Théorie des jeux algorithmique

Didier Lime

École Centrale de Nantes – LS2N

Dernière modification: 15 novembre 2023

Outline

Outline

Le dilemme du prisonnier

Exemple

Alice et Bob sont deux bandits notoires, qui ont été arrêtés pour un délit mineur. Tout le monde sait qu'ils sont aussi coupables de crimes horribles mais il n'y a pas de preuves. Le juge leur propose le marché suivant :

- si vous dénoncez l'autre et que lui ou elle ne vous dénonce pas, il ou elle part en prison pour 10 ans et pour vous on oublie tout
- si chacun dénonce l'autre c'est 8 ans pour tout le monde, avec la clémence due à la coopération avec la justice ;
- si personne ne dénonce l'autre, vous irez tous deux quand même 2 ans en prison pour le délit.

Si Alice et Bob ne se soucient que de leur peine de prison à venir, quelle est leur meilleure stratégie ? et s'ils peuvent se concerter avant ?

Le dilemme du prisonnier

Exemple

Alice et Bob sont deux bandits notoires, qui ont été arrêtés pour un délit mineur. Tout le monde sait qu'ils sont aussi coupables de crimes horribles mais il n'y a pas de preuves. Le juge leur propose le marché suivant :

- si vous dénoncez l'autre et que lui ou elle ne vous dénonce pas, il ou elle part en prison pour 10 ans et pour vous on oublie tout
- si chacun dénonce l'autre c'est 8 ans pour tout le monde, avec la clémence due à la coopération avec la justice ;
- si personne ne dénonce l'autre, vous irez tous deux quand même 2 ans en prison pour le délit.

Si Alice et Bob ne se soucient que de leur peine de prison à venir, quelle est leur meilleure stratégie ? et s'ils peuvent se concerter avant ?

Alice \ Bob	Trahir	Soutenir
Trahir	-8 / -8	0 / -10
Soutenir	-10 / 0	-2 / -2

Jeu non-coopératif en forme normale

Définition

Un jeu fini à n joueurs en forme normale est un triplet (n, A, u), où :

- n est le nombre de **joueurs**. On suppose les joueurs ordonnés arbitrairement ;
- $A = A_1 \times \cdots \times A_n$ où pour tout $i \in [1..n]$, A_i est l'ensembles des **actions** possibles du joueur i. Tout vecteur $(a_1, \ldots, a_n) \in A$ est appelé **profil d'action** (*action profile*) ;
- $u = (u_1, \ldots, u_n)$ où pour tout $i \in [1..n]$, $u_i : A \to \mathbb{R}$ est la fonction d'**utilité** (*utility*), aussi appelée fonction de gain (*payoff*), du joueur i.

- Chaque joueur cherche à **maximiser** son espérance d'utilité ;
- On dit qu'il est **rationnel** si c'est le cas ;
- On peut réprésenter un tel jeu avec n tableaux à n dimensions
  Chaque tableau représente la fonction d'utilité d'un joueur.
- Pour deux joueurs, on a un couple de matrices : une **bimatrice** (*bimatrix*).

Stratégies

Définition (Stratégie)

Une **stratégie (mixte)** du joueur i, est une distribution s_i de probabilité sur A_i.

- S'il existe $a \in A_i$ telle que $s_i(a) = 1$ alors s_i est appelée **stratégie pure** ;
- Le **support** de s_i est l'ensemble d'actions $\{a \in A_i \mid s_i(a) > 0\}$;

Profil de stratégies

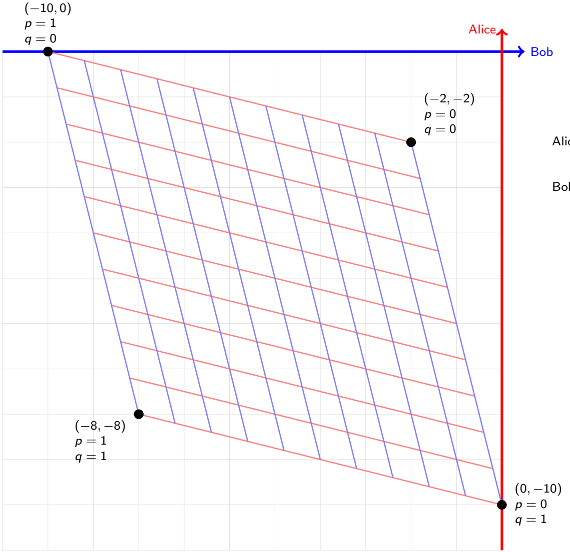
Définition (Profil de stratégies)

Un **profil de stratégies** est un tuple $s = (s_1, \ldots, s_n)$ tel que pour tout $i \in [1..n]$, s_i est une stratégie du joueur i.

- On note s_{-i} le profil partiel obtenu en supprimant la stratégie du joueur i ;
- On note (s'_i, s_{-i}) le profil de stratégies obtenu en remplaçant dans s la stratégie s_i du joueur i par sa stratégie s'_i.
- L'**espérance d'utilité** du joueur i pour le profil de stratégie s est

$$u_i(s) = \sum_{(a_1,\ldots,a_n)\in A} u_i(a_i) \prod_{j=1}^{n} s_j(a_j)$$

Utilité d'un profil de stratégies



Alice : p Trahir + $(1 - p)$ Soutenir

Bob : q Trahir + $(1 - q)$ Soutenir

Matching pennies

Exercice

Alice et Bob ont chacun une pièce de 1 euro. Ils choisissent en secret de la mettre côté pile ou côté face. On révèle ensuite les choix. Si les deux pièces sont du même côté Alice les empoche, sinon c'est Bob qui les prend.

1. Écrire la bimatrice de ce jeu ;
2. Quelle stratégie doit jouer Alice si elle pense que Bob va plus probablement jouer face ?
3. Quelle stratégie doit jouer Alice si elle pense que Bob pense qu'elle pense qu'il va plus probablement jouer face (et qu'il est rationnel) ?
4. Prouver qu'il existe un profil de stratégies pour lequel, pour les deux joueurs, la connaissance de la stratégie adverse n'incite pas à changer la sienne.

Outline

Stratégies strictement dominées

Définition

Une stratégie s_i du joueur i est strictement dominée par une autre de ses stratégies s_i' si pour toutes stratégies s_{-i} des autres joueurs, on a $u(s_i, s_{-i}) < u(s_i', s_{-i})$.

- On a vu qu'un joueur **rationnel** cherche à maximiser l'espérance de son gain ;
- Une stratégie strictement dominée ne sera jamais jouée par un joueur rationnel.

Exemple

1. Dans le dilemme du prisonnier, *soutenir* est strictement dominée par *trahir*
2. Dans la matrice ci-dessous, correspondant aux gains du joueur 1 (sur 2), la 2e ligne correspond à une stratégie strictement dominée, par une stratégie **mixte**.

$$\begin{bmatrix} 4 & 1 \\ 2 & 2 \\ 1 & 4 \end{bmatrix}$$

Stratégies strictement dominées

- On peut décider si une stratégie s_i est dominée en temps **polynomial**, via la programmation linéaire :
  on a domination stricte ssi $\epsilon > 0$

$$\begin{cases} \text{maximiser } \epsilon : \\ \\ \forall \text{ actions } a_{-i}, \sum_{a_k \in A_i} p_k u_i(a_k, a_{-i}) \geq u_i(s_i, a_{-i}) + \epsilon \\ \forall k, p_k \geq 0 \\ \sum_k p_k = 1 \end{cases}$$

- On peut ne regarder que les profils de **stratégies pures** des autres joueurs car pour tout s_{-i}, il existe a_{-i} telle que $u_i(s_i, s_{-i}) \leq u_i(s_i, a_{-i})$:

$$u_i(s_i, s_{-i}) = \sum_{a_{-i} \in A_{-i}} s_{-i}(a_{-i}) u_i(s_i, a_{-i}) \leq \sum_{a_{-i} \in A_{-i}} s_{-i}(a_{-i}) \max_{a'_{-i}} u_i(s_i, a'_{-i}) = \max_{a_{-i}} u_i(s_i, a_{-i})$$

Suppression des stratégies strictement dominées

- On suppose que la rationalité des joueurs est une **connaissance commune** (*common knowledge*) :
  Alice ne jouera pas une stratégie dominée et Bob le sait, et Alice sait que Bob le sait, et ainsi de suite pour les deux
- On peut alors toujours **éliminer** les stratégies strictement dominées ;
- Après élimination d'une stratégies, de **nouvelles relations** de domination sur les stratégies restantes peuvent apparaître :

Exemple

Alice \ Bob	g	d
H	1, 2	2, 1
B	2, 3	1, 2

- On calcule donc **itérativement** un plus grand point fixe.

Suppression des stratégies strictement dominées

- L'**ordre** d'élimination est **indifférent** ;
- On n'élimine que des **stratégies pures** car si s_i n'est pas dominée par s_i' à cause d'une stratégie mixte s_j' $(u(s_i, s_j') \leq u(s_i', s_j'))$, alors il existe une stratégie pure dans le support de s_j qui empêche également la domination. On ne crée donc pas de nouvelle domination en supprimant s_j'.

Exemple

Dilemme du prisonnier : après élimination des stratégies strictement dominées, il ne reste que *trahir* pour chaque joueur.

- Si le jeu est infini, cela peut ne pas terminer ou donner des résultat différents selon ce qu'on élimine à chaque fois

Exemple

Alice et Bob choisissent chacun un nombre entier positif ou nul et reçoivent un gain égal à leur choix.

D'autres variations sur le dilemme du prisonnier

Exercice

Les prisonniers amoureux. Bonnie et Clyde se retrouvent dans la situation d'Alice et Bob face au juge. Contrairement à ceux-ci, ils s'aiment d'un amour tendre et savoir l'autre souffrant en prison leur coûte comme s'ils passaient eux-même la moitié du temps correspondant en prison.

1. Écrire la nouvelle bimatrice de ce jeu.
2. Une stratégie domine-t-elle l'autre ?

D'autres variations sur le dilemme du prisonnier

Exercice

Les prisonniers amoureux. Bonnie et Clyde se retrouvent dans la situation d'Alice et Bob face au juge. Contrairement à ceux-ci, ils s'aiment d'un amour tendre et savoir l'autre souffrant en prison leur coûte comme s'ils passaient eux-même la moitié du temps correspondant en prison.

1. Écrire la nouvelle bimatrice de ce jeu.
2. Une stratégie domine-t-elle l'autre ?

Alice \ Bob	Trahir	Soutenir
Trahir	-12 -12	-10 -5
Soutenir	-5 -10	-3 -3

D'autres variations sur le dilemme du prisonnier

Alice \ Bob	Trahir	Soutenir
Trahir	-12 / -12	-10 / -5
Soutenir	-5 / -10	-3 / -3

Exercice

Les prisonniers amoureux. Bonnie et Clyde se retrouvent dans la situation d'Alice et Bob face au juge. Contrairement à ceux-ci, ils s'aiment d'un amour tendre et savoir l'autre souffrant en prison leur coûte comme s'ils passaient eux-même la moitié du temps correspondant en prison.

1. Écrire la nouvelle bimatrice de ce jeu.
2. Une stratégie domine-t-elle l'autre ?

Exercice

Les prisonniers avec des remords. Supposons qu'Alice et Bob sont tout de même un peu gênés de trahir l'autre. Une trahison unilatérale leur coûte ainsi l'équivalent de 3 ans de prison en affres moraux.

1. Écrire la nouvelle bimatrice de ce jeu.
2. Une stratégie domine-t-elle l'autre ?

D’autres variations sur le dilemme du prisonnier

Exercice

Les prisonniers amoureux. Bonnie et Clyde se retrouvent dans la situation d’Alice et Bob face au juge. Contrairement à ceux-ci, ils s’aiment d’un amour tendre et savoir l’autre souffrant en prison leur coûte comme s’ils passaient eux-même la moitié du temps correspondant en prison.

1. Écrire la nouvelle bimatrice de ce jeu.
2. Une stratégie domine-t-elle l’autre ?

Alice \ Bob	Trahir	Soutenir
Trahir	-12, -12	-5, -10
Soutenir	-10, -5	-3, -3

Exercice

Les prisonniers avec des remords. Supposons qu’Alice et Bob sont tout de même un peu gênés de trahir l’autre. Une trahison unilatérale leur coûte ainsi l’équivalent de 3 ans de prison en affres moraux.

1. Écrire la nouvelle bimatrice de ce jeu.
2. Une stratégie domine-t-elle l’autre ?

Alice \ Bob	Trahir	Soutenir
Trahir	-8, -8	-3, -10
Soutenir	-10, -3	-2, -2

Domination (très) faible

- On peut définir des formes plus faibles de domination :
  - s_i' **domine très faiblement** s_i : pour toutes stratégies des autres joueurs s_{-i}, on a $u(s_i, s_{-i}) \leq u(s_i', s_{-i})$.
  - s_i' **domine faiblement** s_i : s_i' domine très faiblement s_i' et il existe une stratégies des autres joueurs s_{-i}, telle que $u(s_i, s_{-i}) < u(s_i', s_{-i})$.

Exemple

- Si Alice et Bob ont un peu moins de remords alors *soutenir* est faiblement dominée par *trahir*. La stratégie *soutenir* n'est moins bonne que *trahir* que si Bob choisit lui-même *trahir*.
- Après élimination de *soutenir* on retrouve une unique stratégie dominante (faiblement) : *trahir*.
- Même la très faible domination ne permet pas d'éliminer *soutenir* quand les remords sont équivalents à 3 ans.

Alice \ Bob	Trahir	Soutenir
Trahir	-8 -8	-10 -2
Soutenir	-2 -10	-2 -2

Domination (très) faible

- Dans la suppression itérée, supprimer des stratégies (très) faiblement dominées ne conduit en général pas au même ensemble selon l'ordre d'élimination :

Exemple

		Bob	
		Trahir	Soutenir
Alice	Trahir	1, 1	1, 1
	Soutenir	0, 0	1, 1

Outline

Meilleures réponses et équilibre de Nash

Définition (Meilleure réponse)

La stratégie s_i^* du joueur i est une meilleure réponse aux stratégies s_{-i} des autres joueurs s'il n'existe pas une autre stratégie s_i telle que $u_i((s_i, s_{-i})) > u_i((s_i^*, s_{-i}))$.

Définition (Équilibre de Nash)

Le profil de stratégies $s = (s_1, \ldots, s_n)$ est un équilibre de Nash si pour tout joueur i, s_i est une meilleure réponse à s_{-i}.

- Un équilibre de Nash (NE) s est strict si pour tout joueur i, s_i est l'unique meilleure réponse à s_{-i} ;
- Dans un équilibre strict toutes les stratégies sont pures ;
- Sans mécanisme additionnel de coordination, et si la rationalité de tous est une connaissance commune, seuls les équilibres de Nash sont des choix rationnels.

Équilibre de Nash

Exercice

Quels sont les équilibres de Nash dans les trois versions précédentes du dilemme du prisonnier ? Sont-ils stricts ?

Alice \ Bob	Trahir	Soutenir
Trahir	-8 / -8	-10 / 0
Soutenir	0 / -10	-2 / -2

Alice \ Bob	Trahir	Soutenir
Trahir	-12 / -12	-10 / -5
Soutenir	-5 / -10	-3 / -3

Alice \ Bob	Trahir	Soutenir
Trahir	-8 / -8	-10 / -3
Soutenir	-3 / -10	-2 / -2

Équilibre de Nash

Théorème (Meilleure réponse)

Soit s_i^* une stratégie mixte et soit σ_i son support. Si s_i^* est une meilleure réponse à s_{-i} alors toute stratégie pure $a_i \in \sigma_i$ est aussi une meilleure réponse à s_{-i}

On ne peut pas avoir pour tout $a_i \in \sigma_i$, $u_i(a_i, s_{-i}) < u_i(s_i^*, s_{-i})$. Donc il y a forcément une $a_i^+ \in \sigma_i$ telle que $u_i(a_i^+, s_{-i}) \geq u_i(s_i^*, s_{-i})$. Si on avait $u_i(a_i^+, s_{-i}) > u_i(s_i^*, s_{-i})$, on pourrait augmenter la probabilité de a_i^+ pour obtenir mieux que s_i^*.

- Toutes les stratégies **mixtes** de même support sont donc aussi des meilleures réponses !
- La pondération précise des meilleures réponses pures permet d'obtenir un équilibre.
- C'est en utilisant ce résultat qu'on avait trouvé l'unique équilibre de *matching pennies*.

Équilibre de Nash

Exercice

Utiliser le théorème des meilleures réponses pour calculer l'équilibre en stratégies mixtes dans la dernière version du dilemme du prisonnier.

Alice \ Bob	Trahir	Soutenir
Trahir	-8, -8	-3, -10
Soutenir	-10, -3	-2, -2

Équilibre de Nash : résolution par énumération des supports

- Le théorème des meilleures réponses fournit un algorithme pour trouver tous les équilibres de Nash d'un jeu à n-joueurs par **énumération des supports** ;
- Pour un jeu $G = (n, A, u)$, $\sigma_1 \subseteq A_1, \ldots, \sigma_n \subseteq A_n$, on définit $A_{\sigma_1 \cdots \sigma_n} = (n, \sigma_1 \times \cdots \times \sigma_n, (u_{1|\sigma_1}, \ldots, u_{n|\sigma_n}))$ la restriction de G à $\sigma_1, \ldots, \sigma_n$;
- pour tous $\sigma_1, \ldots, \sigma_n$ tels que pour tout i et tout $a_i \in \sigma_i$, a_i n'est pas strictement dominée dans $G_{\sigma_1 \cdots \sigma_n}$, si le programme **non-linéaire** suivant est non-vide, alors toutes ses solutions sont des équilibres de Nash :

$$
\begin{cases}
\text{Étant données les stratégies des autres joueurs, le joueur } i \text{ est indifférent entre toutes les actions de } \sigma_i \\
\forall i \in [1..n], \forall a_i \in \sigma_i, \sum_{a_{-i} \in \sigma_{-i}} \left(\prod_{j \neq i} p_j(a_j)\right) u_i(a_i, a_{-i}) = v_i \\
\text{Étant données les stratégies des autres joueurs, le joueur } i \text{ ne préfère pas une action hors de } \sigma_i \\
\forall i \in [1..n], \forall a_i \notin \sigma_i, \sum_{a_{-i} \in \sigma_{-i}} \left(\prod_{j \neq i} p_j(a_j)\right) u_i(a_i, a_{-i}) \leq v_i \\
p_i \text{ est une distribution de probabilité sur } \sigma_i \\
\forall i \in [1..n], \forall a_i \in \sigma_i, p_i(a_i) \geq 0 \\
\forall i \in [1..n], \forall a_i \notin \sigma_i, p_i(a_i) = 0 \\
\forall i \in [1..n], \sum_{a_i \in \sigma_i} p_i(a_i) = 1
\end{cases}
$$

- Le programme non-linéaire peut-être résolu par les méthodes classiques d'optimisation non-linéaire.

Équilibre de Nash : résolution par énumération des supports

- S'il n'y a que deux joueurs, le programme précédent est **linéaire** :

$$\begin{cases} \text{Étant donnée la stratégie de l'autre joueur, le joueur } i \text{ est indifférent entre toutes les actions de } \sigma_i \\ \forall i \in [1..n], \forall a_i \in \sigma_i, \sum_{a_{-i} \in \sigma_{-i}} p_{-i}(a_{-i}) u_i(a_i, a_{-i}) = v_i \\ \text{Étant donnée la stratégie de l'autre joueur, le joueur } i \text{ ne préfère pas une action hors de } \sigma_i \\ \forall i \in [1..n], \forall a_i \notin \sigma_i, \sum_{a_{-i} \in \sigma_{-i}} p_{-i}(a_{-i}) u_i(a_i, a_{-i}) \leq v_i \\ p_i \text{ est une distribution de probabilité sur } \sigma_i \\ \forall i \in [1..n], \forall a_i \in \sigma_i, p_i(a_i) \geq 0 \\ \forall i \in [1..n], \forall a_i \notin \sigma_i, p_i(a_i) = 0 \\ \forall i \in [1..n], \sum_{a_i \in \sigma_i} p_i(a_i) = 1 \end{cases}$$

Le dilemme du pollueur [2]

Exercice

Chacun des gouvernements des n plus gros pays pollueurs de la planète se demande s'il faut voter une loi anti-pollution. Si la loi n'est pas votée par le gouvernement i, alors il inflige à tous les pays, lui inclus, un coût de 1 sur la période jusqu'aux prochaines élections, lié à la surmortalité, les catastrophes naturelles, etc Si la loi est votée, ce coût disparaît mais la loi elle-même a un coût économique immédiat de 2.

En supposant que k pays votent la loi

1. Quel est le coût pour un pays qui vote la loi ?
2. Quel est le coût pour un pays qui ne vote pas la loi ?
3. En supposant que les gouvernements cherchent uniquement a minimiser leur coût sur le mandat, quelle est leur meilleure stratégie ?
4. Obtient-on un équilibre de Nash ?
5. Quel est alors le coût aggrégé pour tous les pays ? Comparer avec le coût si tout le monde vote la loi ?

La tragédie des biens communs [1]

Exercice

Dix familles d'un même village élèvent des chèvres dans un champ. Une chèvre qui paisse sur une fraction a du champ produit $e^{1-1/10a}$ seaux de lait. Le chef du village décide combien de chèvres allouer à chaque famille.

1. Quel est le nombre de chèvres qui maximise la production de lait ?
2. Combien chaque famille reçoit-elle dans une répartition équitable par le chef ?
3. Le chef meurt dans un malheureux accident et son remplaçant décide de laisser les familles libres de décider de l'exploitation du champ. Si G est le nombre total de chèvres des autres familles, quel est le nombre de chèvres g_i qui maximise la production de lait pour la famille i ?
4. Chaque famille faisant le même raisonnement, quelle est la quantité totale de lait produite par la nouvelle organisation ?
5. S'apercevant de son erreur le nouveau chef décide d'intervenir mais il ne veut pas revenir sur la liberté accordée aux familles de choisir leur nombre de chèvres. Il décide que chaque famille recevra $1/10^{\mathrm{e}}$ du total produit. Comment évolue alors la production ?

Équilibre de Nash : Existence et complexité

Théorème (Nash)

Dans tout jeu fini (nombre de joueurs et nombre d'actions finis), il existe au moins un équilibre de Nash.

- L'existence d'un équilibre en stratégies **pures** n'est pas garantie (cf. *Matching pennies*) ;
- Il peut y avoir une **infinité** d'équilibres de Nash
- Le calcul d'un équilibre de Nash en stratégie mixtes dans un jeu à 2 joueurs est **PPAD-complet**
  Polynomial Parity Arguments on Directed graphs ; inclus dans NP
- L'existence d'un équilibre **avec contraintes** est NP-complète : par exemple, un jeu à deux joueurs a-t-il
  - au moins deux équilibres ?
  - un équilibre avec au moins une certaine valeur pour le joueur 1 ?
  - un équilibre avec au moins une certaine somme pour les deux joueurs ?
  - un équilibre qui inclut au moins n actions ?
  - un équilibre qui inclut l'action a ?
  - un équilibre qui n'inclut pas l'action a ?

Équilibre de Nash : jeux dégénérés

Définition (Jeu dégénéré)

Un jeu à **deux joueurs** est **dégénéré** s'il existe une stratégie mixte dont le support est de taille k, et pour laquelle il existe au moins $k+1$ meilleures réponses pures.

- Un jeu dégénéré peut avoir une infinité d'équilibres de Nash
- Le système linéaire d'équations pour le joueur au plus grand support a plus d'inconnues que d'équations.

Exemple

Alice \ Bob	c	d
a	1, 1	0, 1
b	0, 0	0, 0

- La stratégie pure a d'Alice a 2 meilleures réponses de Bob : c et d ;
- a est une meilleure réponse d'Alice à c et à d ;
- $(a, pc + (1-p)d)$ est un équilibre de Nash pour toute valeur de p.

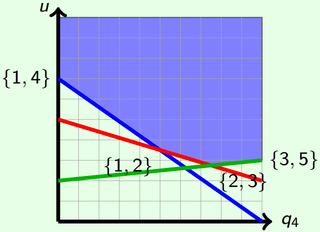
Équilibre de Nash : Polyèdres des meilleures réponses

- On considère un jeu à deux joueurs (non-dégénéré) ;
- On numérote les actions du joueur 1 de 1 à n et celles du joueur 2 de $n+1$ à $n+m$
- Soit $s_2 = (q_{n+1}, \ldots, q_{n+m})$ une stratégie du joueur 2 ;
- Pour une **meilleure réponse** du joueur 1 à s_2, son utilité est égale à $\max_{a_1 \in A_1} u_1(a_1, s_2)$
- On considère l'ensemble $\overline{Q}_1$ des points $(q_{n+1}, \ldots, q_{n+m}, u)$ tels que $u \geq \max_{a_1 \in A_1} u_1(a_1, s_2)$
- C'est un polyèdre convexe de $\mathbb{R}^{m+1}$; on étiquette chacun de ses sommets par :
  - les numéros des actions du joueur 1, qui y réalisent le maximum de u ;
  - les numéros des actions du joueur 2, qui y ont une probabilité nulle.

Exemple

Soit la matrice de gains suivante pour le joueur 1 :

$$\begin{bmatrix} 0 & 7 \\ 2 & 5 \\ 3 & 2 \end{bmatrix}$$



q_4 : proba. que le J2 joue la 1ère colonne

Équilibre de Nash : Polyèdres des meilleures réponses

- On définit de même un polyèdre $\overline{Q}_2$ des meilleures réponses pour le joueur 2.
- Les **équilibres de Nash** sont alors les couples de sommets (v_1, v_2) tels que :
  - $v_1 \in \overline{Q}_1$ et $v_2 \in \overline{Q}_2$;
  - L'union des ensembles d'étiquettes de v_1 et v_2 est de taille $n + m$.
- On peut trouver tous les équilibres par énumération des sommets des deux polyèdres
  en théorie moins coûteux que l'énumération des supports
- L'**algorithme de Lemke-Howson** trouve un équilibre de Nash par un parcours astucieux de ces sommets.

Équilibre de Nash : Polyèdres des meilleures réponses

Exercice

Dilemme de sécurité. L'Alicie et le Bobistan sont deux pays aux relations tendues. Pour se protéger l'un de l'autres, les deux pays envisagent de développer et maintenir une force de dissuasion nucléaire. Le coût d'une telle force est très élevé mais ne pas l'avoir si l'autre l'a est encore pire. Dans l'ordre chaque pays préfère strictement :

1. personne n'a cette force,
2. je l'ai et l'autre non,
3. on l'a tous les deux,
4. je ne l'ai pas et l'autre l'a.

Trouver la bimatrice de ce jeu ainsi que tous les équilibres de Nash, en utilisant les polyèdres des meilleures réponses.

Équilibre de Nash : Polyèdres des meilleures réponses

Exercice

Dilemme de sécurité. L'Alicie et le Bobistan sont deux pays aux relations tendues. Pour se protéger l'un de l'autres, les deux pays envisagent de développer et maintenir une force de dissuasion nucléaire. Le coût d'une telle force est très élevé mais ne pas l'avoir si l'autre l'a est encore pire. Dans l'ordre chaque pays préfère strictement :

1. personne n'a cette force,
2. je l'ai et l'autre non,
3. on l'a tous les deux,
4. je ne l'ai pas et l'autre l'a.

Alicie \ Bobistan	Nukes	No nukes
Nukes	2, 2	3, 1
No nukes	1, 3	4, 4

Trouver la bimatrice de ce jeu ainsi que tous les équilibres de Nash, en utilisant les polyèdres des meilleures réponses.

Équilibre de Nash et stratégies dominées

- Une stratégie dans un équilibre de Nash n'est jamais strictement dominée ;
- Après suppression itérée des stratégies **strictement** dominées, les équilibres de Nash obtenus sont exactement ceux d'origine ;
- Après suppression itérée des stratégies **(très) faiblement** dominées, les équilibres de Nash obtenus sont un sous-ensemble de ceux d'origine ;

Exercice

Alice \ Bob	*g*	*d*
h	2, 2	2, 2
b	2, 2	1, 3

Alice \ Bob	*g*	*d*
h	2, 2	2, 3
b	2, 2	1, 2

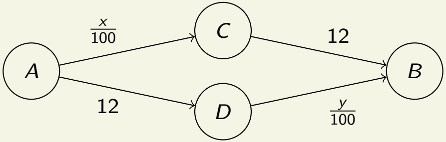
Alice \ Bob	*g*	*d*
h	0, 0	2, 2
b	1, 1	2, 0

1. Calculer les équilibres de Nash purs de ces jeux
2. Peut-on tous les obtenir par élimination des stratégies (très) faiblement dominées ?

Un jeu de congestion

Exercice

Paradoxe de Braess. Chaque jour, 1000 personnes veulent se rendre simultanément d'un point A à un point B pour aller travailler. Il y a deux routes : par C ou par D. Certains tronçons sont des grosses routes et la vitesse dessus est insensible au nombre de personnes les empruntant, d'autres sont plus étroits et la vitesse y dépend du nombre de personnes x dessus.



1. Chaque personne est un joueur qui doit choisir son chemin pour minimiser son temps de trajet de A à B. Quels sont les équilibres de Nash dans ce jeu à 1000 joueurs ? Quelles sont leurs valeurs ?
2. Pour améliorer les conditions de circulation, la mairie décide d'ouvrir un tronçon de C vers D, sur lequel le temps de trajet est 1 (indépendamment du nombre de personnes). Quels sont maintenant les équilibres de Nash, ainsi que leurs valeurs ?

Outline

Efficacité de Pareto

- Chaque équilibre de Nash correspond à des choix rationnels de tous les joueurs
- Lorsqu'il y a plusieurs équilibres de Nash, on ne peut pas toujours facilement choisir
- Un critère, quand il est appliquable, est l'efficacité de Pareto

Définition (Pareto-domination)

Un profil de stratégie s en domine un autre s' au sens de Pareto si pour tout joueur i, $u_i(s) \geq u_i(s')$ et il existe un joueur j pour lequel $u_j(s) > u_j(s')$.

Définition (Pareto-efficacité)

Un profil de stratégies est efficace au sens de Pareto s'il n'est dominé au sens de Pareto par aucun autre.

Exemple

Dans le dilemme du prisonnier original, (soutenir, soutenir) est Pareto-efficace mais pas (trahir, trahir).

Un jeu de coordination et de confiance

Exercice

Chasse au cerf. (*Stag hunt*). Deux villageois partis chasser de quoi nourrir leur famille se recontrent. L'une a aperçu un cerf et l'autre des lapins. Il savent tous deux que, à deux, ils attraperont le cerf sans difficulté.h À deux, ils peuvent aussi attraper un peu plus de lapins que tout seul mais il faudra partager. Seul, attraper le cerf est impossible et, si l'autre fait défection, il restera peu de temps pour les lapins. Ils décident séparément et indépendamment de partir chasser soit le cerf, soit les lapins.

1. Trouver tous les équilibres de Nash de ce jeu ;
2. Y en a-t-il un qui est meilleur que les autres au sens de Pareto ?
3. L'équilibre (L, L) est dit **risque-dominant**. Une façon de le voir est ainsi : ne sachant quel équilibre Bob va choisir, Alice peut considérer qu'il choisit un des deux équilibres purs avec une probabilité p choisie de manière uniforme entre 0 et 1. Quelle stratégie doit-elle alors choisir ? Si toutes ces informations sont des connaissances communes, quel équilibre obtient-on ?
4. Alice a une certaine confiance en Bob et elle estime que la probabilité p qu'il choisisse Cerf est forcément supérieure ou égale à une certaine valeur $r > 0$. Quelle est la plus petite valeur de r pour laquelle elle choisit l'équilibre sur Cerf ?

Alice \ Bob	Cerf	Lapins
Cerf	8, 8	1, 6
Lapins	6, 1	4, 4

Stratégie de sécurité et valeur maxmin

- Si le joueur i n'est pas sûr de l'utilité des autres joueurs, il peut au minimum assurer le gain :

$$\text{maxmin}_i = \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$$

- On l'appelle aussi le **niveau de sécurité** (*security level*) du joueur i ;
- C'est la plus grande valeur qu'il peut assurer s'il ne connaît pas les stratégies des autres ;
- C'est aussi la plus petite valeur que les autres peuvent le forcer à obtenir, connaissant sa stratégie ;
- La stratégie correspondante est dite **stratégie de sécurité** :

$$s_i = \text{argmax}_{s_i} \min s_{-i} u_i(s_i, s_{-i})$$

Caclul de maxmin_i en stratégies pures

- Le min est forcément atteint pour un profil de stratégies pur des autres joueurs
  Même argument que pour les meilleures réponses

$$\text{maxmin}_i = \max_{s_i} \min_{a_{-i}} u_i(s_i, a_{-i})$$

- Si on se restreint aux stratégies pures pour le joueur i, on peut facilement calculer la valeur correspondante par **énumération** :

$$\text{maxmin}_i = \max_{a_i} \min_{a_{-i}} u_i(a_i, a_{-i})$$

Exercice

1. Dans la *Chasse au cerf* quelle est la valeur maxmin et la stratégie correspondante d'Alice, en se restreignant aux stratégies pures ?
2. Et pour l'Alicie dans le *Dilemme de sécurité* ?

Caclul de maxmin_i en stratégies mixtes

- On a vu que :

$$\text{maxmin}_i = \max_{s_i} \min_{a_{-i}} u_i(s_i, a_{-i})$$

- On cherche donc la plus grande valeur u_i^* de u_i et des valeurs de $s_i(a_i)$ pour tout a_i telles que :

$$\forall a_{-i}, u_i(s_i, a_{-i}) \geq u_i^*$$

- On a donc un nombre fini de contraintes linéaires et on peut trouver les valeurs par programmation linéaire :

$$\begin{cases} \text{maximiser } u : \\ \\ \forall a_{-i}, \sum_{a_i} s_i(a_i) u_i(a_i, a_{-i}) \geq u \\ \forall a_i, s_i(a_i) \geq 0 \\ \sum_{a_i} s_i(a_i) = 1 \end{cases}$$

Exercice

1. Quelle est la stratégie mixte de sécurité pour Alice dans la *Chasse au Cerf* ?
2. Et pour l'Alicie dans le *Dilemme de sécurité* ?

Valeur minmax

- On définit la plus grande valeur que le joueur i peut assurer s'il connaît les stratégies des autres joueurs :

$$\mathrm{minmax}_i = \min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$$

- C'est aussi la plus petite valeur que les autres peuvent le forcer à obtenir, ne connaissant **pas** sa stratégie ;
- La stratégie s_{-i} des joueurs $-i$ réalisant la valeur minmax est une **stratégie de punition** du joueur i.

Caclul de minmax_i en stratégies mixtes

- On calcule minmax_i de la même manière que maxmin_i, en utilisant le fait que :

$$\mathrm{minmax}_i = \min_{s_{-i}} \max_{a_i} u_i(a_i, s_{-i})$$

- On cherche donc la plus grande petite valeur u_i^* de u_i et des valeurs de $s_{-i}(a_{-i})$ pour tout a_{-i} telles que :

$$\forall a_i, u_i(a_i, s_{-i}) \leq u_i^*$$

- On a à nouveau un programme linéaire :

$$\begin{cases} \text{minimiser } u : \\ \\ \forall a_i, \sum_{a_{-i}} s_{-i}(a_{-i}) u_i(a_i, a_{-i}) \leq u \\ \forall k \neq i, \forall a_k \in A_k, s_k(a_k) \geq 0 \\ \forall k \neq i, \sum_{a_k \in A_k} s_k(a_k) = 1 \end{cases}$$

Comparaison de maxmin et minmax

- On a toujours :

$$\forall i, \mathrm{maxmin}_i \leq \mathrm{minmax}_i$$

- Dans un jeu à deux joueurs **avec des stratégies mixtes** :

$$\forall i, \mathrm{maxmin}_i = \mathrm{minmax}_i$$

- Dans un jeu à **plus que deux joueurs**, même en stratégies mixtes on peut avoir :

$$\forall i, \mathrm{maxmin}_i < \mathrm{minmax}_i$$

- On a enfin pour tout équilibre de Nash s :

$$\forall i, u_i(s) \geq \mathrm{maxmin}_i \text{ et } u_i(s) \geq \mathrm{minmax}_i$$

maxmin et minmax : exercice

Exercice

1. Caculer le minmax d'Alice dans la *Chasse au cerf* (en se restreignant aux stratégies pures) ;
2. Trouver un jeu à deux joueurs dans lequel $\text{maxmin}_i < \text{minmax}_i$ pour un des deux joueurs (stratégies pures) ;
3. Dans ce jeu, montrer qu'on a bien $\text{minmax}_i = \text{maxmin}_i$ avec des stratégies mixtes.

Exercice

On considère les matrices suivantes représentant les gains du joueur 1 dans un jeu à 3 joueurs (issues de [4]) :

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Le joueur 3 choisit la matrice, le joueur 2 la colonne et le joueur 1 la ligne.

Calculer maxmin_1 et minmax_1 en stratégies mixtes.

Minimisation du regret

- Si Alice utilise sa stratégie de sécurité mais se rend compte que Bob est bien parti chasser le cerf, elle peut avoir des regrets ;
- Le **regret** du joueur i d'avoir choisi s_i, étant données les stratégies s_{-i} des autres est, si s_i^* est une meilleure réponse à s_{-i} :

$$r_i(s_i, s_{-i}) = u_i(s_i^*, s_{-i}) - u_i(s_i, s_{-i})$$

- Ne connaissant pas s_{-i}, le pire regret possible en choisissant s_i est : $\max_{s_{-i}} r_i(s_i, s_{-i}) = \max_{a_{-i}} r_i(s_i, a_{-i})$;
- L'action qui **minimise le regret maximum** est alors :

$$a_i^r = \operatorname*{argmin}_{a_i} \max_{a_{-i}} r_i(s_i, a_{-i})$$

Minimisation du regret

- On peut aussi **éliminer itérativement** les actions qui ne minimisent pas le regret :
- L'ordre d'élimination change le résultat ;
- Peut éliminer des équilibres de Nash (voire tous) ;
- Donne parfois des solutions plus intuitives que les équilibres de Nash.

Minimisation du regret

Exercice

Quelle est l'action (stratégie pure) qui minimise le regret maximum dans la *Chasse au cerf* ?

Minimisation du regret

Exercice

Quelle est l'action (stratégie pure) qui minimise le regret maximum dans la *Chasse au cerf* ?

Exercice

Le dilemme du voyageur. En sortant de l'avion, Alice et Bob se rendent compte que leurs valises respectives ont été perdues. Elles contenaient toutes deux un objet identique de grande valeur. La compagnie leur demande indépendamment la valeur de l'objet (entre 2 et 100 euros) et leur indique qu'elle paiera le **minimum** des deux valeurs à chacun avec un bonus de 2 euros pour le moins disant et un malus de 2 euros pour l'autre. Enfin si les deux donnent la même valeur, il n'y a ni bonus ni malus. Alice et Bob choisissent indépendamment et cherchent à maximiser leur remboursement personnel sans se soucier de l'autre.

1. Quelles sont les stratégies faiblement dominées ?
2. En itérant leur élimination, quel équilibre de Nash trouve-t-on ?
3. Y en a-t-il d'autres ?
4. Quelles sont les actions qui minimisent le regret ?
5. Itérer l'élimination des actions qui ne minimisent pas le regret. Quelle stratégie trouve-t-on ?

Conversation libre (Cheap talk)

- On suppose qu'Alice et Bob peuvent communiquer avant de choisir leur stratégie, et en particulier annoncer leur choix ;
- Ces annonces peuvent ne pas correspondre à ce qu'ils joueront ensuite ;
- Cette discussion ne modifie pas les utilités associées au jeu ;
- Quand Bob suppose que les autres joueurs le croient et jouent en conséquence, l'annonce est :
  - **auto-engageante** (*self-committing*) si la stratégie annoncée est alors optimale (elle correspond à un équilibre de Nash) ;
  - **auto-révélatrice** (*self-signalling*) si, pour tout autre stratégie que celle annoncée, il avait alors une meilleure annonce à faire.

Exemple

- Dans le dilemme du prisonnier, l'annonce par Bob de *Soutenir* n'est pas auto-engageante ;
- Dans la chasse au cerf, l'annonce par Bob de *Cerf* est auto-engageante mais pas auto-révélatrice ;
- Si l'utilité de chasser les lapins seul n'est que de 3, alors *Cerf* est auto-révélateur.

Conversation libre

Exemple

Sortie musicale (*Battle of the sexes*). Alice et Bob veulent sortir ensemble ce soir mais ont des envies différentes : Alice voudrait aller voir un groupe de *Heavy Metal* et Bob préférerait un récital de piano. Malgré leurs avis divergents, ils ne seront heureux que s'ils passent la soirée ensemble.

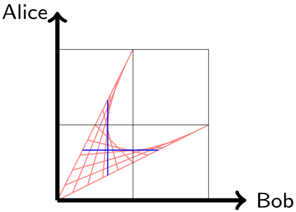
- Il y a deux équilibres de Nash purs : (M, M) qui vaut $(2, 1)$ et (P, P) qui vaut $(1, 2)$
- Il y a un équilibre mixte (non-symétrique) : $(\frac{1}{3}M + \frac{2}{3}P, \frac{2}{3}M + \frac{1}{3}P)$ qui vaut $(\frac{2}{3}, \frac{2}{3})$
- Aucun n'est Pareto-efficace ;
- Comment choisir sans ruiner la soirée ?
- Si Bob dit ≪ je vais voir le groupe de Métal ≫, est-ce auto-engageant ? auto-révélateur ?
- Alice peut le croire mais Bob est-il pleinement content du résultat ?

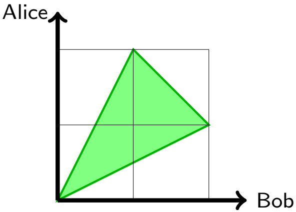
Alice \ Bob	Métal	Piano
Métal	2, 1	0, 0
Piano	0, 0	1, 2

Conversation libre

- Si Alice et Bob ne peuvent pas communiquer ou se faire confiance, ils choisissent leur stratégies indépendamment
  chaque ligne représente la paire de gains pour une stratégie fixée d'Alice (et de Bob car le jeu est symétrique), et toutes les stratégies de Bob

- S'ils peuvent **corréler** leurs choix en toute confiance, ils peuvent obtenir n'importe quel gain dans l'enveloppe convexe des **résultats** du jeu.
  Par exemple, choix de (Metal, Metal) ou (Piano, Piano) avec probabilité 0.5





Conversation libre

Exercice

Poule mouillée *Chicken*. Alice et Bob sont deux automobilistes qui se croisent dans une rue étroite. Si aucun des deux ne ralentit, ils provoquent un accident. Si l'un des deux ralentit et pas l'autre, tout le monde passe sans encombre mais le premier se sent lésé. Si les deux ralentissent, ils perdent tous deux un peu de temps mais leur égo est sauf.

1. Quels sont les équilibres de Nash ?
2. Quelle est la zone de gains atteignable s'ils coopèrent ?
3. Quelle est la zone de gains **théoriquement** atteignable s'ils ne coopèrent pas ?
4. Est-il envisageable **en pratique** d'obtenir $(2, 2)$ s'ils ne coopèrent pas ?

Alice \ Bob	Rapide	Lent
Rapide	-1 / -1	3 / 0
Lent	0 / 3	2 / 2

Conversation libre

- Un accord obtenu par conversation libre ne peut être respecté que s'il est auto-engageant pour les deux joueurs
  $\Rightarrow$ il correspond à un équilibre de Nash
- Les deux joueurs peuvent former des accords probabilistes (corréler leurs choix) si chaque résultat du choix probabiliste conduit à un équilibre de Nash :

Exemple

- Dans *Sortie musicale*, l'accord $\frac{1}{2}(M, M) + \frac{1}{2}(P, P)$ est possible, même si Alice et Bob ne coopèrent pas ;
- Dans *Poule mouillée*, l'accord $\frac{1}{2}(R, L) + \frac{1}{2}(L, R)$ est possible. $\Rightarrow$ Mieux que $(\frac{1}{2}R + \frac{1}{2}L, \frac{1}{2}R + \frac{1}{2}L)$!

- Cela permet d'obtenir des (espérances de) gains dans l'**enveloppe convexe des équilibres de Nash**, pour des joueurs non-coopératifs.
- Cela a des avantages qui permettent de faire accepter l'accord :
  - symétrisation des gains
  - maximisation du gain pour le joueur lésé
  - ...

Équilibres corrélés

- Dans *Poule mouillée*, le ≪ meilleur ≫ accord possible par conversation libre est $\frac{1}{2}(R, L) + \frac{1}{2}(L, R)$ avec un gain symétrique de $(1.5, 1.5)$

Équilibres corrélés

- Dans *Poule mouillée*, le ≪ meilleur ≫ accord possible par conversation libre est $\frac{1}{2}(R, L) + \frac{1}{2}(L, R)$ avec un gain symétrique de $(1.5, 1.5)$
- On peut faire mieux avec un mécanisme d'arbitrage un peu plus complexe :
  1. la stratégie pour chaque joueur est décidée de manière **probabiliste** et **centralisée** ;
  2. chaque joueur reçoit sa stratégie de manière **cachée**.

Équilibres corrélés

- Dans *Poule mouillée*, le ≪ meilleur ≫ accord possible par conversation libre est $\frac{1}{2}(R, L) + \frac{1}{2}(L, R)$ avec un gain symétrique de $(1.5, 1.5)$
- On peut faire mieux avec un mécanisme d'arbitrage un peu plus complexe :
  1. la stratégie pour chaque joueur est décidée de manière **probabiliste** et **centralisée** ;
  2. chaque joueur reçoit sa stratégie de manière **cachée**.

Exercice

Supposons que l'arbitre donne les stratégies suivantes à Alice et Bob : (R, L) avec probabilité $\frac{1}{3}$, (L, R) avec probabilité $\frac{1}{3}$, et (L, L) avec probabilité $\frac{1}{3}$. Les deux joueurs ne connaissent pas la stratégie reçue par l'autre mais connaissent le mécanisme d'arbitrage.

1. Quelle est la probabilité que Bob ait reçu *Rapide* sachant qu'Alice a reçu *Rapide* ? et qu'il ait reçu *Lent* ?
2. A-t-elle intérêt à jouer *Lent* si elle a reçu *Rapide* ?
3. Quelle est la probabilité que Bob ait reçu *Rapide* sachant qu'Alice a reçu *Lent* ? et qu'il ait reçu *Lent* ?
4. A-t-elle intérêt à jouer *Rapide* si elle a reçu *Lent* ?
5. Quelle est l'espérance de gain globale pour Alice ?
6. Peut-on obtenir mieux en augmentant la probabilité de (L, L) ?

Équilibre corrélé

Définition

Un **équilibre corrélé** d'un jeu à n joueurs est une distribution π sur les *profils d'actions* (stratégies pures) telle que pour tout i :

$$\forall a_i, a_i' \in A_i, \sum_{a_{-i} \in A_{-i}} \pi(a_i, a_{-i}) u_i(a_i, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}} \pi(a_i, a_{-i}) u_i(a_i', a_{-i})$$

- Un arbitre tire un profil de stratégies pures pour chacun des joueurs ;
- Il communique son action à chacun des joueurs, mais pas celles des autres ;
- Chaque joueur peut jouer ce qu'il veut ;
- Mais **sachant l'action communiquée**, celle-ci est une **meilleure réponse** aux actions communiquées aux autres.

Exercice

Montrer que tout équilibre de Nash est un équilibre corrélé mais que l'inverse est faux.

Équilibre corrélé

- La valeur de l'équilibre corrélé est **au moins** maxmin_i pour le joueur i (en stratégies mixtes) ;
- Toute combinaison convexe d'équilibres corrélés est un équilibre corrélé, il y en a donc toujours une infinité ;
- Les équilibres obtenus par conversation libre (non coopérative) sont un cas particulier d'équilibre corrélé dans lesquels les stratégies peuvent ne pas être cachées ;
- On peut remplacer le médiateur par un *Multiparty Computation Protocol* (protocole cryptographique) pour implémenter l'équilibre ;
- Ou parfois plus simple : pour *Poule mouillée*, un chapeau avec 2 cartes L et une carte R ;
- Les contraintes de la définition sont affines et définissent un programme linéaire sur π :

$$\begin{cases} \forall i \in [1..n], \forall a_i, a_i' \in A_i, \sum_{a_{-i} \in A_{-i}} \pi(a_i, a_{-i}) u_i(a_i, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}} \pi(a_i, a_{-i}) u_i(a_i', a_{-i}) \\ \forall a \in A, p(a) \geq 0 \\ \sum_{a \in A} p(a) = 1 \end{cases}$$

Équilibre corrélé : exercice

Exercice

On considère le jeu à trois joueurs donné par les matrices suivantes :

G :

	g	d
h	$0, 1, 3$	$0, 0, 0$
b	$1, 1, 1$	$1, 0, 0$

M :

	g	d
h	$2, 2, 2$	$0, 0, 0$
b	$2, 2, 0$	$2, 2, 2$

D :

	g	d
h	$0, 1, 0$	$0, 0, 0$
b	$1, 1, 1$	$1, 0, 3$

1. Montrer quand dans tout équilibre de Nash, le joueur 2 (colonnes) et le joueur 3 (matrice) ne jouent pas simultanément g et M avec probabilité 1 ;
2. Montrer que dans tout équilibre de Nash le joueur 1 (lignes) joue b ;
3. Montrer que dans tout équilibre de Nash le joueur 2 joue g et le joueur 3 ne joue pas M ;
4. En déduire les équilibres de Nash et leurs valeurs ;
5. Il y a un équilibre corrélé dont la valeur domine celle des équilibres de Nash. Trouver cet équilibre.
6. Comment l'implémenter en pratique ?

Outline

Utilité

Exemple

Alice propose à Bob le jeu suivant : elle lance un pièce de monnaie (non biaisée, de manière équitable, etc.).

- Si face sort alors Bob reçoit 1 000 001€ ;
- Si pile sort alors Bob doit payer 1 000 000€.

Bob doit-il accepter de jouer ?

Utilité

Exemple

Alice propose à Bob le jeu suivant : elle lance un pièce de monnaie (non biaisée, de manière équitable, etc.).

- Si face sort alors Bob reçoit 1 000 001€ ;
- Si pile sort alors Bob doit payer 1 000 000€.

Bob doit-il accepter de jouer ?

- La notion d'**utilité** décrit les valeurs que Bob affecte à ses différents niveaux de richesse ;
- Elle permet de modéliser son **attitude au risque** ;
- Cette notion s'applique plus généralement que pour l'argent.

Préférences

- Dans les jeux étudiés précédemment, la valeur des gains n'influe sur l'existence des équilibres (purs) et dominations que par leur ordre relatif ;
- Ces gains représentent les **préférences** des agents ;
- Quand Alice choisit de trahir Bob cela **révèle** qu'elle **préfère** passer 8 ou 0 années en prison plutôt 10 ou 2 ;

Préférences

- Dans les jeux étudiés précédemment, la valeur des gains n'influe sur l'existence des équilibres (purs) et dominations que par leur ordre relatif ;
- Ces gains représentent les **préférences** des agents ;
- Quand Alice choisit de trahir Bob cela **révèle** qu'elle **préfère** passer 8 ou 0 années en prison plutôt 10 ou 2 ;
- Les préférences d'Alice peuvent être modélisées par une relation $\preceq$ telle que :
  $A \preceq B$ ssi Alice préfère B à A

Préférences

- Dans les jeux étudiés précédemment, la valeur des gains n'influe sur l'existence des équilibres (purs) et dominations que par leur ordre relatif ;
- Ces gains représentent les **préférences** des agents ;
- Quand Alice choisit de trahir Bob cela **révèle** qu'elle **préfère** passer 8 ou 0 années en prison plutôt 10 ou 2 ;
- Les préférences d'Alice peuvent être modélisées par une relation $\preceq$ telle que :
  $A \preceq B$ ssi Alice préfère B à A
- On requiert que la relation soit :
  1. **totale** (complète) : pour tout x, on a $x \preceq y$ ou $y \preceq x$ (ou les deux) ;
  2. et **transitive** : pour tout x, y, z, $x \preceq y$ et $y \preceq z$ implique $x \preceq z$
- On définit la relation stricte : $x \prec y$ si $x \preceq y$ et on n'a pas $y \preceq x$
- Et la relation d'**indifférence** : $x \sim y$ si $x \preceq y$ et $y \preceq x$

Préférences

- Dans les jeux étudiés précédemment, la valeur des gains n'influe sur l'existence des équilibres (purs) et dominations que par leur ordre relatif ;
- Ces gains représentent les **préférences** des agents ;
- Quand Alice choisit de trahir Bob cela **révèle** qu'elle **préfère** passer 8 ou 0 années en prison plutôt 10 ou 2 ;
- Les préférences d'Alice peuvent être modélisées par une relation $\preceq$ telle que :
  $A \preceq B$ ssi Alice préfère B à A
- On requiert que la relation soit :
  1. **totale** (complète) : pour tout x, on a $x \preceq y$ ou $y \preceq x$ (ou les deux) ;
  2. et **transitive** : pour tout x, y, z, $x \preceq y$ et $y \preceq z$ implique $x \preceq z$
- On définit la relation stricte : $x \prec y$ si $x \preceq y$ et on n'a pas $y \preceq x$
- Et la relation d'**indifférence** : $x \sim y$ si $x \preceq y$ et $y \preceq x$

Exercice

Si la relation de préférences d'Alice n'est pas transitive, montrer que Bob peut ruiner Alice.

Préférences sur des résulats incertains

- Un joueur sera (le plus) souvent confronté à des choix dont le résultat est **probabiliste** ;
- On modélise ce type de résultats par des **lotteries** : par exemple $pA + qB + (1 - p - q)C$ avec $p \geq 0, q \geq 0$ et $p + q \leq 1$.

Préférences sur des résulats incertains

- Un joueur sera (le plus) souvent confronté à des choix dont le résultat est **probabiliste** ;
- On modélise ce type de résultats par des **lotteries** : par exemple $pA + qB + (1 - p - q)C$
  avec $p \geq 0, q \geq 0$ et $p + q \leq 1$.
- On fait quelques hypothèses sur les préférences sur les lotteries :

Préférences sur des résulats incertains

- Un joueur sera (le plus) souvent confronté à des choix dont le résultat est **probabiliste** ;
- On modélise ce type de résultats par des **lotteries** : par exemple $pA + qB + (1 - p - q)C$
  avec $p \geq 0, q \geq 0$ et $p + q \leq 1$.
- On fait quelques hypothèses sur les préférences sur les lotteries :
  1. **Monotonie** : si Alice préfère A à B alors pour $p > q$ elle préfère $pA + (1 - p)B$ à $qA + (1 - q)B$;

Préférences sur des résulats incertains

- Un joueur sera (le plus) souvent confronté à des choix dont le résultat est **probabiliste** ;
- On modélise ce type de résultats par des **lotteries** : par exemple $pA + qB + (1 - p - q)C$
  avec $p \geq 0, q \geq 0$ et $p + q \leq 1$.
- On fait quelques hypothèses sur les préférences sur les lotteries :
  1. **Monotonie** : si Alice préfère A à B alors pour $p > q$ elle préfère $pA + (1 - p)B$ à $qA + (1 - q)B$;
  2. **Décomposabilité** : si la probabilité **totale** d'obtenir chaque résultat possible dans les lotteries ℓ_1 et ℓ_2 est la même, alors Alice est indifférente entre ℓ_1 et ℓ_2
     par exemple $\frac{1}{3}(\frac{1}{2}A + \frac{1}{2}B) + \frac{2}{3}C \sim \frac{1}{6}A + \frac{1}{6}B + \frac{2}{3}C$

Préférences sur des résulats incertains

- Un joueur sera (le plus) souvent confronté à des choix dont le résultat est **probabiliste** ;
- On modélise ce type de résultats par des **lotteries** : par exemple $pA + qB + (1 - p - q)C$
  avec $p \geq 0, q \geq 0$ et $p + q \leq 1$.
- On fait quelques hypothèses sur les préférences sur les lotteries :
  1. **Monotonie** : si Alice préfère A à B alors pour $p > q$ elle préfère $pA + (1 - p)B$ à $qA + (1 - q)B$;
  2. **Décomposabilité** : si la probabilité **totale** d'obtenir chaque résultat possible dans les lotteries ℓ_1 et ℓ_2 est la même, alors Alice est indifférente entre ℓ_1 et ℓ_2
     par exemple $\frac{1}{3}(\frac{1}{2}A + \frac{1}{2}B) + \frac{2}{3}C \sim \frac{1}{6}A + \frac{1}{6}B + \frac{2}{3}C$
  3. **Substitutabilité** : si Alice est indifférente entre A et B et que ℓ' est obtenue à partir de ℓ en remplaçant A par B alors Alice est indifférente entre ℓ et ℓ'.

Préférences sur des résulats incertains

- Un joueur sera (le plus) souvent confronté à des choix dont le résultat est **probabiliste** ;
- On modélise ce type de résultats par des **lotteries** : par exemple $pA + qB + (1 - p - q)C$
  avec $p \geq 0, q \geq 0$ et $p + q \leq 1$.
- On fait quelques hypothèses sur les préférences sur les lotteries :
  1. **Monotonie** : si Alice préfère A à B alors pour $p > q$ elle préfère $pA + (1 - p)B$ à $qA + (1 - q)B$;
  2. **Décomposabilité** : si la probabilité **totale** d'obtenir chaque résultat possible dans les lotteries ℓ_1 et ℓ_2 est la même, alors Alice est indifférente entre ℓ_1 et ℓ_2
     par exemple $\frac{1}{3}(\frac{1}{2}A + \frac{1}{2}B) + \frac{2}{3}C \sim \frac{1}{6}A + \frac{1}{6}B + \frac{2}{3}C$
  3. **Substitutabilité** : si Alice est indifférente entre A et B et que ℓ' est obtenue à partir de ℓ en remplaçant A par B alors Alice est indifférente entre ℓ et ℓ'.
- Avec ces trois axiomes, on peut prouver que si pour Alice $A \preceq B \preceq C$ alors il existe une probabilité p telle que :
  - pour tout $p' < p$ elle préfère strictement B à $p'A + (1 - p')C$
  - et pour tout $p' > p$ elle préfère strictement $p'A + (1 - p')C$ à B

Préférences sur des résulats incertains

- Un joueur sera (le plus) souvent confronté à des choix dont le résultat est **probabiliste** ;
- On modélise ce type de résultats par des **lotteries** : par exemple $pA + qB + (1-p-q)C$
  avec $p \geq 0, q \geq 0$ et $p + q \leq 1$.
- On fait quelques hypothèses sur les préférences sur les lotteries :
  1. **Monotonie** : si Alice préfère A à B alors pour $p > q$ elle préfère $pA + (1-p)B$ à $qA + (1-q)B$;
  2. **Décomposabilité** : si la probabilité **totale** d'obtenir chaque résultat possible dans les lotteries ℓ_1 et ℓ_2 est la même, alors Alice est indifférente entre ℓ_1 et ℓ_2
     par exemple $\frac{1}{3}(\frac{1}{2}A + \frac{1}{2}B) + \frac{2}{3}C \sim \frac{1}{6}A + \frac{1}{6}B + \frac{2}{3}C$
  3. **Substitutabilité** : si Alice est indifférente entre A et B et que ℓ' est obtenue à partir de ℓ en remplaçant A par B alors Alice est indifférente entre ℓ et ℓ'.
- Avec ces trois axiomes, on peut prouver que si pour Alice $A \preceq B \preceq C$ alors il existe une probabilité p telle que :
  - pour tout $p' < p$ elle préfère strictement B à $p'A + (1-p')C$
  - et pour tout $p' > p$ elle préfère strictement $p'A + (1-p')C$ à B
- Pour obtenir qu'Alice est nécessairement indifférente entre B et $pA + (1-p)C$, il faut un 4e axiome :
  4. **Continuité** : si pour Alice $A \preceq B \preceq C$ alors il existe une probabilité p telle qu'Alice est indifférente entre B et $pA + (1-p)C$.

Utilité de Von Neumann & Morgenstern (VNM)

Théorème

Pour toute relation finie de préférences $\preceq$ satisfaisant les 4 propriétés précédentes, il existe une fonction u à valeurs dans $[0, 1]$ telle que :

1. $u(A) \leq u(B)$ *si et seulement si* $A \preceq B$ *;*
2. $u(p_1 A_1 + p_2 A_2 + \cdots + p_n A_n) = p_1 u(A_1) + p_2 u(A_2) + \cdots + p_n u(A_n)$

- Si u satisfait les conditions du théorème alors pour tout $a, b \in \mathbb{R}$, avec $a > 0$, alors $au + b$ les satisfait aussi ;
- Et aussi dans l'autre sens : si u et v sont deux fonctions satisfaisant les conditions du théorème, alors ils existe $a > 0$ et b tels que $v = au + b$;
- Donc l'utilité VNM est défine de manière unique **à une transformation affine près** ;
- L'unité de mesure de l'utilité est l'*util* ;
- Un joueur **rationnel** maximise l'**espérance de son utilité**, pas les quantités sous-jacentes.

Attitude face au risque

- Soient deux quantités x_1, x_2 d'un bien qui intéresse Bob,
- Soit X_i le résultat dans lequel Bob reçoit la quantité x_i de ce bien. On a $u(X_i) = u(x_i)$.
- Soit la lotterie $X = pX_1 + (1-p)X_2$;
- Si Bob est **averse au risque**, il préfère l'espérance d'une lotterie à la lotterie elle-même.

$$u\big(E(X)\big) \geq u(X)$$

- L'espérance de X est $E(X) = px_1 + (1-p)x_2$;
- L'utilité de X est par définition $u(X) = pu(X_1) + (1-p)u(X_2)$;
- Donc, avec $\forall i, u(X_i) = u(x_i)$:

$$u\big((px_1 + (1-p)x_2)\big) \geq pu(x_1) + (1-p)u(x_2)$$

- Donc l'utilité de X pour Bob est **concave** ;
- Si Bob est un **amoureux du risque**, son utilité pour X sera convexe ;
- Si Bob n'est ni un amoureux du risque, ni averse au risque, il est **neutre face au risque**.

Le paradoxe d'Allais

- On demande de choisir entre les lotteries suivantes :

A :

Gain	$0M€$	$1M€$	$5M€$
Prob.	0	1	0

B :

Gain	$0M€$	$1M€$	$5M€$
Prob.	0.01	0.89	0.1

- Et entre les deux suivantes :

C :

Gain	$0M€$	$1M€$	$5M€$
Prob.	0.89	0.11	0

D :

Gain	$0M€$	$1M€$	$5M€$
Prob.	0.9	0	0.1

Le paradoxe d'Allais

- On demande de choisir entre les lotteries suivantes :

A :

Gain	$0M€$	$1M€$	$5M€$
Prob.	0	1	0

B :

Gain	$0M€$	$1M€$	$5M€$
Prob.	0.01	0.89	0.1

- Et entre les deux suivantes :

C :

Gain	$0M€$	$1M€$	$5M€$
Prob.	0.89	0.11	0

D :

Gain	$0M€$	$1M€$	$5M€$
Prob.	0.9	0	0.1

Exercice

Montrer que les préférences $B \preceq A$ et $C \preceq D$ ne peuvent pas être représentées par une fonction d'utilité à la Von Neumann et Morgenstern.

Le paradoxe de Zeckhauser

- Un pistolet à 6 coups est braqué sur Alice ;
- On lui propose de payer pour enlever une balle du barillet avant d'appuyer sur la gachette ;
- Dans quelle situation est-elle prête à payer le plus si elle est rationnelle ?
  1. le barillet contient 1 balle ;
  2. le barillet contient 4 balles.

Le paradoxe de Zeckhauser

- Un pistolet à 6 coups est braqué sur Alice ;
- On lui propose de payer pour enlever une balle du barillet avant d'appuyer sur la gachette ;
- Dans quelle situation est-elle prête à payer le plus si elle est rationnelle ?
  1. le barillet contient 1 balle ;
  2. le barillet contient 4 balles.

Exercice

Soit X la somme qu'Alice paye dans le premier cas, et soit Y cette somme dans le 2e cas. On suppose que si Alice est morte après le tir elle est indifférente à avoir payé ou pas. Soit L le résultat « en vie après le tir » et D le résultat « mort après le tir ».

1. Quand X est maximum, Alice est indifférente entre deux lotteries. Lesquelles ?
2. Pareil pour Y ;
3. Si Alice préfère ne pas pas mourir et payer le moins possible, en déduire qu'on a forcément $Y_{max} > X_{max}$;
4. Combien Alice serait-elle prête à payer dans le cas où le barillet contient initialement 6 balles ?

Le paradoxe d'Ellsberg

- Une urne contient 300 balles, dont 100 sont rouges ;
- Les 200 autres sont noires ou blanches mais on ne sait pas dans quelle proportion ;
- Alice doit choisir entre ces deux lotteries avant de tirer une balle au hasard :

A :

Résultat	Rouge	Noire	Blanche
Gain	$1M€$	$0M€$	$0M€$

B :

Résultat	Rouge	Noire	Blanche
Gain	$0M€$	$1M€$	$0M€$

- On lui demande de répéter l'exercice avec ces deux lotteries :

C :

Résultat	Rouge	Noire	Blanche
Gain	$0M€$	$1M€$	$1M€$

D :

Résultat	Rouge	Noire	Blanche
Gain	$1M€$	$0M€$	$1M€$

Le paradoxe d'Ellsberg

- Une urne contient 300 balles, dont 100 sont rouges ;
- Les 200 autres sont noires ou blanches mais on ne sait pas dans quelle proportion ;
- Alice doit choisir entre ces deux lotteries avant de tirer une balle au hasard :

A :

Résultat	Rouge	Noire	Blanche
Gain	$1M€$	$0M€$	$0M€$

B :

Résultat	Rouge	Noire	Blanche
Gain	$0M€$	$1M€$	$0M€$

- On lui demande de répéter l'exercice avec ces deux lotteries :

C :

Résultat	Rouge	Noire	Blanche
Gain	$0M€$	$1M€$	$1M€$

D :

Résultat	Rouge	Noire	Blanche
Gain	$1M€$	$0M€$	$1M€$

Exercice

Alice essaie d'estimer le nombre de balles noires. Soit N son estimation.

1. Que dire de N si Alice préfère strictement A à B ?
2. Que dire de N si Alice préfère strictement C à D ?

Le paradoxe de Saint Petersbourg

Exercice

- On considère le jeu suivant :
  1. Alice paie la somme x à Bob ;
  2. Caroline lance une pièce (non biaisée, etc.) jusqu'à obtenir face ;
  3. Bob paie à Alice la somme 2^n où n est le nombre de lancers réalisés.
- Quelle doit être la mise de départ d'Alice pour que le jeu soit équitable ?

Utilité de Von Neumann et Morgenstern

- L'utilité de Von Neumann et Morgenstern, et le principe de maximisation de son espérances, ne modélisent pas toujours bien les **comportements humains** :
  - Les êtres humains n'agissent pas toujours de façon rationnelle ;
  - Phénomènes psychologiques : aversion pour l'incertain, aversion pour le zéro,...
  - L'espérance mathématique n'est pas forcément une bonne représentation pour toutes les distributions.
- Elle reste une approximation utile pour laquelle on peut avoir une analyse mathématique rigoureuse ;
- Elle est une bon choix pour la conception d'**agents artificiels** rationnels.

Outline

Jeu à deux joueurs strictement compétitif

Définition

Le jeu à deux joueurs G est **strictement compétitif** si pour tous profils de stratégies (s_1, s_2) et (s_1', s_2'), on a $u_1(s_1, s_2) \leq u_1(s_1', s_2')$ si et seulement si $u_2(s_1, s_2) \geq u_2(s_1', s_2')$.

Jeu à deux joueurs strictement compétitif

Définition

Le jeu à deux joueurs G est **strictement compétitif** si pour tous profils de stratégies (s_1, s_2) et (s_1', s_2'), on a $u_1(s_1, s_2) \leq u_1(s_1', s_2')$ si et seulement si $u_2(s_1, s_2) \geq u_2(s_1', s_2')$.

- Dans un jeu strictement compétitif, les résultats sont toujours **Pareto-efficaces**
  ce qui constitue une définition équivalente

Jeu à deux joueurs strictement compétitif

Définition

Le jeu à deux joueurs G est **strictement compétitif** si pour tous profils de stratégies (s_1, s_2) et (s_1', s_2'), on a $u_1(s_1, s_2) \leq u_1(s_1', s_2')$ si et seulement si $u_2(s_1, s_2) \geq u_2(s_1', s_2')$.

- Dans un jeu strictement compétitif, les résultats sont toujours **Pareto-efficaces**
  ce qui constitue une définition équivalente

Exemple

Matching pennies, les échecs, le go, le backgammon, etc. sont des jeux strictement compétitifs.

Jeu à deux joueurs strictement compétitif

Exercice

Alice et Bob choisissent chacun un nombre entier entre 0 et 2 inclus.

- Celui qui a choisi le plus grand nombre reçoit 1 et l'autre 0 ;
- S'il y a égalité, tous deux reçoivent 0.

1. Pourquoi n'est-ce pas un jeu strictement compétitif ?
2. Comment modifier le jeu pour qu'il soit strictement compétitif, en en gardant l'esprit ?

Équilibres de Nash dans les jeux strictements compétitifs

- Le joueur 1 peut assurer un gain d'au moins maxmin_1 ;
- Le joueur 2 peut assurer que le gain du joueur n'est pas plus que minmax_1.

Théorème

Dans un jeu strictement compétitif les trois assertions suivantes sont équivalentes :

1. *il y a un équilibre de Nash (s_1^*, s_2^*) en stratégies pures (resp. mixtes) ;*
2. *pour tout $i \in \{1, 2\}$, $\mathrm{maxmin}_i = \mathrm{minmax}_i$ en stratégies pures (resp. mixtes) ;*
3. *pour tout $i \in \{1, 2\}$, s_i^* est une stratégie de sécurité pour le joueur i.*

- Donc tous les équilibres de Nash sont **équivalents** : leur valeur est $(\mathrm{maxmin}_1, \mathrm{maxmin}_2)$;
- Et ils sont **interchangeables** : si (s_1, s_2) et (s_1', s_2') sont des équilibres de Nash alors (s_1', s_2) et (s_1, s_2') aussi.

Équilibres de Nash dans les jeux strictements compétitifs

- En stratégies mixtes, on a vu indépendamment que :
  - il existe toujours un équilibre de Nash
  - on a toujours $\mathrm{maxmin}_i = \mathrm{minmax}_i$
- En stratégies pures, aucun des deux n'est forcément vrai, mais si l'un l'est, l'autre aussi ;
- On peut calculer maxmin_i et une stratégie de sécurité, et donc un équilibre de Nash, en **temps polynomial** par programmation linéaire.

Équilibres de Nash dans les jeux strictements compétitifs

Exercice

Alice et Bob choisissent un nombre entier entre 0 et 2 inclus.

- Celui qui a choisi le plus grand nombre reçoit 3 et l'autre 0 ;
- S'il y a égalité, tous deux reçoivent 1.

1. Calculer les maxmin et minmax en stratégies pures ;
2. Calculer les stratégies (y compris mixtes) de sécurité des deux joueurs et en déduire les équilibres de Nash.

Jeux à somme nulle

Définition

Un jeu à deux joueurs est dit **à somme nulle** si pour tout profil de stratégie s, on a $u_i(s) = -u_{-i}(s)$.

- L'utilité étant définie à une transformation affine près, ils sont plutôt à **somme constante** ;
- On peut ne définir le jeu qu'avec u_1 ;

Jeux à somme nulle

Définition

Un jeu à deux joueurs est dit **à somme nulle** si pour tout profil de stratégie s, on a $u_i(s) = -u_{-i}(s)$.

- L'utilité étant définie à une transformation affine près, ils sont plutôt à **somme constante** ;
- On peut ne définir le jeu qu'avec u_1 ;
- On peut montrer une équivalence avec les jeux strictement compétitifs :
  - un jeu à somme nulle est clairement strictement compétitif ;
  - pour un ≪ bon ≫ choix de fonctions d'utilité, un jeu strictement compétitif est à somme nulle.

Jeux à somme nulle

Définition

Un jeu à deux joueurs est dit **à somme nulle** si pour tout profil de stratégie s, on a $u_i(s) = -u_{-i}(s)$.

- L'utilité étant définie à une transformation affine près, ils sont plutôt à **somme constante** ;
- On peut ne définir le jeu qu'avec u_1 ;
- On peut montrer une équivalence avec les jeux strictement compétitifs :
  - un jeu à somme nulle est clairement strictement compétitif ;
  - pour un ≪ bon ≫ choix de fonctions d'utilité, un jeu strictement compétitif est à somme nulle.
- Le ≪ bon ≫ choix n'est pas forcément réaliste : le poker avec deux joueurs averses au risque est strictement compétitif mais pas à somme nulle

  si u_1 est concave alors $u_2 = -u_1$ est convexe.

Jeux à somme nulle

Définition

Un jeu à deux joueurs est dit **à somme nulle** si pour tout profil de stratégie s, on a $u_i(s) = -u_{-i}(s)$.

- L'utilité étant définie à une transformation affine près, ils sont plutôt à **somme constante** ;
- On peut ne définir le jeu qu'avec u_1 ;
- On peut montrer une équivalence avec les jeux strictement compétitifs :
  - un jeu à somme nulle est clairement strictement compétitif ;
  - pour un ≪ bon ≫ choix de fonctions d'utilité, un jeu strictement compétitif est à somme nulle.
- Le ≪ bon ≫ choix n'est pas forcément réaliste : le poker avec deux joueurs averses au risque est strictement compétitif mais pas à somme nulle
  si u_1 est concave alors $u_2 = -u_1$ est convexe.
- Dans un jeu à deux joueur somme nulle :

$$\forall i, \mathrm{maxmin}_i = -\mathrm{minmax}_{-i}$$

Valeur d'un jeu à somme nulle

Définition (Valeur)

Un jeu à deux joueurs a une **valeur** si $\text{maxmin}_1 = \text{minmax}_1$. La valeur du jeu est alors $v = \text{maxmin}_1 = \text{minmax}_1$.

- Tout jeu fini à somme nulle a donc une valeur en stratégies mixtes, mais pas forcément en stratégies pures.

Point selle

Définition

Un profil de stratégies (s_1^*, s_2^*) est un point selle de la fonction d'utilité u (du joueur 1) si :

$$\begin{cases} \forall s_1 \in S_1, u(s_1^*, s_2^*) \geq u(s_1, s_2^*) \\ \forall s_2 \in S_2, u(s_1^*, s_2^*) \leq u(s_1^*, s_2) \end{cases}$$

- Pour les stratégies pures, $u(s_1^*, s_2^*)$ est la plus grande valeur de sa colonne et la plus petite de sa ligne ;
- (s_1^*, s_2^*) est un point selle de la fonction u du joueur 1 si et seulement si c'est un équilibre de Nash car s_1^* et s_2^* sont des stratégies de sécurité.

Exemple

Si la valeur du jeu d'échecs est ≪ nulle ≫, et qu'on écrit toutes les stratégies possibles de chacun des deux joueurs dans une matrice de jeu en forme normale, alors il existe :

- une ligne qui ne contient que ≪ victoire des blancs ≫ et ≪ nulle ≫ ;
- une colonne qui ne contient que ≪ défaite des blancs ≫ (= ≪ victoire des noirs ≫) et ≪ nulle ≫ ;
- à leur intersection, on a ≪ nulle ≫.

Équilibres corrélés dans les jeux à somme nulle

- Un équilibre corrélé π donne une stratégie à chacun des deux joueurs ;
- Dans un jeu à deux joueurs à somme nulle :
  - Tout équilibre corrélé a la même valeur que les équilibres de Nash (stratégies mixtes) ;
  - Le profil de stratégies s^π où le joueur i joue a_i avec probabilité $\pi(a_i) = \sum_{a_{-i}} \pi(a_i, a_{-i})$ est un équilibre de Nash ;
- La coordination ne permet pas d'améliorer le résultat
  est-ce étonnant ?

Un jeu d'attente : Duel [1]

Exercice

Duel. Bob a insulté Alice et elle demande réparation. Ayant le choix des armes, elle choisit le pistolet (à eau) à 1 coup. Les deux duellistes avancent l'un vers l'autre en partant d'une distance D. Si Alice tire à la distance $0 \leq d \leq D$, elle a une probabilité $p_A(d)$ de toucher Bob et gagner. Si elle le rate, Bob aura alors tout son temps pour arriver à bout portant et gagner. La situation pour Bob est similaire mais sa probabilité de toucher est $p_B(d)$. Les deux fonctions sont décroissantes. Si les deux choisissent de tirer à la même distance, on suppose qu'un joueur qui tire avant l'autre est choisi avec une certaine probabilité $p_=(d)$ et le reste se déroule comme avant.

À quelle distance Alice et Bob doivent-ils tirer pour maximiser leur chance de gagner ?

Outline

Outline

Forme extensive

- Pour représenter un jeu complexe comme les échecs, la forme normale n'est pas pratique ;
- L'aspect **séquentiel** n'est pas présent ;
- On définit un nouveau type de jeux : les jeux sous **forme extensive** ;

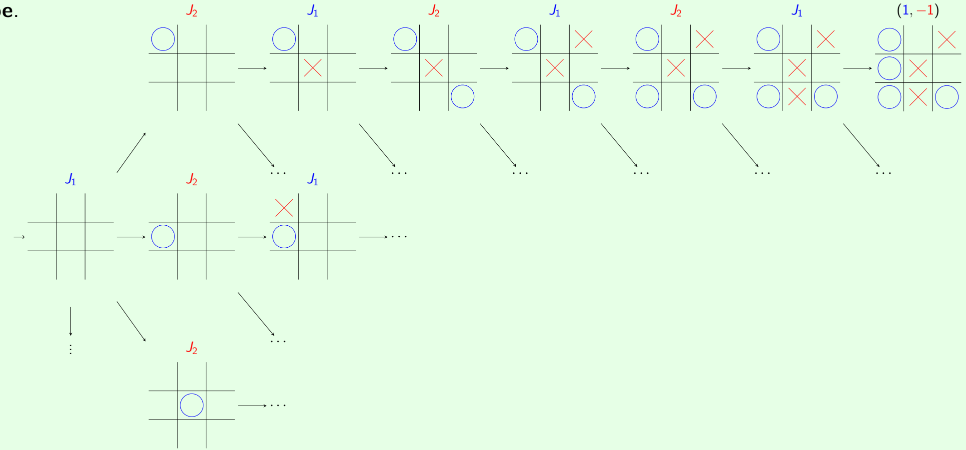
Forme extensive

- Pour représenter un jeu complexe comme les échecs, la forme normale n'est pas pratique ;
- L'aspect **séquentiel** n'est pas présent ;
- On définit un nouveau type de jeux : les jeux sous **forme extensive** ;
- Un jeu est représenté par un **arbre** enraciné (**Arbre de jeu**, *Game tree*) :
  - les sommets sont les positions / états dans le jeu
  - chaque sommet **appartient à un unique joueur** ;
  - la racine est la position initiale ;
  - les arêtes sont les **actions** jouables ;
  - les feuilles donnent le gain / utilité pour chaque joueur.

Forme extensive : exemple

Exemple

Tic-tac-toe.



Forme extensive : exercice

Exercice

Kayles. On considère une rangée de n quilles et deux joueurs. Chaque joueur, à son tour, abat une quille ou deux qui sont adjacentes. Le perdant est celui qui abat la dernière quille et l'autre est le gagnant.

Dessiner la forme extensive de ce jeu pour $n = 4$. On pourra regrouper les configurations issues d'une même configuration et qui sont les mêmes aux symétries près.

Frome extensive : stratégies

Définition

Une **stratégie pure** dans un jeu en forme extensive est une fonction qui à chaque sommet associe une action du joueur possédant le sommet.

- Une fois fixée une stratégie des deux joueurs, on a un **unique chemin** dans l'arbre ;
- La profondeur étant finie, on obtient une unique feuille et donc un **gain d'utilité** unique pour chaque joueur.

Frome extensive : stratégies

Définition

Une **stratégie pure** dans un jeu en forme extensive est une fonction qui à chaque sommet associe une action du joueur possédant le sommet.

- Une fois fixée une stratégie des deux joueurs, on a un **unique chemin** dans l'arbre ;
- La profondeur étant finie, on obtient une unique feuille et donc un **gain d'utilité** unique pour chaque joueur.
- Comme précédemment, on définit une **stratégie mixte** comme une distribution de probabilités sur les stratégies pures.

Frome extensive : stratégies

Définition

Une **stratégie pure** dans un jeu en forme extensive est une fonction qui à chaque sommet associe une action du joueur possédant le sommet.

- Une fois fixée une stratégie des deux joueurs, on a un **unique chemin** dans l'arbre ;
- La profondeur étant finie, on obtient une unique feuille et donc un **gain d'utilité** unique pour chaque joueur.
- Comme précédemment, on définit une **stratégie mixte** comme une distribution de probabilités sur les stratégies pures.

Définition

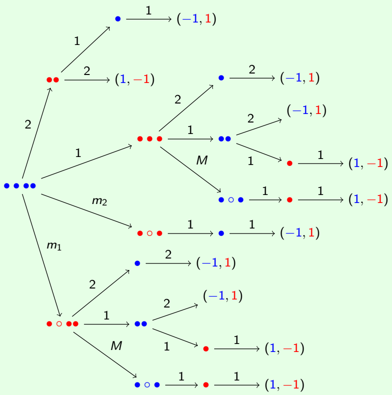
Une **stratégie comportementale** (*behavioral*) dans un jeu en forme extensive est une fonction qui à chaque sommet associe une **distribution de probabilités** sur les actions du joueur possédant le sommet.

- Pour toute stratégie comportementale, on peut construire une stratégie mixte avec, pour toutes les stratégies adverses, les mêmes probabilités d'arriver à chaque état
  et donc qui engendre les mêmes utilités
- Et inversement.

Forme extensive et forme normale

- Pour chaque jeu en forme extensive on peut définir sa **forme normale** ;
- Chaque **action** de la forme normale consiste à choisir une stratégie de la forme extensive ;
- La forme normale est en général de taille **exponentielle** par rapport à la forme extensive.

Exemple



	111	211	121	221	1M1	2M1	112	212	122	222	1M2	2M2	11M	21M	12M	22M	1MM	2MM
2	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1
11	1	1	-1	-1	1	1	1	1	-1	-1	1	1	1	1	-1	-1	1	1
12	-1	-1	-1	-1	1	1	-1	-1	-1	-1	1	1	-1	-1	-1	-1	1	1
m_2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
$m_1 1$	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1
$m_1 2$	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1

- les stratégies pures du joueur 1 sont décrites par ses deux points de choix accessibles (**stratégie réduite**) de haut en bas puis de gauche à droite ;
- les stratégies pures du joueur 2 sont décrites par ses trois points de choix de haut en bas ;
- le jeu est à somme nulle et on n'a donc écrit que la matrice du joueur 1 ;
- le jeu a une **valeur en stratégies pures** qui est -1 ;
- les stratégies d'équilibres sont 122 pour le joueur 2 et n'importe laquelle pour le joueur 1.

Jeux gagnants-perdants

Théorème (Von Neumann, 1928)

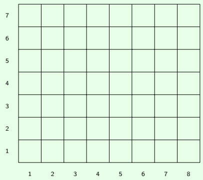
Dans tout jeu à deux joueurs sous forme extensive dans lequel les résultats sont (assimilables à) ≪ le joueur 1 gagne ≫, ≪ le joueur 2 gagne ≫, et ≪ partie nulle ≫, on a :

- *soit le joueur 1 a une stratégie pour gagner à coup sûr ;*
- *soit le joueur 2 a une stratégie pour gagner à coup sûr ;*
- *soit les deux joueurs ont une stratégie pour forcer une partie nulle.*

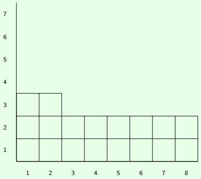
Jeux gagnants-perdants : Chomp

Exemple

Chomp de David Gale. Ce jeu à deux joueurs se joue sur une grille de taille $m \times n$, dont les cases sont représentées par leur coordonnées (i, j) ($i \geq 1$ horizontal, $j \geq 1$ vertical). Chacun son tour, un joueur capture une case (i, j) telle qu'aucune case (i', j') avec $i' \leq i$ et $j' \leq j$ n'a déjà été capturée. Le joueur 1 commence et celui qui capture la case $(1, 1)$ a perdu.



Grille 8×7 initiale



Après capture de $(1, 4)$ et $(3, 3)$

Exercice

1. Exhiber une stratégie gagnante pour le joueur 1 sur une grille carrée $(m = n)$;
2. Montrer qu'en général, si le joueur 2 a une stratégie gagnante, alors le joueur 1 aussi ;
3. Conclure.

Jeux de hasard

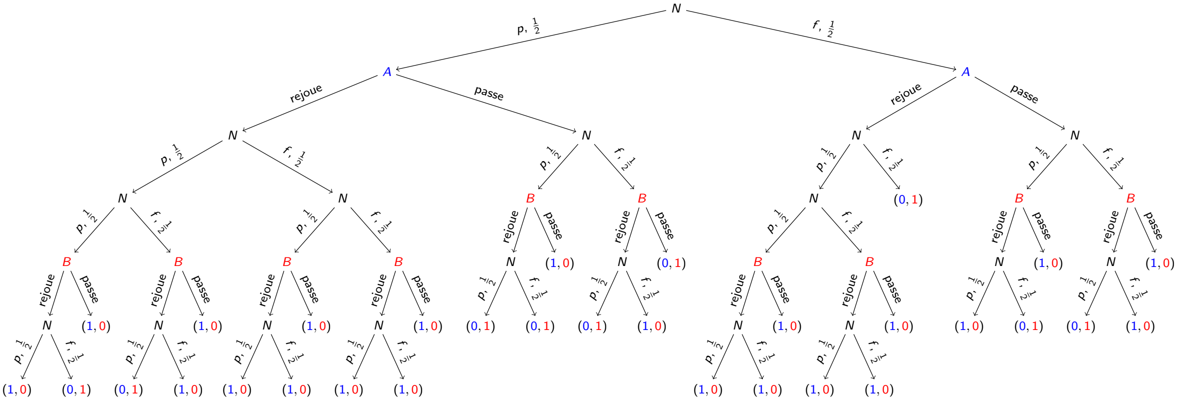
- On peut ajouter des phénomènes **aléatoires** dans les jeux sous-forme extensive ;
- On ajoute un joueur 0, *Chance* ou *Nature*, dont les actions portent des probabilités et sont résolus selon ces probabilités ;
- Pour retrouver la forme normale, on met alors les **espérances de gains** dans la table ;
- Le théorème de Von Neumann n'est plus vrai dans ce cas.

Exercice

Blackjack à pièces. Alice lance une pièce, puis éventuellement une deuxième. Elle compte 1 pour pile et 2 pour face. Si elle obtient 4, Bob gagne. Sinon, Bob lance une pièce, puis éventuellement une deuxième. S'il obtient 3 ou moins mais plus qu'Alice, Bob gagne, sinon c'est Alice qui gagne. Celui ou celle qui gagne obtient 1 et l'autre 0.

Modéliser (au moins partiellement) ce jeu sous forme extensive et déterminer le gain (moyen) des deux joueurs s'ils passent tous les deux tout le temps.

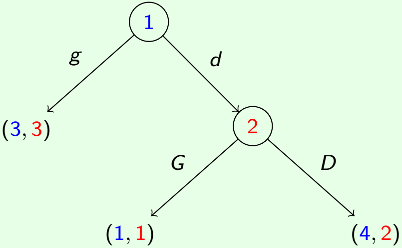
Blackjack à pièces



Outline

Menaces non crédibles

Exemple



J1 \ J2	G	D
g	(3, 3)	(3, 3)
d	(1, 1)	(4, 2)

- gG est un équilibre de Nash ;
- gD n'est pas un équilibre ;
- La **menace** du joueur 2 (G) lui permet d'augmenter son gain ;
- Mais elle n'est pas **crédible** car :
  1. elle concerne une partie du jeu qui n'est pas atteinte ;
  2. si elle était atteinte, cette stratégie ne serait pas optimale pour le joueur 2.

Équilibre parfait en sous-jeux

- La forme normale cache les aspects temporels de la forme extensive ;
- Certains équilibres ne sont pas réalistes ;
- On raffine la notion d'équilibre de Nash par l'**équilibre parfait en sous-jeux** ;
- Un sous-jeu de la représentation extensive est un sous-arbre

Définition

Dans un jeu sous forme extensive, un **équilibre parfait en sous-jeux** (*subgame perfect equilibrium*, SPE) est un équilibre de Nash tel que sa restriction à tout sous-jeu est aussi un équilibre de Nash.

Exemple

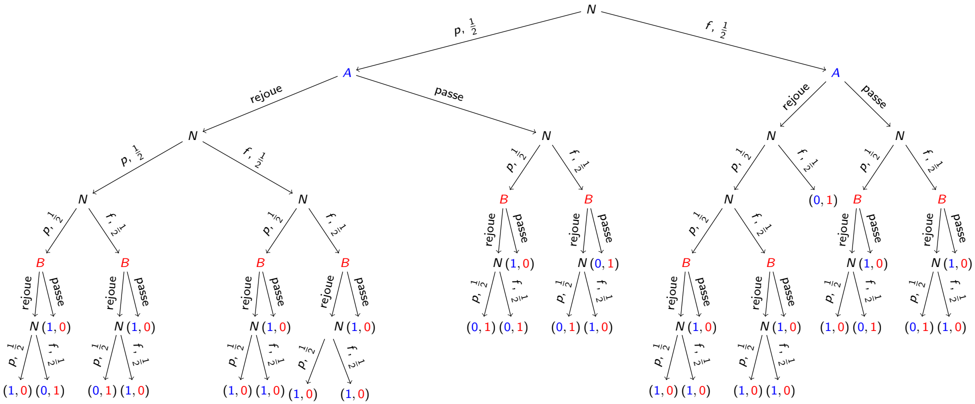
Dans l'exemple précédent :

- gG n'est pas un SPE car G n'est pas un NE du sous-jeu obtenu après d.
- dD est un SPE.

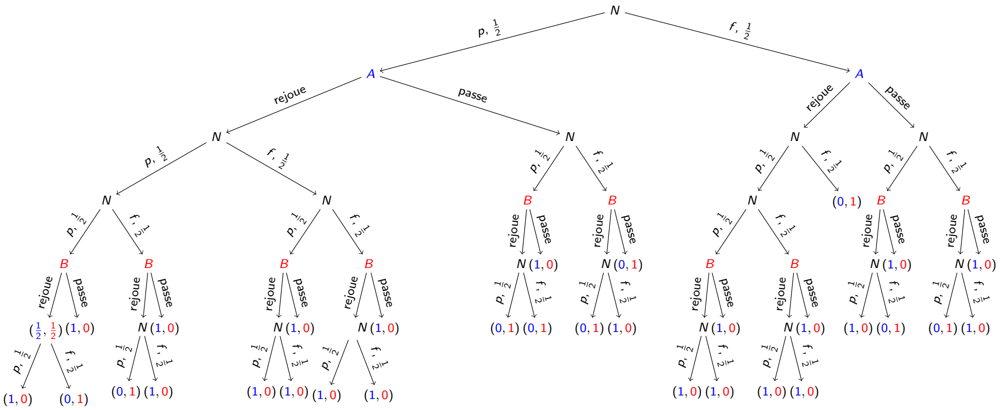
Induction en arrière

- Pour calculer les SPE, on utilise l'algorithme d'**induction arrière**.
- Un jeu réduit à une feuille est un SPE (dégénéré) ;
- Dans un jeu avec un seul sommet (hors *Chance*) qui n'est pas une feuille, et qui est la racine de l'arbre, un NE (qui est aussi un SPE) choisit nécessairement une action qui conduit au meilleur gain (en espérance) ;
- On procède donc en arrière depuis les feuilles :
  1. pour tout sous-jeu avec une racine n'appartenant pas à *Chance* et qui ne contient en plus que des feuilles ou des sommets *Chance*, on le remplace par une feuille dont la valeur est celle de l'un des NE du sous-jeu.
  2. on recommence jusqu'à remonter à la racine du jeu.
- On peut aussi implémenter cela simplement en étiquetant les sommets avec la valeur de leur sous-jeu en remontant.

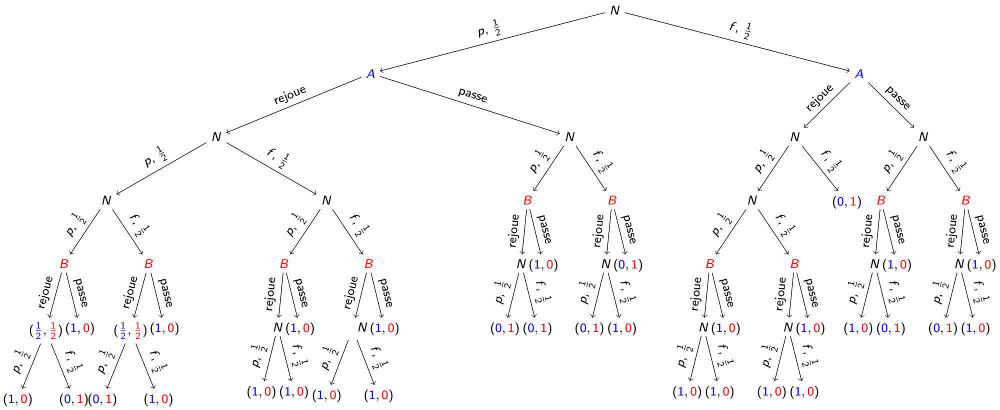
SPEs dans le Blackjack à pièces



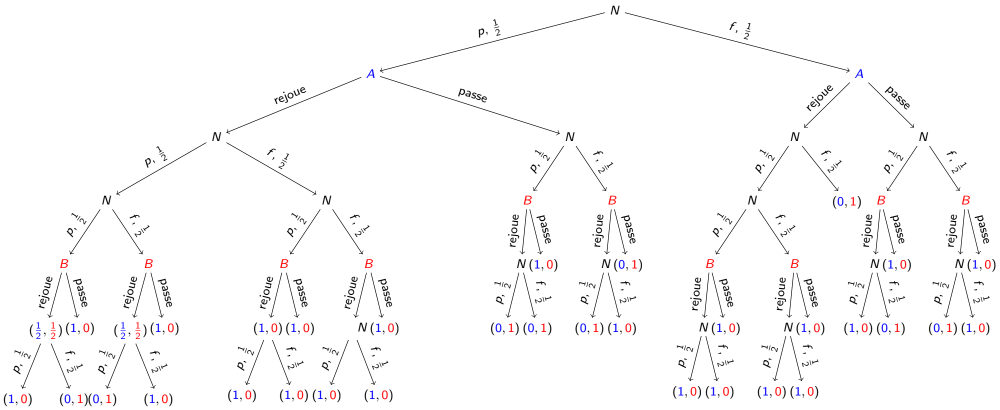
SPEs dans le Blackjack à pièces



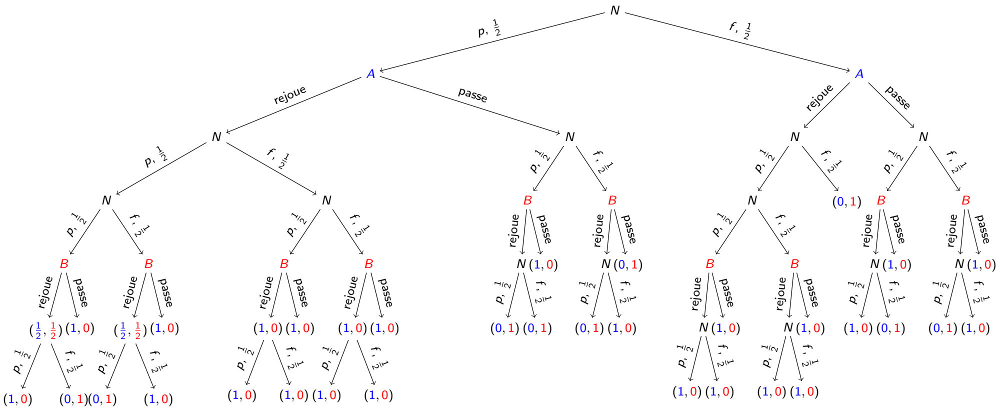
SPEs dans le Blackjack à pièces



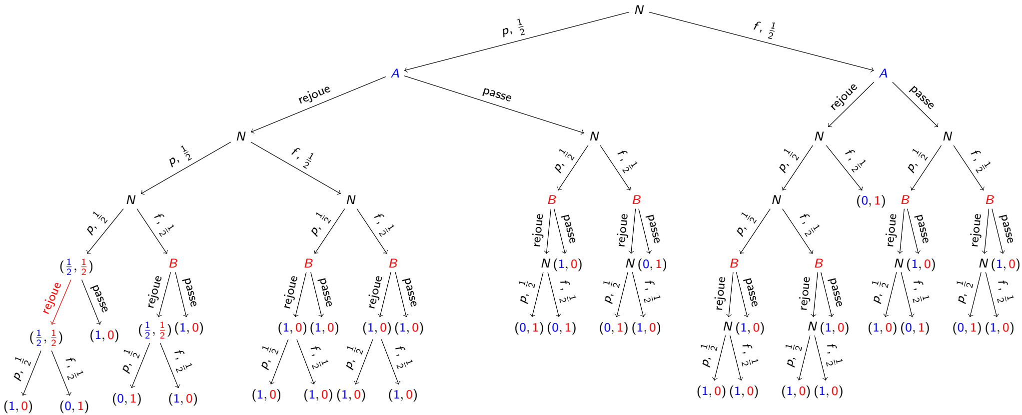
SPEs dans le Blackjack à pièces



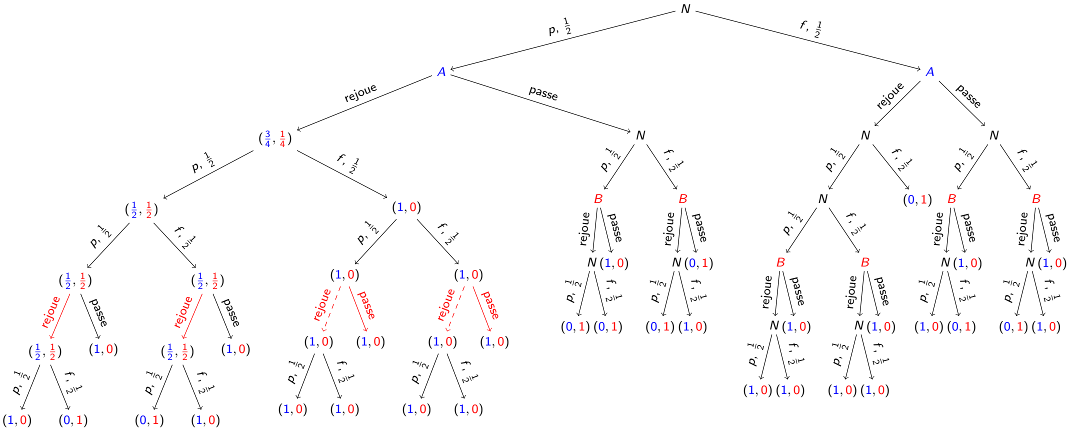
SPEs dans le Blackjack à pièces



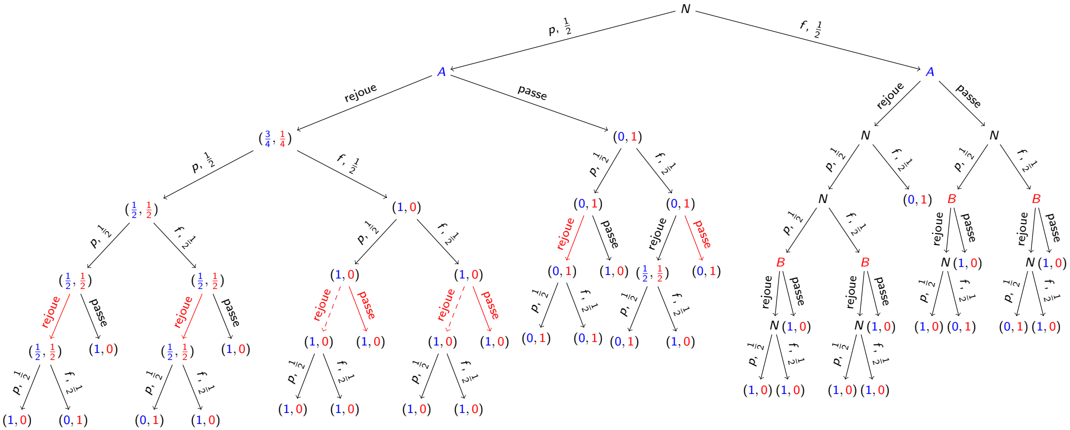
SPEs dans le Blackjack à pièces



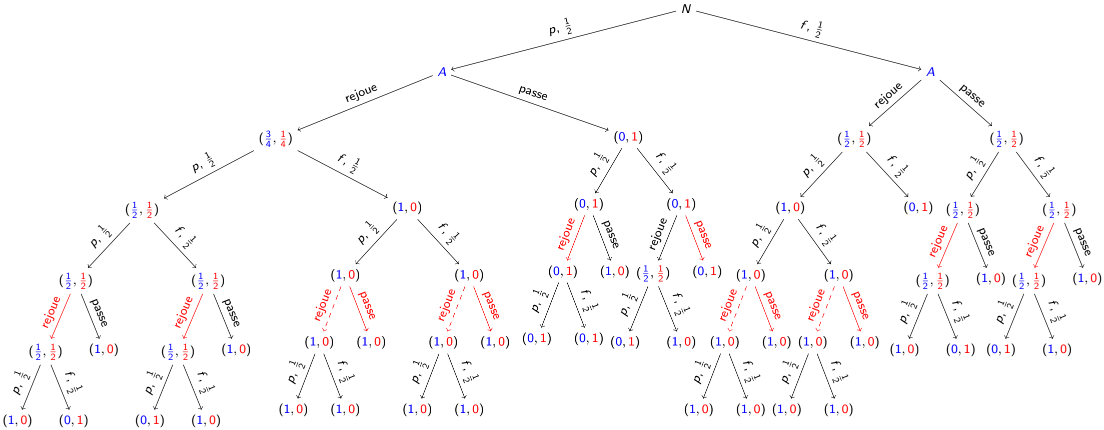
SPEs dans le Blackjack à pièces



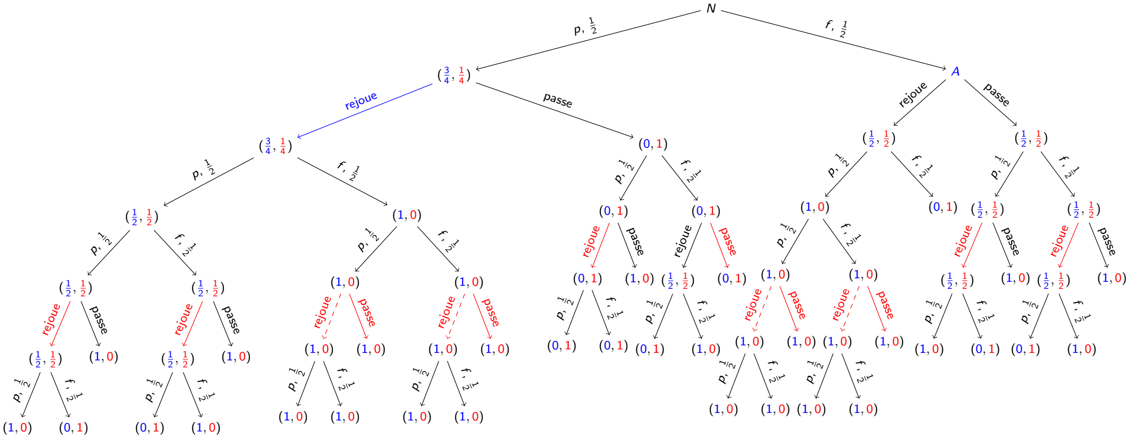
SPEs dans le Blackjack à pièces



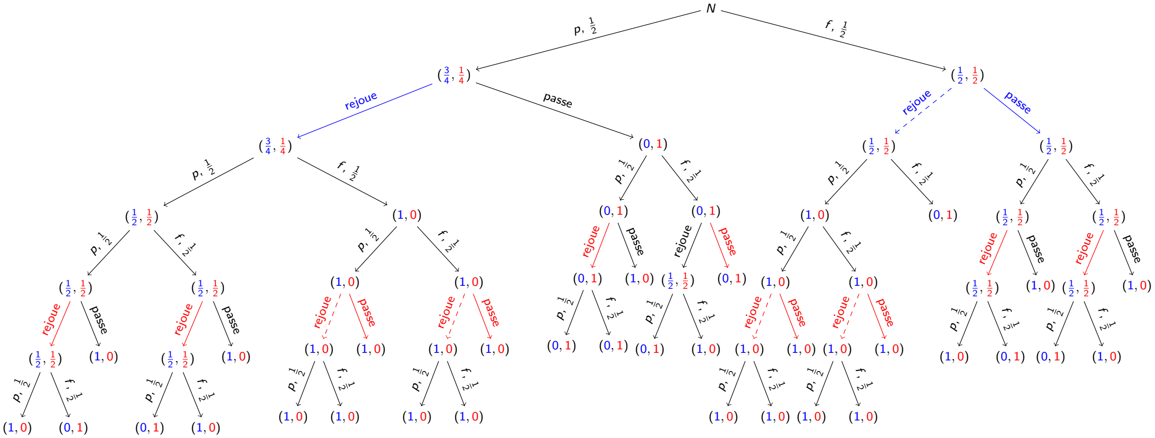
SPEs dans le Blackjack à pièces



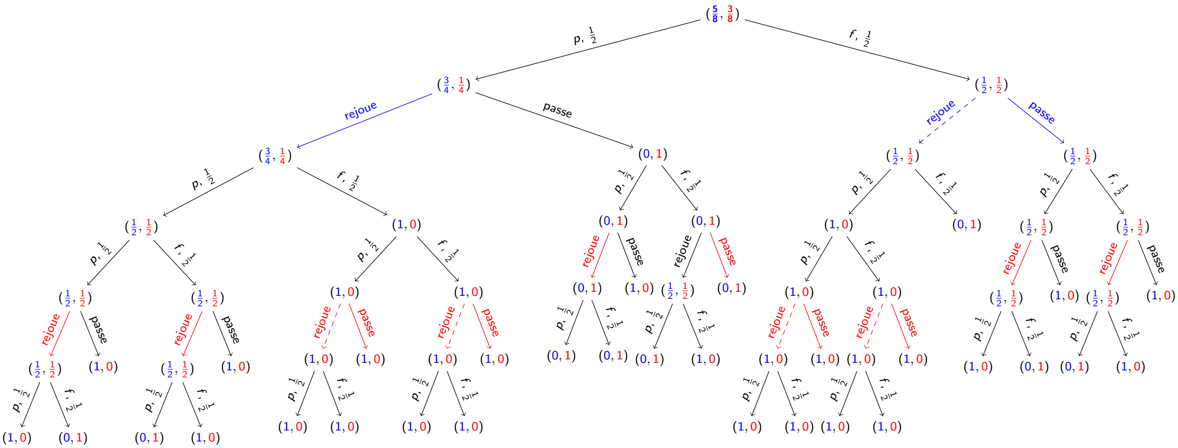
SPEs dans le Blackjack à pièces



SPEs dans le Blackjack à pièces



SPEs dans le Blackjack à pièces



Équilibres de Nash

Théorème

Tout SPE d'un jeu en forme extensive est un équilibre de Nash mais l'inverse n'est pas nécessairement vrai.

Théorème

Tout jeu fini sous forme extensive a une SPE en stratégies pures
et donc il a aussi un équilibre de Nash

SPE : Exercice

Exercice

Mille-patte (Centipede) [1]. C'est le matin de Noël et Alice et Bob découvrent leurs cadeaux. Alice a eu un sac de 100 billes et Bob une boîte de 100 briques en plastique à assembler. Mais Alice préfère 10 fois plus les briques aux billes et Bob préfère 10 fois plus les billes aux briques.
Ils décident donc de procéder à un échange, mais ne se faisant aucune confiance, un objet à la fois : Alice décide d'abord soit d'interrompre l'échange, soit de donner une bille à Bob, puis Bob décide soit d'interrompre l'échange, soit de donner une brique à Alice, et ainsi de suite jusqu'à échange complet.

1. Écrire (partiellement) ce jeu sous forme extensive et trouver l'unique SPE ;
2. Quelles stratégies obtient-on en supprimant celles qui ne minimisent pas le regret ?

ϵ-équilibre (de Nash)

Définition

Un profil de stratégies s est un ϵ-**équilibre** (pour $\epsilon > 0$) si $\forall i, \forall s_i', u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i}) - \epsilon$.

Exemple

Dans *Mille-patte*, échanger tous les objets est un 1-équilibre.

- Un équilibre de Nash est un ϵ-équilibre pour tout $\epsilon > 0$.
- L'utilité pour chaque joueur d'un ϵ-équilibre n'est pas forcément proche à ϵ près du gain d'un équilibre de Nash ;
- Pour qu'il soit possible, il faut que le fait que le joueur est indifférent à ϵ près soit une **connaissance commune**.

Exemple

Alice \ Bob	g	d
h	1, 1	0, 0
b	$1+\epsilon$, 1	100, 100

Exemple adapté de [3].

Calcul d'un SPE dans les gros jeux

- Dans de nombreux problèmes réels le jeu est énorme **énorme** : Backgammon 10^{20} états, Chess 10^{40}, Go 10^{170}, faire voler un hélicoptère $10^{?}$, ...

Calcul d'un SPE dans les gros jeux

- Dans de nombreux problèmes réels le jeu est énorme **énorme** : Backgammon 10^{20} états, Chess 10^{40}, Go 10^{170}, faire voler un hélicoptère $10^{?}$, ...
- Deux idées naturelles pour gérer cette complexité :

Calcul d'un SPE dans les gros jeux

- Dans de nombreux problèmes réels le jeu est énorme **énorme** : Backgammon 10^{20} états, Chess 10^{40}, Go 10^{170}, faire voler un hélicoptère $10^{?}$, ...
- Deux idées naturelles pour gérer cette complexité :
  - réduire la **profondeur** de l'arbre ;

Calcul d'un SPE dans les gros jeux

- Dans de nombreux problèmes réels le jeu est énorme **énorme** : Backgammon 10^{20} états, Chess 10^{40}, Go 10^{170}, faire voler un hélicoptère $10^{?}$, ...
- Deux idées naturelles pour gérer cette complexité :
  - réduire la **profondeur** de l'arbre ;
  - réduire le **branchement** de l'arbre.

Outline

Réduire la profondeur

- On se donne une **profondeur maximale** L ;

Réduire la profondeur

- On se donne une **profondeur maximale** L ;
- On remplace tous les sous-jeux à la profondeur L par un gain direct pour les joueurs ;

Réduire la profondeur

- On se donne une **profondeur maximale** L ;
- On remplace tous les sous-jeux à la profondeur L par un gain direct pour les joueurs ;
- Ce gain est calculé par une fonction d'**évaluation statique** $\tilde{U}$:

Réduire la profondeur

- On se donne une **profondeur maximale** L ;
- On remplace tous les sous-jeux à la profondeur L par un gain direct pour les joueurs ;
- Ce gain est calculé par une fonction d'**évaluation statique** $\tilde{U}$:
  - Obtenue via des experts ;

Réduire la profondeur

- On se donne une **profondeur maximale** L ;
- On remplace tous les sous-jeux à la profondeur L par un gain direct pour les joueurs ;
- Ce gain est calculé par une fonction d'**évaluation statique** $\tilde{U}$:
  - Obtenue via des experts ;
  - Ou par apprentissage (p. ex. apprentissage par renforcement).

Expectiminimax et Minimax

- On considère un **jeu à somme nulle** ;

Expectiminimax et Minimax

- On considère un **jeu à somme nulle** ;
- Étant donné un état s, la profondeur maximum L, et la fonction d'évaluation statique $\tilde{U}$, **Expectiminimax** consiste à :

Expectiminimax et Minimax

- On considère un **jeu à somme nulle** ;
- Étant donné un état s, la profondeur maximum L, et la fonction d'évaluation statique $\tilde{U}$, **Expectiminimax** consiste à :
  1. construire le sous-jeu dont la racine est s et de pronfondeur au plus L ;

Expectiminimax et Minimax

- On considère un **jeu à somme nulle** ;
- Étant donné un état s, la profondeur maximum L, et la fonction d'évaluation statique $\tilde{U}$, **Expectiminimax** consiste à :
  1. construire le sous-jeu dont la racine est s et de pronfondeur au plus L ;
  2. remplacer toutes feuilles s' dans lesquelles un joueur pourrait encore jouer par $\tilde{U}(s')$;

Expectiminimax et Minimax

- On considère un **jeu à somme nulle** ;
- Étant donné un état s, la profondeur maximum L, et la fonction d'évaluation statique $\tilde{U}$, **Expectiminimax** consiste à :
  1. construire le sous-jeu dont la racine est s et de pronfondeur au plus L ;
  2. remplacer toutes feuilles s' dans lesquelles un joueur pourrait encore jouer par $\tilde{U}(s')$;
  3. calculer un SPE sur l'arbre résultant.

Expectiminimax et Minimax

- On considère un **jeu à somme nulle** ;
- Étant donné un état s, la profondeur maximum L, et la fonction d'évaluation statique $\tilde{U}$, **Expectiminimax** consiste à :
  1. construire le sous-jeu dont la racine est s et de pronfondeur au plus L ;
  2. remplacer toutes feuilles s' dans lesquelles un joueur pourrait encore jouer par $\tilde{U}(s')$;
  3. calculer un SPE sur l'arbre résultant.
- La valeur de s dans le SPE est plus précise que $\tilde{U}(s)$ grâce aux L pas d'anticipation ;

Expectiminimax et Minimax

- On considère un **jeu à somme nulle** ;
- Étant donné un état s, la profondeur maximum L, et la fonction d'évaluation statique $\tilde{U}$, **Expectiminimax** consiste à :
  1. construire le sous-jeu dont la racine est s et de pronfondeur au plus L ;
  2. remplacer toutes feuilles s' dans lesquelles un joueur pourrait encore jouer par $\tilde{U}(s')$;
  3. calculer un SPE sur l'arbre résultant.
- La valeur de s dans le SPE est plus précise que $\tilde{U}(s)$ grâce aux L pas d'anticipation ;
- S'il n'y a pas de nœud où *Chance* joue, Expectiminimax est appelé **Minimax** (historiquement proposé avant).

Expectiminimax et Minimax

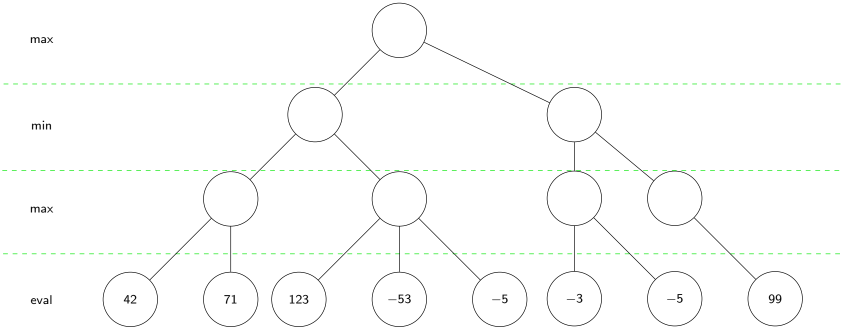
- On considère un **jeu à somme nulle** ;
- Étant donné un état s, la profondeur maximum L, et la fonction d'évaluation statique $\tilde{U}$, **Expectiminimax** consiste à :
  1. construire le sous-jeu dont la racine est s et de pronfondeur au plus L ;
  2. remplacer toutes feuilles s' dans lesquelles un joueur pourrait encore jouer par $\tilde{U}(s')$;
  3. calculer un SPE sur l'arbre résultant.
- La valeur de s dans le SPE est plus précise que $\tilde{U}(s)$ grâce aux L pas d'anticipation ;
- S'il n'y a pas de nœud où *Chance* joue, EXPECTIMINIMAX est appelé **Minimax** (historiquement proposé avant).
- En pratique on ne construit pas l'arbre explicitement : on en réalise un **parcours un profondeur**.

Minimax (jeu à somme nulle)

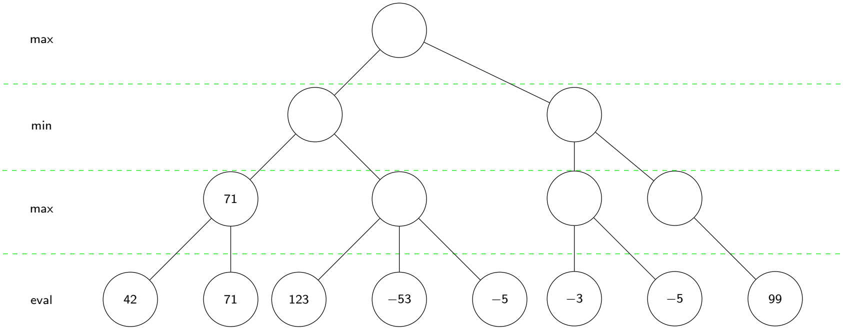
```
input:  s          // état source
        maximizer  // joueur: vrai pour le joueur 1 (qui maximise)
        d          // profondeur (distance à L)
output: U          // valeur de l'équilibre en s

if d = 0: // Feuille
    U ← evaluate(s)
else :
    if maximizer: // Le joueur~1 joue
        v ← −∞
        foreach child s′ of s:
            v ← max(v, MINIMAX(s′, not maximizer, d − 1))
    else : // Le joueur~2 joue
        v ← +∞
        foreach child s′ of s:
            v ← min(v, MINIMAX(s′, not maximizer, d − 1))
    U ← v
```

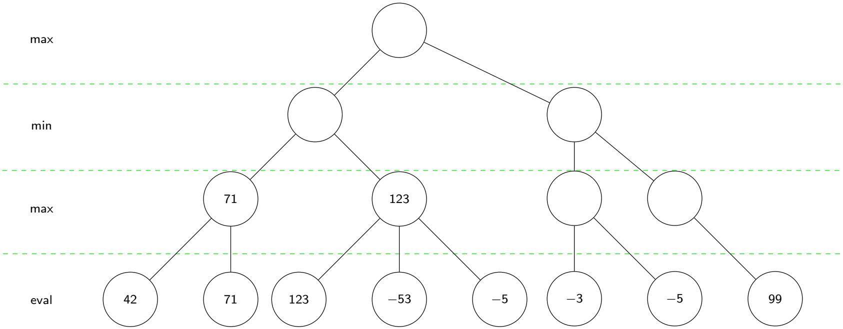
Minimax



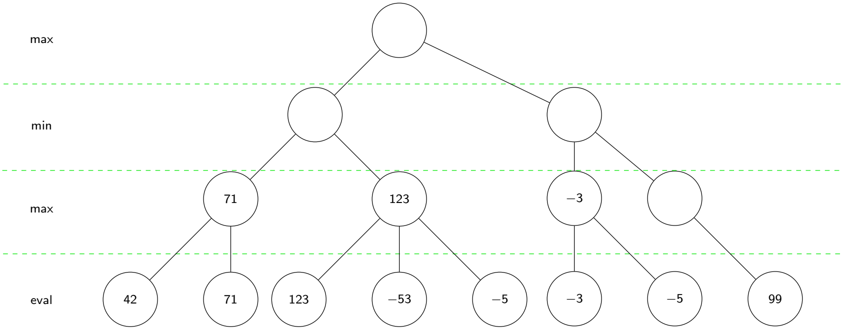
Minimax



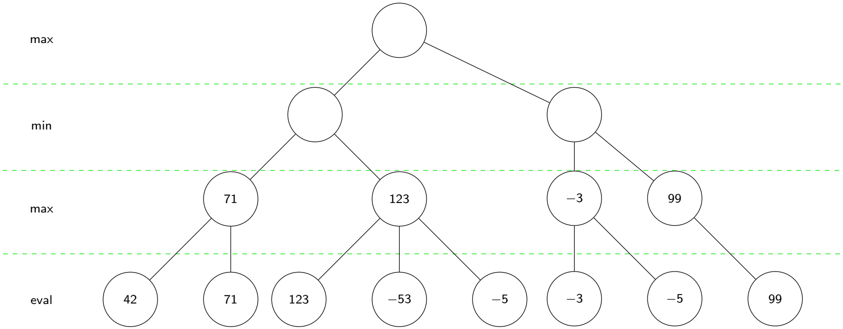
Minimax



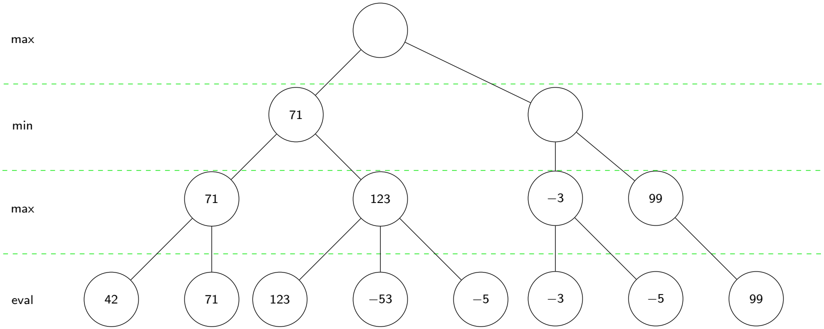
Minimax



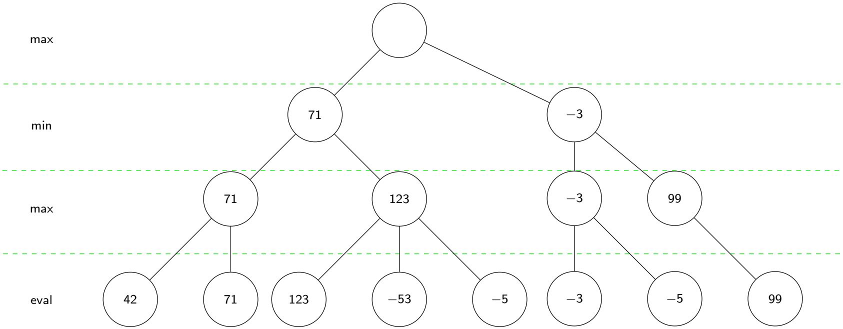
Minimax



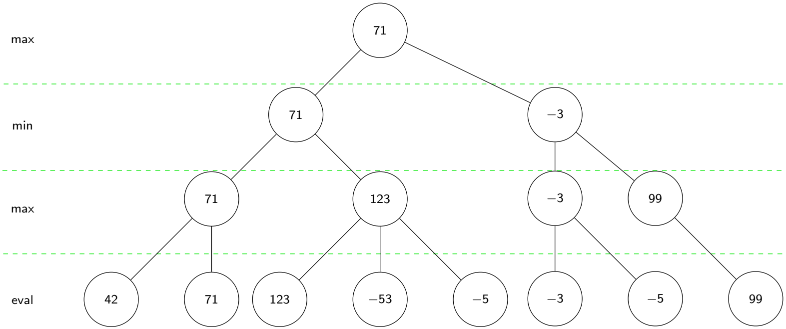
Minimax



Minimax



Minimax



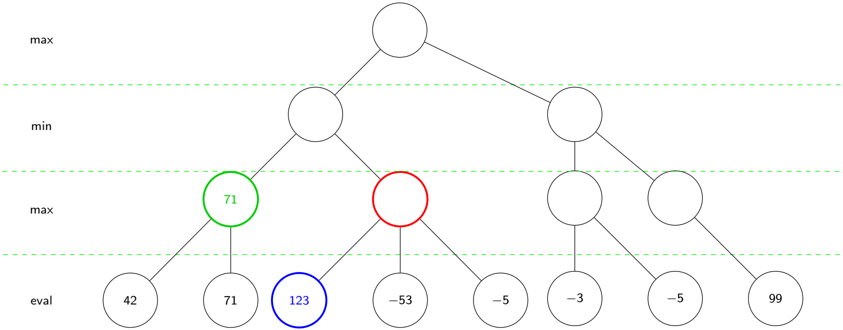
Negamax

- On peut faire maximiser le joueur 2 aussi en utilisant $\min(a, b) = -\max(-a, -b)$

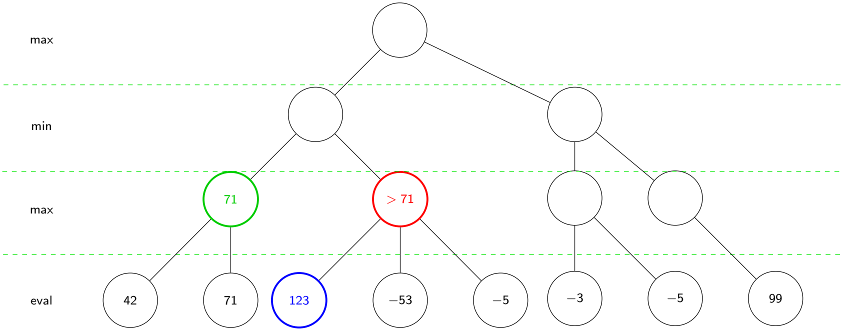
```
input:  s // état source
        p // qui joue: +1 pour le joueur 1, −1 pour le joueur 2
        d // profondeur (distance à L)
output: U // valeur de l'équilibre en s

if d = 0: // Feuille
    U ← evaluate(s)
else:
    v ← −∞
    foreach child s' of s:
        v ← max(v, p ∗ NEGAMAX(s', −p, d − 1))
    U ← p ∗ v
```

Élagage Alpha-beta (Pruning

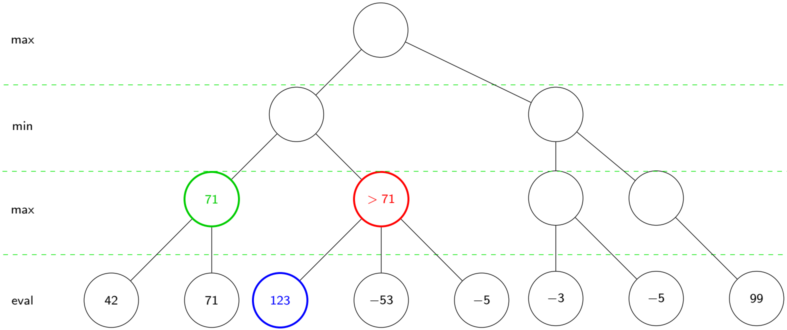


Élagage Alpha-beta (Pruning



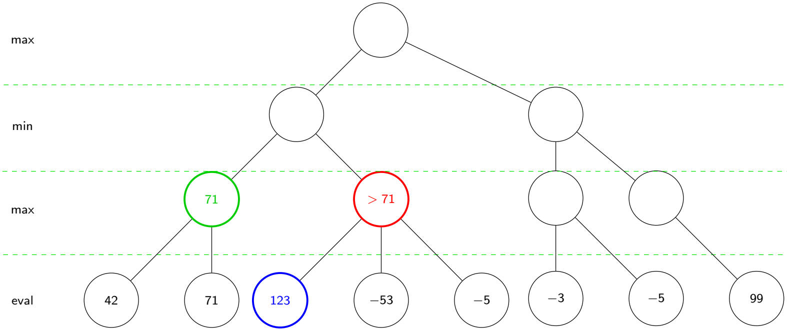
- Après l'exploration de , on sait que la valeur de est **au moins** 123 ;
- Comme le parent de minimise, ne sera jamais préféré à ;

Élagage Alpha-beta (Pruning



- Après l'exploration de ◯, on sait que la valeur de ◯ est **au moins** 123 ;
- Comme le parent de ◯ minimise, ◯ ne sera jamais préféré à ◯ ;
- Il est donc **futile** de continuer à explorer ses successeurs : **coupe** β (β-*cutoff*) ;

Élagage Alpha-beta (Pruning



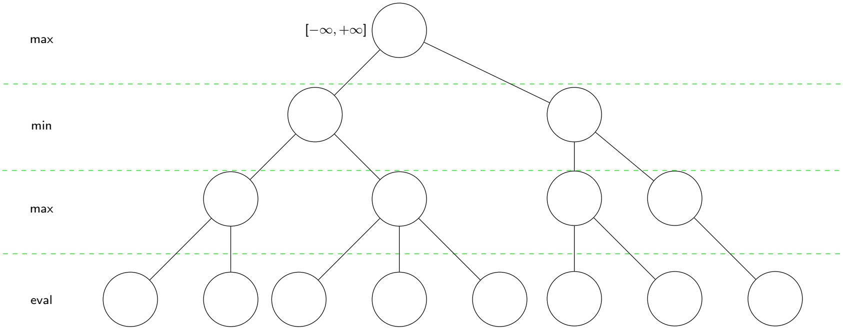
- Après l'exploration de ○, on sait que la valeur de ○ est **au moins** 123 ;
- Comme le parent de ○ minimise, ○ ne sera jamais préféré à ○ ;
- Il est donc **futile** de continuer à explorer ses successeurs : **coupe** β (β-*cutoff*) ;
- La même situation peut se produire avec le successeur d'un nœud max node dont la valeur est au forcément inférieure à une autre : **coupe** α**.**

AlphaBeta

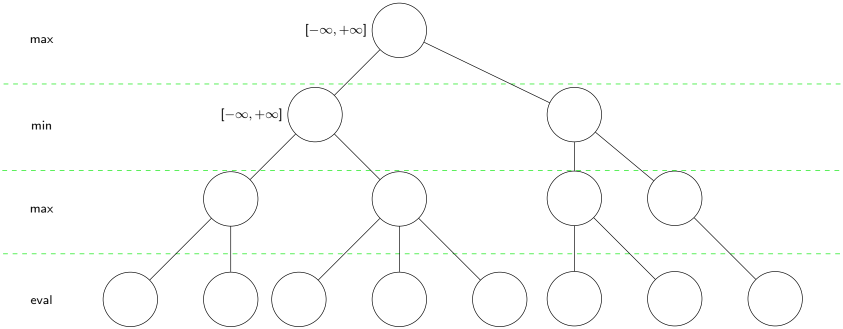
```
input:  s          // état source
        maximizer  // joueur qui joue: vrai si joueur 1 (maximise)
        d          // profondeur (distance à L)
        α          // borne inférieure de la valeur (initialement −∞)
        β          // borne supérieure de la valeur (initialement +∞)
output: U          // valeur de s

if d = 0: // Feuille
    U ← evaluate(s)
else :
    if maximizer: // Joueur 1
        v ← −∞
        foreach child s′ of s while α ≤ β:
            v ← max(v, AlphaBeta(s′, not maximizer, d − 1, α, β))
            α ← max(α, v)
    else : // Joueur 2
        v ← +∞
        foreach child s′ of s while α ≤ β:
            v ← min(v, AlphaBeta(s′, not maximizer, d − 1, α, β))
            β ← min(β, v)
    U ← v
```

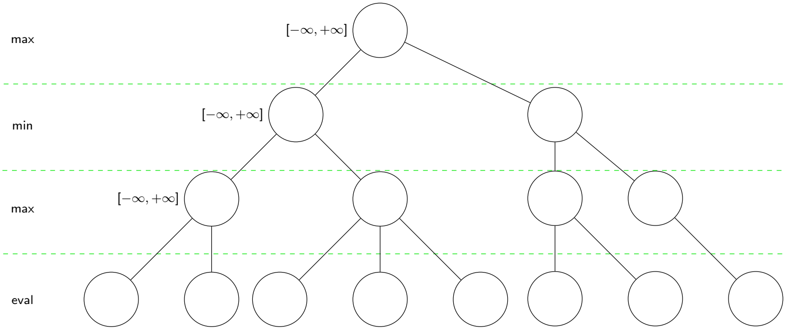
Élagage Alpha-beta



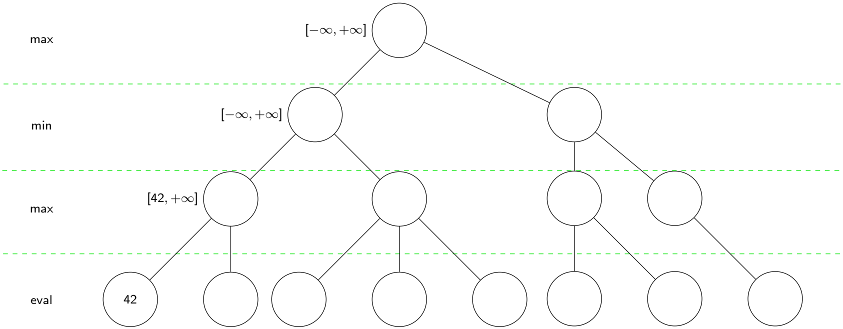
Élagage Alpha-beta



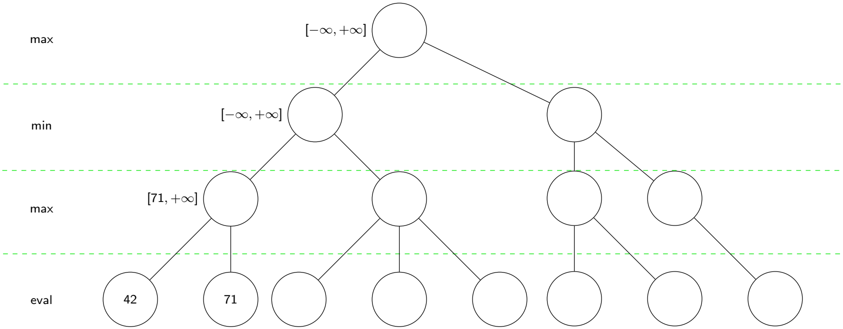
Élagage Alpha-beta



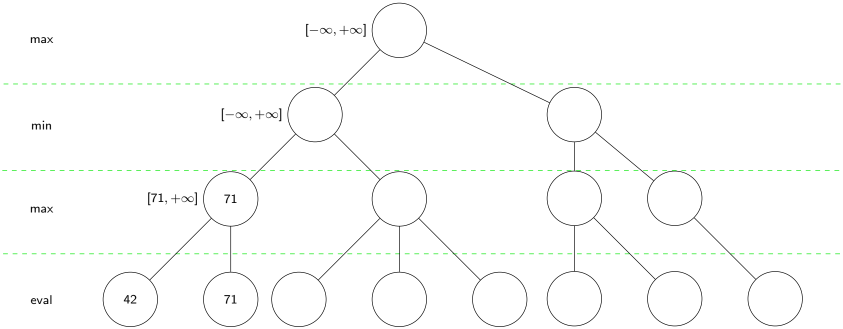
Élagage Alpha-beta



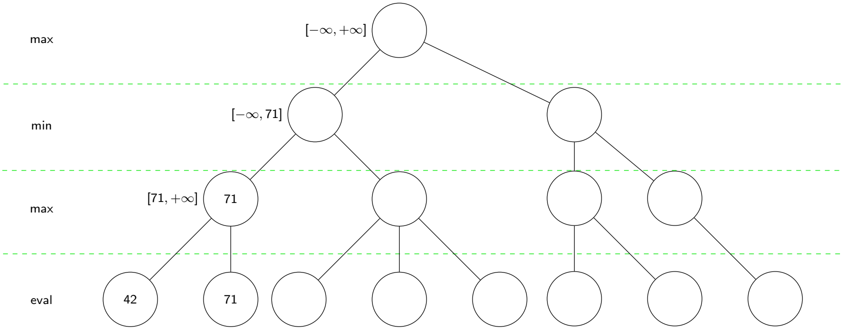
Élagage Alpha-beta



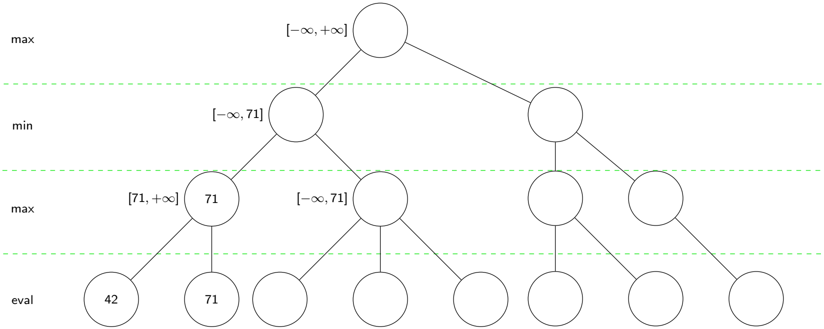
Élagage Alpha-beta



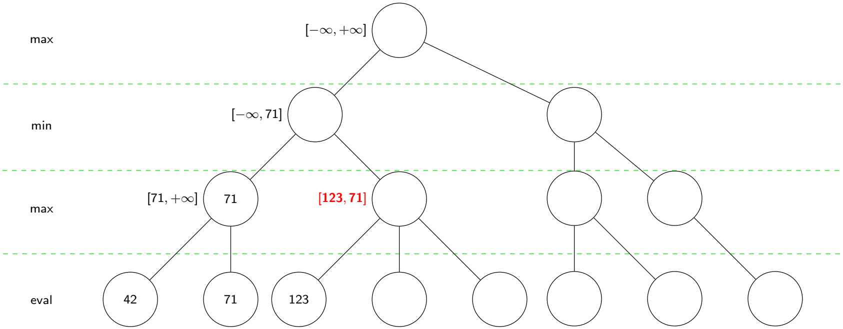
Élagage Alpha-beta



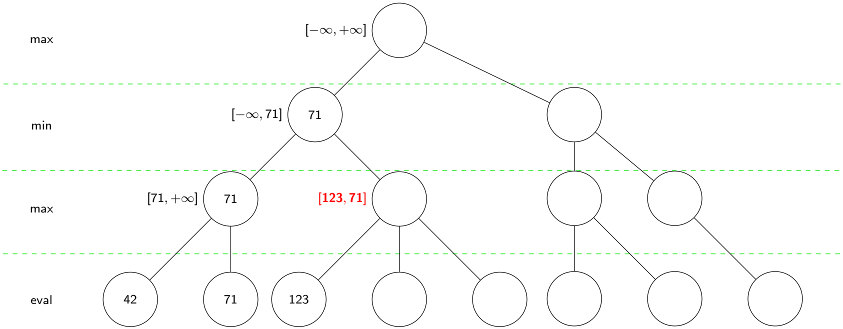
Élagage Alpha-beta



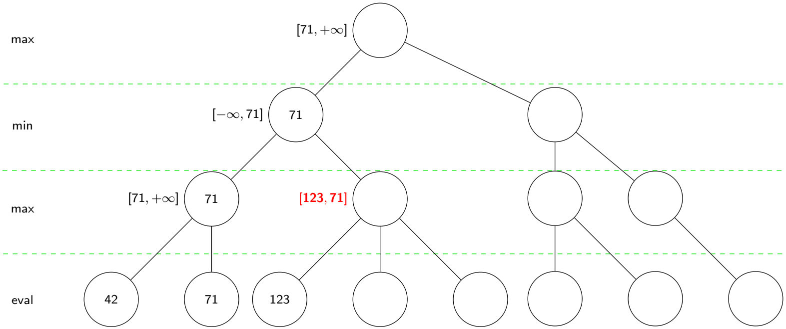
Élagage Alpha-beta



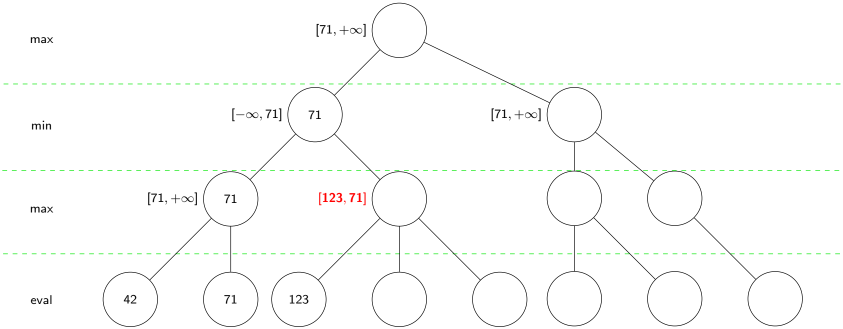
Élagage Alpha-beta



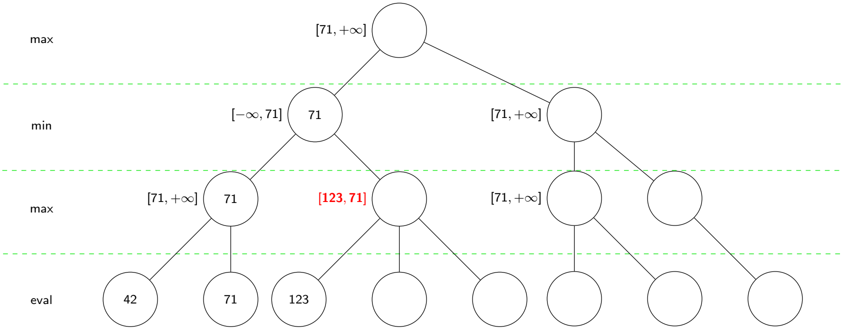
Élagage Alpha-beta



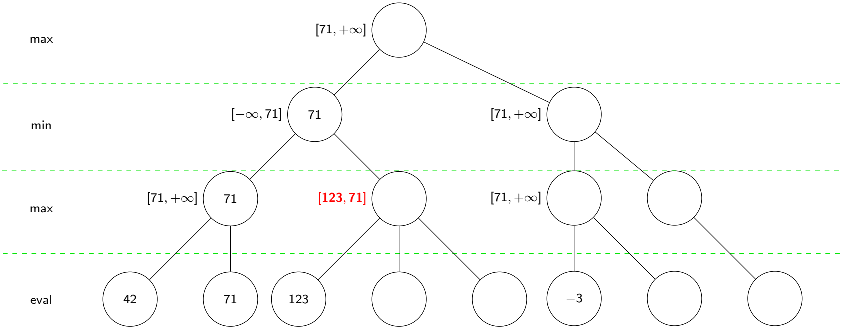
Élagage Alpha-beta



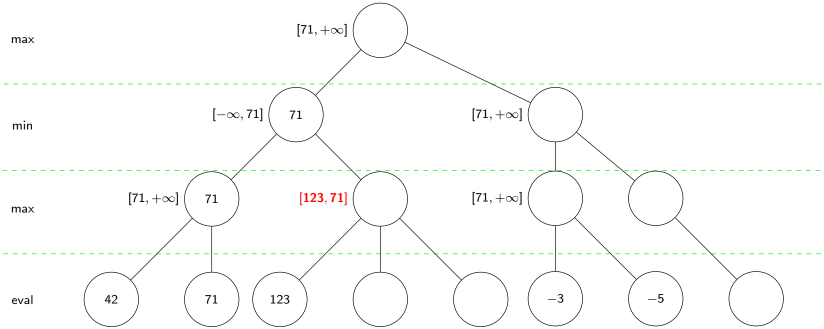
Élagage Alpha-beta



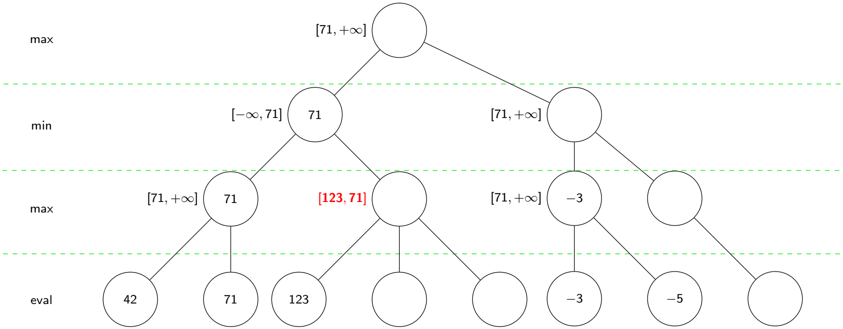
Élagage Alpha-beta



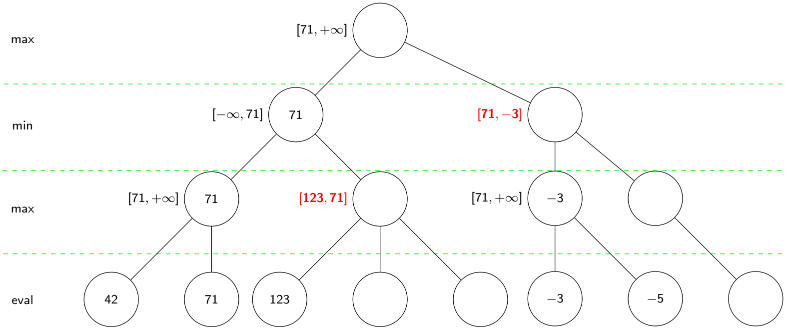
Élagage Alpha-beta



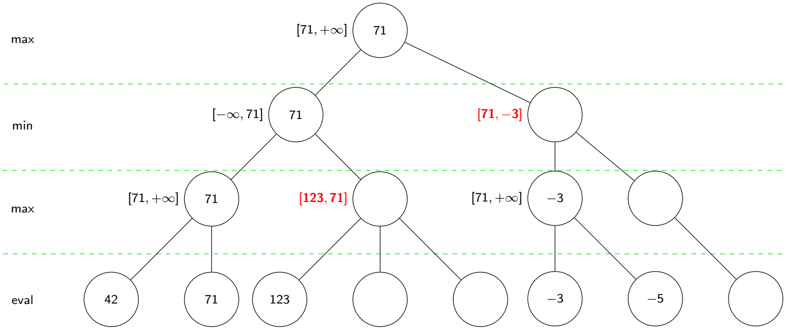
Élagage Alpha-beta



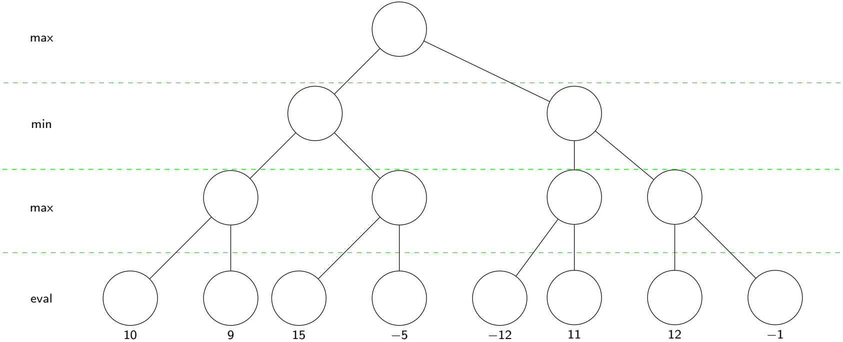
Élagage Alpha-beta



Élagage Alpha-beta



Élagage Alpha-beta



Exercice

1. Appliquer l'algorithme AlphaBeta à l'arbre ci-dessus, en explorant toujours le fils gauche d'abord.
2. Recommencer en explorant toujours le fils droit d'abord.

Élagage Alpha-beta

- Avec l'élagage alpha-beta, on peut explorer exponentiellement moins d'états ;
- L'ordre d'exploration des actions est critique : les **meilleures actions** de chaque joueur doivent être explorés en priorité
  Que se passe-t-il dans la première question de l'exercice précédent en changeant le 11 en 9 ?
- Déterminer l'ordre d'exploration peut être réalisé via des explorations moins profondes en particulier en utilisant l'**approfondissement itératif** (*iterative deepening*).

Outline

Monte-Carlo Tree Search (MCTS)

- On cherche à réduire le branchement en n'explorant qu'une partie des chemins ;

Monte-Carlo Tree Search (MCTS)

- On cherche à réduire le branchement en n'explorant qu'une partie des chemins ;
- Une première idée est de **tirer aléatoirement** un ≪ petit ≫ ensemble de chemins commençant après chaque action initiale ;

Monte-Carlo Tree Search (MCTS)

- On cherche à réduire le branchement en n'explorant qu'une partie des chemins ;
- Une première idée est de **tirer aléatoirement** un ≪ petit ≫ ensemble de chemins commençant après chaque action initiale ;
- Chaque chemin va jusqu'à une feuille du jeu initial ;

Monte-Carlo Tree Search (MCTS)

- On cherche à réduire le branchement en n'explorant qu'une partie des chemins ;
- Une première idée est de **tirer aléatoirement** un ≪ petit ≫ ensemble de chemins commençant après chaque action initiale ;
- Chaque chemin va jusqu'à une feuille du jeu initial ;
- On applique EXPECTIMINIMAX sur l'arbre résultant ;

Monte-Carlo Tree Search (MCTS)

- On cherche à réduire le branchement en n'explorant qu'une partie des chemins ;
- Une première idée est de **tirer aléatoirement** un ≪ petit ≫ ensemble de chemins commençant après chaque action initiale ;
- Chaque chemin va jusqu'à une feuille du jeu initial ;
- On applique EXPECTIMINIMAX sur l'arbre résultant ;
- La **valeur moyenne** de l'équilibre obtenue converge vers la vraie valeur mais très lentement ;

Monte-Carlo Tree Search (MCTS)

- On cherche à réduire le branchement en n'explorant qu'une partie des chemins ;
- Une première idée est de **tirer aléatoirement** un « petit » ensemble de chemins commençant après chaque action initiale ;
- Chaque chemin va jusqu'à une feuille du jeu initial ;
- On applique EXPECTIMINIMAX sur l'arbre résultant ;
- La **valeur moyenne** de l'équilibre obtenue converge vers la vraie valeur mais très lentement ;
- On explore de nombreux chemins qui ne seraient jamais choisis par une stratégie optimale.

Monte-Carlo Tree Search (MCTS)

- Une meilleure idée est de combiner cette approche avec une **construction incrémentale** de l'arbre pour diriger la recherche vers les actions les plus prometteuses ;

Monte-Carlo Tree Search (MCTS)

- Une meilleure idée est de combiner cette approche avec une **construction incrémentale** de l'arbre pour diriger la recherche vers les actions les plus prometteuses ;
- À chaque fois qu'on ajoute un nœud à l'arbre, on effectue une **partie complète aléatoire** (*random playout*) depuis ce nœud ;

Monte-Carlo Tree Search (MCTS)

- Une meilleure idée est de combiner cette approche avec une **construction incrémentale** de l'arbre pour diriger la recherche vers les actions les plus prometteuses ;
- À chaque fois qu'on ajoute un nœud à l'arbre, on effectue une **partie complète aléatoire** (*random playout*) depuis ce nœud ;
- On **propage en arrière** le résultat obtenu tout le long du chemin jusqu'à la racine ;

Monte-Carlo Tree Search (MCTS)

- Une meilleure idée est de combiner cette approche avec une **construction incrémentale** de l'arbre pour diriger la recherche vers les actions les plus prometteuses ;
- À chaque fois qu'on ajoute un nœud à l'arbre, on effectue une **partie complète aléatoire** (*random playout*) depuis ce nœud ;
- On **propage en arrière** le résultat obtenu tout le long du chemin jusqu'à la racine ;
- On calcule et mémorise la récompense moyenne et le nombre de parties aléatoires réalisées dans chaque nœud ;

Monte-Carlo Tree Search (MCTS)

- Une meilleure idée est de combiner cette approche avec une **construction incrémentale** de l'arbre pour diriger la recherche vers les actions les plus prometteuses ;
- À chaque fois qu'on ajoute un nœud à l'arbre, on effectue une **partie complète aléatoire** (*random playout*) depuis ce nœud ;
- On **propage en arrière** le résultat obtenu tout le long du chemin jusqu'à la racine ;
- On calcule et mémorise la récompense moyenne et le nombre de parties aléatoires réalisées dans chaque nœud ;
- C'est un algorithme d'**apprentissage par renforcement** ;

Monte-Carlo Tree Search (MCTS)

- Une meilleure idée est de combiner cette approche avec une **construction incrémentale** de l'arbre pour diriger la recherche vers les actions les plus prometteuses ;
- À chaque fois qu'on ajoute un nœud à l'arbre, on effectue une **partie complète aléatoire** (*random playout*) depuis ce nœud ;
- On **propage en arrière** le résultat obtenu tout le long du chemin jusqu'à la racine ;
- On calcule et mémorise la récompense moyenne et le nombre de parties aléatoires réalisées dans chaque nœud ;
- C'est un algorithme d'**apprentissage par renforcement** ;
- Le défi principal est, à chaque étape, de choisir entre :

Monte-Carlo Tree Search (MCTS)

- Une meilleure idée est de combiner cette approche avec une **construction incrémentale** de l'arbre pour diriger la recherche vers les actions les plus prometteuses ;
- À chaque fois qu'on ajoute un nœud à l'arbre, on effectue une **partie complète aléatoire** (*random playout*) depuis ce nœud ;
- On **propage en arrière** le résultat obtenu tout le long du chemin jusqu'à la racine ;
- On calcule et mémorise la récompense moyenne et le nombre de parties aléatoires réalisées dans chaque nœud ;
- C'est un algorithme d'**apprentissage par renforcement** ;
- Le défi principal est, à chaque étape, de choisir entre :
  - **Exploration** : commencer à explorer une nouvelle branche ;

Monte-Carlo Tree Search (MCTS)

- Une meilleure idée est de combiner cette approche avec une **construction incrémentale** de l'arbre pour diriger la recherche vers les actions les plus prometteuses ;
- À chaque fois qu'on ajoute un nœud à l'arbre, on effectue une **partie complète aléatoire** (*random playout*) depuis ce nœud ;
- On **propage en arrière** le résultat obtenu tout le long du chemin jusqu'à la racine ;
- On calcule et mémorise la récompense moyenne et le nombre de parties aléatoires réalisées dans chaque nœud ;
- C'est un algorithme d'**apprentissage par renforcement** ;
- Le défi principal est, à chaque étape, de choisir entre :
  - **Exploration** : commencer à explorer une nouvelle branche ;
  - **Exploitation** : étendre une branche existante prometteuse.

Machines à sous et UCB1

- On considère n machines à sous avec différentes espérances de gains **inconnues** ;

Machines à sous et UCB1

- On considère n machines à sous avec différentes espérances de gains **inconnues** ;
- Si on gagne sur une machine, doit-on continuer à jouer dessus ou en essayer une autre qui pourrait être encore meilleure ?

Machines à sous et UCB1

- On considère n machines à sous avec différentes espérances de gains **inconnues** ;
- Si on gagne sur une machine, doit-on continuer à jouer dessus ou en essayer une autre qui pourrait être encore meilleure ?
- Une stratégie possible est **UCB1** (pour *Upper Confidence Bound*) :

  Jouer sur la machine j qui maximise $\bar{r}_j + \sqrt{\frac{2\ln(n)}{n_j}}$ où :

Machines à sous et UCB1

- On considère n machines à sous avec différentes espérances de gains **inconnues** ;
- Si on gagne sur une machine, doit-on continuer à jouer dessus ou en essayer une autre qui pourrait être encore meilleure ?
- Une stratégie possible est **UCB1** (pour *Upper Confidence Bound*) :

  Jouer sur la machine j qui maximise $\bar{r}_j + \sqrt{\frac{2\ln(n)}{n_j}}$ où :

  - $\bar{r}_j$ is le gain empirique moyen sur la machine j jusqu'à maintenant ;

Machines à sous et UCB1

- On considère n machines à sous avec différentes espérances de gains **inconnues** ;
- Si on gagne sur une machine, doit-on continuer à jouer dessus ou en essayer une autre qui pourrait être encore meilleure ?
- Une stratégie possible est **UCB1** (pour *Upper Confidence Bound*) :

  Jouer sur la machine j qui maximise $\bar{r}_j + \sqrt{\frac{2\ln(n)}{n_j}}$ où :

  - $\bar{r}_j$ is le gain empirique moyen sur la machine j jusqu'à maintenant ;
  - n est le nombre total de parties jouées ;

Machines à sous et UCB1

- On considère n machines à sous avec différentes espérances de gains **inconnues** ;
- Si on gagne sur une machine, doit-on continuer à jouer dessus ou en essayer une autre qui pourrait être encore meilleure ?
- Une stratégie possible est **UCB1** (pour *Upper Confidence Bound*) :

  Jouer sur la machine j qui maximise $\bar{r}_j + \sqrt{\frac{2\ln(n)}{n_j}}$ où :

  - $\bar{r}_j$ is le gain empirique moyen sur la machine j jusqu'à maintenant ;
  - n est le nombre total de parties jouées ;
  - n_j est le nombre de parties jouées sur la machine j.

Machines à sous et UCB1

- On considère n machines à sous avec différentes espérances de gains **inconnues** ;
- Si on gagne sur une machine, doit-on continuer à jouer dessus ou en essayer une autre qui pourrait être encore meilleure ?
- Une stratégie possible est **UCB1** (pour *Upper Confidence Bound*) :

  Jouer sur la machine j qui maximise $\bar{r}_j + \sqrt{\frac{2\ln(n)}{n_j}}$ où :

  - $\bar{r}_j$ is le gain empirique moyen sur la machine j jusqu'à maintenant ;
  - n est le nombre total de parties jouées ;
  - n_j est le nombre de parties jouées sur la machine j.
- $\sqrt{\frac{2\ln(n)}{n_j}}$ est un **majorant** de la taille de l'**intervalle de confiance** pour le gain moyen tel que l'espérance de la vraie valeur est dedans avec une très grande probabilité ;

Machines à sous et UCB1

- On considère n machines à sous avec différentes espérances de gains **inconnues** ;
- Si on gagne sur une machine, doit-on continuer à jouer dessus ou en essayer une autre qui pourrait être encore meilleure ?
- Une stratégie possible est **UCB1** (pour *Upper Confidence Bound*) :

  Jouer sur la machine j qui maximise $\bar{r}_j + \sqrt{\frac{2\ln(n)}{n_j}}$ où :

  - $\bar{r}_j$ is le gain empirique moyen sur la machine j jusqu'à maintenant ;
  - n est le nombre total de parties jouées ;
  - n_j est le nombre de parties jouées sur la machine j.
- $\sqrt{\frac{2\ln(n)}{n_j}}$ est un **majorant** de la taille de l'**intervalle de confiance** pour le gain moyen tel que l'espérance de la vraie valeur est dedans avec une très grande probabilité ;
- Moins une machine est jouée, plus cette valeur est grande.

Upper Confidence on Trees (UCT)

On adapte UCB1 aux arbres de jeux :

- Globalement :

Upper Confidence on Trees (UCT)

On adapte UCB1 aux arbres de jeux :

- Globalement :
  1. Réaliser autant de **rounds** que voulu (voir ci-dessous) ;

Upper Confidence on Trees (UCT)

On adapte UCB1 aux arbres de jeux :

- Globalement :
    1. Réaliser autant de **rounds** que voulu (voir ci-dessous) ;
    2. Sélectionner l'action qui mène au nœud j avec la plus grande valeur $p * r_j$, où $p \in \{-1, 1\}$ représente le joueur qui joue.

Upper Confidence on Trees (UCT)

On adapte UCB1 aux arbres de jeux :

- Globalement :
  1. Réaliser autant de **rounds** que voulu (voir ci-dessous) ;
  2. Sélectionner l'action qui mène au nœud j avec la plus grande valeur $p * r_j$, où $p \in \{-1, 1\}$ représente le joueur qui joue.
- Dans chaque round, en commençant à la racine, et étant donné un nœud :

Upper Confidence on Trees (UCT)

On adapte UCB1 aux arbres de jeux :

- Globalement :
  1. Réaliser autant de **rounds** que voulu (voir ci-dessous) ;
  2. Sélectionner l'action qui mène au nœud j avec la plus grande valeur $p * r_j$, où $p \in \{-1, 1\}$ représente le joueur qui joue.
- Dans chaque round, en commençant à la racine, et étant donné un nœud :
  1. si un successeur direct n'a jamais été exploré ($n_j = 0$) :

Upper Confidence on Trees (UCT)

On adapte UCB1 aux arbres de jeux :

- Globalement :
  1. Réaliser autant de **rounds** que voulu (voir ci-dessous) ;
  2. Sélectionner l'action qui mène au nœud j avec la plus grande valeur $p * r_j$, où $p \in \{-1, 1\}$ représente le joueur qui joue.
- Dans chaque round, en commençant à la racine, et étant donné un nœud :
  1. si un successeur direct n'a jamais été exploré ($n_j = 0$) :
     - L'ajouter à l'arbre ;

Upper Confidence on Trees (UCT)

On adapte UCB1 aux arbres de jeux :

- Globalement :
  1. Réaliser autant de **rounds** que voulu (voir ci-dessous) ;
  2. Sélectionner l'action qui mène au nœud j avec la plus grande valeur $p * r_j$, où $p \in \{-1, 1\}$ représente le joueur qui joue.
- Dans chaque round, en commençant à la racine, et étant donné un nœud :
  1. si un successeur direct n'a jamais été exploré ($n_j = 0$) :
     - L'ajouter à l'arbre ;
     - Réaliser une partie aléatoire depuis ce nouveau nœud ;

Upper Confidence on Trees (UCT)

On adapte UCB1 aux arbres de jeux :

- Globalement :
  1. Réaliser autant de **rounds** que voulu (voir ci-dessous) ;
  2. Sélectionner l'action qui mène au nœud j avec la plus grande valeur $p * r_j$, où $p \in \{-1, 1\}$ représente le joueur qui joue.
- Dans chaque round, en commençant à la racine, et étant donné un nœud :
  1. si un successeur direct n'a jamais été exploré ($n_j = 0$) :
     - L'ajouter à l'arbre ;
     - Réaliser une partie aléatoire depuis ce nouveau nœud ;
     - Propager le résultat en arrière jusqu'à la racine an l'ajoutant au gain total de tous les nœuds rencontrés sur le chemin ;

Upper Confidence on Trees (UCT)

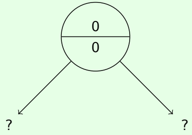
On adapte UCB1 aux arbres de jeux :

- Globalement :
  1. Réaliser autant de **rounds** que voulu (voir ci-dessous) ;
  2. Sélectionner l'action qui mène au nœud j avec la plus grande valeur $p * r_j$, où $p \in \{-1, 1\}$ représente le joueur qui joue.
- Dans chaque round, en commençant à la racine, et étant donné un nœud :
  1. si un successeur direct n'a jamais été exploré ($n_j = 0$) :
     - L'ajouter à l'arbre ;
     - Réaliser une partie aléatoire depuis ce nouveau nœud ;
     - Propager le résultat en arrière jusqu'à la racine an l'ajoutant au gain total de tous les nœuds rencontrés sur le chemin ;
  2. sinon descendre **récursivement** dans l'arbre en sélectionnant le successeur avec UCB1 : celui qui maximise $p * \left(\frac{r_j}{n_j} + p * \sqrt{\frac{2\ln(n)}{n_j}}\right)$.

     n_j est le nombre de fois qu'on a fait une partie depuis le nœud j et n est la somme des n_j pour tous les nœuds j que l'on compare.

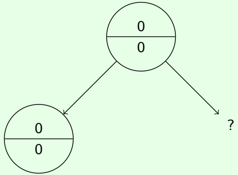
Upper Confidence on Trees (UCT)

Exemple



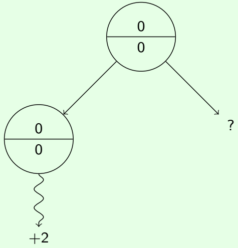
Upper Confidence on Trees (UCT)

Exemple



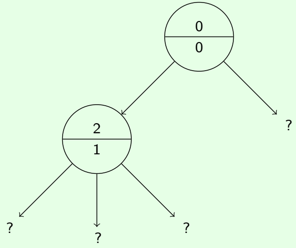
Upper Confidence on Trees (UCT)

Exemple



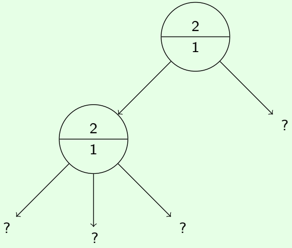
Upper Confidence on Trees (UCT)

Exemple



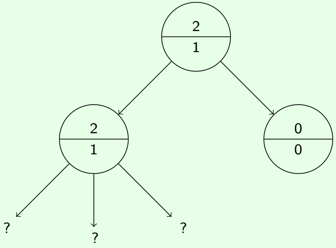
Upper Confidence on Trees (UCT)

Exemple



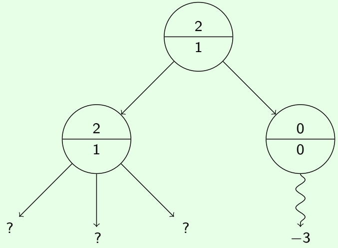
Upper Confidence on Trees (UCT)

Exemple



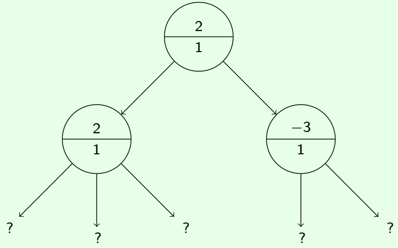
Upper Confidence on Trees (UCT)

Exemple



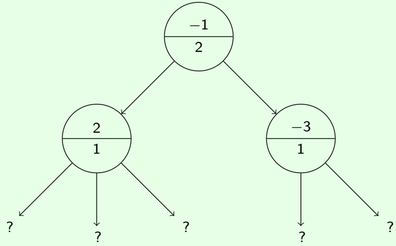
Upper Confidence on Trees (UCT)

Exemple



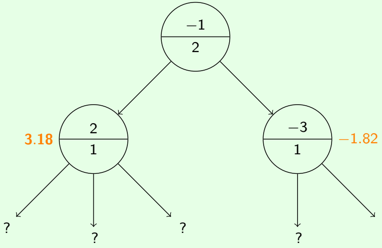
Upper Confidence on Trees (UCT)

Exemple



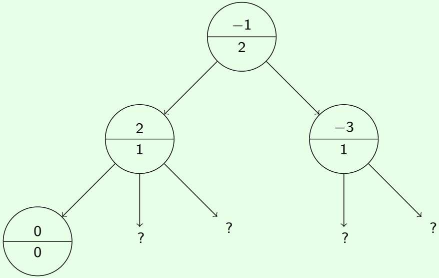
Upper Confidence on Trees (UCT)

Exemple



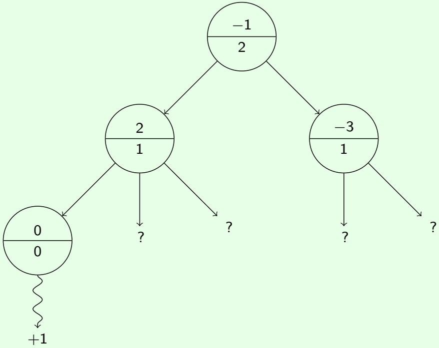
Upper Confidence on Trees (UCT)

Exemple



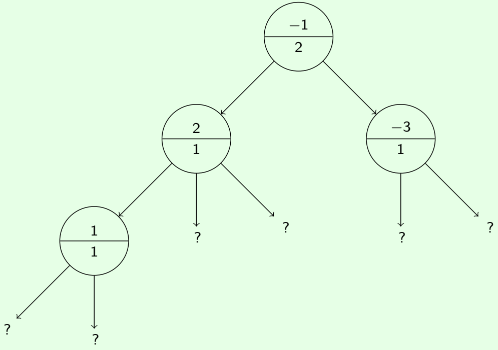
Upper Confidence on Trees (UCT)

Exemple



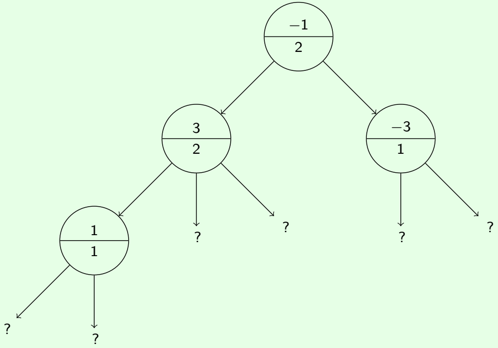
Upper Confidence on Trees (UCT)

Exemple



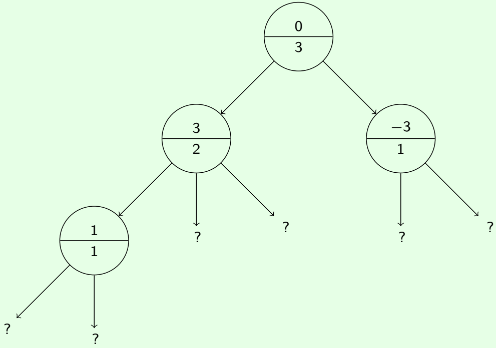
Upper Confidence on Trees (UCT)

Exemple



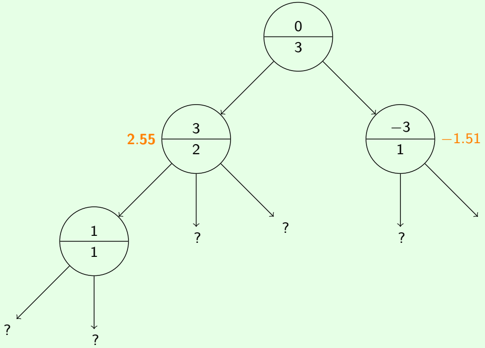
Upper Confidence on Trees (UCT)

Exemple



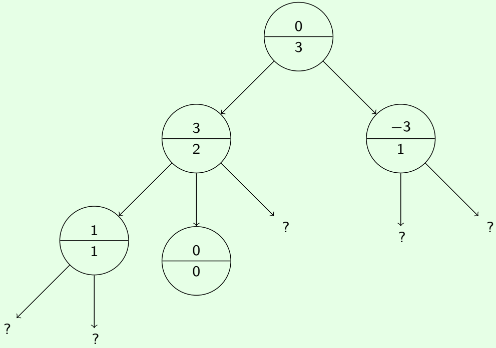
Upper Confidence on Trees (UCT)

Exemple



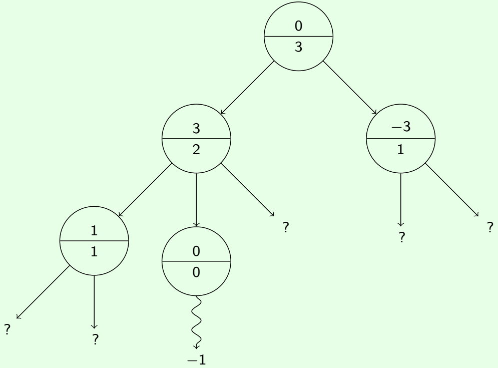
Upper Confidence on Trees (UCT)

Exemple



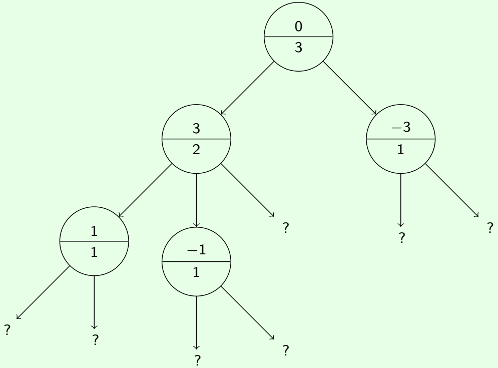
Upper Confidence on Trees (UCT)

Exemple



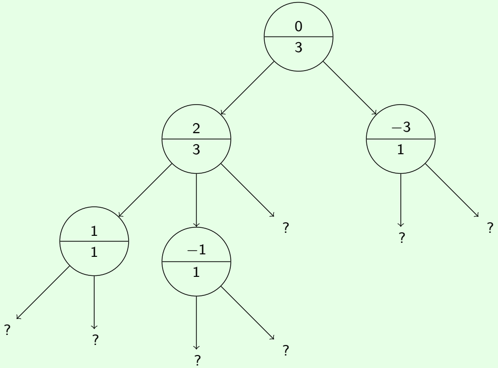
Upper Confidence on Trees (UCT)

Exemple



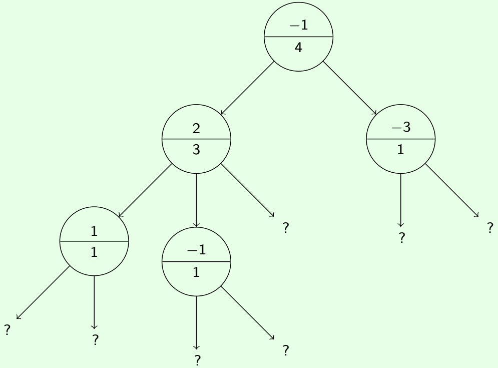
Upper Confidence on Trees (UCT)

Exemple



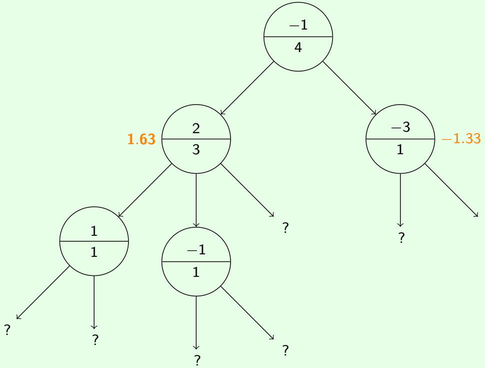
Upper Confidence on Trees (UCT)

Exemple



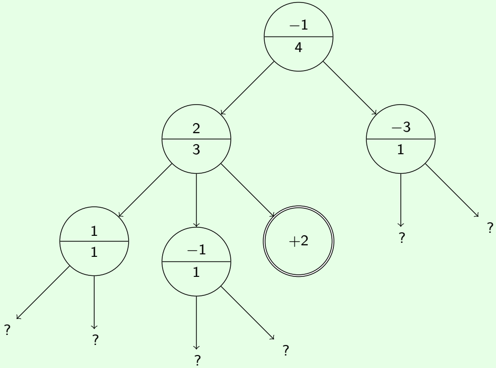
Upper Confidence on Trees (UCT)

Exemple



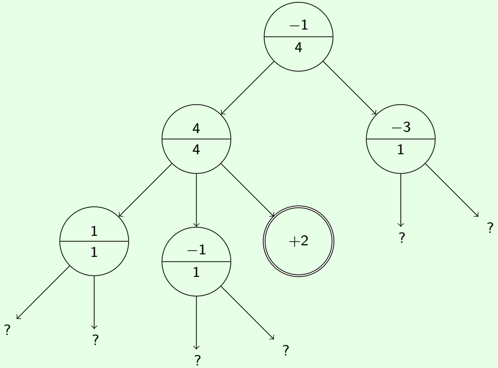
Upper Confidence on Trees (UCT)

Exemple



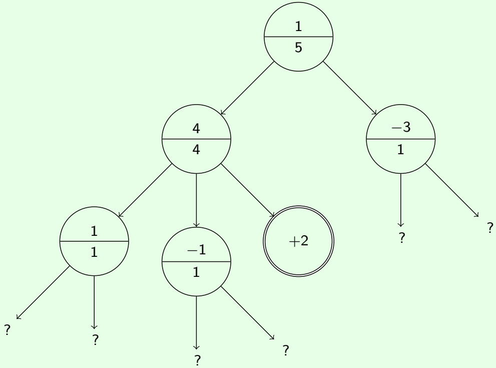
Upper Confidence on Trees (UCT)

Exemple



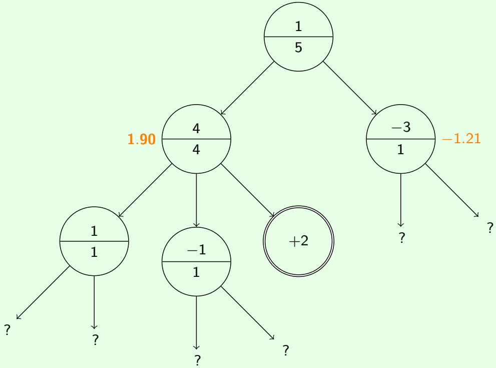
Upper Confidence on Trees (UCT)

Exemple



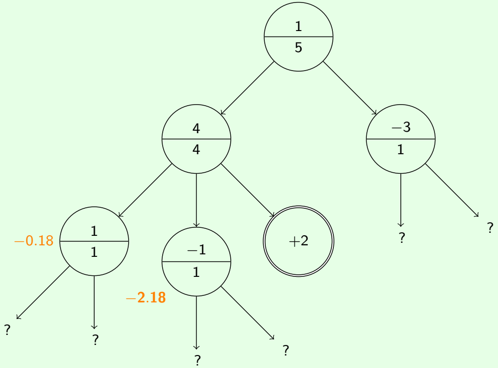
Upper Confidence on Trees (UCT)

Exemple



Upper Confidence on Trees (UCT)

Exemple



UCT

- Quand le nombre de simulations croit vers l'infini, UCT **converge** vers MINIMAX ;
- La convergence est plus rapide si l'on tire aléatoire avec un biais favorisant les **meilleurs coups** de chaque joueur ;
- Les meilleurs coups potentiels dans chaque nœuds peuvent être **appris** ;
- UCT peut être combiné avec une limitation de la profondeur ;
- AlphaGo utilise MCTS avec une limite de profondeur, et des réseaux de neurones profonds pour l'évaluation statique des positions (remplace les parties aléatoires) et pour trouver les bons coups potentiels.

Outline

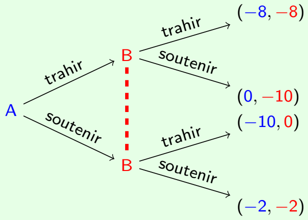
Jeu à information imparfaite

- Les jeux sous forme extensive ne permettent pas de modéliser le *Dilemme du prisonnier*, le *Poker*, *Stratego*, etc.
- Dans ces jeux, l'état des autres joueurs n'est que **partiellement connu** ;
- On définit un jeu sous forme extensive **à information imparfaite**, en ajoutant pour chaque joueur un ensemble de sous-ensembles de son ensemble d'états ;
- Chaque sous-ensemble est appelé **ensemble d'information** (*information set*)
- Un ensemble d'information du joueur i contient les états qui sont **indistingables** par ce joueur ;
- Tous les états d'un ensemble d'information doivent avoir les **mêmes actions** possibles ;
- Le joueur *Chance* a toujours des ensembles d'information qui sont des singletons ;
- Si tous les ensembles d'informations sont des singletons, on a un jeu **à information parfaite**.

Jeu à information imparfaite

Exemple

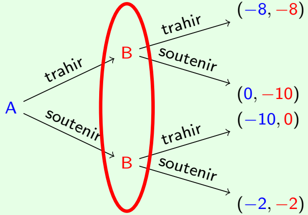
Dilemme du prisonnier (forme extensive).



Jeu à information imparfaite

Exemple

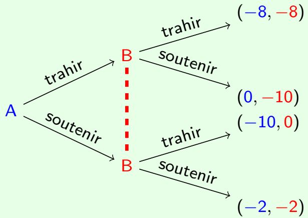
Dilemme du prisonnier (forme extensive).



Forme extensive pour les jeux en forme normale

Exemple

Dilemme du prisonnier (forme extensive).



Pour tout jeu en forme normale on peut construire une forme extensive **canonique** dans laquelle :

- Le joueur 1 joue en premier ;
- Tous les états obtenus après que le joueur i a joué forment un **unique** ensemble d'information pour le joueur $i+1$;
- Quand tous les joueurs ont joué on a les gains associés : chaque chemin dans l'arbre obtenu représente un profil de stratégies pures.

Stratégies dans les jeu sous forme extensive à information imparfaite

- Une stratégie pure affecte une action à chaque **ensemble d'information** ;
- Chaque joueur doit jouer la **même action** dans les états d'un même ensemble d'information
- On modifie les définitions de stratégies mixtes et comportement de manière similaire ;
- On n'a **plus d'équivalence** entre les deux :
  mêmes probabilités d'atteindre les états pour toutes stratégies adverses

Exercice

1. Trouver un jeu sous forme extensive et une stratégie comportementale qui n'a pas d'équivalent en stratégie mixte ;
2. Trouver un jeu sous forme extensive et une stratégie mixte qui n'a pas d'équivalent en stratégie comportementale.

Jeux avec mémoire parfaite

- Dans l'exercice précédent, soit le joueur a oublié quelle action il a faite, soit même s'il en fait une ;
- On définit une sous-classe dans laquelle ce n'est pas possible

Définition

Dans un jeu sous forme extensive à information imparfaite, le joueur i a une mémoire parfaite (*perfect recall*) si :

1. chaque chemin issu de la racine passe par tout ensemble d'information du joueur i au plus une fois
   Sinon, on a soit l'exemple précédent du conducteur, soit on pourrait distinguer les deux avec un état intermédiaire
2. si deux chemins issus de la racine passent par le même ensemble d'information, alors avant cela :
   1. ils sont passés par les mêmes ensembles d'information du joueur i,
   2. dans le même ordre,
   3. dans chacun de ces ensembles d'information, la même action a été effectuée sur les deux chemins.

   les deux chemins sont indistingables pour le joueur qui se souvient de ce qu'il a fait et où il était

Si tous les joueurs ont une mémoire parfaite, alors le jeu est dit à mémoire parfaite.

Jeux avec mémoire parfaite

- Dans l'exercice précédent, soit le joueur a oublié quelle action il a faite, soit même s'il en fait une ;
- On définit une sous-classe dans laquelle ce n'est pas possible

Définition

Dans un jeu sous forme extensive à information imparfaite, le joueur i a une mémoire parfaite (*perfect recall*) si :

1. chaque chemin issu de la racine passe par tout ensemble d'information du joueur i au plus une fois
   Sinon, on a soit l'exemple précédent du conducteur, soit on pourrait distinguer les deux avec un état intermédiaire
2. si deux chemins issus de la racine passent par le même ensemble d'information, alors avant cela :
   1. ils sont passés par les mêmes ensembles d'information du joueur i,
   2. dans le même ordre,
   3. dans chacun de ces ensembles d'information, la même action a été effectuée sur les deux chemins.

   les deux chemins sont indistingables pour le joueur qui se souvient de ce qu'il a fait et où il était

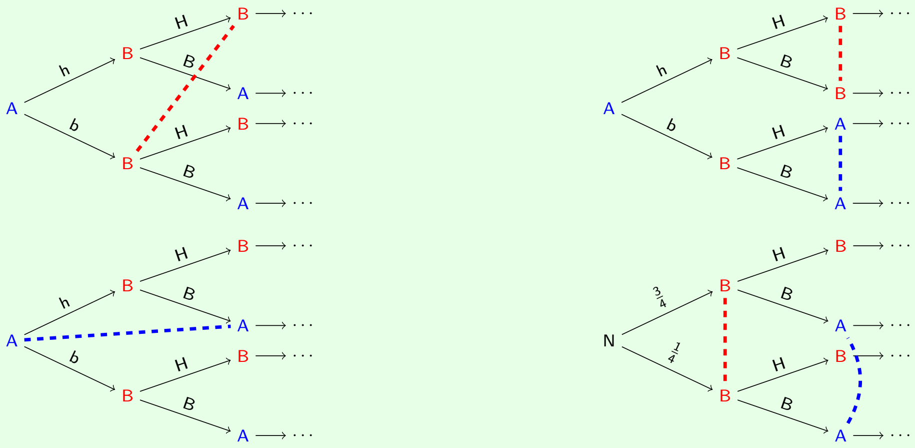
Si tous les joueurs ont une mémoire parfaite, alors le jeu est dit à mémoire parfaite.

Théorème (Kuhn, 1953)

Dans un jeu à mémoire parfaite, pour toute stratégie comportementale, on peut construire une stratégie mixte avec, pour toutes les stratégies adverses, les mêmes probabilités d'arriver à chaque état et inversement.

Mémoire parfaite : (contre-)exemples

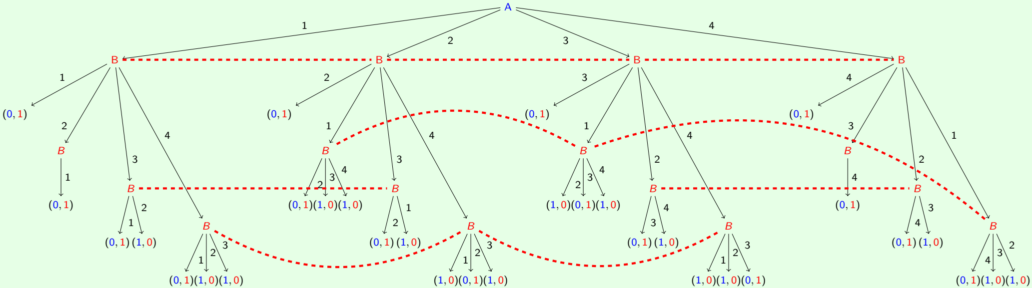
Exemple



Jeux à information imparfaite : exemple

Exemple

Alice choisit un nombre entre 1 et 4. Bob a deux essais pour le trouver. Après le premier essai, s'il n'a pas trouvé, Alice lui dit si le nombre cherché est plus grand ou plus petit que sa proposition. Alice gagne si Bob ne trouve pas, et perd sinon et on a l'inverse pour Bob.



Forme normale pour les jeux à information imparfaite

- On définit la forme normale comme pour les jeux à information parfaite, en utilisant les ensembles d'information plutôt que les sommets :

Exemple

- Alice a 1 ensemble d'information, Bob 7 :
  - a1234 (après le choix initial d'Alice), b2a1 (Bob a déjà joué 2 alors qu'Alice a choisi 1), b2a34, b3a4, b3a12, b1a234, et b4a123
  - dans b2a1 et b3a4 la stratégie est forcée
  - une stratégie pure réduite combine un choix dans a1234 et un dans l'ensemble d'information résultant du premier choix dans lequel la stratégie n'est pas forcée ;
  - on a donc les stratégies pures : 12 (joue 1 puis 2), 13, 14, 23, 24, 32, 31, 41, 42, et 43.
- Alice a 4 stratégies pures, Bob 10 (sous forme réduite) :

	12	13	14	23	24	31	32	41	42	43
1	0	0	0	0	0	0	1	0	1	1
2	0	1	1	0	0	1	0	1	0	1
3	1	0	1	0	1	0	0	1	1	0
4	1	1	0	1	0	0	0	0	0	0

Forme normale pour les jeux à information imparfaite

Exemple

	12	13	14	23	24	31	32	41	42	43
1	0	0	0	0	0	0	1	0	1	1
2	0	1	1	0	0	1	0	1	0	1
3	1	0	1	0	1	0	0	1	1	0
4	1	1	0	1	0	0	0	0	0	0

- Il n'y a pas d'équilibre en stratégies pures ;
- Pour Alice, le jeu étant strictement compétitif, on cherche une stratégie de sécurité :

$$\begin{aligned}\text{maxmin}_A &= \max_{p_1,p_2,p_3} \min\{1-p_1-p_2, 1-p_1-p_3, p_2+p_3, 1-p_1-p_2-p_3, p_3, p_2, p_1, p_2+p_3, p_1+p_3, p_1+p_2\} \\ &= \max_{p_1,p_2,p_3} \min\{1-p_1-p_2-p_3, p_3, p_2, p_1\} = \frac{1}{4} \text{ quand tous sont égaux}\end{aligned}$$

- Pour Bob, par interchangeabilité des équilibres, on peut se restreindre aux meilleures réponses à $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$:
- Ce sont les colonnes qui réalisent le max (pour Bob donc en inversant 0 et 1) avec cette stratégie : 23, 24, 31, 32
- Et donc $\text{maxmin}_B = \max_{q_1,q_2,q_3} \min\{q_1+q_2+q_3, 1-q_3, 1-q_2, 1-q_1\} = \frac{3}{4}$ quand tous sont égaux.

SPE dans les jeux à information imparfaite

- On peut calculer des équilibres parfaits en sous-jeux dans un jeu à information imparfaite ;
- Un **sous-jeu** dans ce cas doit contenir **tous les états** d'un même ensemble d'information ;

Exemple

- dans le jeu de devinette précédent, les seuls sous-jeux non-triviaux sont les réponses forcées de Bob (Alice a choisi 1 et Bob dit 2 initialement, par exemple) ;
- dans la forme extensive canonique d'un jeu en forme normale, il n'y a qu'un seul sous-jeu, le jeu lui-même.

Forme séquence pour les jeux à information imparfaite

- Recalculer la forme normale à partir d'un jeu sous forme extensive a un coût exponentiel en général ;
- La **forme séquence** permet de formuler l'existence d'un équilibre de Nash via un système de contraintes presque linéaires directement à partir de la fome extensive ;
- Le nombre de contraintes est linéaire dans le nombre d'ensembles d'information ;
- L'**algorithme de Lemke** (pas Howson) permet de résoudre ce système de contrainte en un temps exponentiel ;
- Pour les jeux strictement compétitifs (ou les stratégies de sécurité) les contraintes sont linéaires et le système peut être résolu en temps polynomial par la programmation linéaire.

Information imparfaite : Exercices

Exercice

Le problème de Monty Hall. Alice participe à un jeu télévisé. L'animateur, Monty, lui présente trois portes fermées. Derrière deux d'entre-elles se trouve une chèvre, et derrière la troisième une voiture (on suppose qu'Alice préfère la voiture).
Alice doit choisir une porte, puis avant de l'ouvrir Monty en ouvre une autre. Il sait où se trouve la voiture et n'ouvre ni la porte choisie par Alice, ni celle menant à la voiture.
Alice a ensuite la possibilité de confirmer son choix initial ou de changer de porte. Enfin la porte choisie est ouverte et les gains distribués.
On suppose que l'utilité de Monty est opposée à celle d'Alice.

1. Modéliser ce jeu sous forme extensive. On pourra ne développer que le cas où Alice choisit la porte du milieu, les autres étant symétriques.
2. Trouver les équilibres de Nash dans ce jeu.

Information imparfaite : Exercices

Exercice

Poker de Von Neumann simplifié [1]. Alice et Bob jouent au poker avec un jeu de trois cartes : valet, dame, roi de cœur. Le jeu se déroule ainsi :

1. Le paquet est mélangé puis Alice reçoit la première carte, et Bob la seconde.
2. Alice et Bob misent alors 1 chacun.
3. Alice peut surenchérir (*raise*, ajouter 1 à la mise) ou *passer* (*check*) ;
4. Puis Bob peut ou *abandonner* (*fold*) ou, soit passer si Alice a passé, soit égaliser (*call*, ajouter 1 à sa mise aussi) si elle a surenchéri ;
5. si Bob n'abandonne pas, on révèle les deux cartes et le joueur avec la meilleure carte emporte la mise totale. Si Bob abandonne, c'est Alice qui emporte la mise totale sans que les cartes soient révélées.

Modéliser ce jeu sous forme extensive et trouver l'unique équilibre de Nash. Notons que si Alice passe, Bob n'a jamais intérêt à abandonner donc on peut supposer que le jeu s'arrête tout de suite et les cartes sont révélées.

Outline

Outline

Dilemme du prisonnier répété

Exemple

On suppose qu'Alice et Bob jouent au *Dilemme du prisonnier* deux fois de suite.

- Quelles sont les stratégies pures d'Alice ?
  1. T puis T
  2. S puis S
  3. S puis T
  4. S puis S

Alice \ Bob	Trahir	Soutenir
Trahir	-8 / -8	-10 / 0
Soutenir	0 / -10	-2 / -2

Dilemme du prisonnier répété

Exemple

On suppose qu'Alice et Bob jouent au *Dilemme du prisonnier* deux fois de suite.

- Quelles sont les stratégies pures d'Alice ?
  1. T puis T
  2. S puis S
  3. S puis T
  4. S puis S
  5. S puis si Bob S alors S
  6. S puis si Bob S alors T
  7. S puis si Bob T alors S
  8. …

Alice \ Bob	Trahir	Soutenir
Trahir	-8, -8	0, -10
Soutenir	-10, 0	-2, -2

Dilemme du prisonnier répété

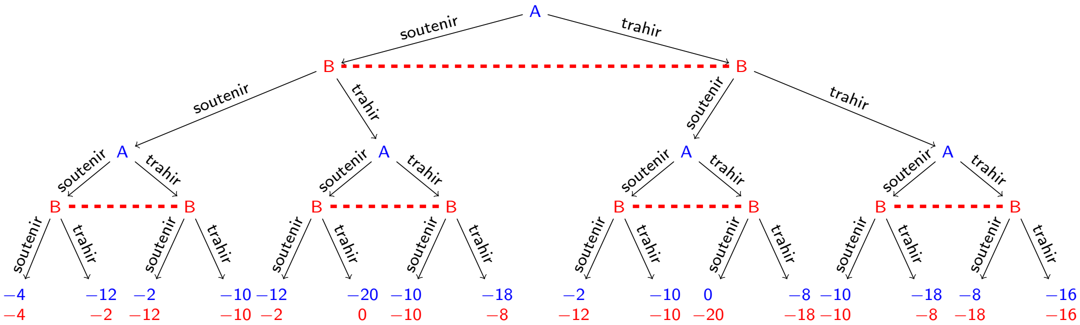
Exemple

On suppose qu'Alice et Bob jouent au *Dilemme du prisonnier* deux fois de suite.

- Quelles sont les stratégies pures d'Alice ?
  1. T puis T
  2. S puis S
  3. S puis T
  4. S puis S
  5. S puis si Bob S alors S
  6. S puis si Bob S alors T
  7. S puis si Bob T alors S
  8. ...
- On a un jeu sous forme extensive.

Alice \ Bob	Trahir	Soutenir
Trahir	-8, -8	0, -10
Soutenir	-10, 0	-2, -2

Dilemme du prisonnier répété



Jeux répétés et stratégies

- Un **jeu répété** G^N est défini par un *jeu de base* (*stage game*) en forme normale G et un nombre de répétition N ;
- Les gains des joueurs peuvent être définis comme :
  - le total des gains : $\sum_{k=0}^{N} u_i(s^k)$
  - ou la moyenne par tour des gains : $\frac{1}{N}\sum_{k=0}^{N} u_i(s^k)$
  - ou la somme actualisée (*discounted*) par un facteur $0 < \gamma < 1$: $\frac{1}{N}\sum_{k=0}^{N} \gamma^k u_i(s^k)$.
- L'**historique** des coups h_k du jeu au tour k est une séquence donnant l'alternance des coups de chaque joueur jusque là
  par exemple, au tour 3, $tTtT$ signifie Alice et Bob ont tous les deux trahi lors des deux premiers tours
- Une **stratégie pure** du joueur i dans un jeu répété est une fonction qui à tout historique associe une action $a_i \in A_i$
- Une stratégie mixte donne une distribution sur les stratégies pures ;
- Une stratégie comportementale donne une distribution sur A_i pour tout historique.

Dilemme du prisonnier répété

Exemple

$t(s,t)$ indique qu'Alice trahit puis soutient si elle est trahie et trahit si elle est soutenue.

	T(T,T)	T(S,T)	T(T,S)	T(S,S)	S(T,T)	S(S,T)	S(T,S)	S(S,S)
t(t,t)	(-16, -16)	(-8, -18)	(-16, -16)	(-8, -18)	(-8, -18)	(0, -20)	(-8, -18)	(0, -20)
t(s,t)	(-18, -8)	(-10, -10)	(-18, -8)	(-10, -10)	(-8, -18)	(0, -20)	(-8, -18)	(0, -20)
t(t,s)	(-16, -16)	(-8, -18)	(-16, -6)	(-8, -18)	(-10, -10)	(-2, -12)	(-10, -10)	(-2, -12)
t(s,s)	(-18, -8)	(-10, -10)	(-18, -8)	(-10, -10)	(-10, -10)	(-2, -12)	(-10, -10)	(-2, -12)
s(t,t)	(-18, -8)	(-18, -10)	(-10, -10)	(-10, -10)	(-10, -10)	(-10, -10)	(-2, -12)	(-2, -12)
s(s,t)	(-20, 0)	(-20, 0)	(-12, -2)	(-12, -2)	(-10, -10)	(-10, -10)	(-8, -18)	(-8, -18)
s(t,s)	(-18, -8)	(-18, -8)	(-10, -10)	(-10, -10)	(-8, -18)	(-8, -18)	(-4, -4)	(-4, -4)
s(s,s)	(-20, 0)	(-20, 0)	(-12, -2)	(-12, -2)	(-12, -2)	(-12, -2)	(-4, -4)	(-4, -4)

Exercice

Quels sont les équilibres de Nash en stratégies pures de ce jeu ?

Dilemme du prisonnier répété

Exemple

$t(s,t)$ indique qu'Alice trahit puis soutient si elle est trahie et trahit si elle est soutenue.

	T(T,T)	T(S,T)	T(T,S)	T(S,S)	S(T,T)	S(S,T)	S(T,S)	S(S,S)
t(t,t)	($\boxed{-16}$, $\boxed{-16}$)	($\boxed{-8}$, -18)	(-16, $\boxed{-16}$)	($\boxed{-8}$, -18)	($\boxed{-8}$, -18)	($\boxed{0}$, -20)	(-8, -18)	($\boxed{0}$, -20)
t(s,t)	(-18, $\boxed{-8}$)	(-10, -10)	(-18, $\boxed{-8}$)	(-10, -10)	($\boxed{-8}$, -18)	($\boxed{0}$, -20)	(-8, -18)	($\boxed{0}$, -20)
t(t,s)	($\boxed{-16}$, -16)	($\boxed{-8}$, -18)	(-16, -16)	($\boxed{-8}$, -18)	(-10, $\boxed{-10}$)	(-2, -12)	(-10, $\boxed{-10}$)	(-2, -12)
t(s,s)	(-18, $\boxed{-8}$)	(-10, -10)	(-18, $\boxed{-8}$)	(-10, -10)	(-10, -10)	(-2, -12)	(-10, -10)	(-2, -12)
s(t,t)	(-18, $\boxed{-8}$)	(-18, $\boxed{-8}$)	($\boxed{-10}$, -10)	(-10, -10)	(-10, -10)	(-10, -10)	($\boxed{-2}$, -12)	(-2, -12)
s(s,t)	(-20, $\boxed{0}$)	(-20, $\boxed{0}$)	(-12, -2)	(-12, -2)	(-10, -10)	(-10, -10)	(-8, -18)	(-8, -18)
s(t,s)	(-18, -8)	(-18, -8)	($\boxed{-10}$, -10)	(-10, -10)	($\boxed{-8}$, -18)	(-8, -18)	(-4, $\boxed{-4}$)	(-4, $\boxed{-4}$)
s(s,s)	(-20, $\boxed{0}$)	(-20, $\boxed{0}$)	(-12, -2)	(-12, -2)	(-12, -2)	(-12, -2)	(-4, -4)	(-4, -4)

Exercice

Quels sont les équilibres de Nash en stratégies pures de ce jeu ?

Équilibres de Nash dans les jeux répétés

- Si les deux joueurs jouent le même équilibre de Nash à chaque tour, on obtient un équilibre de Nash ;
- Plus généralement, c'est le cas s'ils jouent un équilibre de Nash quelconque à chaque tour.

Exemple

- (Métal, Piano) (*Piano*, *Métal*) dans *Sortie Musicale* joué deux fois ;
- (Lent, Rapide) (*Rapide*, *Lent*) dans *Poule Mouillée* joué deux fois.

Folk theorem

Exercice

Répétons *Poule mouillée* :

1. Le gain $(2, 2)$ est-il un équilibre de Nash du jeu de base ?
2. Quelle est la valeur minimale qu'Alice peut forcer Bob à obtenir à chaque étape ? Avec quelle stratégie ?

Supposons qu'Alice et Bob jouent deux fois et se mettent d'accord pour jouer (Lent, Lent) :

3. Comment Alice peut-elle empêcher Bob de jouer *Rapide* au premier tour ?
4. Alice peut-elle empêcher Bob de jouer *Rapide* au deuxième tour ?
5. Supposons qu'ils jouent $N > 3$ fois. Comment Alice peut-elle empêcher Bob de jouer *Rapide* à tous les tours sauf les trois derniers ?
6. Dans les deux derniers tours, on joue des équilibres de Nash. Lesquels ? Comment cela empêche-t-il Bob de dévier dans les **trois** derniers tours ?
7. Quel est le gain moyen de l'équilibre de Nash obtenu (en fonction de N) ?

Alice \ Bob	Rapide	Lent
Rapide	-1, -1	3, 0
Lent	0, 3	2, 2

Folk theorem dans les jeux répétés finis

- Soit G un jeu tel que pour tout joueur i il existe un équilibre de Nash s^* avec $u_i(s^*) > \mathrm{minmax}_i$;
- Le gain $\vec{x}$ est **faisable** (*feasible*) s'il est dans l'**enveloppe convexe** des résultats de G ;
- Il est **individuellement rationnel** si $\forall i, x_i \geq \mathrm{minmax}_i$;
- Alors pour tout $\epsilon > 0$, il existe $N \in \mathbb{N}$ tel que pour tout x faisable et individuellement rationnel et tout $n \geq N$, il existe un équilibre de Nash s^* du jeu répété G^n dont le gain moyen de chaque joueur i est proche à ϵ près de x_i.
- Il existe une variante pour les coûts actualisés.

Folk theorem dans les jeux répétés infinis

- Si le jeu est répété à l'infini, alors on n'a plus besoin de prévoir un *coda* pour gérer les déviations tardives ;
- Et on obtient donc un résultat exact ;
- Dans les jeux répétés infinis, pour tout gain x faisable est individuellement rationnel, il existe un équilibre de Nash de valeur x.
- Il existe une variante un peu plus complexe pour les coûts actualisés.

Exercice

Donner un équilibre de Nash de valeur $(\frac{7}{4}, \frac{7}{4})$ dans *Poule mouillée* répété à l'infini.

Folk Theorem et équilibres parfaits en sous-jeux

- La peur de la punition doit suffire à ce que Bob ne dévie pas ;
- Mais s'il dévie ? Alice est-elle prête à (potentiellement) se sacrifier pour mettre en œuvre la punition ?
- Il existe une version du *Folk theorem* qui permet d'obtenir des équilibres parfaits en sous-jeux :
  - Un sous-ensemble de joueurs punit la déviation de Bob ;
  - Un autre sous-ensemble punit les justiciers qui n'appliquent pas la punition ;
  - etc.

Outline

Minimisation du regret dans les jeux en forme normale

- Alice joue répétitivement à un jeu sous forme normale ;
- À chaque tour t :
  - elle joue selon une stratégie s^t
  - elle observe les gains de **toutes** les actions ;
  - Elle maintient un **regret cumulé** pour chaque action a au tour n :
    (a^t_{-A}) est le profil d'actions des autres joueurs au tour t

$$R^n[a] = \sum_{t=1}^{n} \left(u_A(a, a^t_{-A}) - u_A(s^t, a^t_{-A}) \right)$$

  - Elle choisit sa stratégie s^{n+1} pour (essayer de) minimiser son **regret externe** total :
    regret de ne pas avoir toujours joué la meilleure action globalement

$$R = \sum_{t=1}^{n} \left(\max_a u_A(a, a^t_{-A}) - u_A(s^t, a^t_{-A}) \right)$$

Par exemple, jouer à chaque tour l'action avec le plus petit regret cumulé

Regret matching dans les jeux en forme normale

- Une stratégie efficace est le **regret matching** ;
- La stratégie *regret matching* par rapport à R d'Alice au tour n est alors :

$$\forall a, s_A^n(a) = \max\left(0, \frac{R^n[a]}{\sum_{b, R^n[b]>0} R^n[b]}\right)$$

  s'il n'y a pas d'action au regret strictement positif on prend une stratégie uniforme
- On utilise en général une forme **contrefactuelle** des regrets cumulés :
  Alice joue selon s^t mais calcule le regret comme si elle avait voulu dès le départ jouer a^t l'action obtenue au final

$$R_c^n[a] = \sum_{t=1}^{n} \left(u_A(a, a_{-A}^t) - u_A(\mathbf{a^t}, a_{-A}^t)\right)$$

- On peut aussi tronquer les regrets cumulés au dessous de zéro (*regret-matching+*) :

$$R_+^n[a] = \max(0, R_c^n[a])$$

Regret matching dans les jeux en forme normale

- Si Alice joue simplement sa stratégie *regret matching* à chaque tour et que Bob le devine, il peut choisir une stratégie pour optimiser ses réponses et exploiter cette faiblesse ;
- Pour empêcher cela, Alice apprend d'abord à jouer contre elle-même (*self-play*) avant de jouer contre Bob.
- Si tous les joueurs jouent leur stratégie *regret-matching* pendant n tours alors la distribution empirique sur les profils d'actions joués est un **équilibre corrélé** à un certain ϵ près avec $\epsilon \to 0$ quand $n \to +\infty$.
- Le **profil de stratégies moyen** $\frac{1}{n}\sum_{t=1}^{n} s^t$ et la distribution empirique des profils d'actions convergent vers la même valeur ;
- Dans les jeux **strictement compétitifs** :
  - si le regret est inférieur ϵ, alors le profil de stratégies moyen est un équilibre de Nash approché à 2ϵ
  - la stratégie moyenne pour le joueur i, $\frac{1}{n}\sum_{t=1}^{n} s_i^t$, est alors une **stratégie de sécurité** (à 2ϵ près).

Exemple

Démo : Alice apprend à jouer à *Matching Pennies* en jouant contre elle-même avec des stratégies *regret-matching*.

Regret matching et jeux sous forme extensive

- On applique la même idée dans les jeux sous forme **extensive** à somme nulle (avec mémoire parfaite) ;
- On utilise *regret-matching* dans chaque ensemble d'information ;
- On calcule ainsi une **stratégie comportementale**
- Deux problème principaux :
  1. Comment calculer la stratégie moyenne ?
  2. Comment calculer les résultats de chaque action ?

Regret matching et jeux sous forme extensive : histoires

- L'histoire h d'un **sommet** est la séquence des actions effectuée jusqu'à arriver à lui
  y compris les actions de Chance
- Elle est unique pour un état donné et on pourra identifier l'un et l'autre ;
- En suivant le profil de stratégies s^t, on peut calculer la contribution $\pi_i^{s^t}(h)$ du joueur i à la probabilité de réalisation de l'histoire h :
  en notant I_h l'ensemble d'information auquel h mène

$$\pi_i^{s^t}(\emptyset) = 1 \text{ et } \pi_i^{s^t}(h.a) = \begin{cases} \pi_i^{s^t}(h)s_i^t(I_h)(a) \text{ si } i \text{ doit jouer après } h \\ \pi_i^{s^t}(h) \text{ sinon} \end{cases}$$

- On note $\pi_c(h)$ la contribution de *Chance* à cette probabilité ;
- La probabilité totale de h pour le profil s^t est alors :

$$\pi^{s^t}(h) = \pi_1^{s^t}(h)\pi_2^{s^t}(h)\pi_c(h)$$

- Enfin, on note $\pi_{-1}^{s^t}(h) = \pi_2^{s^t}(h)\pi_c(h)$ et $\pi_{-2}^{s^t} = \pi_1^{s^t}(h)\pi_c(h)$.

Regret matching et jeux sous forme extensive : stratégie moyenne

- En suivant le profil de stratégies s^k, on peut calculer la probabilité $\pi_i^{s^t}(I)$ de d'atteindre l'ensemble d'information I du joueur i :

$$\pi_i^{s^t}(I) = \sum_{h \in I} \pi_i^{s^t}(h)$$

- On définit la **stratégie moyenne** après n tours par :

$$\overline{s}_i(I)(a) = \frac{\sum_{t=1}^{n} \pi_i^{s^t}(I) s^t(I)(a)}{\sum_{t=1}^{n} \pi_i^{s^t}(I)}$$

Regret matching et jeux sous forme extensive : valeurs des sommets

- La valeur d'un sommet dépend du futur :
  $\pi^{s_t}(h, h')$ est la probabilité selon s^t et les coups de *Chance* de compléter h en h'

$$v_i^{s^t}(h) = \sum_{h' \text{ terminale} \sqsupseteq h} \pi^{s^t}(h, h') u_i(h')$$

- La contribution de h à la valeur de la racine est alors $\pi^{s^t}(h) v_i^{s^t}(h)$;
- Si on fait comme si le joueur i avait joué pour maximiser la réalisation de h, on a une contribution **contrefactuelle** $\pi_{-i}^{s^t}(h) v_i^{s^t}(h)$.

Regret matching et jeux sous forme extensive : regrets contrefactuels

- Le **regret contrefactuel** (immédiat) en h entre appliquer s^t normalement ou changer pour a dans l'ensemble d'information I_h est alors :
  $s^t_{I_h \to a}$ est une copie de s^t dans laquelle l'action jouée en I_h a été changée pour a

$$r_i^{s^t}(h, a) = \pi^{s^t}_{-i}(h) \left(v_i^{s^t_{I_h \to a}}(h) - v_i^{s^t}(h) \right)$$

- Cela s'étend pour un ensemble d'information :

$$r_i^{s^t}(I, a) = \sum_{h \in I} r_i^{s^t}(h, a)$$

- Et le regret contrefactuel cumulé au tour n :

$$R_i^n(I, a) = \sum_{t=1}^{n} r_i^{s^t}(I, a)$$

Regret matching et jeux sous forme extensive : regrets contrefactuels

- En minimisant les regrets contrefactuels cumulés dans chaque ensemble d'information, on peut montrer qu'on minimise le regret total moyen :

$$R_{\text{total}} = \frac{1}{n} \max_{s_i^* \in S_i} \sum_{t=1}^{n} \left(u_i(s_i^*, s_{-i}^t) - u_i(s^t) \right)$$

- Si le regret total moyen est inférieur à ϵ, alors le profil de stratégies comportementales moyen $\overline{s}$ est un équilibre de Nash approché à 2ϵ près ;
- On minimise le regrets contrefactuels cumulés dans chaque ensemble d'information en y appliquant **regret matching**.

Algorithme CFR (counterfactual regret minimisation)

- On répète n fois, à tour de rôle pour les joueurs 1 et 2 :
  1. Pour le joueur i, on calcule la valeur de chaque sommet par un parcours en profondeur d'abord postfixe de l'arbre de jeu ;
  2. En remontant (postfixe) on met aussi à jour les regrets cumulés et la stratégie moyenne pour le joueur i ;
  3. Pour cela, il faut calculer en descendant $\pi_i^s(h)$ et $\pi_{-i}^s(h)$ pour la stratégie *regret matching*.
- La stratégie moyenne obtenue est un équilibre de Nash approché.

Algorithme CFR (counterfactual regret minimisation)

- L'algorithme CFR est la **référence** pour les « gros » jeux à somme nulle sous forme extensive et à information imparfaite ;
- On peut le coupler avec des techniques de type **Monte Carlo** ;
  - échantillonner les coups de *Chance* plutôt que de calculer l'espérance ;
  - échantilloner les valeurs et les regrets de sous-arbres ;
- On peut représenter les stratégies en utilisant l'apprentissage statistique (**Deep CFR**).

Conclusion générale

- Nous avons abordé :
  - des représentations : formes **normales** et **extensives** ;
  - des **concepts de solutions** : équilibres de Nash, corrélés, stratégies dominées, regret ;
  - des **algorithmes** pour calculer ces solutions.
- Pour continuer l'étude :
  - les algorithmes mentionnés : Lemke, Lemke-Howson, forme séquence.
  - théorie des jeux évolutionnaire ;
  - jeux à information incomplète, jeux et équilibres Bayésiens ;
  - jeux coopératifs, répartition des gains, *Core*, valeur de Shapley, algorithmes de mariages stables ;
  - conception de mécanismes : quel jeu créer pour engendre un comportement souhaiter (p. ex. enchères)

Références

- Ken Binmore. *Playing for Real : A Text on Game Theory*, OUP USA, 2007.
- Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani (eds). *Algorithmic Game Theory*, Cambridge University Press, 2007.
- Yoav Shoham and Kevin Leyton-Brown.*Multiagent Systems : Algorithmic, Game-Theoretic, and Logical Foundations*, Cambridge University Press, 2009.
- Michael Maschler, Eilon Solan, Shmuel Zamir. *Game Theory*, Cambridge University Press, 2013.