stratégie gagnante

(iii) chaque joueur a une stratégie qui lui garantie au moins l'égalité.

Il est facile de voir que (i) est équivalente à $\underline{v} \geq 1$. Puisque 1 et la plus grande valeur du jeu, on déduit que nécessairement $\overline{v} \leq 1$. Puisque $\underline{v}$ est toujours inférieur à $\overline{v}$ on déduit donc que (i) est équivalente à (i') où:

(i') la valeur du jeu d'échec existe et est égale à 1.

Par symétrie, (ii) est équivalente à (ii') où:

(ii') la valeur du jeu d'échec existe et est égale à -1.

Enfin, (iii) est équivalente à $\underline{v} \geq 0$ (le joueur Blanc garantie l'égalité) et $\overline{v} \leq 0$. D'où (iii) est équivalente à (iii') où:

(iii') la valeur du jeu d'échec existe et est égale à 0.

D'où, le théorème de Von Neumann peut être reformulé de la manière suivante:

Théorème 11 *Le jeu d'échec a une valeur.*

Remarquons que le théorème de Von Neumann ne dit rien sur le montant de la valeur du jeu d'échec (celle-ci pouvant aussi bien être -1, 0 ou 1). Elle est égale à une et une seule de ces valeurs, mais on ne sait pas laquelle.

Le théorème de Von Neumann en général prend donc la forme suivante :

Théorème 12 *Tout jeu sous forme extensive, à information parfaite, à deux joueurs et à somme nulle, admet une valeur.*

Chapter 10

Stratégies mixtes et applications

Les stratégies mixtes formalisent l'idée que dans certaines situations réelles, les agents doivent être imprévisible. En effet, si un joueur de tennis servait toujours du même côté, alors il serait beaucoup plus facile pour son adversaire de renvoyer la balle. De même, si un joueur de football utilise toujours le côté droit lors de la séance des tirs au but, cette information publique peut servir à n'importe quel gardien de but pour arrêter le tir. On pourrait dire de même d'un entraîneur, qui doit non seulement établir une stratégie pour son équipe mais aussi veiller à ce que cela reste une surprise pour l'adversaire. Un homme politique doit aussi veiller à ce que son programme soit imprévisible avant son entré en campagne, il l'a même intérêt à être flou en début de campagne pour pouvoir s'adapter au programme de son adversaire et à aux attentes des l'électeurs...

Si la stratégie consistant à *être imprévisible* est clairement utilisée par certains agents c'est parce qu'il est optimal pour eux d'agir ainsi.

Au niveau théorique, sans introduire la prossibilité d'être imprévisible, on a vu qu'il n'y a pas toujours un équilibre de Nash et ce même dans les jeux à somme nulle (par exemple le jeu de Pile ou Face ou encore le jeu de Pierre-Papier-Ciseaux). On a donc besoin d'introduire cette possibilité.

10.1 Théorème de Nash, jeux à somme non nulle

Ici, on supposera, pour simplifier, que S est au plus dénombrable et que r est bornée (sinon il faut ajouter des hypothèses de mesurabilité).

Notation : Pour un ensemble dénombrable Z, soit $\Delta(Z)$ l'ensemble des mesures de probabilités sur Z.

Définition 26 *Soit un jeu sous forme stratégique* $\Gamma = \left(N, (S_i)_{i\in N}, (r_i)_{i\in N}\right)$. *L'extension mixte de* Γ *est le jeu sous forme stratégique* $\Gamma_m = \left(N, (\Sigma_i)_{i\in N}, (R_i)_{i\in N}\right)$ *où* $\forall i \in N$:

- $\Sigma_i = \Delta(S_i)$;
- $\forall j \in N, \forall p_j \in \Delta(S_j)$:

$$\begin{aligned} R_i(p_1, ..., p_n) &= \sum_{\widetilde{s}_1,...,\widetilde{s}_n} r_i(\widetilde{s}_1, ..., \widetilde{s}_n)p_1(\widetilde{s}_1)...p_n(\widetilde{s}_n) \\ &= E_{p_1\otimes p_2...\otimes p_n}\left[r_i(\widetilde{s}_1, ..., \widetilde{s}_n)\right] \end{aligned}$$

$R_i(p_1, ..., p_n)$ *est l'espérance d'utilité du joueur* i *si chaque joueur* j *compte choisir une stratégie dans* S_j

en utilisant la loterie p_j. $p_i \in \Delta(S_i)$ est appelée une stratégie mixte du joueur i, et $\Delta(S_i)$ est l'ensemble des stratégies mixtes du joueur i. $s_i \in S_i$ est appelée une stratégie pure du joueur i et elle est identifiée à la masse de Dirac δ_{s_i}.
S_i est appelée l'ensemble des stratégies pures.

On veut analyser le jeu de tir au but entre un joueur de football et un gardien de but. On simplifie le jeu de la manière suivante: l'attaquant a le choix entre tirer à droite où à gauche et le gardien décide de sauter vers la droite où la gauche. S'ils choisissent le même côté alors le gardien gagne, sinon c'est l'attaquant qui gagne. Ceci décrit tout simplement le jeu de Pile ou Face:

$$\begin{array}{cc} & \begin{array}{cc} \text{P} & \text{F} \end{array} \\ \begin{array}{c} \text{P} \\ \\ \text{F} \end{array} & \begin{pmatrix} 1 & -1 \\ & \\ -1 & 1 \end{pmatrix} \end{array}$$

Figure 10.1:

Son extension mixte est donc:

$$\Gamma_m = (\{I, II\}, X, Y, R)$$

avec: : $X = Y = [0,1]$ et:

$$\begin{aligned} R(x,y) &= xy + (1-x)(1-y) - x(1-y) - y(1-x) \\ &= 1 - 2x - 2y + 4xy \end{aligned}$$

$x \in X$ doit être interprété comme étant la stratégie mixte pour laquelle $proba(H) = x, proba(T) = 1 - x$. La stratégie mixte du joueur colonne, y, doit être interprétée d'une manière similaire.

Pour résoudre ce jeu à somme nulle, on calcule les maxmin et minmax:

$$\begin{aligned} \min_y R(x,y) &= \min_y [(4x-2)y + (1-2x)] \\ &= \begin{cases} 2x - 1 & x \leq \frac{1}{2} \\ 1 - 2x & x \geq \frac{1}{2} \end{cases} \end{aligned}$$

donc: $\underline{v} = \max_x \min_y R(x,y) = 0$, et l'unique stratégie maxmin du joueur I est $x^* = \frac{1}{2}$. De même:

$$\begin{aligned} \max_x R(x,y) &= \max_x [(4y-2)x + (1-2y)] \\ &= \begin{cases} 2y - 1 & y \leq \frac{1}{2} \\ 1 - 2y & y \geq \frac{1}{2} \end{cases} \end{aligned}$$

Donc: $\overline{v} = \min_y \max_x R(x, y) = 0$, et $y^* = \frac{1}{2}$ est l'unique stratégie minmax du joueur II. On conclut que l'extension mixte du jeu de Pile ou Face admet une valeur $v = 0$ et que l'unique vecteur de stratégies optimales est $(x^*, y^*) = (\frac{1}{2}, \frac{1}{2})$. Donc $(\frac{1}{2}, \frac{1}{2})$ est l'unique équilibre de Nash du jeu.

Donc, la stratégie optimale du footballeur pour un tir au but est de tirer à droite avec probabilité 1/2 et à gauche avec une probabilité 1/2.

En fait, ceci est un cas particulier du théorème suivant qui est un des plus fondamentaux de la théorie des jeux :

Théorème 13 ***(Le théorème de Nash, 1951)***
Soit Γ un jeu fini sous forme stratégique. Alors, l'extension mixte Γ_m de Γ admet un équilibre de Nash (autrement: dit tout jeu fini admet un équilibre en stratégies mixtes).

C'est une conséquence du théorème de point fixe suivant de Kakutani, qu'on admettra.

Théorème 14 *(Kakutani)*
Soit X un compact convexe de R^n et soit f une correspondance de X dans X (i.e. pour tout x dans X, $f(x)$ est un sous ensemble de X). ***Si***
(i) pour tout x dans X, $f(x)$ est convexe et non vide;
(i)) le graphe de f est fermé (i.e. si $x_n \to x$, $y_n \to y$ et $y_n \in f(x_n)$ alors $y \in f(x)$);
Alors *il existe $x^* \in X$ tel que $x^* \in f(x^*)$*

Prouvons maintenant le théorème de Nash.

Preuve. Soit $X = \prod_i \Delta(S_i)$. X est alors convexe et compact. Soit $f = (f_i)_{i \in N}$ la correspondance suivante sur X:

$$f_i(x_1, ..., x_n) = \arg \max_{p \in \Delta(S_i)} R_i(x_1, ..., x_{i-1}, p, x_{i+1}, ..., x_n)$$

Puisque $\Delta(S_i)$ est convexe et compact et puisque $p \to R_i(x_1, ..., x_{i-1}, p, x_{i+1}, ..., x_n)$ est une fonction linéaire (donc en particulier concave) $\arg\max_{p \in \Delta(S_i)} R_i(x_1, ..., x_{i-1}, p, x_{i+1}, ..., x_n)$ est non vide et convexe. Le point (i) est vérifié. Le point (ii) est une conséquence immédiate de la continuité jointe de la fonction $x \to R(x)$ (car c'est une fonction multi-linéaire en dimension finie). ■

10.2 Théorème de von Neumann, jeux à somme nulle

Corollaire 5 ***(Le théorème de Von Neumann, 1923)***
Tout jeu fini à deux joueurs et à somme nulle admet une valeur en stratégies mixtes.

10.3 L'indifférence à l'équilibre

En pratique, la propriété suivante est très utile pour le calcul d'un équilibre de Nash pour un jeu fini :

Proposition 12 ***(Le principe d'indifférence)***
$p = (p_i)_{i \in N} \in \Pi_{i \in N} \Delta(S_i)$ est un équilibre en stratégies mixtes du jeu si et seulement si:

$$\forall s_i \in S_i : p_i(s_i) > 0 \Rightarrow R_i(s_i, p_{-i}) = \max_{t_i \in S_i} R_i(t_i, p_{-i})$$

Preuve. Par définition

$$R_i(p_i, p_{-i}) = \sum_{s_i \in S_i} p(s_i) R_i(s_i, p_{-i}) = \max_{t_i \in S_i} R_i(t_i, p_{-i}),$$

d'où:

$$\sum_{s_i \in S_i} p(s_i) \left[R_i(s_i, p_{-i}) - \max_{t_i \in S_i} R_i(t_i, p_{-i}) \right] = 0$$

Puisque $R_i(s_i, p_{-i}) - \max_{t_i \in S_i} R_i(t_i, p_{-i}) \leq 0$, on déduit que si $p_i(s_i) > 0$ alors $R_i(s_i, p_{-i}) = \max_{t_i \in S_i} R_i(t_i, p_{-i})$.

Pour la réciproque il suffit de remonter la preuve précédente. ■

Ainsi, dans un équilibre de Nash p, seules les stratégies pures qui sont une meilleure réponse aux stratégies des autres, sont jouées avec une probabilité strictement positive et toutes ces stratégies donnent la même utilité espérée. Ceci caractérise donc complètement l'ensemble des équilibres de Nash (mixte) pour un jeu fini. En pratique, le principe d'indifférence est très utile pour le calcul des équilibres de Nash, en voici un exemple.

Exemple:

On considère le jeu à somme nulle suivant:

$$\begin{array}{cc} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \\ \text{B} \end{array} & \begin{pmatrix} 3 & 1 \\ & \\ 0 & 4 \end{pmatrix} \end{array}$$

Figure 10.2:

Dès lors qu'on a vérifié qu'il n'y a aucun équilibre dans lequel au moins un joueur joue en stratégie pure, on sait (par le théorème de Nash ou de Von Neumann) que les stratégies optimales (x^*, y^*) sont des stratégies complètement mixtes: $0 < x^* < 1$ et $0 < y^* < 1$. Donc, chaque joueur joue, à l'équilibre, chacune de ses stratégies (pures) avec une probabilité strictement positive. Donc pour chaque joueur, toutes les stratégies pures ont le même paiement espéré. Ainsi, pour le joueur ligne, étant donné que le paiement espéré s'il joue T est $3y^* + (1 - y^*)$ et que son paiement espéré s'il joue B est $4(1 - y^*)$, y^* doit satisfaire:

$$3y^* + (1 - y^*) = 4(1 - y^*)$$

d'où $y^* = \frac{1}{2}$.

De même, pour satisfaire le principe d'indifférence pour le joueur 2, x^* doit satisfaire:

$$3x^* = x^* + 4(1 - x^*)$$

et donc: $x^* = \frac{2}{3}$. Maintenant, il est facile d'obtenir la valeur, qui est tout simplement le paiement espéré (du joueur 1) quand les stratégies optimales sont jouées. Dans ce cas on a $v = 2$.

Chapter 11

Stratégies de comportement

11.1 Définition

Les stratégies pures des joueurs dans un jeu sous forme extensive, ont été définies plus haut. Il s'agit d'un plan d'actions qui dicte à chaque joueur l'action (déterministe) qu'il compte entreprendre à chacun de ces ensembles d'information.

Pour chaque joueur i, soit S_i l'ensemble de toutes ses stratégies pures. On a vu comment transformer un jeu sous forme extensive en forme stratégique. On a vu aussi comment définir les stratégies mixtes pour un jeu sous forme stratégique. Ceci nous permet de définir l'ensemble des stratégies mixtes pour le joueur i comme étant: $\Sigma_i = \Delta(S_i)$.

L'interprétation d'une stratégie mixte pour un joueur dans un jeu sous forme extensive consiste à tirer au hasard au début du jeu un plan d'actions pures (une stratégie pure) parmi tous les plans d'actions possibles, puis d'exécuter ce plan d'actions pures et ne plus randomiser (c'est-à-dire jouer d'une manière aléatoire).

Si dans un jeu sous forme stratégique qui contient seulement une étape, il s'agit de l'unique possibilité de randomiser pour un joueur, ce n'est pas l'unique manière de randomiser pour les jeux sous forme extensive (et ce n'est pas non plus la plus naturelle). En effet, on pourrait imaginer une autre façon de randomiser, qui respecte plus l'aspect dynamique (séquentiel) du jeu (et l'existence de plusieurs points de décision). Un joueur peut avoir le comportement suivant: chaque fois qu'il se trouve à l'ensemble d'information U (un point de décision) il randomise. On définit ainsi une stratégie de comportement: elle consiste en un plan d'actions aléatoires qui, à chaque ensemble d'information, associe la probabilité suivant laquelle un joueur va faire son choix. Ainsi, les randomisations sont faites au fur et à mesure et ce en chaque point de décision. Dans les stratégies mixtes, la randomisation est faite une fois pour toutes juste avant le début du jeu.

Définition 27 *Une stratégie de comportement pour un joueur dans un jeu sous forme extensive est une fonction qui associe à chaque ensemble d'information une probabilité sur ses actions à cet ensemble d'information.*
Plus précisément: si $\left\{U_1^i, U_2^i, \ldots\right\}$, *est l'ensemble d'information du joueur i et si* $A\left(U_j^i\right)$ *est l'ensemble d'actions possibles à* U_j^i *et si* $A^i = \cup_j A\left(U_j^i\right)$, *alors une stratégie de comportement* q_i *du joueur i et une fonction:* $q_i : \left\{U_1^i, U_2^i, \ldots\right\} \rightarrow \Delta\left[A^i\right]$ *telle que* $\forall j$:

$$q_i\left(U_j^i\right) \in \Delta\left[A\left(U_j^i\right)\right].$$

11.2 Comparaison stratégies mixtes - stratégies de comportement

Considérons l'exemple suivant:

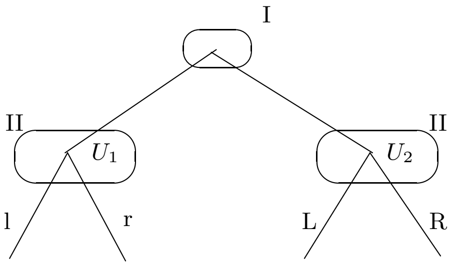

Figure 11.1:

Dans ce jeu, le joueur II a deux ensembles d'information U_1 et U_2 et deux actions possibles pour chaque ensemble d'information. Il a donc quatre stratégies pures: $s_1 = (l; L)$, $s_2 = (l, R)$, $s_3 = (r, L)$ et $s_4 = (r, R)$. Une stratégie mixte du joueur II est donc un vecteur $p = (p_1, p_2, p_3, p_4)$ tel que p_i est la probabilité de jouer s_i. L'ensemble des stratégies mixtes est donc:

$$\Sigma = \left\{ p = (p_1, p_2, p_3, p_4) \middle| p_i \geq 0, \sum_{i=1}^{4} p_i = 1 \right\}$$

Dans une stratégie de comportement, le joueur II n'est amené à randomiser seulement quand il est amené à jouer. Donc, quand il se trouve à U_1 il choisit de jouer l'action l avec probabilité α et l'action r avec probabilité $(1 - \alpha)$ et quand il se trouve à U_2 il choisit de jouer l'action L avec probabilité β et l'action R avec probabilité $(1 - \beta)$. D'où l'ensemble des stratégies de comportement pour le joueur II est:

$$B = \left\{ q = (\alpha, \beta) \middle| \begin{array}{c} 0 \leq \alpha \leq 1 \\ 0 \leq \beta \leq 1 \end{array} \right\}$$

On voit donc que Σ est un objet mathématique de dimension trois alors que B est de dimension deux.

Quelles sont les relations entre les deux ensembles de stratégies ?. On remarque d'abord qu'une stratégie de comportement définit d'une unique façon une distribution de probabilité sur les stratégies pures. Dans notre exemple :

$$p(l, L) = \alpha\beta; \quad p(l, R) = \alpha(1 - \beta); \quad p(r, L) = (1 - \alpha)\beta; \quad p(r, R) = (1 - \alpha)(1 - \beta)$$

En d'autres termes, jouer la stratégie de comportement $q = (\alpha, \beta)$ est équivalent à joueur la stratégie mixte $\sigma(q) = (\alpha\beta, \alpha(1 - \beta), (1 - \alpha)\beta, (1 - \alpha)(1 - \beta))$. L'interprétation d'une stratégie mixte représentant une stratégie de comportement est tout naturellement telle que les randomisations dans les différents ensembles d'information sont indépendantes. La différence entre q et $\sigma(q)$ est seulement dans la façon dont les randomisations sont faites: en utilisant q les randomisations sont faites indépendamment dans deux endroits différents alors que pour $\sigma(q)$ une seule randomisation est faite avant le début du jeu. q et $\sigma(q)$ *sont équivalentes* dans le sens où: étant donnée n'importe quelles stratégies des autres joueurs, q et

$\sigma(q)$ induisent la même distribution de probabilité sur l'ensemble des parties (et donc sur l'ensemble des paiements).

L'application $\sigma : B \to \Sigma$ n'est pas surjective. En effet, la stratégie mixte $\left(\frac{1}{2}, 0, 0, \frac{1}{2}\right)$ n'est pas dans l'image de σ car il n'y a aucun $q = (\alpha, \beta)$ tel que $\sigma(q) = \left(\frac{1}{2}, 0, 0, \frac{1}{2}\right)$. Donc formellement, l'ensemble des stratégies mixtes est plus large que l'ensemble des stratégies de comportement.

Les stratégies de comportement paraissent souvent plus attractives que les stratégies mixtes non seulement parce qu'elles sont beaucoup moins nombreuses mais aussi beaucoup plus adaptées à l'aspect multi-étapes: si on a 10 points de décisions et que dans chaque point il a deux actions possibles, il paraît plus raisonnable de penser qu'on décide ce qu'on doit faire au moment où on se trouve à un point de décision plutôt que de considérer au début du jeu les 2^{10} plans d'actions possibles (les stratégies pures) et puis de choisir entre elles (peut être au hasard) un plan d'action que l'on va exécuter jusqu'à la fin du jeu. Une telle stratégie est non seulement non intuitive mais aussi mathématiquement trop complexe.

La question importante est la suivante : que peut perdre un joueur en se restreignant à ses stratégies de comportement ?

Pour répondre à cette question, revenons à l'exemple précédent. On a vu que le joueur II, s'il se restreint aux stratégies de comportement, ne peut pas générer la stratégie mixte $p = \left(\frac{1}{2}, 0, 0, \frac{1}{2}\right)$. Cependant, supposons que le joueur I choisisse de joueur a avec probabilité θ et b avec probabilité $(1-\theta)$. Ainsi, si le joueur II utilise la stratégie mixte $p = \left(\frac{1}{2}, 0, 0, \frac{1}{2}\right)$ alors la distribution de probabilité sur l'ensemble des résultats (o_1, o_2, o_3, o_4) est $(\frac{1}{2}\theta, \frac{1}{2}\theta, \frac{1}{2}(1-\theta), \frac{1}{2}(1-\theta))$. Mais nous vérifions facilement qu'on obtiendrait exactement la même distribution si le joueur II utilisait la stratégie de comportement $q = \left(\frac{1}{2}, \frac{1}{2}\right)$. En d'autres termes, p et q (qui ne définissent pas la même distribution de probabilité sur l'ensemble des stratégies pures) satisfont quand même la relation suivante :

- Etant donné n'importe quelle stratégie τ du joueur I, (τ, p) et (τ, q) induisent la même distribution de probabilité sur l'ensemble des parties (et donc sur l'ensemble de tout les résultats possibles).

Cette relation entre p et q est appelée: *équivalence en terme de paiement* ou tout simplement *équivalence*. Ainsi, dans notre dernier exemple, il est facile de voir que toute stratégie mixte $p = (p_1, p_2, p_3, p_4)$ est équivalente à la stratégie de comportement $q = (\alpha, \beta)$ où α est la distribution marginale de p sur $\{l, r\}$ et β est la distribution marginale de p sur $\{L, R\}$:

$$\begin{aligned} \alpha &= p_1 + p_2 \\ \beta &= p_1 + p_3 \end{aligned}$$

Cet exemple est cependant un peu particulier car, pour chaque partie, seulement un ensemble d'information est atteint.

Considérons maintenant l'exemple suivant:

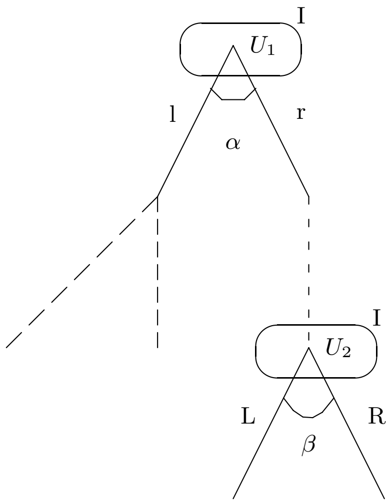

Figure 11.2:

Dans cette exemple, Le joueur I a deux ensembles d'information qui peuvent tout les deux être atteints dans une même partie. De nouveau, il existe des stratégies mixtes qui ne peuvent pas être générées par les stratégies de comportement, comme par exemple $\left(\frac{1}{3}, 0, 0, \frac{2}{3}\right)$. Cependant, celle-ci est équivalente à la stratégie de comportement $\left(\frac{1}{3}, 0\right)$.

Ainsi, dans un jeu où pour un certain joueur toute stratégie mixte est équivalente (en terme de paiement) à une certaine stratégie de comportement, il se peut que le jeu reste essentiellement inchangé si on restreint ce joueur à utiliser seulement ses stratégies de comportement. Cela a été le cas dans nos deux derniers exemples. Cependant ce n'est pas toujours le cas.

Considérons le jeu à somme nulle présenté dans la figure qui suit:

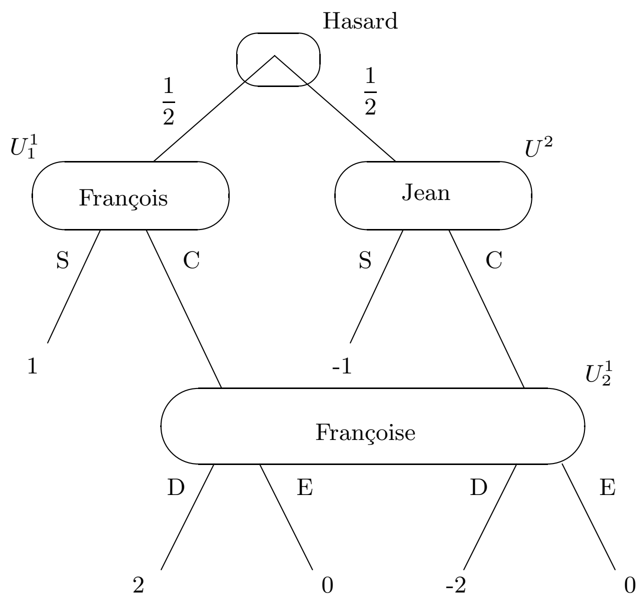

Figure 11.3:

Ce jeu à deux joueurs peut être illustré par l'histoire suivante: il y a deux joueurs dans ce jeu. Le joueur I est en fait un couple (François et Françoise) et le joueur II est Jean. Il y a deux cartes: Grande et Petite qui sont distribuées avec une égale probabilité à Jean et à François. L'arbitre demande à celui qui a obtenu la plus grande carte de décider de continuer le jeu où de stopper. Dans le cas où celui qui a la plus grande carte décide de stopper le jeu, celui-ci reçoit une unité de l'autre joueur. Sinon, le jeu continue. Dans ce cas Françoise (qui est dans une autre chambre et qui ne sait pas qui a continué le jeu: François ou Jean) doit choisir entre E (ce qui implique un paiement nul pour tout le monde) où D (ce qui double la mise du détenteur de la plus grande carte: il gagne 2 unités et l'autre joueur perd 2 unités).

Ici, le joueur II a deux stratégies pures $\{s, c\}$ et le joueur I a 4 stratégies pures $\{(S, D), (S, E), (C, D), (C, E)\}$. Ainsi, le jeu sous forme stratégique est le suivant:

Joueur II

Joueur I		S	C
	S,D	0	$-\frac{1}{2}$
	S,E	0	$\frac{1}{2}$
	C,D	$\frac{1}{2}$	0
	C,E	$-\frac{1}{2}$	0

Figure 11.4:

La valeur de ce jeu est $v = \frac{1}{4}$ et l'unique stratégie optimale du joueur I est $\left(0, \frac{1}{2}, \frac{1}{2}, 0\right)$ et pour le joueur II $\left(\frac{1}{2}, \frac{1}{2}\right)$. Ceci implique en particulier que le joueur I peut garantir un paiement de $\frac{1}{4}$ en utilisant la stratégie mixte (S, E) avec une probabilité $\frac{1}{2}$ et (C, D) avec une probabilité $\frac{1}{2}$. Pour implémenter cette stratégie, il faut qu'avant le début du jeu, François et Françoise se concertent en tirant au hasard entre les deux plans d'action (S, E) et (C, D) puis décident, ensemble, lequel des deux ils vont exécuter. En effet, ils ne pourront plus discuter plus tard pour se coordonner.

Le joueur I peut-il se garantir le même montant en utilisant les stratégies de comportement ?

Supposons que le joueur I utilise la stratégie de comportement (α, β) : François stoppe avec probabilité α et Françoise double avec probabilité β et par définition des stratégies de comportement, les deux randomisations sont faites indépendamment.

Dans ce cas, le paiement espéré du joueur I est:

- Si Jean joue s:

$$\frac{1}{2}\left[\alpha + (1-\alpha)\left\{2\beta + 0(1-\beta)\right\}\right] + \frac{1}{2}(-1) = (1-\alpha)\left(\beta - \frac{1}{2}\right)$$

- Si Jean joue c:

$$\frac{1}{2}\left[\alpha + (1-\alpha)\left\{2\beta + 0(1-\beta)\right\}\right] + \frac{1}{2}\left[\beta(-2) + 0(1-\beta)\right] = \alpha\left(\frac{1}{2} - \beta\right)$$

Ainsi, le joueur I garantit, en se restreignant à ces stratégies de comportement, seulement:

$$\underline{v}_{comp} = \max_{\alpha,\beta} \min\left\{(1-\alpha)\left(\beta - \frac{1}{2}\right), \alpha\left(\frac{1}{2} - \beta\right)\right\} = 0$$

En effet ou bien $\beta - \frac{1}{2} \leq 0$ ou bien $\beta - \frac{1}{2} \geq 0$.
Ainsi, le joueur I, dans ce jeu, peut faire strictement plus en utilisant les stratégies mixtes que s'il utilisait seulement les stratégies de comportement. On déduit donc que la stratégie mixte $\left(0, \frac{1}{2}, \frac{1}{2}, 0\right)$ n'est équivalente (en terme de paiement) à aucune stratégie de comportement. Comme on va le voir, ceci est lié au fait qu'ici le joueur I ne se sait plus s'il a joué ou non dans le passé. On dit que ce jeu n'est pas à mémoire parfaite.

11.3 Mémoire parfaite

La notion de mémoire parfaite veut tout simplement dire qu'un joueur donné ne perd pas d'information durant le jeu. Celle-ci peut être décomposée en deux propriétés:

- **P1) le joueur ne perd pas d'information le concernant durant le jeu:** il se rappelle de tout ce qu'il a fait ou pas fait dans le passé. Dans l'exemple précédent, ceci n'est pas vrai car dans U_2^1, le joueur I ne se rappelle s'il a déjà joué ou pas dans le passé.
- **P2) le joueur ne perd pas d'information durant le jeu sur ce qu'il savait concernant les autres joueurs:** à tout moment il se rappel de tout ce qu'il a su concernant les autres. Ceci n'est pas vérifié dans l'exemple avec François et Françoise le joueur I, qui connaît le résultat de la loterie au noeud U_1^1, ne le sait plus quand il est successeur de U_1^1 qui est dans U_2^1.

La question est donc de savoir comment exprimer mathématiquement sur une forme extensive cette idée suivant laquelle un joueur ne perd pas d'information durant le jeu. D'abord, remarquons qu'on devrait exclure les situations telles que la suivante:

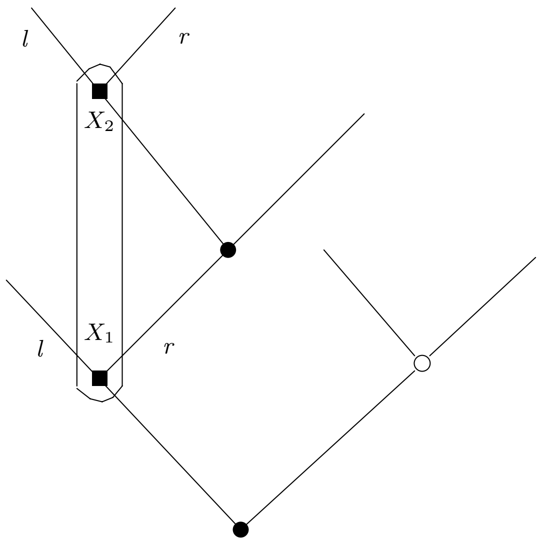

Figure 11.5:

En effet, dans cet exemple les noeuds x_1 et x_2 appartiennent à un même ensemble d'information alors qu'ils correspondent à des noeuds chronologiquement différents. Le joueur concerné ne se rappelle donc plus quand il est au noeud x_2 qu'il a déjà joué au noeud x_1. Ainsi, pour garantir la mémoire parfaite, on a besoin de la linéarité:

Définition 28 *Un jeu sous forme extensive est dit linéaire pour le joueur i s'il n'existe aucune partie qui intersecte plus d'une fois un ensemble d'information du joueur i (pour tous noeuds distincts u et v dans un même ensemble d'information du joueur i, il n'existe aucun chemin dirigé de u vers v).*

En effet, s'il existait un chemin dirigé qui passe par u puis par v, où u et v sont dans un même ensemble d'information du joueur i, cela signifie que le joueur i va jouer au noeud u une action qui va mener plus tard au noeud v. Ne distinguant pas entre u et v, cela veut dire qu'il a oublié ce qu'il a fait dans le passé.

Ceci peut être vu autrement: quand il est au noeud u, le joueur i ne distinguant pas entre u et v, il ne sait pas (1) s'il a déjà joué dans le passé (dans ce cas il devrait être en v) ou (2) s'il n'a pas déjà joué (et dans ce cas il devrait être en u).

La linéarité est donc une condition nécessaire pour garantir qu'un joueur se rappelle ses propres actions, cependant il est facile de voir qu'elle n'est pas suffisante. En effet, le jeu de la section précédente est un jeu linéaire pour le joueur I (François-Françoise) cependant dans U_2^1 le joueur 1 ne se rappelle plus s'il a joué ou pas. Ainsi, pour satisfaire les conditions $P1$ et $P2$ on a besoin d'une condition beaucoup plus forte que la simple linéarité. Par exemple, il faudrait exclure les situations telles que la suivante:

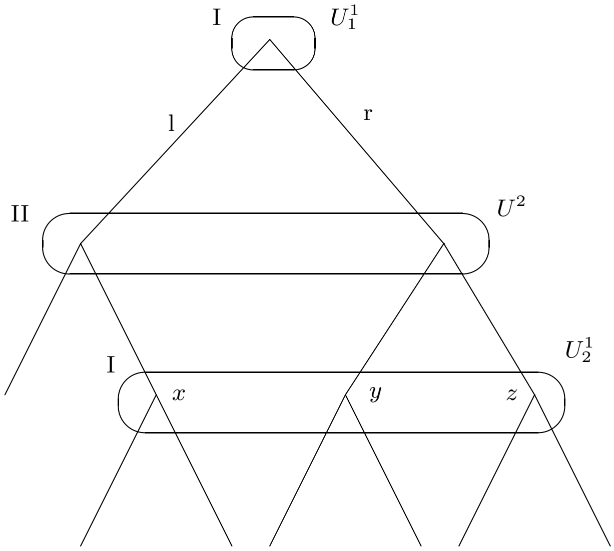

Figure 11.6:

En effet, dans cet exemple, si le joueur I se rappelle ses propres mouvements, x ne peut pas être dans le même ensemble d'information que y et z: si le joueur I a joué l en U_1^1 alors, quand il est dans l'ensemble d'information U_2^1, il devrait savoir qu'il est en x. Donc dans ce jeu, le joueur I n'a pas de mémoire parfaite.

Définition 29 *(Kuhn)*
Le jeu est dit à mémoire parfaite pour le joueur i si est seulement s'il vérifie les propriétés $PR1$ et $PR2$ ci-dessous :

- *(PR1) il est linéaire pour i;*
- *(PR2) pour tout couple (x, y) dans un même ensemble d'information U_i^j, si x' est un prédécesseur de x appartenant à l'ensemble d'information $U_i^{j'}$ avec $U_i^{j'} \neq U_i^j$ alors:*
    - *il existe y' un prédécesseur de y tel que $y' \in U_i^{j'}$.*
    - *l'action qui mène de x' à x est la même que celle qui mène de y' à y.*

11.4 Théorème de Kuhn

Théorème 15 *(Kuhn, 1953)*
Dans un jeu sous forme extensive, ***si*** *le joueur i a:*

- *un nombre fini d'actions dans chacun de ses ensembles d'informations;*
- *une mémoire parfaite;*

alors: *toute stratégie mixte du joueur i est équivalente (en terme de paiement) à une stratégie de comportement de i.*

Corollaire 6 *Tout jeu fini, sous forme extensive et à mémoire parfaite pour tous les joueurs, est essentiellement inchangé en restreignant les joueurs à utiliser seulement leurs stratégies de comportement. En particulier, le jeu admet un équilibre de Nash en stratégie de comportement.*

Ce résultat de Kuhn est d'un très grand intérêt en pratique car, dans beaucoup d'applications, les jeux sont à mémoire parfaite. Ceci ne veut pas dire que les jeux qui ne sont pas à mémoire parfaite ne sont pas concevables car on peut imaginer qu'un joueur dans le modèle représente en fait plusieurs décideurs dans une même entreprise (où encore une équipe ou un couple) qui ne peuvent pas toujours communiquer entre eux (par exemple les décisions doivent se prendre simultanément). L'analyse du jeu avec François, Françoise et Jean (figure 11.3) montre l'importance de considérer parfois des jeux où certains joueurs n'ont pas de mémoire parfaite....

Chapter 12

Raffinement

Ce qui est considéré comme un défaut par certains (une qualité pour d'autres) c'est la multiplicité des équilibres. En effet, le but d'un scientifique, au delà de la simple compréhension des phénomènes, c'est la prédiction. D'où la nécessité de classifier, d'éliminer et de sélectionner.

D'abord, est-ce vraiment un défaut que d'avoir une multiplicité d'équilibres ? Remarquons qu'un jeu n'est autre que la description de tout ce qui est possible lors d'une interaction. Un équilibre décrit une issue raisonnable parmi toutes les issues possibles. La multiplicité des équilibres nous indique qu'il n'y a pas une vérité, un seul chemin mais plusieurs possibilités. On reste toujours confronté au problème de choix, même si l'ensemble des choix a été considérablement réduit (génériquement, tout jeu matriciel admet un nombre fini, impair, d'équilibres).

Cependant, comme en mécanique, il y a des équilibres stables et des équilibres instables. Il y a des équilibres possédant certaines propriétés raisonnables, d'autres non. D'où la théorie du raffinement.

12.1 Induction en amont (anticiper le comportement future)

Considérons le jeu matriciel suivant:

$$\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} A \\ \\ B \end{array} & \begin{bmatrix} 0,0 & 2,1 \\ & \\ 1,2 & 1,2 \end{bmatrix} \end{array}$$

Ce jeu possède deux équilibres de Nash: (A, R) et (B, L). Jouer B est justifié pour le joueur I (le joueur ligne) car cela lui procure une sécurité (il est sûr d'obtenir 1). (B, L) est préféré par le joueur II car son gain est maximal. (B, L) peut donc sembler une issue raisonnable. En fait, (A, R) peut être justifié de la même manière, R sécurise une utilité de 1 pour le joueur II et (A, R) procure une paiement maximal pour le joueur I. Que faire ?

Imaginons que ce jeu soit la forme normale du jeu sous forme extensive suivant. Le joueur I joue en premier et a le choix entre A et B. S'il choisit B, le jeu s'arrête et les joueurs gagnent $(1, 2)$. Si par contre il choisit A, alors le joueur II doit jouer. Celui-ci a le choix entre R et L. Si II choisit L, le jeu s'arrête et le vecteur de paiement est $(0, 0)$. Si par contre II choisit R, le jeu s'arrête et le vecteur de paiement est $(2, 1)$. Supposons que les joueurs soient rationnels et que la rationalité soit une connaissance commune entre les joueurs. Si le joueur I sait que le joueur II est rationnel, alors il prédit que, s'il est amené à jouer, celui-ci aura le choix entre jouer L et gagner 0 ou R et gagner 1. Celui-ci devrait choisir, rationnellement, R. Dans ces conditions, le choix rationnel du joueur I est A. (A, R) est donc l'unique issue rationnelle du jeu. En effet, l'équilibre (B, L) est basé sur une menace non crédible de la part du joueur II. Celui-ci veut forcer le joueur I à jouer B, et si jamais celui-ci dévie et joue A, alors il le menace de jouer L. Le problème c'est que II s'auto-punit en faisant cela. Rappelez vous que, puisqu'on a affaire à des utililités, toute satisfaction éventuelle à punir son adversaire est déjà intégrée aux paiements.

Quel est le raisonnement correspondant dans le jeu sous forme normale ? Remarquons que l'action R domine faiblement l'action L pour le joueur II. (A, R) est donc l'issue résultante d'une élimination répétée des stratégies faiblement dominées. Ceci constitue un résultat général.

Définition 30 *Dans un jeu sous forme extensive, un équilibre est dit en sous jeux parfait s'il induit un équilibre dans tous les sous jeux.*

L'algorithme de Kuhn. Pour déterminer l'équilibre sous jeux parfait dans un jeu sous forme extensive fini à information parfaite, on commence par déterminer l'action rationnelle pour chaque joueur i concerné et ce pour tout noeud a ayant comme successeur que des noeuds terminaux. Appelons un tel noeud un noeud avant terminal. En effet, le joueur jouant à un noeud avant terminal fait face à un problème de décision individuel sans aucune interaction stratégique car tous ses choix arrêtent le jeu. Si ce joueur est rationnel, il devrait jouer l'action qui maximise son gain. Si cette action est unique et si la rationalité entre les joueurs est connaissance commune alors les joueurs jouant avant a considéreraient que l'issue du jeu est complètement déterminée si jamais le jeu atteint le noeud a. Si l'action rationnelle n'est pas unique au noeud a alors tout choix rationnel induirait un équilibre sous jeux parfait différent (et dans ce cas il y a une multiplicité d'équilibre sous jeu parfait). En cas de multiplicité, l'algorithme choisit n'importe quelle action rationnelle parmi les possibles. Si le noeud de décision a correspond au joueur hasard, les joueurs avant a vont considérer que leur paiement si le jeu atteint ce noeud n'est autre que l'espérance d'utilité. On pourrait donc refaire une autre étape de l'algorithme en assimilant le noeud a à un noeud terminal auquel en associe le vecteur de paiement rationnel anticipé par les autres joueurs. Ceci nous permet donc de remonter l'arbre partant des noeuds terminaux et finissant, après un certain nombre d'étapes, par atteindre la racine de l'arbre en parcourant tous les noeuds. Cet algorithme montre l'existence d'un équilibre pur sous jeux parfait pour tout jeu sous forme extensive fini et à information parfaite. Il procure aussi une méthode pour déterminer tous les équilibres sous jeux parfaits et montre l'unicité dans le cas où les paiements de n'importe quel joueur correspondant à deux parties différentes sont différents.

Appliquons cette algorithme à l'exemple suivant.

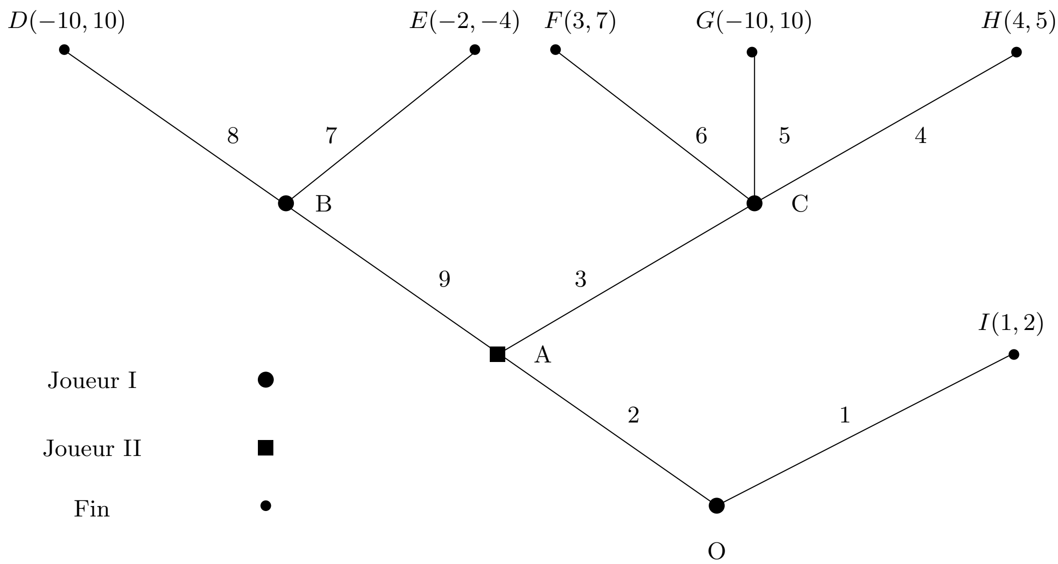

Figure 12.1:

Les noeuds avant terminaux sont B et C. Au noeud B (resp. C) l'action rationnelle du joueur I est de jouer l'action 7 (resp. l'action 4). Connaissant cela, au noeud A, le joueur II doit choisir l'action 3. Connaissant cela, il est rationnel pour le joueur I au noeud O de choisir l'action 2. L'issue de la confrontation des deux stratégies est le chemin O,A,C,H avec comme vecteur de paiement final (4,5).

Théorème 16 *Tout jeu sous forme extensive fini et* ***à information parfaite*** *admet un équilibre sous jeux parfait et en stratégies pures.*

Preuve. Une conséquence immédiate de l'algorithme de Kuhn. ■

L'élimination des stratégies faiblement dominées dans un jeu sous forme extensive correspond en général (mais pas toujours car il y a toujours des cas pathologiques) à la perfection en sous jeux. Dans le cas d'un jeu sous forme normale, l'élimination d'une stratégie faiblement dominée peut être justifiée à l'aide de la main tremblante de Selten.

Dans le cas d'un jeu à information imparfaite, si le jeu est à mémoire parfaite, tout sous jeu est aussi à mémoire parfaite, et par le théorème de Kuhn, tout sous jeu admet un équilibre en stratégies de comportement. Une raisonnement en amont comme précédemment permet là encore de construire un équilibre sous jeux parfait. D'où le résultat qui suit.

Théorème 17 *Tout jeu sous forme extensive fini et à* ***mémoire parfaite*** *admet un équilibre sous jeux parfait, en stratégies de comportement (non nécessairement pures).*

Parfois, l'élimination des stratégies faiblement dominées est plus puissante que le concept d'équilibre sous jeux parfait. En effet, considérons une petite variante du jeu sous forme extensive décrit plus haut. Le joueur I joue en premier et a le choix entre A et B. Le joueur II joue juste après mais ne sais pas ce que le joueur 1 a choisi. Le joueur II a le choix entre R et L. Si le joueur I choisit B, le vecteur de gain est $(1, 2)$ et ce quelque soit l'action du joueur II. Si par contre il choisit A, alors si II choisit L, le vecteur de paiement est $(0, 0)$, sinon c'est $(2, 1)$. Quels sont les sous jeux de ce jeu ? Seulement le jeu lui même. Dans ce cas, tous les équilibres (Ici (A, R) et (B, L)) sont des équilibres sous jeux parfaits. Quelle est la forme normale de ce jeu ? C'est le jeu matriciel décrit plus haut dans lequel (A, R) est le seul équilibre restant après élimination des stratégies faiblement dominées.

12.2 Induction en aval (interpréter le comportement passé)

Dans la définition du concept d'équilibre sous jeux parfait, les joueurs regardent toujours vers l'avant et n'essayent jamais d'interpréter les choix effectuées dans le passé. Chaque joueur essaye d'anticiper, pour chacune des ses actions aujourd'hui, les conséquences d'un comportement rationnel des autres joueurs et de lui même dans la suite du jeu. Puis il utilise ses anticipations pour faire un meilleur choix aujourd'hui. Son raisonnement est indépendamment des choix effectués dans le passé.

L'induction en aval, introduite par Kohlberg et Mertens, consiste à interpréter les actions passées afin de sélectionner parmi les différents équilibres aujourd'hui. Ceci peut s'avérer parfois très puissant.

Commençons d'abord par analyser l'exemple suivant contenant (au plus) trois étapes. A l'étape 1, le joueur I a le choix entre a1 (arrêter) ou c1 (continuer).

- L'étape 1: si le joueur I arrête, le vecteur de gain est (2,1) et s'il continue, le jeu atteint l'étape 2.
- A l'étape 2, le joueur II a le choix entre a2 (arrêter) ou c2 (continuer). Si le joueur II arrête, le vecteur de gain est (3,1) et s'il continue le jeu atteint alors l'étape 3.

- A l'étape 3, le joueur I a le choix entre D ou G. S'il choisit D, le vecteur de gain est (0,0) et s'il choisit G, le vecteur de gain est (1,2).

Quel est l'équilibre sous jeux parfait ? L'unique noeud avant terminal est l'étape 3. A cette étape, l'action rationnelle du joueur I est G. Anticipant cela, l'action rationnelle du joueur II à l'étape 2 est de continuer. Dans ces conditions, l'unique choix rationnel du joueur I à l'étape 1 est d'arrêter (a1). Le paiement de l'unique équilibre sous jeux parfait est donc (2,1). Le chemin d'équilibre sous jeux parfait est l'arrêt dès l'étape 1 du jeu. En particulier, si les joueurs sont rationnels, le jeu ne devrait jamais atteindre l'étape 2.

Le problème philosophique qui se pose maintenant est le suivant : le choix d'arrêter le jeu à l'étape 1 est fondé sur l'anticipation d'un comportement rationnel dans la suite du jeu alors même qu'atteindre l'étape 2 est un comportement irrationnel ! le joueur I devrait se poser donc la question suivante: que va penser le joueur II si le jeu atteint l'étape 2. II va peut être penser que le joueur I est irrationnel ? dans ce cas, qu'est ce qui empêche II de penser que I va agir encore irrationnellement à l'étape 3 en choisissant D au lieu de G avec comme conséquence le vecteur de paiement (0,0) au lieu de (1,2). Dans ces conditions, il est peut être plus raisonnable pour le joueur II d'arrêter le jeu à l'étape 2 et d'assurer un gain de 1 pour lui au lieu de prendre le risque de recevoir 0 si le joueur I d'une manière irrationnelle ?. Mais, dans ces conditions, il est maintenant dans l'intérêt du joueur I de continuer à l'étape 1 car cela risque de lui procure un paiement de 3 correspondant à un arrêt du jeu à l'étape 2 (ce qui rend finalement son action de continuer rationnelle étant donnée son anticipation). Mais dans ce cas le joueur II peut aussi penser que le joueur I est rationnel et que son but est justement l'obtention du paiement 3. Cela illustre un peu les bases du raisonnement de l'induction en aval : comment justifier rationnellement les actions passées d'un joueur ?

Dans notre exemple, l'unique justification rationnelle de c1 est le désir du joueur I d'obtenir un paiement plus élevé que 2, paiement qu'il aurait obtenu en arrêtant à l'étape 1. Dans le jeu qui commence à l'étape 2, le joueur II doit choisir entre deux équilibres (a2,D) et (c2,G). Seul l'équilibre (a2,D) procure un paiement pour le joueur I plus élevé que 2. Est-ce qu'on peut conclure que le joueur 2 arrête effectivement le jeu à l'étape 2 ? Ceci n'est pas évident car cet équilibre reste basé sur une menace non crédible de la part du joueur I : celle de jouer D ...

L'idéal est de pouvoir marier la rationalité, l'induction en amont et l'induction en aval. L'analyse précédente montre que l'induction en aval peut parfois être en contradiction avec l'hypothèse de rationnalité commune.

Cependant, cela peut être possible. En effet, considérons une petite variante de l'exemple précédent. Le jeu maintenant se déroule en seulement deux étapes. La première étape est la même que précédemment: le joueur I a le choix entre arrêter et continuer. S'il le joueur I arrête il obtient 2. S'il continue, le jeu atteint la deuxième étape. La deuxième étape est un vrai jeu simultané possédant exactement deux équilibres de Nash. Le premier équilibre procure au joueur I un paiement de 3 et le deuxième équilibre lui procure un paiement de 1. Ainsi, si le jeu atteint la deuxième étape, un raisonnement en aval montre que l'unique équilibre rationalisable en deuxième étape celui qui procure le paiement de 3 au joueur I. Un raisonnement en amont maintenant permet de rationaliser l'action c1. L'induction en aval permet ainsi au joueur de se coordonner sur un équilibre en deuxième étape d'une manière endogène.

12.3 Paradoxe : le jeu du mille-pattes

Le jeu du mille-pattes est un jeu en 100 étapes. Aux étapes impaires, $t = 1, 3, ..., 99$ le joueur I a le choix entre arrêter le jeu où continuer. S'il arrête à l'étape t, le vecteur de paiement est $(t, t-1)$ sinon le jeu continue à l'étape $t+1$. Aux étapes paires, $t = 2, 4, ..., 100$ le joueur II a le choix entre arrêter où continuer. Si II arrête le vecteur de paiement est $(t-2, t+1)$ sinon le jeu continue. Si aucun joueur ne prend la décision d'arrêter le jeu, le vecteur de paiement est $(100, 100)$.

Ainsi, si v_t dénote le vecteur de paiement si le jeu s'arrête à l'étape t alors : $v_1 = (1, 0)$, $v_2 = (0, 3)$, ...,

$v_{49} = (49, 48)$, $v_{50} = (48, 51)$, ..., $v_{98} = (96, 99)$, $v_{99} = (99, 98)$, $v_{100} = (98, 101)$. Les deux joueurs ont donc un intérêt mutuel à ce que le jeu dure le plus longtemps possible (au moins jusqu'à la 90ième étape, voir plus.)

Examinons l'équilibre sous jeux parfait. A l'étape 100, le joueur II a le choix entre arrêter et gagner 101 ou continuer et gagner 100. Rationnellement, il devrait arrêter. Connaissant cela, à l'étape 99, le joueur I devrait rationnellement arrêter le jeu. En continuant ce raisonnement, on voit que l'unique équilibre sous jeux parfait est d'arrêter le jeu à chaque étape. En effet, aux étapes t impaires, le joueur I a le choix entre arrêter et gagner t ou continuer et gagner $t + 1 - 2$ et aux étapes t paires, le joueur II a le choix entre arrêter et gagner $t + 1$ et continuer et gagner $t + 1 - 1$. Est-ce raisonnable comme prédiction ? les résultats expérimentaux montrent que les joueurs continuent le jeu jusqu'à atteindre un niveau de paiement suffisamment élevé avant d'arrêter... Est-ce que cela signifie que les joueurs sont irrationnels ?

Si la rationalité pour un joueur est d'essayer d'augmenter son paiement alors manifestement on ne peut pas conclure que ce comportement est irrationnel puisqu'il permet aux joueurs d'obtenir plus que le paiement de l'équilibre sous jeux parfait... Comment donc justifier ce comportement sans violer complètement l'hypothèse de rationalité ? Remarquons qu'une rationalité absolue implique un arrêt à la date 100, et donc nécessairement un arrêt à la date 99,.... Arrêter à la date 1 est donc l'unique issue de tout équilibre de Nash (sous jeux parfait ou pas). On est donc obligé de supposer un minimum d'irrationalité pour comprendre les résultats expérimentaux. Voyons ce que nous dit un raisonnement en aval avec un *peu d'irrationalité* de la part des joueurs. Si le jeu atteint l'étape 2, le joueur II va conclure qu'avec une grande probabilité le joueur I espère que le joueur II va continuer le jeu aussi car sinon il aurait arrêté avant. Aussi, si le joueur I ne se contente pas d'un paiement de 1, ce n'est pas un paiement de 3 suite à un arrêt du joueur I à l'étape 3 qui le satisfera, le joueur II est donc incité à continuer à l'étape 2. Tous ces bons sentiments renforcent donc la coopération entre les joueurs pour continuer le jeu tant que la date butoir (la date 100) est encore loin. A l'approche de la fin, chaque joueur commence à redouter un arrêt (sous l'hypothèse de rationalité) ce qui fini bien par arrêter le jeu effectivement un peu avant la date butoir. L'induction en aval mariée à un petit peu d'irrationalité permet donc de comprendre partiellement ces résultats expérimentaux.

En fait, dès lors qu'un jeu est fini, un problème de bord (due à la date butoir) peut rendre impossible la coopération rationnelle pour obtenir des paiements plus élevés (c'est ce qui arrive avec le dilemme du prisonnier répété n fois). Dans le cas où un jeu est joué un nombre infini de fois le problème de bord disparaît et la coopération redevient rationnelle (voir le chapitre qui suit).

Chapter 13

Jeux répétés, phénomène de coopération

Souvent, les agents économiques sont confrontés à des problèmes de prises de décisions similaires et ce de manière récurrente. Si le comportement d'un agent dépend naturellement de son expérience passée, il est aussi obligé de prendre en considération les conséquences de son comportement sur le comportement futur de ses concurrents. Les jeux répétés constituent un cadre théorique idéal pour analyser ce genre de situations. Cela permet de mettre en évidence des phénomènes nouveaux tels que la coopération, la punition, la coordination...

Un jeu répété consiste en la répétition un nombre fini ou infini de fois d'un jeu sous forme normale. A chaque étape du jeu, les agents ont une observation complète ou partielle des actions des autres joueurs dans le passé. Un joueur peut alors adopter un comportement très sophistiqué en faisant dépendre sa stratégie du passé. Dans un jeu de coordination par exemple, les joueurs peuvent utiliser leur passé commun pour mieux se coordonner (par exemple, ils peuvent décider de se coordonner sur la première action du premier joueur ou encore jouer au hasard jusqu'à la première coordination). Dans un jeu tel que le dilemme du prisonnier, ils peuvent se menacer mutuellement d'avouer toujours dans le futur si jamais l'autre joueur a avoué dans le passé. Ce genre de modèles permet aussi de comprendre les raisons pour lesquelles une entreprise a intérêt à faire une guerre de prix où à mettre des barrières pour se prévenir d'un entrant potentiel, même si cela engendre des coûts à court terme.

Dans cette partie, on va se limiter aux jeux escomptés, répétés infiniment de fois et à observation parfaite des actions passées. La question qu'on va se poser est la suivante: est-ce que la répétition du jeu permet d'atteindre des paiements Pareto-Optimaux à l'équilibre ? En d'autres termes, est-ce que le paradoxe soulevé par le dilemme du prisonnier peut être contourné ? En effet, dans le dilemme du prisonnier (un peu modifié).

$$\begin{array}{ccc} & & \text{Joueur II} \\ & & \begin{array}{cc} T \quad\quad & \quad\quad C \end{array} \\ \text{Joueur I} & \begin{array}{c} T \\ \\ C \end{array} & \begin{pmatrix} (1,1) & (3,0) \\ & \\ (0,3) & (2,2) \end{pmatrix} \end{array}$$

Figure 13.1:

deux joueurs I et II ont le choix entre coopérer (C) ou trahir (T). Il est strictement dominant pour les deux joueurs de trahir, ce qui procure une utilité de 1 à chaque joueur. Cependant, en coopérant, ils ont un paiement plus élevé (une utilité de 2). La répétition favorise-t-elle la coopération ?

13.1 Le modèle du jeu escompté

On se donne un jeu sous forme normal $\Gamma = \left(N, (S_i)_{i\in N}, (r_i)_{i\in N}\right)$.

Le jeu infiniment répété est défini comme suit. Il se joue en temps discret, $t = 1, 2, ...$:

- **Etape t = 1:** chaque joueur $i \in N$ choisit indépendamment des autres joueurs une action $s_i^1 \in S_i$. On note par $s^1 = \left(s_i^1\right)_{i\in N}$ le vecteur des actions jouées à l'étape 1. A la fin de cette étape, s^1 est révélée à tous les joueurs.
- **Etape t ≥ 2:** par récurrence. Connaissant l'histoire $h^t = (s^1, ..., s^{t-1})$ des actions jouées dans le passé, chaque joueur $i \in N$ choisit indépendamment des autres joueurs une action $s_i^t \in S_i$.
- La description du jeu est connaissance commune et les joueurs sont supposés se souvenir de toutes les actions observées dans le passé.

$H^t = S^{t-1}$ dénote l'ensemble des histoires possible à l'étape t. Par exemple, dans le dilemme du prisonnier,

$$H^2 = \{(C,C), (C,T), (T,C), (T,T)\}$$

et

$$H^3 = \{(C,C), (C,T), (T,C), (T,T)\} \times \{(C,C), (C,T), (T,C), (T,T)\}.$$

Une partie est une histoire de longueur infinie et prend la forme $\left(s^1, s^2, ...\right)$. H^∞ dénote l'ensemble des parties (des histoires de longueur infinie.)

Une stratégie de comportement σ_i du joueur i est une suite $\sigma_i = (\sigma_i^t)_{t=1,2,...}$ telle que σ_i^t est une fonction des informations H_t à l'étape t vers $\Delta(S_i)$, l'ensemble des actions mixtes du joueur i. Si σ_i^t est à valeur dans S_i alors la stratégie de comportement est dite pure. Ainsi, dans le dilemme du Prisonnier, une stratégie pure du joueur I pourrait être :

$$\begin{aligned} \sigma_1^1(\emptyset) &= C \\ \sigma_1^t(s^1, ..., s^{t-1}) &= \begin{cases} C & si \quad \forall \tau \le t-1 : s_2^\tau = C \\ T & si \quad \exists \tau \le t-1 : s_2^\tau = T \end{cases} \end{aligned}$$

Suivant cette stratégie, le joueur I commence par coopérer puis il continue à coopérer tant que l'adversaire fait de même. Le jour où l'adversaire le trahit il arrête de coopérer et décide de trahir pour toutes les étapes futures. Een quelque sorte, notre joueur est initialement sympathique mais très rancunier. S'il est trahi une fois, il se venge pour toujours. Cette stratégie est appelée **grim-trigger** (stratégie à seuil). Elle illustre bien le concept de punition, essentiel pour l'étude des équilibres dans les jeux répétés.

Soit Σ_i l'ensemble des stratégies pures du joueur i et soit Σ l'ensemble des profils de stratégies ($\Sigma = \prod_i \Sigma_i$). Chaque profil de stratégies pures $\sigma \in \Sigma$ induit donc une partie unique. Par exemple, si les deux joueurs dans le dilemme du prisonnier adoptent la stratégie **grim-trigger**, cela va induire une coopération à chaque période (les joueurs n'avouent jamais).

Il nous reste à définir la fonction de paiement. Pour cela, il faut déterminer comment les joueurs évaluent le résultat d'une histoire de longueur infinie. En effet, si $\left(s^1, s^2, ...\right) \in H^\infty$, le joueur i reçoit à l'étape t le paiement $r_i(s^t)$. On supposera ici que le joueur donne plus de poids à une unité d'utilité reçue aujourd'hui par rapport à une unité d'utilité reçue demain. Pour modéliser cela, on utilise généralement en économie un taux d'escompte $\delta \in]0,1[$. Ainsi, 1 unité d'utilité reçue à l'étape 2 vaut seulement δ unité d'utilité à l'étape 1 et une unité d'utilité reçue à l'étape t, vaut δ^{t-1} à l'étape 1.

Dans ce contexte, l'utilité du joueur i pour la partie $h = \left(s^1, s^2, ...\right)$ évaluée (dit aussi escomptée) à la date 1 est $g_i^\delta(h) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}r_i(s^t)$. Le facteur $(1-\delta)$ est tout simplement un facteur de normalisation qui permet de ramener à une même unité le jeu en une étape et le jeu infiniment répété. Par exemple dans le dilemme du prisonnier, le paiement escompté associé à l'histoire infinie où les joueurs coopèrent toujours est $(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1} \times 2 = 2$.

Le jeu répété Γ^δ est maintenant entièrement défini sous forme normale : chaque profil de stratégies pures $\sigma \in \Sigma$ induit une partie unique, $h = (s^1, s^2, ...)$ et donc un paiement le joueur i égal à $g_i^\delta(h) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}r_i(s^t)$.

Remarquez que, puisque le jeu est à mémoire parfaite, par le théorème de Kuhn, il nous suffit d'étudier le jeu à l'aide des stratégies de comportement. Soit donc Θ_i l'ensemble des stratégies de comportement (non nécessairement pures) du joueur i. Chaque profil de stratégies de comportement $\sigma = (\sigma_i)_{i\in N}$ défini, par le théorème de Kolmogorov, une probabilité unique, $P(\sigma)$ sur H^∞ (munie de la topologie naturelle induite par les cylindres). On notera par E_σ l'opérateur de l'espérance mathématique associée à $P(\sigma)$. Ainsi, si le profil de stratégies σ est choisi, l'espérance d'utilité du joueur i est $G_i^\delta(\sigma) = E_\sigma\left[g_i^\delta(\widetilde{h})\right]$.

13.2 Etude du dilemme du prisonnier répété

On se limitera ici aux stratégies pures. Remarquez que plus δ est proche de 1, plus les joueurs attribuent de la valeur aujourd'hui aux paiements futurs (on dit que les joueurs sont plus patients). Un cas extrême qui intéresse beaucoup les économistes est la limite quand δ tend vers 1. Il est encore plus intéressant de déterminer les profils σ qui constituent un équilibre pour tout jeu Γ^δ où les joueurs sont suffisamment patients. C'est-à-dire qu'il existe $\delta_0 \in]0,1[$ tel que σ est un équilibre de Nash de Γ^δ pour tout $\delta \in]\delta_0, 1[$. En effet, même si on sait qu'un acteur économique possède un taux d'escompte proche de 1, il est généralement difficile de le déterminer avec exactitude. Il est donc intéressant de considérer les comportements indépendant de δ.

Définition 31 *Un vecteur de stratégies pures $\sigma = (\sigma_i)_{i\in N}$ est un équilibre de Nash uniforme du jeu Γ^δ s'il existe $\delta_0 \in]0,1[$ tel que σ est un équilibre de Γ^δ pour tout $\delta \in]\delta_0, 1[$.*

Trahir toujours. Jouer T toujours est un équilibre de Nash uniforme du jeu. En effet, si l'autre joueur joue l'action T tout le temps, l'unique meilleure réponse est de joueur aussi T pour toujours. Ceci est donc un équilibre du jeu $\delta-$escompté, et ce quelque soit $\delta \in]0,1[$.

Grim-trigger. On va étudier quand grim-trigger définit un équilibre de Nash. Supposons que le joueur 2 adopte la stratégie grim-trigger. Si le joueur 1 adopte aussi la stratégie grim-trigger, alors l'issue de l'interaction est la coopération pour toujours, soit une séquence de paiement pour le joueur 1 de 2 unités à chaque période, donc un paiement escompté total, $(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1} \times 2$, égal aussi à 2. Maintenant, considérons la meilleure déviation possible du joueur 1. Ceci veut dire qu'à une certaine période t_0 le joueur 1 dévie est joue T. Supposons que t_0 est la première période où le joueur 1 joue T. En conséquence, toutes les périodes $t_0+1, t_0+2, ...$, le joueur 2 va jouer T (car il joue la stratégie grim-trigger). En conséquence, la meilleure déviation possible du joueur 1 lui dicte de jouer aussi T pendant toutes les périodes $t_0+1, t_0+2, ...$ Donc, pendant les t_0-1 premières périodes, le joueur 1 reçoit 2, à la t_0 ième période il reçoit 3, puis il reçoit 1 pour toujours. Son paiement escompté total est donc

$$\begin{aligned}
& (1-\delta)\left[2+\delta 2+...+\delta^{t_0-2}2+3\delta^{t_0-1}+\delta^{t_0}+\delta^{t_0+1}+....\right] \\
= \quad & 2+\delta^{t_0-1}-2\delta^{t_0} = 2+2\delta^{t_0-1}\left(\frac{1}{2}-\delta\right)
\end{aligned}$$

Ainsi, la déviation est profitable si est seulement si $\delta < \frac{1}{2}$. En conséquence, si les joueurs sont suffisamment patient ($\delta \geq \frac{1}{2}$) la paire de stratégies (grim-trigger, grim-trigger) est un équilibre de Nash pour tout jeu Γ^{δ}.

Les équilibres de Nash dans ce contexte peuvent contenir des menaces non crédibles, car en punissant l'autre joueur, un joueur peut s'auto-punir. D'où la nécessité d'étudier les équilibres sous jeux parfaits.

Définition 32 *Un profil de stratégies pures σ est un équilibre sous jeux parfait du jeu répété si pour toute histoire $h^t \in H$, $\sigma(\cdot\,|h^t)$ est un équilibre de Nash du jeu escompté qui commence à l'instant t. Il est uniforme si cette propriété est vraie pour tout taux d'escompte suffisamment proche de 1*

Il est facile de montrer que pour qu'un équilibre σ soit sous jeux parfait, il suffit de se limiter à vérifier qu'il n'existe de *déviation momentanée* dans aucun sous jeu. C'est-à-dire qu'il n'existe aucune histoire h^t et aucun joueur i tels que i possède une déviation gagnante $\widetilde{\sigma}_i$ dans le sous jeu escompté à la date t ayant pour histoire h^t et telle que $\widetilde{\sigma}_i$ est égale à σ_i partout sauf momentanément à la date t.

Pour illustrer cela, on va revenir au dilemme du prisonnier.

Trahir toujours est un équilibre sous jeux parfait. En effet, si le joueur 2 trahit toujours, alors quelque soit l'instant t et l'histoire du jeu h^t, si le joueur 1 dévie et coopère à l'instant t, alors il obtient 0 (car l'autre trahit) au lieu de 1 s'il trahissait. Donc il n'a intérêt à dévier à aucun moment.

Grim-trigger. Considérons l'histoire (T, C) (i.e. à la période 1, le joueur 1 dévie et trahit). Et supposons que le joueur 1 va continuer le jeu suivant la stratégie grim-trigger. Etant donné que le joueur 2 joue selon grim-trigger, on déduit que le joueur 1 obtient la suite de paiement suivant, à partir de la période 2 : $0, 1, 1, ...$ ce qui donne un paiement total escompté à la date 2 égal à δ. Si maintenant le joueur 1 dévie de cette stratégie et joue T à la seconde période, puis la stratégie grim-trigger après, il obtient le paiement 1 à chaque période (à compter de la période 2) et donc un paiement total escompté à la période 2 égal à 1. Puisque $\delta < 1$, cette déviation est profitable et grim-trigger n'est donc pas un équilibre sous jeux parfait.

Grim-trigger II. Maintenant on va modifier un peu la stratégie Grim-trigger pour obtenir un équilibre sous jeux parfait. Soit la stratégie suivante pour le joueur i. Le joueur i commence par coopérer. Puisque dans le passé les deux joueurs ont coopéré, alors le joueur i coopère. Si au moins un des deux joueurs trahit (même si c'était lui le traître) alors i trahit pour toujours. Il est facile de montrer que ceci est un équilibre sous jeux parfait pour tout $\delta \geq \frac{1}{2}$. En effet, considérons une histoire h^t du type $\{(C, C) ..., (C, C)\}$ alors, le paiement conditionnel du joueur 1 est donc

$$G_1^{\delta}(\sigma\,|h^t) = (1 - \delta) \sum_{\tau=t}^{\infty} \delta^{\tau - t} \times 2 = 2$$

Une déviation momentanée à l'instant t consiste à joueur T. Celle-ci donne alors le paiement suivant au joueur 1

$$(1 - \delta)\,(r_1(T, C) + \delta + \delta^2 ...) = 3 - 2\delta = 3\left(\frac{1}{2} - \delta\right)$$

et donc si $\delta \geq \frac{1}{2}$ la déviation n'est pas profitable. Soit maintenant une histoire h^t différente de $\{(C, C) ..., (C, C)\}$. En suivant les recommandations de la stratégie grim-trigger II, le joueur 1 obtient un paiement escompté à l'instant t égal à $(1 - \delta) \sum_{\tau=t}^{\infty} \delta^{\tau - t} \times 1 = 1$. S'il dévie à l'instant t, il joue C, il obtient

$$(1 - \delta)\,(r_1(C, T) + \delta + \delta^2 ...) = \delta < 1$$

Cependant, les stratégies grim-trigger sont trop violentes dans leurs punitions. En fait, il est possible de les modifier pour obtenir une coopération entre les joueurs même après une trahison.

Forgiving-trigger (stratégie à seuil avec oubli). Dans cette nouvelle stratégie, un joueur i commence par coopérer. Il continue à coopérer tant que les autres joueurs coopèrent dans le passé. Si à un moment donné un des joueurs (n'importe lequel) trahit, le joueur i trahit (joue T) pendant k périodes consécutives. Après ces k périodes de punition, le joueur i recommence à coopérer jusqu'à une nouvelle trahison. Ainsi, après une trahison, un joueur est rancunier seulement pendant un certain laps de temps, puis il oubli (ou surpasse) sa rancoeur et recommence à coopérer avec l'adversaire (ou le partenaire). Etudions sous quelles conditions ceci est un équilibre sous jeux parfait. Considérons une histoire h^t finissant par k périodes consécutives de coopération entre les joueurs. Le paiement escompté conditionnel du joueur 1 est 2. Une déviation momentanée à l'instant t consiste à jouer T puis à suivre les recommandations de la stratégie forgiving-trigger. Il obtient le paiement suivant

$$(1-\delta)\left(3+\delta+\delta^2...+\delta^k+2\delta^{k+1}+2\delta^{k+2}....\right)=3-2\delta+\delta^{k+1}$$

la déviation n'est pas profitable si est seulement si

$$3-2\delta+\delta^{k+1}\leq 2$$

Si $k=1$, ceci est impossible car $\delta<1$. Si $k=2$ alors la déviation n'est pas profitable si $\delta\geq 0.62$ et quand $k\to\infty$ la borne inférieure de δ tend vers $\frac{1}{2}$. Il reste à vérifier qu'après une histoire finissant par un certain nombre de trahisons, il n'existe pas de déviation momentanée profitable. Ceci est laissé au lecteur.

13.3 Les "folk theorems"

13.3.1 Paiements réalisables

Ici, nous montrons que la répétition du jeu permet d'agrandir l'ensemble des paiements réalisables. Ensuite, nous introduisons les paiements minmax.

Soit $PR(\Gamma)\subset R^n$ l'ensemble des paiements réalisables dans le jeu en une étape Γ, à partir des stratégies pures. $PR(\Gamma)$ est donc la collection d'au plus $|S|=\Pi_i\,|S_i|$ points. $\widetilde{PR}(\Gamma)$ dénotera l'ensemble des paiements réalisables en stratégies mixtes. Remarquez que $\widetilde{PR}(\Gamma)\subset conv\{PR(\Gamma)\}$ où $conv\{PR(\Gamma)\}$ dénote l'enveloppe convexe $PR(\Gamma)$. Dans le dilemme du prisonnier, $\widetilde{PR}(\Gamma)=conv\{PR(\Gamma)\}$. Ceci n'est pas toujours le cas, par exemple dans la bataille des sexes

$$\begin{array}{cc} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} (1,3) & (0,0) \\ (0,0) & (3,1) \end{pmatrix} \end{array}$$

Figure 13.2:

on a

$$PR(\Gamma)=\{(1,3);(3,1);(0,0)\}\,,$$

et

$$\widetilde{PR}(\Gamma)=\cup_{(x,y)\in[0,1]^2}\left\{(xy+3(1-x)(1-y),3xy+(1-x)(1-y)\right\}.$$

Ici, $\widetilde{PR}(\Gamma)$ n'est pas un ensemble convexe (par exemple, $(2,2)=\frac{1}{2}(3,1)+\frac{1}{2}(1,3)\notin\widetilde{PR}(\Gamma)$.

$PR(\Gamma^\delta)$ est l'ensemble des paiements réalisables dans le jeu escompté Γ^δ. Il est clair que $PR(\Gamma^\delta)\subset conv\{PR(\Gamma)\}$ car le paiement de chaque période est dans $conv\{PR(\Gamma)\}$ et le paiement escompté est une combinaison convexe des paiements de chaque étape. En fait l'autre inclusion est vraie quand $\delta\to 1$:

Proposition 13 $\lim_{\delta\to 1}PR(\Gamma^\delta)=conv\{PR(\Gamma)\}$

Preuve. Soit $x \in conv\{PR(\Gamma)\}$. Donc, $x = \sum_{k=1}^{|S|} \alpha^k r(s^k)$ où $\alpha^k \geq 0$, $\sum_k \alpha^k = 1$ et $s^k \in S$. Supposons d'abord que les coefficients α_k sont rationnels, réduits au même dénominateur commun T : $\alpha^k = \frac{t^k}{T}$ avec $\sum_k t^k = T$. Soit le profil de stratégies, périodique avec période T où les joueurs jouent pendant t^1 étapes consécutives le profil s^1, puis pendant t^2 étapes s^2...., $t^{|S|}$ étape $s^{|S|}$. Puisque la limite des moyennes de Césaro, $\lim_{t_0 \to \infty} \frac{1}{t_0} \sum_{t=1}^{t_0} r(s_t)$, de la suite des paiements induite par ce profil de stratégies périodique est x, ceci implique que la limite des moyennes d'Abel, $\lim_{\delta \to 1}(1-\delta) \sum_t \delta^t r(s_t)$, est aussi égale à x. En effet, si $(a_t)_{t \geq 1}$ est une suite bornée à valeurs réelles alors

$$\begin{aligned} \lim \inf_{t_0 \to \infty} \frac{1}{t_0} \sum_{t=1}^{t_0} a_t &\leq \lim \inf_{\delta \to 1} (1-\delta) \sum_t \delta^t a_t \\ &\leq \lim \sup_{\delta \to 1} (1-\delta) \sum_t \delta^t a_t \leq \lim \sup_{t_0 \to \infty} \frac{1}{t_0} \sum_{t=1}^{t_0} a_t. \end{aligned}$$

Quand x n'est pas rationnel, on considère une suite $\{x_m\}_m$ telle que $x_m \to x$ avec x_m rationnel. A chaque x_m est associé une période T_m avec $T_m \to \infty$. On définit maintenant des cycles consécutifs de longueurs respectives T_1, T_2, ..., T_m, ... Dans chaque cycle, les joueurs jouent la stratégie correspondante... c'est-à-dire pendant t_m^k périodes, ils jouent s_m^k, $k = 1, ..., |S|$. Ainsi, le paiement moyen (de Césaro) du cycle de longueur T_m est x_m. Puisque la limite de x_m est x on déduit que la limite des moyennes de Césaro des suites de paiement est aussi x et donc il est de même pour la limite des moyennes d'Abel. ■

13.3.2 Paiement de punition

Comme le dilemme du prisonnier l'illustre bien, les équilibres dans les jeux répétés qui nécessitent une certaine coopération entre les joueurs sont souvent basés sur les punitions et les menaces. Ainsi, il est naturel d'introduire le niveau maximal d'une punition.

Définition 33 *Le vecteur de punition minmax $\beta = (\beta_i)_{i \in N}$ du jeu Γ est définit par*

$$\beta_i = \min_{(s_j)_{j \neq i} \in \prod_{j \neq i} \Delta(S_j)} \max_{s_i \in \Delta(S_i)} R_i (s_1, ..., s_n)$$

On dénote par s_{-i}^* une stratégie de punition des joueurs autres que i :

$$s_{-i}^* \in \arg \min_{(s_j)_{j \neq i} \in \prod_{j \neq i}} \max_{s_i \in \Delta(S_i)} R_i (s_1, ..., s_n)$$

Ainsi, β_i correspond à la punition maximale que les joueurs autres que i sont sûrs d'infliger en stratégies mixtes au joueur i, même si leurs stratégies sont connues par i. Il est clair que n'importe quel équilibre de Nash dans Γ^δ ou dans Γ procure un paiement au joueur i au moins égal à β_i (prouvez le). C'est pour cela que tout vecteur $x = (x_i)_{i \in N} \in R^n$ tel que pour tout i, $x_i \geq \beta_i$ est appelé individuellement rationnel.

Définition 34 *$PIR(\Gamma)$ est l'ensemble des paiements individuellement rationnels. C'est l'ensemble des $x \in R^n$ tels que pour tout i, $x_i \geq \beta_i$.*

Remarquons que dans le cas de deux joueurs, par le théorème de l'existence de la valeur pour les jeux à somme nulle, $\beta_i = \underline{v}_i$, où $\underline{v}_i$ est la valeur maxmin du joueur i (le montant minimal que peut garantir le joueur i même si son adversaire connaît sa stratégie). Ceci est faux pour plus de deux joueurs (pourquoi ? donner un exemple).

13.3.3 Les "folk theorems" : coopération, punition et récompense des punisseurs

Notons $PEN(\Gamma^\delta)$ l'ensemble des paiements d'équilibre de Nash du jeu Γ^δ. Il est clair que tout paiement d'équilibre de Γ^δ doit être un paiement réalisable mais aussi individuellement rationnel. L'autre inclusion est un des théorèmes centraux en théorie des jeux répétés.

Théorème 18 *(Folk Theorem 1 : Robert Aumann)*
$conv\{PR(\Gamma)\} \cap PIR(\Gamma) = \lim_{\delta \to 1} PEN(\Gamma^{\delta})$.

Preuve. La preuve est largement inspirée de la stratégie **grim-trigger** dans le dilemme du prisonnier. Soit $x \in conv\{PR(\Gamma)\}$ tel que $x_i \geq \beta_i$ pour tout joueur i. Nous supposons les joueurs rationnels. Comme pour la preuve précédente, $x = \sum_{k=1}^{|S|} \frac{t_k}{T} r(s^k)$, $\sum_k t_k = T$ et $s^k \in S$. Considérons encore la stratégie pure et périodique de période T où le joueur i joue à l'instant $t_k + mT$, $m \geq 0$, l'action pure s_i^k et ce tant que les autres joueurs ont **coopéré** en respectant cette recommandation dans le passé. Si un joueur i dévie, alors les joueurs autres que i **le punissent** et ce jusqu'à la fin des temps à son niveau minmax β_i. Plus précisément, les joueurs $-i$ jouent i.i.d. la stratégie de punition s^*_{-i}. Ainsi, si le joueur i dévie, par la loi des grands nombres, la limite de la moyenne de Césaro (et donc d'Abel) de ces paiements est au plus β_i. Si le joueur i ne dévie pas, il reçoit x car la limite de la moyenne d'Abel d'une suite convergente vers x n'est autre que x. Puisque $x_i \geq \beta_i$, le joueur i n'a pas intérêt à dévier, d'où $x \in \lim_{\delta \to 1} PEN(\Gamma^{\delta})$. Quand x n'est pas rationnel, on procède comme précédemment dans la preuve des paiements réalisables (en construisant des cycles de plus en plus long...) ■

Remarques :

- Si $x \in conv\{PR(\Gamma)\}$ est tel que $x_i > \beta_i$ pour tout i alors il existe $\delta_0 \in]0, 1[$ tel que le profil de stratégies grim-trigger défini dans la preuve précédente constitue un équilibre de Γ^{δ} et ce pour tout $\delta \in]\delta_0, 1[$, autrement dit x est un paiement d'équilibre uniforme.

- Le théorème précédent est vrai avec des équilibres sous jeux parfaits (le Folk theorem 2). Pour cela, il suffit de considérer des stratégies du type forgiving-trigger, où les punitions ne durent pas indéfiniment mais seulement un certain nombre de périodes k, permettant ainsi de compenser le gain instantané du joueur déviant. Dans ce cas, même si les punisseurs s'autopunissent pendant ces k périodes, ceci n'a pas beaucoup d'incidence sur leur paiement total escompté (quand δ est suffisamment proche de 1) car la période de punition dure seulement un nombre borné d'étapes après, on recommence à coopérer : c'est la **récompense des punisseurs**.

- Le théorème précédent est vrai avec des taux d'escomptes différents pour les joueurs ou encore si on remplace la limite de la moyenne d'Abel par la limite de la moyenne de Césaro.

- Dans le dilemme du prisonnier répété t_0 fois avec comme paiement total la moyenne de Césaro des paiements reçus pendant les t_0 périodes, on peut montrer que l'unique équilibre de Nash (sous jeux parfait ou pas) est la trahison à chaque étape (montrer-le). Ceci est en contradiction flagrante avec l'expérimentation. Cet effet de bord montre que le modèle du jeu escompté infiniment répété est meilleur pour analyser les interactions de longues durées, même si, dans la réalité, elles ne durent jamais indéfiniment. C'est une des raisons essentielles de son succès auprès des économistes. L'effet de bord peut aussi être éliminé autrement en supposant que la durée du jeu est incertaine et que chaque joueur a une information privée partielle sur cette durée. Par exemple que le joueur 1 pense que c'est la durée est soit 99 soit 100 et le joueur 2 pense que c'est soit 100 ou 101.

Chapter 14

Modélisation de l'information incomplète

14.1 Information privée, secrète, hiérarchie de croyances

Jusqu'à présent, on a souvent assumé que le jeu et ses paramètres sont une connaissance commune entre les joueurs. Dans la réalité, un agent économique (entreprise, consommateur, etc) dispose d'informations privées que lui seul détient. Par exemple, quand vous voulez vendre votre voiture ou une marchandise, vous êtes le seul à connaître les éventuels défauts. Mais aussi, si une entreprise connaît bien ses coûts elle ne connaît pas ceux de son concurrent ni sa situation financière... Ainsi, dans un jeu donné, un joueur peut avoir un doute sur *le type* de son partenaire, sur ses stratégies ou encore sur sa fonction d'utilité.

L'approche bayésienne (Harsanyi) consiste à attacher une distribution de probabilité à tout élément incertain. Après réception d'une nouvelle information, un agent révise ses croyances à l'aide de la *règle de Bayes* pour les probabilités conditionnelles. Par exemple, considérons deux joueurs I et II qui jouent un jeu matriciel 2×2 dans lequel les paiements dépendent d'un paramètre inconnu $k \in \{1, 2\}$. Soit donc p (resp. q) la probabilité que le joueur I (resp. II) affecte à l'évènement $\{k = 1\}$. Le joueur I peut être incertain sur la croyance du joueur II concernant l'état de la nature. L'approche bayésienne nous recommande donc que I devrait posséder une probabilité sur la croyance de II. Par exemple, I peut croire qu'avec probabilité $\frac{1}{3}$ le joueur II croit que la probabilité de $\{k = 1\}$ est q_1 et qu'avec probabilité $\frac{2}{3}$ c'est q_2. Mais dans ce cas, le joueur I devrait aussi posséder une croyance sur la croyance de la croyance de l'autre sur lui, etc. Ceci induit naturellement une hiérarchie infinie de croyances.

A titre d'exemple, supposons que $p = q = \frac{1}{2}$ et que les matrices des paiements, dépendant de l'état de la nature, sont :

$$si \quad k=1 \qquad \begin{array}{c c} & \begin{array}{cc} G & D \end{array} \\ \begin{array}{c} H \\ \\ B \end{array} & \begin{bmatrix} 1,0 & 4,-5 \\ \\ 2,-2 & -4,3 \end{bmatrix} \end{array}$$

et

$$si \quad k=2 \qquad \begin{array}{c c} & \begin{array}{cc} G & D \end{array} \\ \begin{array}{c} H \\ \\ B \end{array} & \begin{bmatrix} -1,0 & 0,-1 \\ \\ -4,0 & 0,-7 \end{bmatrix} \end{array}$$

Si aucun des joueurs ne sait quel état a été tiré alors cela revient à jouer dans le jeu (dit moyen) suivant :

$$\frac{1}{2}\begin{bmatrix} 1,0 & 4,-5 \\ \\ 2,-2 & -4,3 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} -1,0 & 0,-1 \\ \\ -4,0 & 0,-7 \end{bmatrix} = \begin{bmatrix} 0,0 & 2,-3 \\ \\ -1,-1 & -2,-2 \end{bmatrix}.$$

L'unique équilibre de Nash après élimination des stratégies strictement dominées est (H,G) avec un vecteur de paiement moyen espéré égal à $(0,0)$. Considérons les variantes suivantes dans lesquelles le joueur I est informé de l'état de la nature mais pas le joueur II :

1) I découvre secrètement quel est le jeu joué. Il prévoit aussi que le joueur II va jouer sa stratégie d'équilibre G. Dans ce cas il change secrètement son comportement quand $k = 1$. En effet, sa meilleure réponse et de continuer à jouer H si $k = 2$ par contre il doit jouer B s'il sait que $s = 1$. Le vecteur de paiement moyen espéré dans le jeu est $\frac{1}{2} \times (2, -2) + \frac{1}{2} \times (-1, 0) = \left(\frac{1}{2}, -1\right)$.

2) Si II sait que I connaît le jeu mais que I ne le sait pas. II prévoit donc que I va jouer B si $k = 1$ et qu'il va jouer H si $k = 2$. Dans ce cas II va jouer D plutôt que G. Le vecteur de paiement moyen espéré devient donc $(-2, 1)$.

3) Si I sait que II sait que I connaît le jeu, I prévoit que II va jouer D. Dans ce cas, celui-ci joue H et le paiement espéré est donc $(2, -3)$.

4) Si II sait que I sait que II sait que I connaît le jeu, alors II prévoit que I va jouer H et sa meilleure réponse est de jouer G avec un vecteur paiement moyen de (0,0).

5) Si I sait que II sait que I sait que II sait que I connaît le jeu, I prévoit que II joue G. Retour à l'étape 1...

6) Si I connaît l'état de la nature et que cela est une connaissance commune entre les joueurs (chacun sait que l'autre sait que... que I sait que II sait.... que I connaît l'état de la nature). Dans ce cas, le jeu sous forme normale est le suivant :

$$\begin{array}{c} \\ HH \\ HB \\ BH \\ BB \end{array} \begin{array}{c} \begin{array}{cc} G & D \end{array} \\ \begin{bmatrix} 0,0 & 2,-3 \\ -\frac{3}{2},0 & 2,-6 \\ \frac{1}{2},-1 & -2,1 \\ -1,-1 & -2,-2 \end{bmatrix} \end{array}$$

où HB pour le joueur I correspond à la stratégie de jouer H si $k = 1$ et de jouer B si $k = 2$. Le vecteur de paiement $\left(-\frac{3}{2}, 0\right)$ est la moyenne entre le vecteur $(1, 0)$ (l'action H contre l'action G dans l'état $k = 1$) et le vecteur $(-4, 0)$ (l'action B contre l'action G dans l'état $k = 2$). L'unique équilibre de ce jeu est mixte. A l'équilibre, le joueur I joue la stratégie mixte $\left(\frac{2}{5}, 0, \frac{3}{5}, 0\right)$ et le joueur II la stratégie mixte $\left(\frac{8}{9}, \frac{1}{9}\right)$ avec comme vecteur de paiement espéré $\left(\frac{2}{9}, -\frac{3}{5}\right)$.

Cette exemple montre que pour analyser le jeu avec incertitude sur les paiements, la différence d'information sur l'incertitude doit être prise en compte dans l'analyse du jeu. Dans le cas où un joueur a une information et que les autres le savent, cette information est dite privée. Sinon, celle-ci est dite secrète.

14.2 Sur la valeur de l'information

Il est toujours avantageux d'avoir une information secrète. En effet, si les autres joueurs ne savent pas que je détiens une information, je peux toujours jouer comme si je n'en avais pas si je n'ai aucun avantage à l'utiliser. Dans le cas de l'information privé, l'exemple de la section précédente montre que parfois le joueur informé (ici le joueur I) obtient un paiement supérieur à son paiement lorsqu'il n'est pas informé (en l'occurrence, $\frac{2}{9}$ au lieu de 0). Cette intuition selon laquelle avec une information privée on obtient plus est fausse en général (par contre elle est toujours vraie pour les jeux à somme nulle.) En effet, supposons que la matrice des paiements si $k = 1$ est la suivante :

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} a & b \end{array} \\ \begin{bmatrix} -2,10 & 8,-8 \\ 10,-4 & 10,-8 \end{bmatrix} \end{array}$$

et que les paiements pour $k = 2$ sont

$$\begin{array}{cc} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} A \\ \\ B \end{array} & \begin{bmatrix} 10, -2 & -4, 10 \\ \\ -8, 8 & -8, 10 \end{bmatrix} \end{array}$$

Supposons en plus que la probabilité de l'état $k = 1$ est $\frac{1}{2}$.

Cas 1 : aucun joueur n'est informé du vrai état et les deux joueurs le savent. Dans ce cas, les joueurs jouent dans le jeu moyen suivant

$$\frac{1}{2}\begin{bmatrix} -2, 10 & 8, -8 \\ \\ 10, -4 & 10, -8 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 10, -2 & -4, 10 \\ \\ -8, 8 & -8, 10 \end{bmatrix} = \begin{bmatrix} 4, 4 & 2, 1 \\ \\ 1, 2 & 1, 1 \end{bmatrix}.$$

L'unique équilibre après élimination des stratégies strictement dominées est (A,a). Le paiement moyen associé à cet équilibre est $(4, 4)$.

Cas 2 : le joueur I connaît l'état et II sait que I connaît l'état. Il est donc strictement dominant pour I de jouer B dans l'état $k = 1$ et strictement dominant de joueur A dans l'état $k = 2$. Dans ce cas, l'unique meilleure réponse pour II c'est de joueur l'action b. En effet, s'il joue a, il espère gagner $\frac{1}{2} \times -4 + \frac{1}{2} \times -2 = -3$ alors qu'en jouant b il espère gagner $\frac{1}{2} \times -8 + \frac{1}{2} \times 10 = 1$. Le paiement espéré du joueur I est donc $\frac{1}{2} \times 10 + \frac{1}{2} \times -4 = 3$: le joueur I gagne moins que sans information !

Ceci montre un aspect très important des jeux à information incomplète : il est possible qu'un joueur préfère ne pas être informé car si les autres savaient qu'il est informé, cela risquerait de changer leur comportement en sa défaveur. Intuitivement, si vous saviez qu'un investisseur n'a pas de problème de trésorerie et qu'il possède une information privée crédible sur la valeur d'un titre financier, vous n'avez aucun intérêt à faire une transaction avec lui. Ce paradoxe est connu dans la littérature sous le nom de *paradoxe de non-échange* ou encore *être d'accord pour ne pas être d'accord* (Aumann).

En fait il y a une équivalence entre le fait qu'il n'est pas toujours avantageux d'avoir plus d'information et le fait qu'il n'est pas toujours avantageux d'avoir plus de stratégies. En effet, avoir une information privée a comme effet d'augmenter les possibilités stratégiques d'un joueur (car il peut jouer en fonction de son information). Maintenant il est plus simple de comprendre pourquoi augmenter les possibilités stratégiques d'un joueur (ou de tous les joueurs) peut avoir un effet désastreux. En effet, si on rajoute à un individu (resp. un trader) la possibilité de retrait d'argent (resp. d'investissement) illimitée ou encore si on autorise la vente d'armes aux particuliers etc, les effets peuvent être très négatifs ...

14.3 Modèles à la Harsanyi, équilibre bayésien

Harsanyi en 1967-68 a été le premier à introduire la notion de type. Un type, dans le point de vue de Harsanyi, est une sorte d'état d'esprit. Il contient non seulement la croyance du joueur sur l'état de la nature mais aussi sur sa croyance sur la croyance des autres, et la croyance du joueur sur la croyance des autres sur sa croyance, etc. Cependant, Harsanyi ne donna pas une construction rigoureuse de cette idée. Cela a été fait une vingtaine d'années plus tard (1985) par Mertens et Zamir. Les ingrédients du modèle de Harsanyi standard sont les suivants :

- $N = \{1, 2, ..., i, .., n\}$ est l'ensemble des joueurs.
- T_i est l'ensemble des types pour le joueur i.
- $T = \prod_{i \in N} T_i$ dénote l'ensemble des profiles des types.
- Pour tout $t \in T$, est associé un jeu en forme normale $\Gamma_t = \left(N, (S_i(t_i))_{i \in N}, (r_i(\cdot; t))_{i \in N}\right)$ où $S_i(t_i)$ dénote l'ensemble des stratégies pures du joueur i si son type est t_i et $u_i(\cdot; t)$ dénote la fonction d'utilité du joueur i si le profil des types est $t = (t_1, ..., t_n)$.

- $p \in \Delta(T)$ est une distribution de probabilité sur T.

Un jeu bayésien standard (dit aussi consistant) se déroule de la manière suivante :

- La nature utilise la probabilité p pour choisir un profil de type $t = (t_1, ..., t_n) \in T$.
- Chaque joueur $i \in N$ est informé de son propre type t_i.
- Les joueurs choisissent simultanément leurs actions : chaque joueur i choisit une action $s_i \in S_i(t_i)$.
- Chaque joueur i reçoit le paiement $u_i(s, t)$.

Quand un joueur i sait que son type est t_i, il peut calculer, à l'aide de la règle de Bayes, la probabilité conditionnelle, $p_{t_i} = p(\cdot\,|t_i) \in \Delta\left(\times_{j \neq i} T_j\right)$, sur les types possibles des autres joueurs. Le modèle de Harsanyi standard est appelé consistant car toutes les probabilités conditionnelles sont obtenues par la règle de Bayes d'une même distribution de probabilité, p, sur T. Harsanyi suppose donc que les croyances des joueurs dérivent toutes d'un même **mécanisme objectif** qui assigne une information à chaque joueur. Le modèle standard de Harsanyi est aussi appelé le jeu bayésien car il définit vraiment un jeu (un joueur peut évaluer *ex-ante* et d'une manière objective son paiement espéré).

Il est possible de définir un modèle encore plus général que celui de Harsanyi, en considérant une famille de probabilité $(p_{t_i})_{i \in N, t_i \in T_i}$. Quand le joueur i sait que son type est t_i alors il considère que les types des autres joueurs sont distribués selon la probabilité p_{t_i}. Ces probabilités ne dérivent pas nécessairement toutes d'une même distribution jointe sur l'ensemble des types. Ce modèle bayésien est appelé non (nécessairement) consistant ou encore le modèle bayésien universel. Dans un tel modèle, les probabilités sont subjectives et non objectives. En effet, un joueur i peut croire que la probabilité d'un évènement est α (par exemple avec une probabilité $> \frac{1}{2}$ le CAC 40 va monter) et en même temps, le joueur j peut croire que cette probabilité est $\beta \neq \alpha$ (par exemple avec une probabilité $< \frac{1}{2}$ le CAC 40 va baisser). Cependant, ce modèle ne décrit pas vraiment un jeu car le mécanisme utilisé pour informer les joueurs n'est pas spécifié. Les paiements ne peuvent pas être évalués, pour un profil de stratégies, avant que les joueurs ne soient informés de leurs types. Cependant, pour chaque joueur i, il existe plusieurs probabilités subjectives *ex-ante* p_i sur T compatibles avec la famille de probabilités *ex-poste* p_{t_i}. Rien ne nous empêche de supposer que le joueur i croit que la vraie probabilité est l'une d'entre elles. Ceci permet donc à chaque joueur i d'avoir une espérance subjective *ex-ante* sur ces paiements. Cependant, celle-ci ne correspondra pas toujours au paiement réel reçu.

Une stratégie mixte σ_i du joueur i dans le modèle bayésien universel est une application qui associe à chaque $t_i \in T_i$ une probabilité $\sigma_i(t_i) \in \Delta\left(S_i(t_i)\right)$. Si le type d'un joueur est t_i alors il va jouer la stratégie $\sigma_i(t_i)$. Si les joueurs autres que i jouent le profil σ_{-i}, alors le joueur i de type t_i espère obtenir le paiement suivant (en supposant pour simplifier que tout est fini):

$$R_i((\sigma_i(t_i), \sigma_{-i})\,; t_i) := \sum_{t_{-i} \in T_{-i}} p_{t_i}(t_{-i}) \sum_{s \in [\times_{i \in N} S_j(t_j)]} \prod_{j \in N} \sigma_j(t_j)(s_j) r_i(s; t)$$

où $t = (t_i, t_{-i})$.

Définition 35 *Un profil de stratégies mixtes σ est un équilibre bayésien si pour tout i et pour type $t_i \in T_i$, $R_i(\sigma_i(t_i); \sigma_{-i}; t_i) = \max_{\widetilde{\sigma}_i \in \Delta(S_i(t_i))} R_i(\widetilde{\sigma}_i; \sigma_{-i}; t_i)$*

Harsanyi souleva l'importance d'étudier le jeu *ex-ante* : le joueur i doit prendre en considération dans sa stratégie, même après connaissance de son type, de ce qu'il aurait fait si son type avait été différent. Remarquons d'abord que décider ce qu'il doit faire avant de connaître son type ne restreint aucunement le joueur i car dans tout équilibre de Nash, le joueur i va jouer une meilleure réponse aux stratégies des autres joueurs, et ce connaissant son type. Enfin, le joueur i sait que les autres joueurs ont une incertitude concernant son type. Ainsi, si le comportement de i dépend de son type, cela affecte les stratégies des autres joueurs. i doit prendre cela en considération.

Théorème 19 *(Harsanyi)*
Si tous les ensembles (des joueurs, types et actions) sont finis alors le modèle bayésien (consistant ou non) admet un équilibre bayésien en stratégies mixtes.

Preuve. On applique le théorème de Nash. Pour cela, il suffit de remarquer qu'un équilibre bayésien n'est autre qu'un équilibre de Nash du jeu fini en forme stratégique suivant :

- l'ensemble des joueurs est l'ensemble des (i, t_i) où $i \in N$ et $t_i \in T_i$.
- l'ensemble des stratégies pures du joueur (i, t_i) est $S_i(t_i)$.
- si chaque joueur (j, t_j) choisit la stratégie pure $s_j(t_j)$ alors le paiement de (i, t_i) est :

$$\sum_{t_{-i} \in T_{-i}} p_{t_i}(t_{-i}) r_i(\Big(s_i(t_i), [s_j(t_j)]_{j \neq i}\Big) ; (t_i, t_{-i}))$$

Remarquez que dans cette preuve, on n'a pas supposé que le modèle bayésien était consistant. Ceci permet donc d'étendre le résultat de Harsanyi.

■

14.3.1 Application 1 : une justification des stratégies mixtes

Les équilibres en stratégies mixtes sont parfois difficiles à interpréter. En effet, considérons l'exemple suivant :

$$\begin{array}{cc} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} A \\ \\ B \end{array} & \begin{bmatrix} 0,0 & 0,-1 \\ & \\ 1,0 & -1,3 \end{bmatrix} \end{array}$$

L'unique équilibre est mixte. Le joueur I utilise la loterie $\left(\frac{3}{4}, \frac{1}{4}\right)$ et le joueur II la loterie $\left(\frac{1}{2}, \frac{1}{2}\right)$. Quand le joueur II joue la stratégie $\left(\frac{1}{2}, \frac{1}{2}\right)$, le joueur I est indifférent entre A et B. Pourquoi alors il utiliserait une loterie pour choisir entre les deux ? et même si on admettait cela, pourquoi $\left(\frac{3}{4}, \frac{1}{4}\right)$ et pas une autre loterie ?

Harsanyi a montré, en 1973, que les équilibres de Nash mixtes peuvent être interprétés comme limites d'équilibres bayésiens dans lesquels les joueurs jouent presque toujours en stratégies pures. L'idée est de modifier un tout petit peu le jeu en introduisant de l'information privée. Dans l'exemple précédent, cela donne :

$$\begin{array}{cc} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} A \\ \\ B \end{array} & \begin{bmatrix} \varepsilon\widetilde{\alpha}, \varepsilon\widetilde{\beta} & \varepsilon\widetilde{\beta}, -1 \\ & \\ 1, \varepsilon\widetilde{\alpha} & -1, 3 \end{bmatrix} \end{array}$$

avec $\varepsilon > 0$ très petit, $\widetilde{\alpha}$ et $\widetilde{\beta}$ sont deux variables aléatoires uniformément distribuées dans $[0, 1]$. La nature tire $\widetilde{\alpha}$ et $\widetilde{\beta}$, le joueur I est informé de $\widetilde{\alpha}$ et le joueur II est informé de $\widetilde{\beta}$. Etant donné que ε est très petit, cette modification a une très petite influence sur les paiements. Etant donné ε, on peut montrer qu'il existe essentiellement un unique équilibre bayésien. Si le joueur I observe que $\widetilde{\alpha} > \frac{2+\varepsilon}{8+\varepsilon^2}$ alors il joue A, sinon il joue B. Si $\widetilde{\beta} > \frac{4-\varepsilon}{8+\varepsilon^2}$ le joueur II joue a, sinon il joue b. Il suffit maintenant de remarquer que $\lim_{\varepsilon \to 0} proba(\widetilde{\alpha} > \frac{2+\varepsilon}{8+\varepsilon^2}) = \frac{3}{4}$ et que $\lim_{\varepsilon \to 0} proba(\widetilde{\beta} > \frac{4-\varepsilon}{8+\varepsilon^2}) = \frac{1}{2}$.

14.3.2 Application 2 : une modélisation de l'irrationalité

L'information incomplète peut servir pour modéliser le fait qu'un joueur puissent être irrationnel avec une certaine probabilité et/ou être incertain sur la rationalité de l'adversaise. Tout le problème consiste à définir ce que signifie être irrationnel. Cela est plus ou moins clair selon les situations. Par exemple, il est irrationnel qu'un joueur puisse jouer une stratégie strictement dominée. Il est irrationnel dans le dilemme du prisonnier répété un nombre fini de fois de coopérer à une période donnée. Aussi, il est irrationnel dans le jeu du mille-pattes de continuer le jeu à un moment donné.

14.4 L'espace de croyances de Mertens-Zamir

14.4.1 Croyances hiérarchiques implicites dans le modèle de Harsanyi

Considérons pour simplifier deux joueurs I et II. L'ensemble des types du joueur I est $T_I = \{t_{I_1}, t_{I_2}\}$ et l'ensemble des types de II est $T_{II} = \{t_{II_1}, t_{II_2}\}$. Ainsi, $T = \{a = (t_{I_1}, t_{II_1}), b = (t_{I_1}, t_{II_2}), c = (t_{I_2}, t_{II_1}), d = (t_{I_2}, t_{II_2})\}$. La probabilité, t, sur les types est la suivante

$$\begin{array}{cc} & \begin{array}{cc} t_{II_1} & t_{II_2} \end{array} \\ \begin{array}{c} t_{I_1} \\ \\ t_{I_2} \end{array} & \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ & \\ \frac{1}{3} & \frac{1}{6} \end{bmatrix} \end{array}$$

a, b, c et d sont les différents états possibles. Supposons que l'état soit a, quelle est la structure d'information des joueurs ?

- **Niveau 1 des croyances :**
  - Type I_1 : avec probabilité $\frac{1}{2}$ l'état est a et avec probabilité $\frac{1}{2}$ l'état est b.
  - Type I_2 : avec probabilité $\frac{2}{3}$ l'état est c et avec probabilité $\frac{1}{3}$ l'état est d.
  - Type II_1 : avec probabilité $\frac{3}{7}$ l'état est a et avec probabilité $\frac{4}{7}$ l'état est c.
  - Type II_2 : avec probabilité $\frac{3}{5}$ l'état est b et avec probabilité $\frac{2}{5}$ l'état est d.
- **Niveau 2 des croyances :**
  - Type I_1 : le joueur I croit qu'avec probabilité $\frac{1}{2}$ le joueur II croit qu'avec probabilité $\frac{3}{7}$ l'état est a et avec probabilité $\frac{4}{7}$ l'état est c, et le joueur I croit qu'avec probabilité $\frac{1}{2}$ le joueur II croit qu'avec probabilité $\frac{3}{5}$ l'état est b et avec probabilité $\frac{2}{5}$ l'état est d.
  - Type I_2 : le joueur I croit qu'avec probabilité $\frac{2}{3}$ le joueur II croit qu'avec probabilité $\frac{3}{7}$ l'état est a et avec probabilité $\frac{4}{7}$ l'état est c, et le joueur I croit qu'avec probabilité $\frac{1}{3}$ le joueur II croit qu'avec probabilité $\frac{3}{5}$ l'état est b et avec probabilité $\frac{2}{5}$ l'état est d.
  - etc

Ainsi, tout modèle de Harsanyi cache une structure hiérarchique infinie de croyances, en concordance donc avec l'intuition originelle de Harsanyi. La question que se sont posée Mertens et Zamir (1985) est la suivante : est-ce que toute hiérarchie de croyances peut être décrite par un modèle de type Harsanyi ?

14.4.2 Construction de l'espace universel de croyances

Mertens et Zamir se sont donné comme objectif la construction d'un espace universel de croyances. C'est-à-dire, étant donné un ensemble de joueurs et un ensemble d'états de la nature, quel est l'ensemble minimal de toutes les structures mutuelles de croyances. Cela permet de justifier et de donner un sens au modèle abstrait de Harsanyi. Mertens et Zamir ont permis de comprendre la subtilité et de marquer la distinction entre entre état de la nature, état du monde et état d'esprit (i.e. type). Par exemple, on peut avoir un ensemble d'états de la nature fini mais un ensemble d'états du monde (esprit) infini.

On supposera que l'ensemble des joueurs $N = \{1, 2, ..., n\}$ est fini et que l'ensemble K des états de la nature est un compact métrique (par exemple un ensemble fini). On dénotera pour un compact métrique A, par $\Delta(A)$ l'ensemble des probabilités boréliennes sur A. En munissant $\Delta(A)$ de la topologie faible étoilée, $\Delta(A)$ est aussi compact et métrique. Chaque état de la nature peut correspondre à un jeu sous forme normale.

Introduisons d'abord les ensembles suivants pour un joueur i :

- $X_i^1 = \Delta(K)$: l'espace de toutes les croyances possibles d'ordre (ou de degré) 1 pour le joueur i.
- $X_i^2 = X_i^1 \times \Delta(K \times (\times_{j \neq i} X_j^1))$: l'espace de toutes les croyances possibles jusqu'à l'ordre 2. En effet, dans un jeu, le comportement du joueur i dépend de ce qu'il pense que les autres vont faire. Mais ce que les autres vont faire dépend de ce qu'ils pensent que i croit. Ainsi, le joueur i doit avoir une croyance sur le premier ordre de croyance des autres joueurs. Si $(\mu_1, \mu_2) \in X_i^2$, alors μ_1 dénote la croyance (de niveau 1) du joueur i sur K, et μ_2 dénote la croyance de i sur K et sur ce que croient les $n-1$ autres joueurs sur K. Dans notre définition de X_i^2, la croyance de premier ordre du joueur i y est implicitement incluse.
- $X_i^3 = X_i^2 \times \Delta(K \times (\times_{j \neq i} X_j^2))$: l'espace de toutes les croyances possibles jusqu'à l'ordre 3.
- ...
- $X_i^{k+1} = X_i^k \times \Delta(K \times (\times_{j \neq i} X_j^k))$: l'espace des croyances jusqu'à l'ordre $k+1$.
- ...

Remarquons que nous nous intéressons seulement aux *croyances cohérentes* et npn à toutes les croyances possibles. Par exemple, si $(\mu_1, \mu_2) \in X_2$ où $\mu_1 \in \Delta(K)$ et $\mu_2 \in \Delta(K \times (\times_{j \neq i} X_j^1))$ alors la marginale de μ_2 sur K est par définition la croyance (d'ordre 1) du joueur i sur K, celle-ci doit donc coïncider avec μ_1. D'où la contrainte suivante :

La contrainte de cohérence : Tout événement admet la même probabilité et ce pour toute mesure de probabilité définie sur cet évènement (voir les exemples plus bas).

Imposent cette contrainte à toute la construction induit donc une suite projective d'espaces C_i^1, C_i^2, C_i^3,... qui satisfait :

- $C_i^1 = \Delta(K)$.
- $C_i^2 \subset C_i^1 \times \Delta(K \times (\times_{j \neq i} C_j^1))$.
- ...
- $C_i^{k+1} \subset C_i^k \times \Delta(K \times (\times_{j \neq i} C_j^k))$.
- ...

Soit $C_i = C_i^\infty$ la limite projective de cette séquence (on peut montrer qu'elle est non vide et compacte). $C = \times_i C_i$ est appelé l'espace universel des types. Il est déterminé seulement par K et n. Les espaces C_i satisfont :

$$C_i = \Delta(K \times (\times_{j \neq i} C_j))$$

et l'espace universel des croyances Ω est défini par

$$\Omega = K \times C.$$

Un élément $\omega \in \Omega$ est appelé un *état du monde* (noté la différence avec un état de la nature) :

$$\omega = (k; c_1(\omega), ..., c_n(\omega))$$

où k est l'état de la nature et où $c_i(\omega)$ est la croyance (le type) du joueur i à l'état du monde ω (aussi appelé un *état d'esprit* du joueur i à l'état du monde ω). Remarquez que $c_i(\omega)$ peut aussi être vue comme une distribution de probabilité sur Ω qui a comme marginale sur C_i la masse de Dirac au point $c_i(\omega)$.

Un état du monde est un $(n+1)$-uple contenant en un état de la nature et pour chaque joueur, une distribution de probabilité sur l'ensemble des états du monde.

Définition 36 *Un sous espace de croyance est un sous ensemble fermé Y de Ω tel que pour tout $\omega \in Y$ et tout $i \in N$, $c_i(\omega)[Y] = 1$.*

Y est donc fermé pour les croyances : aucune croyance d'un joueur, quelque soit son ordre, ne fait appel à un état du monde en dehors de Y. Etant donné un sous espace Y, l'espace des types du joueur i est tout simplement $T_i = \{c_i(\omega), \omega \in Y\}$. Généralement, $T_i \neq T_j$ pour $i \neq j$. Notez que par construction $Y \subset K \times_{i \in N} T_i$ et que $T_i \subset \Delta(K \times (\times_{j \neq i} T_j))$. On notera dans la suite les éléments de T_i par t_i.

14.4.3 Exemples, propriétés et lien avec les modèles à la Harsanyi

Les exemples qui suivent montrent la richesse du modèle de Mertens-Zamir qui permet de modéliser plusieurs situations réelles.

- **Jeu à information complète**

  Ceci est le cas si et seulement si le sous espace de croyances Y est réduit à un singleton : $Y = \{\omega = (k; (1), ..., (1))\}$. L'unique état de la nature, k, est une connaissance commune entre les joueurs.

- **Incertitude commune sur l'état de la nature**

  Supposons que l'état est soit k_1 avec probabilité $\frac{1}{3}$ soit k_2 avec probabilité $\frac{2}{3}$ et que cela soit une connaissance commune entre les joueurs. Dans ce cas :

  - $Y = \{\omega_1, \omega_2\}$
  - $K = \{k_1, k_2\}$
  - $\omega_1 = \{k_1; \left(\frac{1}{3}, \frac{2}{3}\right), ..., \left(\frac{1}{3}, \frac{2}{3}\right)\}$
  - $\omega_2 = \{k_2; \left(\frac{1}{3}, \frac{2}{3}\right), ..., \left(\frac{1}{3}, \frac{2}{3}\right)\}$

- **Jeux à information incomplète d'un côté**

  - $N = \{I, II\}$
  - $K = \{k_1, k_2\}$
  - $Y = \{\omega_1, \omega_2\}$
  - $\omega_1 = \{k_1; (1, 0), (p, 1-p)\}$
  - $\omega_2 = \{k_2; (0, 1), (p, 1-p)\}$

  Ainsi, l'état de la nature est soit k_1 (avec probabilité p), soit k_2 (avec probabilité $1-p$). L'état de la nature est connu seulement par le joueur I mais pas par II. Dans la description à la Harsanyi, le joueur I a deux types $(1, 0)$ et $(0, 1)$. Le joueur II a un seul type $(p, 1-p)$.

- **Jeux à information incomplète d'un côté et demi**

  - $N = \{I, II\}$
  - $K = \{k_1, k_2\}$
  - $Y = \{\omega_1, \omega_2, \omega_3\}$
  - $\omega_1 = \{k_1; \left(\frac{1}{2}, \frac{1}{2}, 0\right), (1, 0, 0)\}$
  - $\omega_2 = \{k_1; \left(\frac{1}{2}, \frac{1}{2}, 0\right), \left(0, \frac{1}{3}, \frac{2}{3}\right)\}$

- $\omega_3 = \{k_2; (0,0,1), \left(0, \frac{1}{3}, \frac{2}{3}\right)\}$

Dans cet exemple, les deux types du joueur I ($t_{I_1} = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$ et $t_{I_2} = (0,0,1)$) connaissent l'état de la nature. En effet, si le joueur I est de type t_{I_1}, il sait que l'état de la nature est k_1 et vice versa, dans tout état du monde où l'état est k_1 le joueur I le sait. L'incertitude du joueur I repose seulement sur l'information qu'a le joueur II.

Ici, le joueur II n'est pas informé de l'état de la nature dans les états du monde ω_2 et ω_3. Dans ces deux états, II croit que k_1 a la probabilité $\frac{1}{3}$ et que k_2 a la probabilité $\frac{2}{3}$ et dans ces deux états, le joueur I le sait. Par contre dans l'état du monde ω_1 le joueur II est informé de l'état de la nature mais le joueur I ne le sait pas. Le joueur I attribue une probabilité $\frac{1}{2}$ à ce que II ne sache pas avec exactitude l'état de la nature. Ce modèle peut être décrit comme un jeu à la Harsanyi (un modèle standard). La nature tire au hasard avec probabilité $\frac{1}{3}$ l'état k_1 et avec probabilité $\frac{2}{3}$ l'état k_2. Le joueur I est informé de l'état. Si l'état est k_1 alors avec probabilité $\frac{1}{2}$ le joueur II est informé. Sinon, II n'est pas informé. La description du jeu est une connaissance commune.

Remarquez que dans le cas fini, la consistance se traduit de la manière suivante. Dans tout état du monde ω tel que la probabilité de ω pour le type t_i du joueur i est non nulle, le type du joueur i dans l'état ω doit précisément être t_i. Ainsi, dans cet exemple, puisque le type t_{I_1} attribue une probabilité positive à ω_1 et à ω_2, la condition de cohérence implique que le type du joueur I dans ω_1 (et ω_2) est t_{I_1}.

- **Jeux à information incomplète des deux côtés : le cas indépendant**
  - $N = \{I, II\}$
  - $K = \{k_{11}, k_{12}, k_{21}, k_{22}\}$
  - $Y = \{\omega_{11}, \omega_{12}, \omega_{21}, \omega_{22}\}$
  - $\omega_{11} = \{k_{11}; \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right), \left(\frac{1}{3}, 0, \frac{2}{3}, 0\right)\}$
  - $\omega_{12} = \{k_{12}; \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right), \left(0, \frac{1}{3}, 0, \frac{2}{3}\right)\}$
  - $\omega_{21} = \{k_{21}; \left(0, 0, \frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{3}, 0, \frac{2}{3}, 0\right)\}$
  - $\omega_{22} = \{k_{22}; \left(0, 0, \frac{1}{2}, \frac{1}{2}\right), \left(0, \frac{1}{3}, 0, \frac{2}{3}\right)\}$

L'ensemble des types du joueur I est $T_I = \{t_{I_1} = \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right), t_{I_2} = \left(0, 0, \frac{1}{2}, \frac{1}{2}\right)\}$ et celui du joueur II est $T_{II} = \{t_{II_1} = \left(\frac{1}{3}, 0, \frac{2}{3}, 0\right), t_{II_2} = \left(0, \frac{1}{3}, 0, \frac{2}{3}\right)\}$. Quand le type du joueur I est t_{I_1} il croit qu'avec probabilité $\frac{1}{2}$ l'état du monde est ω_{11} (et que donc le type du joueur II est t_{II_1}) et il croit qu'avec probabilité $\frac{1}{2}$ l'état du monde est ω_{12} (et que donc le type du joueur II est t_{II_2}). Quand le type du joueur I est t_{I_2} il a la même croyance sur le type du joueur II. Dans cet exemple, la croyance de I sur le type de II est indépendante de son type. De même, quelque soit son type, le joueur II croit qu'avec probabilité $\frac{1}{3}$ le type de I est t_{I_1} et qu'avec probabilité $\frac{2}{3}$ son type est t_{I_2}. Ceci est en fait un jeu à la Harsanyi (un modèle standard) où la probabilité objective, p, est une probabilité produit (la probabilité jointe sur les types est le produit des marginales) : $p = \left(\frac{1}{2}, \frac{1}{2}\right) \otimes \left(\frac{1}{3}, \frac{2}{3}\right)$. Plus précisément, avec probabilité $\left(\frac{1}{2}, \frac{1}{2}\right)$ la nature tire au hasard un type pour le joueur I, et en même temps et indépendamment, la nature tire au hasard un type pour le joueur II. Chaque joueur est informé de son propre type mais pas du type de l'autre. La description est une connaissance commune.

- **Jeux à information incomplète des deux côtés : le cas dépendant**
  - $N = \{I, II\}$
  - $K = \{k_{11}, k_{12}, k_{21}, k_{22}\}$
  - $Y = \{\omega_{11}, \omega_{12}, \omega_{21}, \omega_{22}\}$
  - $\omega_{11} = \{k_{11}; \left(\frac{3}{7}, \frac{4}{7}, 0, 0\right), \left(\frac{3}{5}, 0, \frac{2}{5}, 0\right)\}$
  - $\omega_{12} = \{k_{12}; \left(\frac{3}{7}, \frac{4}{7}, 0, 0\right), \left(0, \frac{4}{5}, 0, \frac{1}{5}\right)\}$
  - $\omega_{21} = \{k_{21}; \left(0, 0, \frac{2}{3}, \frac{1}{3}\right), \left(\frac{3}{5}, 0, \frac{2}{5}, 0\right)\}$
  - $\omega_{22} = \{k_{22}; \left(0, 0, \frac{2}{3}, \frac{1}{3}\right), \left(0, \frac{4}{5}, 0, \frac{1}{5}\right)\}$

L'ensemble des types du joueur I est $T_I = \{t_{I_1} = \left(\frac{3}{7}, \frac{4}{7}, 0, 0\right), t_{I_2} = \left(0, 0, \frac{2}{3}, \frac{1}{3}\right)\}$ et celui du joueur II est $T_{II} = \{t_{II_1} = \left(\frac{3}{5}, 0, \frac{2}{5}, 0\right), t_{II_2} = \left(0, \frac{4}{5}, 0, \frac{1}{5}\right)\}$. Quand le type du joueur I est t_{I_1} il croit qu'avec probabilité $\frac{3}{7}$ le type du joueur II est t_{II_1} et qu'avec probabilité $\frac{4}{7}$ le type du joueur II est t_{II_2}. Par contre, quand le type du joueur I est t_{I_2}, sa croyance est différente : il croit qu'avec probabilité $\frac{2}{3}$ le type du joueur II est t_{II_1} et qu'avec probabilité $\frac{1}{3}$ le type du joueur II est t_{II_2}. Ici la croyance d'un joueur sur le type de l'autre dépend du type. Cet exemple peut aussi être décrit comme un jeu à la Harsanyi standard avec comme distribution de probabilité objective sur les types, p, la suivante

$$p = \begin{array}{c c} & \begin{array}{cc} t_{II_1} & t_{II_2} \end{array} \\ \begin{array}{c} t_{I_1} \\ \\ t_{I_2} \end{array} & \begin{bmatrix} \frac{3}{10} & \frac{4}{10} \\ & \\ \frac{2}{10} & \frac{1}{10} \end{bmatrix} \end{array}$$

En effet, quand par exemple le joueur I est informé que son type est t_{I_1}, il croit qu'avec probabilité $\frac{\frac{3}{10}}{\frac{3}{10}+\frac{4}{10}} = \frac{3}{7}$, le type du joueur II est t_{II_1}. Si le joueur II sait que son type est t_{II_1}, il croit qu'avec probabilité $\frac{\frac{3}{10}}{\frac{3}{10}+\frac{2}{10}} = \frac{3}{5}$ le type de I est t_{I_1}.

Propriété: Remarquons que pour n'importe quel état du monde ω, chaque joueur est capable de retrouver Y et de calculer la probabilité objective p. En effet, supposons que l'état du monde est ω_{12}. Le premier ordre de croyances du joueur I invoque les états $\{\omega_{11}, \omega_{12}\}$. Le premier ordre de croyances du joueur II dans l'état du monde ω_{12} invoque $\{\omega_{12}, \omega_{22}\}$. Le second ordre de croyances des deux joueurs invoque Y en entier. Partant de Y on reste toujours dans Y. Y est donc un sous espace fermé de croyances qui peut être considéré comme une connaissance commune entre les joueurs (chaque joueur peut le reconstituer et sait que les autres le peuvent aussi et ce quelque soit le joueur et quelque soit son type). En plus, partant de son type, chaque joueur peut calculer la probabilité p. Cette propriété est vraie pour tout espace de croyances qui peut être décrit comme un jeu de Harsanyi standard (c'est-à-dire possédant une probabilité objective).

- **Croyances inconsistantes**

  On considère le même exemple que le précédent en remplaçant le type $t_{II_2} = \left(0, \frac{4}{5}, 0, \frac{1}{5}\right)$ du joueur II par $t_{II_2} = \left(0, \frac{1}{2}, 0, \frac{1}{2}\right)$. Il est facile de vérifier qu'il est impossible de trouver une distribution de probabilité jointe sur les types p pour laquelle cet exemple dérive d'un jeu de Harsanyi standard.

  Propriété : Ceci permet de voir que les modèles consistants ont une mesure nulle dans l'espace des croyances de Mertens-Zamir car il suffit de perturber un peut un modèle consistant pour aboutir à modèle inconsistant.

  Propriété: Cependant, on peut toujours écrire cet exemple comme un modèle de Harsanyi universel inconsistant (qui n'est pas un jeu). En effet chaque type de joueur calculer (et donc possède) une distribution (dépendant de son type) sur les types possibles des autres joueurs. Ceci implique donc le résultat suivant :

  Théorème 20 *Supposons que chaque état de la nature correspond à un jeu sous forme normale et à information parfaite. Dans ce cas, tout modèle de Mertens-Zamir peut s'écrire comme un modèle bayésien universel à la Harsanyi. En particulier, si le sous espace de croyances Y est fini et que dans tout état de la nature, le jeu sous forme normale est fini, un équilibre bayésien en stratégies mixtes existe.*

- **Y est une connaissance commune**

  Dans le dernier exemple chaque joueur peut et chaque joueur sait que les autres peuvent (etc) reconstituer Y, et ce partant de n'importe quel état du monde ω. Y est donc ici une connaissance commune entre les joueurs. Les joueurs savent chacun que le modèle est inconsistant : c'est-à-dire que leurs probabilités sont subjectives et non objectives. Ils admettent donc qu'ils puissent être en désaccord. Ceci contraste avec le célèbre article de Aumann ¨agree to disagree" (être d'accord pour ne pas être d'accord) dans lequel il considère un modèle bayésien consistant.

- **Y n'est pas une connaissance commune**

  – $N = \{I, II\}$

- $K = \{k_1, k_2, k_3\}$
- $Y = \{\omega_1, \omega_2, \omega_3\}$
- $\omega_1 = \{k_1; (1,0,0), (0,1,0)\}$
- $\omega_2 = \{k_2; (1,0,0), (0,1,0)\}$
- $\omega_3 = \{k_3; (1,0,0), (0,1,0)\}$

Dans tout état du monde, le joueur I croit que l'état de la nature est k_1 et croit que le joueur II croit que l'état de la nature est k_2. Le joueur II croit que l'état de la nature est k_2 et croit que I croit que l'état de la nature est k_1. Mais, si l'état du monde est ω_1 ou ω_2, aucun joueur ne peut retrouver ω_3.

Propriété : Cependant, on peut toujours utiliser cette notion de croyance pour dériver ce que chaque joueur i croit être une connaissance commune. Ici, partant de n'importe quel état du monde, ce que les joueurs considèrent comme étant une connaissance commune est l'ensemble $\{\omega_1, \omega_2\}$. En particulier, si l'état du monde est ω_3, les deux joueurs considèrent que l'évènement $\{\omega_1, \omega_2\}$ est (subjectivement) certain.

- **Information secrète**

Parfois, il est possible que dans certains états du monde, les joueurs ne soient pas d'accord sur l'ensemble des états du monde considérés comme connaissance commune. Ceci est illustré à l'aide de l'exemple suivant :

- $N = \{I, II\}$
- $K = \{k_1, k_2\}$
- $Y = \{\omega_1, \omega_2\}$
- $\omega_1 = \{k_1; (1,0), (1,0)\}$
- $\omega_2 = \{k_2; (1,0), (0,1)\}$

Dans ce modèle, le joueur II est toujours au courant de l'état du monde. Maintenant, si l'état du monde est ω_2, le type du joueur I est $(1,0)$. I est donc convaincu que l'état du monde est ω_1. Dans ω_1, les deux joueurs connaissent l'état de la nature et cette information est une connaissance commune entre eux. Partant donc de ω_2, le joueur I croit qu'il est connaissance commune que l'état du monde est ω_1. Cependant, si l'état du monde est ω_2 le joueur II le sait mais il sait aussi que l'autre joueur ne le sait pas et qu'il croit que l'état du monde commun est ω_1. k_2 constitue donc une information secrète pour le joueur II.

Part III

Jeux coopératifs

Chapter 15

Marchandage à deux joueurs

15.1 Introduction

Le concept de coopération est très important en théorie des jeux mais reste difficile à cerner. Coopérer veut dire agir ensemble dans un intérêt commun. Cependant, pour que deux joueurs ou plus agissent ensemble pour un intérêt commun, il est nécessaire de se séparer des utilités individuelles pour définir une sorte d'utilité commune qui va déterminer leur comportement commun. Ceci est en contradiction avec l'hypothèse essentielle suivant laquelle le but d'un agent est de maximiser son espérance d'utilité étant donné sa croyance. On a donc besoin de modéliser les comportements de coopération sans violer les fondements de la théorie des jeux.

Nash (1951) puis d'autres chercheurs ont permi de mettre en évidence que la coopération peut être étudiée avec les mêmes concepts d'équilibre étudiés dans les chapitres précédents. En effet, un profil de stratégies coopératives peut être obtenu comme un équilibre sous jeux parfait de Nash d'un processus de marchandage entre les joueurs.

Une transformation coopérative est une application qui associe à un jeu Γ (sous forme extensive, normale ou bayésienne) un autre jeu $\Psi(\Gamma)$ dans lequel chaque joueur possède, en plus des options existantes, d'autres options pour marchander avec les autres joueurs. Celles-ci pouvent prendre la forme d'un contrat signé entre les joueurs ou résulter d'un processus de communication... Par exemple, supposons que dans le dilemme du prisonnier, un intervenant extérieur (par exemple un avocat appelé généralement un arbitre) présente le contrat suivant aux joueurs : ¨moi, le signataire, promet de coopérer si ce contrat est signé par l'autre joueur aussi, sinon je trahirai". En supposant que les joueurs sont obligés de respecter leurs engagements contractuels sous peine de poursuites judiciaires, signer un contrat est un équilibre du jeu ainsi modifié, avec comme résultat : une coopération entre les joueurs.

Dans des situations réelles où des joueurs peuvent réellement contracter, il y a en général une multitude de contrats possibles. En plus, le choix d'un de ces contrats se fait à la fin d'un long processus de marchandage. Maintenant si un arbitre il y a (c'est-à-dire une sorte d'autorité extérieure qui à chaque jeu associe un contrat), celui-ci devrait être impartial. Il devrait baser sa sélection seulement sur des critères objectifs et équitables. C'est l'approche normative.

15.2 La solution normative de Nash

La théorie du marchandage a été introduite par Nash en (1950, 1953). Sa formulation pour le problème de marchandage à deux joueurs contient l'hypothèse suivante. Si deux joueurs négocient, ou encore un intervenant extérieur impartial arbitre, le paiement que les joueurs reçoivent doit dépendre seulement des paiements joints possibles pour les deux joueurs et de ce qu'un joueur reçoit si la négociation n'aboutit pas (en cas de désaccord). La justification est la suivante : pour l'équité, le gain d'un joueur par rapport à ce qu'il aurait eu sans accord devrait être comparable à ce que l'autre joueur gagne par rapport à ce qu'il aurait eu sans accord.

Le problème de marchandage consiste en une paire (F, v) où F est un ensemble convexe fermé de R^2 et $v = (v_1, v_2)$ est un vecteur de R^2. On suppose en plus que $FIR := F \cap \{(x_1,x_2) \in R^2 : x_1 \geq v_1, x_2 \geq v_2\}$ est non vide et borné.

F représente l'ensemble de tous les paiements (ou allocations) réalisables et v représente le vecteur de paiement en cas de désaccord (v_1 pour le joueur 1 et v_2 pour le joueur 2). La convexité de F est justifiée pas le fait que les joueurs peuvent se mettre d'accord de tirer au hasard entre deux allocations réalisables (d'une manière corrélée avec l'intervention peut être d'un arbitre) puis de mettre en place l'allocation qui résulte du tirage aléatoire. La non vacuité de FIR implique que l'ensemble des paiements réalisables contient des paiements au moins aussi bien que le point de désaccord (car sinon, au moins un des joueurs n'a pas intérêt à conclure un accord). La bornitude de FIR traduit l'idée selon laquelle il n'existe pas de paiement infiniment plus élevé que le point de désaccord.

Définition : Le problème de marchandage est dit *essentiel* s'il existe au moins une allocation $(x_1, x_2) \in F$ qui soit strictement préférée pas les deux joueurs au point de désaccord (i.e. $x_1 > v_1$ et $x_2 > v_2$).

Nash a d'abord traité ce problème d'une manière normative. Il a formulé une liste de propriétés raisonnables que la solution du problème de marchandage devrait satisfaire. Plus précisément, une solution est une application Φ qui associe à chaque problème de marchandage (F, v) un vecteur de paiement $\Phi(F, v) = (\Phi_1(F, v), \Phi_2(F, v))$. Si x et y sont deux vecteurs dans R^2, on notera $x \geq y$ si $x_1 \geq y_1$ et $x_2 \geq y_2$ et on notera $x > y$ si $x_1 > y_1$ et $x_2 > y_2$.

Les axiomes pour la solution sont les suivants:

(A1) **Pareto optimalité forte**

$\Phi(F, v) \in F$ et $\forall x \in F$, si $x \geq \Phi(F, v)$ alors $x = \Phi(F, v)$.

(A2) **Rationalité individuelle**

$\Phi(F, v) \geq v$.

(A3) **Invariance par translation et changement d'échelle**

Pour tous nombres réels $\alpha_1 > 0$, $\alpha_2 > 0$, β_1 et β_2 :

si $G = \{(\alpha_1 x_1 + \beta_1, \alpha_2 x_2 + \beta_2) : (x_1, x_2) \in F\}$ et $w = (\alpha_1 v_1 + \beta_1, \alpha_2 v_2 + \beta_2)$

alors $\Phi(G, w) = (\alpha_1 \Phi_1(F, v) + \beta_1, \alpha_2 \Phi_2(F, v) + \beta_2)$.

(A4) **Indépendance par rapport aux alternatives non pertinentes**

Pour tout convexe fermé $G \subset F$, si $\Phi(F, v) \in G$ alors $\Phi(G, v) = \Phi(F, v)$.

(A5) **Symétrie**

Si $v_1 = v_2$ et si $\{(x_2, x_1) : (x_1, x_2) \in F\} = F$ alors $\Phi_1(F, v) = \Phi_2(F, v)$.

L'axiome A1 exige que la solution d'un problème de marchandage doit être réalisable et Pareto-optimale: il ne doit y avoir aucune autre alternative qui est strictement préférée pour un des joueurs et faiblement préférée pour l'autre. L'axiome A2 exige qu'aucun joueur ne doit recevoir moins que son paiement de désaccord (car sinon, pourquoi accepterait-il un accord ?). A3 exige que si on modifie les utilités par une transformation linéaire croissante, la solution du problème transformé devrait résulter de la précédente solution par la même transformation. Cela implique donc que le résultat de la négociation ne dépend pas de la représentation de la fonction d'utilité vNM des joueurs (avec l'interprétation que les paiements dans F sont les utilités de vNM associées à des alternatives, des actions ou des stratégies comme produire, vendre, réduire les coûts, baisser le prix, etc). L'axiome A4 exige que si on élimine des alternatives qui n'auraient pas été choisies dans le problème de marchandage (F, v) (sans modifier le point de désaccord) cela ne devrait pas affecter la solution. Enfin, A5 exige que si les positions des joueurs sont complètement symétriques, la solution devrait être symétrique aussi (les joueurs doivent être traités équitablement). Le résultat remarquable de Nash est qu'il existe une et une seule solution satisfaisant tous ces axiomes. Elle est appelée la solution de Nash pour le problème de marchandage (en anglais : the Nash barganing solution).

Théorème 21 *Il existe une unique solution $\Phi(F, v)$ satisfaisant les axiomes A1 à A5. Cette solution est l'unique vecteur Pareto-optimal qui satisfait :*

$$\Phi(F, v) \in \arg \max_{x \in F, x \geq v} (x_1 - v_1)(x_2 - v_2)$$

Quand le problème de marchandage est essentiel, la solution individuellement rationnelle qui maximise le produit de Nash : $(x_1 - v_1)(x_2 - v_2)$ est unique.

Preuve. Fixons un problème de marchandage **essentiel** (F, v). Il existe alors un point $x \in F$ tel que $x > v$. Soit alors (y_1, y_2) **l'unique** solution du produit de Nash. En effet, cette solution **existe** par compacité et non vacuité de FIR et par continuité de la fonction $(x_1, x_2) \to (x_1 - v_1)(x_2 - v_2)$.

$y > v$ car sinon, le produit de Nash serait nul (ce qui est contradictoire avec la propriété du maximum). Enfin, elle est **unique** car maximiser le produit de Nash est équivalent à maximiser son logarithme, qui est une fonction strictement concave dans R^2. Posons :

$$\alpha_i = \frac{1}{y_i - v_i}, \qquad \beta_i = -\frac{v_i}{y_i - v_i}, \qquad i = 1, 2$$

Soit (G, w) le problème obtenu par la transformation linéaire croissante de (F, v) associé à $(\alpha_i, \beta_i)_{i=1,2}$ (comme dans l'axiome A3). Dans ce cas, $w = (0, 0)$. Puisque y maximise le produit de Nash du problème (F, v), il est facile de vérifier que son image, z, par la transformation linéaire : $z = (1, 1) = (\alpha_1 y_1 + \beta_1, \alpha_2 y_2 + \beta_2)$ maximise le produit de Nash $x_1 \times x_2$ du problème (G, w).

Puisque l'hyperbole $\{x : x_1 x_2 = 1\}$ doit être tangente (par optimalité) au convexe G au point $z = (1, 1)$, on conclut que la droite $\{x : x_1 + x_2 = 2\}$ (la droite tangente à l'hyperbole) est aussi la tangente à G au point z. D'où, par le théorème de séparation et la convexité de G, $G \subset E := \{(x_1, x_2) : x_1 + x_2 \leq 2\}$.

Quelle est la solution de Nash associée à (G, w) ? Pour satisfaire A1 et A5, on doit avoir $\Phi(E, w) = (1, 1)$. Pour A4, il est nécessaire que $\Phi(G, w) = (1, 1) = z$.

Revenons alors au problème (F, v). Pour satisfaire A3, et puisque (F,v) n'est autre que la transformation affine croissante de (G, w), il est nécessaire que $\Phi(F, v) = y$. Ainsi, pour satisfaire les axiomes de Nash, Φ doit sélectionner l'allocation qui maximise le produit de Nash, parmi toutes les allocations réalisables et individuellement rationnelles.

Supposons maintenant que (F, v) n'est pas essentiel. Il n'existe donc aucun point $x \in F$ tel que $x > v$. Par convexité, il existe au moins un joueur i tel que pour tout $x \in F$, si $x \geq v$ alors $x_i = v_i$. Soit y l'allocation dans F qui est la mieux pour l'autre joueur, sous la contrainte que $y_i = v_i$. y est donc l'unique vecteur qui satisfait A1 et A2. Ainsi, pour satisfaire A1 et A2, on doit avoir $\Phi(F, v) = y$. Evidement y atteint le maximum pour le produit de Nash qui par ailleurs est égal à 0.

On a donc montré que les cinq axiomes ne peuvent être satisfait que par une unique solution. Il est facile de vérifier que, inversement, cette solution satisfait les cinq axiomes. ■

15.3 Autres solutions, le pouvoir de négociation

L'Axiome de symétrie A5 a été utilisé dans la preuve du théorème précédent, seulement pour prouver que : quand $E = \{z : z_1 + z_2 \leq 2\}$, la solution du problème de marchandage $(E, (0, 0))$ devrait être $(1, 1)$. On peut enlever cet axiome et le remplacer le par l'axiome suivant (avec $\alpha + \beta = 2$) :

A5(α, β) **Dissymétrie**

La solution du problème de référence $(E, (0, 0))$ est le point de la frontière efficiente (α, β).

Dans ce cas, la même preuve adaptée montre que l'unique solution du problème de marchandage (F, v), en remplaçant A5 par $A5_{(\alpha,\beta)}$, est le vecteur individuellement rationnel et Pareto-Optimal qui maximise le produit de Nash généralisé :

$$(x_1 - v_1)^{\alpha} \times (x_2 - v_2)^{\beta}$$

Cette solution, $\Phi_{(\alpha,\beta)}$ est appelée, la solution de Nash non symétrique.

Pour comprendre comment cette solution peut être utilisée, supposons que l'état doive verser une sorte de prime (ou allocation) de 1000 Euros, à diviser entre deux familles. La famille 1 contient 4 personnes alors que la famille 2 est constituée seulement d'un couple de deux personnes. Une manière équitable serait donc de donner $\frac{4}{6} \times 1000$ Euros à la famille nombreuse et $\frac{2}{6} \times 1000$ Euros à la seconde famille. Dans ce problème de partage, il est donc naturel de prendre $\frac{\alpha}{\beta} = \frac{4}{2} = 2$ car la première famille pèse deux fois plus que l'autre.

En général, le rapport $\frac{\alpha}{\beta}$ mesure le pouvoir de négociation (le rapport de pouvoir) entre les deux joueurs. Plus il est grand, plus le premier joueur a du poids par rapport à l'autre. Quand ce rapport tend vers ∞, le premier joueur est le seul décideur (une sorte de dictateur), quand celui-ci tend vers 0, c'est l'autre qui a tous les pouvoirs.

15.4 Menaces rationnelles

Il est facile de vérifier que à F et v_1 fixés, plus le paiement de désaccord du joueur II diminue, plus le paiement attribué par la solution de Nash au joueur I augmente : la possibilité d'infliger une plus grande perte au joueur II en cas de désaccord augmente le paiement d'un joueur. Le point de désaccord est, en quelque sorte, une variable stratégique. Chacun essaie de se présenter, avant le début de la négociation, le mieux possible.

Soit $\Gamma = (\{1,2\}, (S_1, S_2), (r_1, r_2))$ un jeu sous forme stratégique et soit F l'ensemble des paiements réalisables avec des stratégies mixtes corrélées : $F = \{(r_1(\mu), r_2(\mu)) : \mu \in \Delta(S_1 \times S_2)\}$. Supposons que les joueurs ont la possibilité de négocier afin de coopérer dans ce jeu (concrètement, à la fin de la négociation, ils signent un contrat les obligeant à respecter le compromis). Avant de rentrer dans la négociation, chaque joueur i doit choisir une stratégie de menace $\sigma_i \in \Delta(S_i)$; ce sera sa stratégie de désaccord. $v_i = r_i(\sigma_1, \sigma_2)$ est donc le paiement espéré du joueur i en cas de désaccord. Pour que les stratégies de désaccord soient crédibles, il faudrait aussi que (σ_1, σ_2) forme un équilibre de Nash du jeu Γ, sinon, au moins un des joueurs a intérêt à dévier, ceci peut se faire par la signature d'un contrat avant le début de la négociation qui oblige chaque joueur à respecter son engagement de jouer σ_i en cas de désaccord. Il résulte donc que le paiement de désaccord d'un joueur dépend de sa stratégie mais aussi de la stratégie de désaccord de l'autre. On est donc en présence d'une intéraction stratégique.

Supposons qu'à l'issue de la négociation, les joueurs finissent de toutes manières par coopèrer en choisissant une solution de coopération (par exemple la solution de Nash). Leur vecteur de paiement sera donc $w(\sigma_1, \sigma_2) = \Phi[F, (r_1(\sigma_1, \sigma_2), r_2(\sigma_1, \sigma_2))]$. Dans ce cas, avant le début de la négociation, chaque joueur i essayera de choisir sa stratégie de désaccord σ_i pour avoir le plus d'impact sur l'issue de la négociation. On dira donc que (σ_1, σ_2) forme une paire de menaces rationnelles si $\forall (\tau_1, \tau_2) \in \Delta(S_1) \times \Delta(S_2)$:

$$\begin{aligned} w_1(\sigma_1, \sigma_2) &\geq w_1(\tau_1, \sigma_2) \\ w_2(\sigma_1, \sigma_2) &\geq w_2(\sigma_1, \tau_2) \end{aligned}$$

Les menaces rationnelles sont tout simplement les équilibres de Nash du jeu sous forme normale $(\{1,2\}, (S_1, S_2), (w_1, w_2))$. L'existence d'un équilibre (en stratégies mixtes) peut être prouvé par le théorème de Kakutani, quand par exemple les ensembles de stratégies pures sont finis.

15.5 Un jeu de marchandage répété, lien avec les solutions normatives

Stahl (1972) et Rubinstein (1982) ont considéré le modèle de marchandage suivant. Deux joueurs proposent alternativement des offres jusqu'à ce qu'un des joueurs accepte. Plus précisément, si à l'étape t c'est le joueur $i \in \{1, 2\}$ qui est actif, celui-ci peut proposer n'importe quelle alternative $x \in F$. Si l'autre joueur accepte cette offre, le jeu s'arrête, et le vecteur de paiement est x. Sinon, si c'est le joueur 1 (resp. 2) qui a fait la proposition alors avec probabilité $p_1 \in]0, 1[$ (resp. $p_2 \in]0, 1[$) le jeu s'arrête sur un désaccord avec comme vecteur de paiement v et avec une probabilité $(1 - p_1)$ (resp. $(1 - p_2)$) le jeu continue encore une étape et c'est au tour de l'autre joueur de faire une offre.

p_1 modélise le degré d'impatience du joueur 1 ($(1 - p_1)$ peut être interprété comme le taux d'escompte pour le joueur 1). Rubinstein a montré que si le jeu de marchandage est régulier (voir ci dessous) alors il existe un unique équilibre sous jeux parfait. Binmore a montré que si $p_1 = p_2 \to 0$, le paiement d'équilibre converge vers la solution de Nash.

Définition 37 (F, v) *est régulier si et seulement s'il est essentiel et pour tout vecteur* $y \in F$:

- *si* $y_1 > v_1$ *alors* $\exists z \in F$ *tel que* $v_1 \leq z_1 < y_1$ *et* $z_2 > y_2$
- *si* $y_2 > v_2$ *alors* $\exists z \in F$ *tel que* $v_2 \leq z_2 < y_2$ *et* $z_1 > y_1$

En d'autres termes, un jeu est régulier si (1) il existe une allocation strictement préférée au paiement de désaccord pour les deux joueurs et (2) en toute allocation où un joueur obtient strictement plus que son paiement de désaccord, il a la possibilité de réduire strictement son propre paiement et d'augmenter strictement celui de l'autre (faites un dessin pour voir plus clair).

Pour un problème régulier (F, v) et pour tout $v_2 \leq z_2 \leq \max_{y \in FIR} y_2$, dénotons par $h_1(z_2)$ le paiement maximal possible que peut obtenir le joueur 1 : $h_1(z_2) = \max \{z_1 : (z_1, z_2) \in F\}$. Définissons de même $h_2(z_1)$. Ces deux fonctions dépendent évidemment de F et v.

Théorème 22 *Si le problème de marchandage est régulier alors le jeu de marchandage escompté admet un unique équilibre sous jeux parfait dans lequel le joueur 1 propose tout le temps l'allocation* x^**, le joueur 2 propose tout le temps* y^**, le joueur 1 accepte une offre si et seulement si elle lui offre un paiement supérieure ou égal à* y_1^* *et le joueur 2 accepte une offre si et seulement si elle lui procure un paiement supérieur ou égal à* x_2^**. Les vecteurs* x^* *et* y^* *sont fortement Pareto-optimaux et satisfont :*

$$\begin{aligned} y_1^* &= (1 - p_2)(x_1^* - v_1) + v_1 \\ x_1^* &= (1 - p_1)(y_2^* - v_2) + v_2 \end{aligned}$$

Quand $p_1 = p_2 \to 0$, x^* *et* y^* *convergent vers la solution de Nash.*

Preuve. Remarquez que tous les sous jeux dans lesquels le joueur 1 commence le jeu ont les mêmes équilibres sous jeux parfaits. Soit $\overline{x}_1$ le paiement maximal qu'obtient le joueur 1 dans un équilibre sous jeux parfait dans lequel il commence et soit $\underline{x}_1$ son paiement minimal. De même, définissons $\overline{y}_2$ et $\underline{y}_2$ comme étant respectivement les paiements maximal et minimal dans un équilibre sous jeux parfait quand le joueur 2 commence.

Dans tout équilibre sous jeux parfait, le joueur 2 accepte toute offre lui donnant plus que $(1 - p_1)\,\overline{y}_2 + p_1 v_2$ car c'est le paiement maximal qu'il peut espérer en rejetant une offre. Dans ce cas, dans tout équilibre sous jeux parfait, le joueur 1 ne peut pas faire moins bien que $h_1((1 - p_1)\,\overline{y}_2 + p_1 v_2)$. En plus, il existe un équilibre sous jeux parfait dans lequel le joueur 2 peut garantir, à ε près, le paiement $\overline{y_2}$ dans le sous jeu commençant à l'étape 2. Dans ce cas, le joueur 2 peut se garantir, dans au moins un équilibre sous jeux parfait, un paiement espéré au moins égal à $(1 - p_1)\,(\overline{y}_2 - \varepsilon) + p_1 v_2$ pour le jeu commençant à l'étape 1. Ainsi, le minimum que le joueur 1 peut espérer dans tout sous jeu où il joue en premier est $\underline{x}_1 = h_1((1 - p_1)\,\overline{y}_2 + p_1 v_2)$.

Un argument similaire montre que $\underline{y}_2 = h_2((1-p_2)\,\overline{x}_1 + p_2 v_1)$.

De même, le joueur 2 ne peut jamais accepter une offre lui donnant moins que $(1-p_1)\,\underline{y}_2 + p_1 v_2$ et donc, le joueur 1 ne peut pas espérer obtenir plus que $h_1((1-p_1)\,\underline{y}_2 + p_1 v_2)$. Puisqu'il existe au moins un sous jeu dans lequel ce paiement est possible pour le joueur 2, on conclut que $\overline{x}_1 = h_1((1-p_1)\,\underline{y}_2 + p_1 v_2)$.

Un argument similaire montre que $\overline{y}_2 = h_2((1-p_2)\,\underline{x}_1 + p_2 v_1)$.

Complétons la définition des vecteurs $\overline{x}$, $\overline{y}$, $\underline{x}$ et $\underline{y}$ en posant :

$$\overline{x}_2 = h_2(\overline{x}_1), \qquad \underline{x}_2 = h_2(\underline{x}_1), \qquad \overline{y}_1 = h_1(\overline{y}_2), \qquad \underline{y}_1 = h_1(\underline{y}_2)$$

Ces vecteurs sont donc tous Pareto-optimaux. La régularité du problème de marchandage et les équations précédentes impliquent donc que (prouvez le) :

$$\begin{cases} \underline{x}_2 = (1-p_1)\,\overline{y}_2 + p_1 v_2 \\ \underline{y}_1 = (1-p_2)\,\overline{x}_1 + p_2 v_1 \\ \overline{x}_2 = (1-p_1)\,\underline{y}_2 + p_1 v_2 \\ \overline{y}_1 = (1-p_2)\,\underline{x}_1 + p_2 v_1 \end{cases}$$

Par ailleurs, il est relativement facile de montrer (prouvez-le) qu'il existe une unique paire (x^*, y^*) individuellement rationnelle et Pareto-optimale dans F satisfaisant :

$$\begin{aligned} x_2^* &= (1-p_1)\,y_2^* + p_1 v_2 \\ y_1^* &= (1-p_2)\,x_1^* + p_2 v_1 \end{aligned}$$

On déduit donc que $\underline{x} = \overline{x} = x^*$ et que $\underline{y} = \overline{y} = y^*$. x^* est le paiement dans tout sous jeu où le joueur 1 commence et y^* est le paiement d'équilibre où dans tout sous jeu où le joueur 2 commence.

Supposons maintenant que $p_1 = p_2 = p$. On déduit donc que les deux vecteurs x^* et y^* ont le même produit de Nash :

$$\begin{aligned} (x_1^* - v_1) \times (x_2^* - v_2) &= (x_1^* - v_1) \times (1-p) \times (y_2^* - v_2) \\ &= (y_1^* - v_1) \times (y_2^* - v_2) \end{aligned}$$

d'où, les points x^* et y^* sont dans une même courbe d'indifférence pour l'hyperbole définie par le produit de Nash et appartiennent tous les deux à la frontière Pareto-optimale. Ils encerclent donc la solution de Nash (il suffit de faire un dessin pour s'en convaincre). Par ailleurs, les équations d'optimalité montrent que plus p est proche de 0, plus ces deux points sont proches. On déduit qu'ils convergent tous les deux vers la solution de Nash. ■

Analyse :

- Si $p = p_1 = p_2$, $x_1^* \geq y_1^*$ et $x_2^* \leq y_2^*$. Ceci peut s'interpréter de la manière suivante : si les joueurs ont le même degré de patience, il y a un avantage à commencer une négociation (à cause du coût associé à la valeur du temps qui passe). Cet avantage disparaît au fur et à mesure que les joueurs sont de plus en plus patients.

- L'équilibre précédent dépend du rapport $\frac{p_1}{p_2}$. En fait, ce rapport peut être déterminant dans la négociation, même si $|p_1 - p_2|$ est très petit. En effet, si $1 - p_1 = (1-\varepsilon)^\alpha$ et que $1 - p_2 = (1-\varepsilon)^\beta$, les équations d'optimalité dans le théorème précédent impliquent que:

$$\begin{aligned} (x_1^* - v_1)^\alpha \times (x_2^* - v_2)^\beta &= (x_1^* - v_1)^\alpha \times (1-\varepsilon)^{\alpha\times\beta} \times (y_2^* - v_2)^\beta \\ &= (y_1^* - v_1)^\alpha \times (y_2^* - v_2)^\beta \end{aligned}$$

  Ce qui prouve que les deux solutions x^* et y^* entourent la solution de Nash non symétrique. Quand $\varepsilon \to 0$, les deux solutions convergent vers la solution de Nash non symétrique de paramètres (α, β). Remarquez que pour $\varepsilon \approx 0$, $p_1 \approx \alpha\varepsilon$, $p_2 \approx \beta\varepsilon$ et que $\frac{\alpha}{\beta} \approx \frac{p_1}{p_2}$. Cette analyse montre qu'un bon négociateur a intérêt de faire croire à l'autre qu'il y a une probabilité relativement élevée que la négociation s'arrête sur un désaccord si l'autre joueur refuse son offre et qu'au contraire il existe une probabilité suffisamment élevée que le marchandage continue si c'est lui qui refuse l'offre.

Chapter 16

Coalitions dans les jeux coopératifs à utilité transférable

Dans le chapitre précédent, on a analysé le marchandage et la coopération pour deux joueurs. Il est facile de généraliser le problème de Marchandage et la solution de Nash aux jeux avec plus de deux joueurs. Si $\{1, ..., n\}$ est l'ensemble de joueurs, F est l'ensemble d'allocations réalisables et si v dénote le vecteur dans R^n des paiements de désaccord, alors la solution de Nash pour un problème de marchandage est définie comme étant l'unique vecteur réalisable dans F, Pareto-optimal et individuellement rationnel qui maximise le produit de Nash $\prod_{i=1}^{n} (x_i - v_i)$. Il n'est pas difficile de vérifier que cette solution satisfait une version généralisée des axiomes de Nash. Cependant, quand $n \geq 3$, cette solution est rarement utilisée dans l'analyse des jeux coopératifs. En effet, celle-ci ne prend pas en considération la possibilité qu'un sous-groupe de joueurs (appelé coalition) coopèrent entre eux.

Par exemple, soit un jeu sous forme normale où trois joueurs $\{1, 2, 3\}$ ont le même ensemble de stratégies $S^i = \{(x_1, x_2, x_3) : x_i \geq 0, x_1 + x_2 + x_3 \leq 300\}$. Supposons que, simultanément, chaque joueur i propose une allocation $s^i \in S^i$. On va considérer deux situations :

Jeu 1 si les trois allocations sont différentes, chaque joueur reçoit 0, et si elles sont toutes égales à s, celle-ci est réalisée (le joueur i reçoit s_i)

Jeu 2 si les joueurs 1 et 2 proposent la même allocation s, elle est réalisée et sinon tout le monde reçoit 0.

Dans les deux jeux, la solution de Nash pour le problème de Marchandage est $(100, 100, 100)$ (car l'ensemble des allocations réalisables est symétrique et chaque joueur ne peut garantir, tout seul, que le paiement 0). Cependant, dans le deuxième cas, cette allocation n'est pas raisonnable car le joueur 3 n'a en fait aucun pouvoir de décision, les joueurs 1 et 2 suffisent à eux seul pour déterminer une allocation. Il est plus raisonnable pour eux de prendre chacun un paiement de 150 et ne rien laisser au joueur 3. Si les joueurs 1 et 2 ignorent le joueur 3 et font comme s'ils étaient dans un jeu de marchandage à deux joueurs, la solution de Nash leur recommande justement 150 chacun. Pour cela, il est nécessaire que les joueurs puissent vraiment négocier (discuter) car si dans le jeu 2, les joueurs ne peuvent pas communiquer l'allocation (100,100,100) ne peut pas être rejetée car elle constitue un équilibre de Nash. Si le joueur 1 pense que le joueur 2 va proposer un partage équitable de 100 chacun, il a intérêt à faire de même.

Ce qui sépare donc les jeux coopératifs des jeux non coopératifs est la possibilité de communication, de marchandage. Il est souhaitable alors qu'à la fin des négociations, il n'y ait pas de possibilité pour une coalition de rejeter ou d'inclure un individu pour obtenir plus. C'est le concept de coeur. Cette condition n'est pas toujours possible. En effet, considérons la variante suivante du jeu précédent :

Jeu 3 si au moins deux joueurs (n'importe lesquels) proposent une même allocation, celle-ci est réalisée, sinon tout le monde reçoit 0.

Dans le jeu 3, deux joueurs suffisent pour prendre une décision. Ce jeu est connu aussi sous le nom de jeu de majorité à trois joueurs. En effet, si une majorité de personnes est d'accord pour exécuter un plan, alors leur décision est adoptée. La symétrie du problème suggère qu'il est plus raisonnable de prédire l'allocation équitable. Cependant, celle-ci est instable. En effet, si les joueurs peuvent négocier, les joueurs 1 et 2 peuvent trouver cela profitable de se séparer de 3 et proposer ensemble l'allocation (150,150,0). Dans ce cas, rien n'empêche le joueur 3 d'essayer de persuader le joueur 1 de trahir 2 et de proposer avec lui l'allocation (200,0,100). Dans ce cas, le joueur 2 va persuader le joueur 3 de s'accorder afin d'obtenir l'allocation (0,100,200)... On voit donc que dans tout équilibre, il y a toujours une paire de joueurs qui va essayer d'écarter l'autre joueur.

Une possibilité de s'en sortir serait de supposer que les accords ne sont pas définitifs (un peu comme dans un forum). Les joueurs font le tour des coalitions déjà en place, chaque fois qu'un joueur rejoint une coalition, il lui apporte de la valeur ajoutée (qui peut être négatives)... à la fin de la journée, chacun se fait une idée moyenne de sa valeur ajouté : c'est l'idée derrière la valeur de Shapley.

Une autre possibilité pour éliminer le problème du jeu 3 est de supposer qu'une fois qu'un accord est conclu au sein d'une coalition, il n'y a plus la possibilité pour un joueur de re-conclure un autre accord avec une autre coalition qui ne contient pas tous les membres de sa coalition initiale. Dans ce cas, si la grande coalition se met d'accord sur l'allocation (100,100,100) aucun joueur ne peut modifier individuelement cet accord sans faire éclater la coalition. Aussi, si les joueurs 1 et 2 se mettent d'accord sur (150,150,0) le joueur 3, ne peut rien apporter de plus à la coalition jointe de 1 et 2 et puisqu'il ne peut négocier séparément avec un des joueurs 1 ou 2 ce plan est exécuté. On voit donc que l'ordre suivant lequel les coalitions peuvent se rencontrer et négocier est très important, avec un avantage pour les coalitions qui négocient plus tôt. Ceci peut être analysé par le biais de la structure coalitionnelle.

Ici, pour la simplicité, on ne traitera que les jeux à utilités transférables. Toutes les notions développées peuvent être généralisées aux jeux avec utilités non-transférables.

16.1 Utilité transférable et fonction caractéristique

La notion d'utilité transférable est très importante dans la théorie des jeux coopératifs. Ceci suppose que les utilités des joueurs ont une même unité (disons monétaire) c'est-à-dire que tous les joueurs évaluent les conséquences de leurs actes en utilisant la même échelle utilitaire.

Dans ce cas, on peut ajouter à chaque joueur la possibilité (dans ses stratégies) de transférer de la monnaie (de l'utilité) vers un autre joueur. A quoi cela peut-il servir ? Supposons qu'il existe une stratégie mauvaise pour le joueur 1 mais qui permet au joueur 2 de gagner beaucoup. Le joueur 2 peut alors inciter le joueur 1 à choisir cette stratégie quitte à lui transférer une partie de ces gains (de son utilité) pour le dédommager. Les joueurs ont donc intérêt à choisir les stratégies coopératives qui maximisent leurs gains additionnés puis de partager entre eux ce gain total, quitte à faire des transferts monétaires.

Supposons que chaque coalition C peut (espérer) obtenir (ou garantir) en coopérant un gain total égal à $v(C)$ (appelé le poids de la coalition C).

Définition 38 *Une fonction caractéristique (appelée un jeu coalitionnel) est une fonction qui associe à chaque coalition $C \subset \{1, ..., n\}$ une valeur $v(C) \in R$ avec la convention que $v(\emptyset) = 0$.*

Exemples :

- Dans le jeu 1, la fonction caractéristique est
  $v(\{1,2,3\}) = 300$, $v(\{1,2\}) = v(\{2,3\}) = v(\{1,3\}) = 0$ et $v(\{1\}) = v(\{2\}) = v(\{3\}) = 0$.
- Dans le jeu 2, la fonction caractéristique est
  $v(\{1,2,3\}) = 300 = v(\{1,2\})$, $v(\{2,3\}) = v(\{1,3\}) = 0$ et $v(\{1\}) = v(\{2\}) = v(\{3\}) = 0$.

- Dans le jeu 3, la fonction caractéristique est
  $v(\{1,2,3\}) = 300 = v(\{1,2\}) = v(\{2,3\}) = v(\{1,3\})$ et $v(\{1\}) = v(\{2\}) = v(\{3\}) = 0$.

16.2 Approche stratégique, le concept de coeur

Soit un jeu coalitionnel avec utilité transférable $\{v(C)\}_{C\subset N}$, où $N = \{1, ..., n\}$ est l'ensemble des joueurs. Un vecteur de paiement $x = (x_1, ..., x_n)$ (appelé une allocation) est un vecteur de R^n où x_i correspond au paiement reçu par le joueur i dans l'allocation x. On dit que l'allocation x est réalisable pour une coalition C si $\sum_{i\in C} x_i \leq v(C)$. Ainsi, les joueurs peuvent dans C, en coopérant, recevoir un gain total $v(C)$ qu'ils peuvent diviser entre eux en donnant à chaque joueur $i \in C$ le montant x_i. On dira qu'une allocation x est réalisable si elle est réalisable pour la grande coalition N.

Une coalition $C \neq \emptyset$ rejette l'allocation x si $\sum_{i\in C} x_i < v(C)$. En effet, en coopérant entre eux, les membre de C peuvent garantir strictement plus que $\sum_{i\in C} x_i$, il est donc irrationnel pour eux d'accepter x.

Définition 39 *Une allocation x est dans le coeur si elle est réalisable ($\sum_{i\in N} x_i = v(N)$) et s'il n'existe aucune coalition $C \neq \emptyset$ qui rejette x (il faut donc que $\forall C \subset N\backslash \{\emptyset\}$, $\sum_{i\in C} x_i \geq v(C)$)*

Ainsi, si x n'est pas dans le coeur, il existe une coalition de joueurs C qui va rejeter x, coopérer et obtenir strictement plus que $\sum_{i\in C} x$.

Exemples :

- Dans le jeu 1 où les joueurs obtiennent 300 seulement s'ils coopèrent tous, le coeur est
  $\left\{x = (x_1, x_2, x_3) \in R^3_+ : x_1 + x_2 + x_3 = 300\right\}$.
- Dans le jeu 2 où les joueurs 1 et 2 suffisent pour garantir 300, le coeur est
  $\left\{x = (x_1, x_2, x_3) \in R^3_+ : x_1 + x_2 = 300, x_3 = 0\right\}$.
- Dans le jeu 3 où deux joueurs suffisent pour garantir 300, le coeur est vide.

Ce concept très attractif, présente cependant beaucoup de défauts. Il est souvent très grand (exemple 1 et 2) ou vide (exemple 3). La vacuité du coeur indique cependant l'importance de la dynamique de la négociation (comme dans le jeu 3, voir l'introduction).

Le coeur présente aussi un autre défaut : l'instabilité et la possibilité de suggérer des solutions extrêmes. Par exemple, soit le jeu suivant :

Jeu 4 : Soit 2.000.001 joueurs. 1.000.000 d'entre eux, noté D, possèdent chacun une chaussure droite (une technologie d) toutes identiques et 1.000.001d'entre eux, noté G, possèdent chacun la chaussure gauche (une technologie g) toutes identiques. Chaque paire de chaussure (une gauche, et une droite) vaut 1 Euro. Le poids (la valeur) d'une coalition est définie par le nombre de paire de chaussure qu'elle peut constituer.

L'unique allocation du coeur du jeu 4 donne à chaque joueur dans D 1 Euro et donne 0 Euro pour chaque joueur dans G. Ainsi, la rareté relative (même minime) des joueurs dans D par rapport aux joueurs dans G, fait que le coeur leur donne tout le pouvoir de négociation. Cette prédiction du coeur est en plus très instable, car il suffit de rajouter deux malheureux petits joueurs dans D pour changer le résultat complètement en faveur des joueurs dans G (qui deviennent minoritaires). Dans cet exemple, cela paraît très inadapté car il est impossible d'imaginer ces 2.000.001 joueurs négocier et le fait qu'un type soit légèrement rare ne devrait pas influencer trop la solution. Etant donné que la probabilité qu'un joueur rencontre un complémentaire est à peu prés la même pour les deux groupes (avec un léger avantage pour les joueurs dans D), une solution raisonnable devrait prédire un gain légèrement supérieur à 0.5 pour les joueurs dans D et légèrement inférieur à 0.5 pour les joueurs dans G. C'est ce que prédit la valeur de Shapley.

16.3 Approche normative : la valeur de Shapley

Le fait que le coeur soit trop grand, vide ou instable diminue les possibilités de son application comme outil de prédiction. Le mieux qu'on puisse espérer c'est de définir un concept de solution qui, à tout jeu coalitionnel, associe une espérance de gain pour chaque joueur. C'est-à-dire une fonction qui, à chaque jeu coalitionnel v associe un vecteur de paiement $\Psi(v) \in R^n$ où $\Psi_i(v)$ est l'espérance de paiement du joueur i dans le jeu coalitionnel v.

Shapley (1953) a traité ce problème d'une manière axiomatique. Il a définit les bonnes propriétés qu'une solution devrait satisfaire, puis il a montré l'unicité et caractérisé la solution.

Une permutation est une bijection $\pi : N \to N$. Etant donné une permutation π et un jeu coalitionnel v, πv est le jeu coalitionnel obtenu à partir de v en permutant les noms des joueurs :

$$\pi v(C) = v\left(\pi^{-1}(C)\right) \forall \emptyset \neq C \subset N.$$

On dira qu'une coalition D est porteuse pour le jeu v si pour toute autre coalition C, $v(C \cap D) = v(C)$. Ainsi, si un joueur n'est pas dans D, il est inutile : il ne rajoute (et ne retire) rien en rejoignant une coalition. En plus, une coalition constituée seulement de joueurs inutiles a une valeur nulle.

Au lieu du troisième axiome d'additivité de Shapley (qu'on introduit pas ici) on considère ici un concept très proche : celui de la linéarité. A N fixé, et pour tous jeux coalitionnels v et w et pour tout nombre $p \in [0,1]$, soit $pv + (1-p)w$ le jeu coalitionnel où pour toute coalition C, $[pv + (1-p)w](C) = pv(C) + (1-p)w(C)$.

Voici les axiomes de Shapley :

(A1) **Symétrie**

Pour tout jeu coalitionnel v, toute permutation π et tout joueur i, $\Psi_i(\pi(v)) = \Psi_{\pi^{-1}(i)}(v)$.

(A2) **Coalition porteuse**

Pour tout v et toute coalition porteuse D, $\sum_{i \in R} \Psi_i(v) = v(D)$.

(A3) **Linéarité**

Pour tous jeux coalitionnels v et w, tout $p \in [0,1]$ et tout joueur i, $\Psi_i(pv + (1-p)w) = p\Psi_i(v) + (1-p)\Psi_i(w)$.

Le premier axiome de Shapley dit que seulement le rôle stratégique d'un joueur compte mais pas son nom. Le second axiome veut que seul les joueurs dans une coalition porteuse D se partage les gains (ou les pertes selon le contexte) total(e)s $v(D)$. Pour interpréter le troisième axiome, supposons que les joueurs doivent jouer demain soit le jeu coalitionnel v, soit le jeu coalitionnel w (dépendant de la réalisation d'un événement aléatoire exogène observable seulement demain comme par exemple a=il pleut ou b=il ne pleut pas). Soit p la probabilité de jouer v demain et $(1-p)$ celle de jouer w demain. Si les joueurs décident de marchander dès aujourd'hui sur ce qu'ils comptent faire demain, dans chaque état de la nature, ceci définit un jeu équivalent à $pv + (1-p)w$. Si maintenant $\Psi_i(z)$ est interprété comme étant l'espérance de gain du joueur i dans le jeu z, alors on devrait satisfaire l'axiome A3.

Théorème 23 *Il existe exactement une et une seule fonction Ψ satisfaisant les axiomes de Shapley A1, A2 et A3. Cette fonction est la suivante :*

$$\Psi_i(v) = \sum_{S \subset N \setminus \{i\}} \frac{|S|! \times (|N| - |S| - 1)!}{|N|!} [v(S \cup \{i\}) - v(S)]$$

Où, $|S|$ est le cardinal de S et $k! = 1 \times 2 \times ...k$ et $0! = 1$.

Cette formule peut être interprétée. Supposons qu'il est convenu que les joueurs vont rentrer successivement (un par un) dans une salle. Pour cela, ils sont appelés à faire la queue (ce qui définie une permutation (un ordre)). Il y a $|N|!$ permutations possibles. Maintenant, pour tout ensemble S ne contenant pas i, il y a $|S|! \times (|N| - |S| - 1)!$ permutations dans lesquelles les joueurs dans S se trouvent tous devant dans la queue, et après eux se trouve le joueur i. Si les permutations sont distribuées uniformément (aucun joueur ou coalition n'est prioritaire), $\frac{|S|! \times (|N|-|S|-1)!}{|N|!}$ est la probabilité selon laquelle, quand le joueur i rentre dans la salle, il se trouve devant la coalition S. Puisque sa contribution marginale au moment où il rentre dans la salle est $[v(S \cup \{i\}) - v(S)]$, la valeur de Shapley pour i est tout simplement l'espérance des contributions marginales du joueur i sous l'hypothèse d'équiprobabilité de formation des coalitions.

Preuve. Après l'interprétation probabiliste, il est facile de montrer que les axiomes de Shapley sont vérifiés pour la formule du théorème. Il reste à montrer l'unicité. L'axiome 2 implique que $\Psi_i(\mathbf{0}) = 0$, où $\mathbf{0}$ est le jeu coalitionnel identiquement nul pour toute coalition. Ceci et l'axiome 3 impliquent que Ψ est une application linéaire de l'espace des jeux coalitionnels vers R^n.

Pour toute coalition D, soit v_D le jeu coalitionnel tel que $v_D(S) = 1$ si $D \subset S$ et $v_D(S) = 0$ sinon. Par l'axiome 2, et l'axiome 3 on déduit que

$$\Psi_i(v_D) = \frac{1}{|D|} 1_{\{i \in D\}}$$

Remarquez qu'il y a 2^n jeux différents v_D (un pour chaque coalition D). De plus la famille de jeux $\{v_D\}_{D \subset N}$ est linéairement indépendante (i.e. si $\sum_{D \subset N} \alpha_D v_D = 0$ alors pour tout D, $\alpha_D = 0$). Sinon, soit S une coalition de cardinal minimal telle que $\alpha_S \neq 0$. Dans ce cas,

$$0 = \sum_{D \subset N} \alpha_D v_D(S) = \sum_{D \subset S, D \neq \emptyset} \alpha_D = \alpha_S,$$

une contradiction. Donc $\{v_D\}_{D \subset N}$ est une base de l'espace des jeux coalitionnels (car cet espace est de dimension 2^n). Puisqu'une application linéaire est complètement déterminée par son comportement pour une base de l'espace. On a l'unicité. ■

La valeur de Shapley est une très bonne mesure des rapports de pouvoirs entre les joueurs dans un jeu coalitionnel. Dans le jeu 1 où l'unanimité est nécessaire, la valeur de Shapley est (100,100,100). Pour le jeu 2, où seulement les joueurs 1 et 2 ont le pouvoir de décision, sa prédiction est (150,150,0). Pour le jeu 3 où toute paire de joueurs suffit pour prendre une décision, sa valeur est (100,100,100). Pour le jeu 4, Aumann a calculé que la valeur de Shapley est égale à 0,4999557 pour les joueurs majoritaires (dans G) et elle est égale à 0,500443 pour les joueurs minoritaires (dans D).

16.4 Structure coopérative

Souvent des facteurs exogènes font que certaines coalitions sont plus probables que d'autres. La valeur de Shapley, qui les traitent ¨équiprobablement" est dans ce cas inadaptée (axiome A1). En effet, si les joueurs désignent des partis politiques et que le poids d'une coalition est le nombre de sièges obtenus par la coalition à l'assemblé nationale, il est souvent absurde dans certains pays d'imaginer une coalition entre un parti de droite et un parti de gauche. Il est donc plus pertinent de mesurer les rapports de pouvoir au sein d'une même famille politique. Ceci pourrait servir pour désigner le nombre de ministre que devrait avoir chaque parti dans la majorité.

La plus simple approche est de supposer que chaque joueur est associé naturellement à une certaine coalition. Dans ce contexte, la structure coopérative du jeu peut être décrite par une partition $F = \{F_1, ..., F_l\}$ de l'ensemble des joueurs. Chaque $C \in F$ est appelée une coalition active. Ainsi, étant donné un jeu coalitionnel v, une coalition active C et un joueur $i \in C$, on doit spécifier le paiement espéré du joueur i, $\Psi_i(v, C)$. Harsanyi (1963) puis Aumann et Dreze (1974) ont suggéré que ce paiement doit être la valeur de Shapley restreinte aux sous ensembles de C :

$$\Psi_i(v, C) = \sum_{S \subset C \setminus \{i\}} \frac{|S|! \times (|C| - |S| - 1)!}{|C|!} [v(S \cup \{i\}) - v(S)]$$

Plus tard, Myerson (1980) a montré que cette solution est l'unique solution du système :

$$\begin{cases} \sum_{i \in F_k} \Psi_i(v, C) = v(C) & \forall C \subset N \\ \Psi_i(v, C) - \Psi_i(v, C \setminus \{j\}) = \Psi_j(v, C) - \Psi_j(v, C \setminus \{i\}) & \forall C \subset N, \forall (i, j) \in C^2 \end{cases}$$

La première équation est une condition d'efficience : elle traduit le fait que la somme des poids individuels attribués à des membres d'une coalition active devrait être le poids total de la coalition active. La seconde équation est une condition d'équité : elle dit que le gain additionnel du joueur i si le joueur j quitte la coalition active doit être le même que le gain de j si i quitte la coalition active.

Exemples :

- Dans le jeu 1 : $\Psi_i(v, N) = 100$ et $\Psi_i(v, C) = 0$ si $|C| \leq 2$.
- Dans le jeu 3 : $\Psi_i(v, N) = 100$, $\Psi_i(v, C) = 150$ si $|C| = 2$ et $\Psi_i(v, C) = 0$ si $|C| = 1$. Il y a donc une incitation des joueurs dans la grande coalition N à essayer d'exclure un des joueurs pour former une coalition active plus petite.

Références

Aumann, Zamir & Tauman. Lecture notes on game theory. *Everyman's University, Israel (Livre en Hébreu).*

Demange & Ponssard, Théorie des jeux et analyse économique. *PUF, 1994.*

Kreps. Notes on the theory of choice. *Westview Press, 1988.*

Gale et Shapley. College admissions and the stability of mariage. *American Mathematical Monthly, Volume 69, 1962.*

Myerson. Game theory analysis of conflict. *Harvard University Press, 1991.*