Jeux à deux joueurs et positions gagnantes

1) Présentation

a) On s'intéresse ici aux jeux où deux joueurs **antagonistes jouent alternativement**. Le jeu consiste en un certain nombre, généralement fini, de positions, et une position particulière est appelée position de départ. Les règles précisent clairement les coups qu'un joueur peut réaliser à partir d'une position donnée. Les positions accessibles par un joueur sont appelées ses options. Les deux joueurs connaissent parfaitement l'état du jeu, c'est-à-dire que le jeu est à information complète.

On s'intéresse aux jeux dont les règles assurent que **toute partie se termine en un nombre fini** de coups (ce qui est appelé la condition de terminaison).

S'il existe des positions avec des options différentes pour les deux joueurs, le jeu est dit **partisan**, et si les options disponibles sont toujours les mêmes pour les deux joueurs, le jeu est dit **impartial**.

Remarque : Les jeux comme les échecs ou les dames n'appartiennent pas à cette catégorie de jeux puisque des parties peuvent être nulles ou infinies.

c) **Graphe de jeu.** On représente un jeu à deux joueurs A et B par son graphe, c'est-à-dire le graphe dont les sommets sont de deux types : les états appartenant au joueur A et ceux appartenant au joueur B.

Quand un joueur possède un état, c'est lui qui choisit l'état suivant parmi les états possibles. Les objectifs sont pour le joueur d'arriver à une position gagnante pour lui.

$$\begin{array}{ccccccc}
 & & a0 & \longrightarrow & \boxed{b1} & & \\
 & \swarrow & & \searrow & & & \\
b0 & \longrightarrow & a1 & \longrightarrow & b2 & \longrightarrow & \boxed{a3} \\
 & \searrow & & \nearrow & & & \\
 & & a2 & \longrightarrow & \boxed{b3} & &
\end{array}$$

Sur le schéma, les positions **finales** sont encadrées : $a3$ ici gagnante pour le joueur A, $b1$ et $b3$ pour B.

Partant de $a0$, le joueur A a une stratégie gagnante en considérant les transitions : $a0 \to b0$, $a1 \to b2$ et $a2 \to b2$. En fait, ici, toutes les positions non finales sont gagnantes pour le joueur A.

b) Formellement, la donnée d'un jeu est la donnée d'une arène et d'un objectif :

- Une arène A est essentiellement **un graphe biparti** acyclique (S, E) : S est défini comme la réunion disjointe $S = S_a \sqcup S_b$ (sommets), et on a $E \subset S_a \times S_b \cup S_b \times S_a$ (arêtes).
Les sommets sont les états possibles du jeu.
S_a est l'ensemble des états contrôlés par le joueur A, et de même pour S_b.

- L'objectif $\Omega = \Omega_a \sqcup \Omega_b$ est l'ensemble des positions finales partagées entre celles qui sont gagnantes pour A (qu'on note Ω_a) et celles qui sont gagnantes pour B (qu'on note Ω_b).

Une **stratégie** (sans mémoire) pour le joueur A est une fonction $f_a : S_a \to S_b$ qui à chaque état contrôlé par le joueur A associe une transition à partir de cet état, de sorte que $\forall x \in S_a$, $(u, f_a(u)) \in E$.

Une **configuration de jeu** est un chemin dans le graphe. Elle est dite associée à la stratégie f_a lorsqu'elle est de la forme :
$x_0 \to f_a(x_0) \to x_1 \to f_a(x_1) \to x_2 \to f_a(x_2) \to \ldots$ avec $x_j \in S_a$
(ou bien de la forme : $y \to x_0 \to f_a(x_0) \to x_1 \to f_a(x_1) \to \ldots$, où ici la position initiale $y \in S_b$.

Une stratégie f_a pour le joueur A est dite gagnante ssi les configurations possibles associées à f_a se terminent en Ω_a (et de même pour le joueur B). Autrement dit, une stratégie est dite gagnante pour l'un des joueurs si **quelle que soit** la stratégie de l'autre joueur, le joueur atteint un objectif, c'est-à-dire une position (finale) gagnante pour lui. Même si l'autre joueur joue au mieux pour contrecarrer les choix du premier joueur, celui-ci pourra aboutir à une position finale gagnante.

Le premier problème qu'on considère est de "résoudre le jeu", c'est-à-dire de décider si le joueur j gagne depuis un état initial fixé. **L'attracteur** de A est l'ensemble des positions initiales pour lesquelles le joueur A admet au moins une stratégie gagnante.

2) Attracteurs

a) Pour déterminer **les stratégies gagnantes** du joueur A, on procède par un **calcul d'attracteur**.

Pour $k \in \mathbb{N}$, considérons l'ensemble W_k des états du jeu à partir desquels le joueur A possède une stratégie qui l'amène à une position gagnante en au plus k étapes. Le calcul se fait récursivement sur k :

- $W_0 = \Omega_a$ ensemble des positions gagnantes du joueur A (qui peuvent appartenir à S_a ou à S_n).
- Au rang $k+1$, on obtient W_{k+1} en ajoutant à W_k :

- d'une part les états contrôlés par le joueur A pour lesquels **au moins** un successeur appartient à W_k

- d'autre part les états contrôlés par le joueur B pour lesquels **tous** les successeurs appartiennent à W_k.

La suite $(W_k)_{k \geq 0}$ est croissante dans S au sens de l'inclusion, donc stationnaire (car V est fini).
Sa limite est notée ici W, appelé "l'attracteur pour Ω_a", et est atteinte en au plus $n-1$ itérations.

Prop :

(i) Depuis tout sommet de W, il existe une stratégie pour le joueur A l'amenant dans Ω_a.

(ii) Depuis tout sommet de $S \setminus W$, il existe une stratégie pour le joueur B de rester dans $S \setminus W$, et donc en particulier de ne jamais visiter Ω_a.

Remarque :

Comme les parties sont supposées finies, elles se terminent par un état gagnant pour le joueur B. Ainsi, dans un jeu codé par un graphe acyclique, toute position est gagnante pour l'un des deux joueurs.

Dans le cas où le graphe n'est pas supposé acyclique (les parties peuvent être infinies), les sommets de $S \setminus W$ ne sont pas nécessairement gagnantes pour B (mais il existe une stratégie lui permettent de ne pas perdre).

Preuve de la prop :

(i) Soit $x \in W$.Par définition, il existe un unique $k \in \mathbb{N}$ (et même $k \leq n-1$) tel que $x \in \Delta_k$.
Si $x \in S_a$, alors il existe $v \in \Delta_{k-1}$ telle que $x \to y$. On pose $f(x) = y$.
On définit ainsi une fonction $f : W \cap S_a \to S_b$ qu'on prolonge arbitrairement en $f : S_a \to S_b$.
Alors f est une stratégie gagnante pour le joueur A en considérant une configuration associée à f_a : on construit un chemin $x_0 \to f(x_0) \to x_1 \to f(x_1) \to ...$ jusqu'à aboutir à une position finale ; tous les états visités appartiennent à W et en particulier, l'état final est dans $W \cap \Omega = \Omega_a$.

(ii) On procède en fait par "dualité". Soit $x \notin W$, c'est-à-dire $x \in V = S \setminus W$.

Par contraposée, on a pour tout $x \in V \cap S_b$, il existe $y \in S_a$ tel que $(x, y) \in E$ et $y \in V$. En effet, sinon, on aurait $y \in W \cap S_a$, donc $y \in W_k \cap S_a$ pour un certain k, et ainsi $x \in W_{k+1}$, ce qui contredit $x \notin W$.

Toujours par contraposée, soit $x \in V \cap S_a$. Alors $y \in V$ pour tout $(x, y) \in E$.

En effet, sinon, on aurait $y \in W_k \cap S_b$ pour un certain k, et ainsi $x \in W_{k+1}$, ce qui contredit $x \notin W$.

On obtient ainsi exactement la même situation qu'au (i), mais pour le joueur B (au lieu du joueur A) et pour $V = S \setminus W$ au lieu de W. D'où le résultat.

Important : Algorithme de calcul de l'attracteur

(1) On calcule le graphe inversé, c'est-à-dire on détermine les antécédents de chaque position

(2) On calcule et on stocke (par un tableau ou un dictionnaire) les degrés sortants de chaque position.

Cette structure comptera le nombre de successeurs de chaque sommet qui ne sont pas des sommets gagnants pour le joueur A. Lorsque ce compteur atteint 0 pour un sommet, on sait que ce sommet est gagnant pour le joueur A.

(3) On crée un attracteur vide

(4) On commence le calcul sur les positions qu'on sait être des sommets gagnants pour le joueur A.

Pour chaque sommet qu'on sait être gagnant et non encore rencontré (c'est-à-dire qu'il n'appartient pas à l'attracteur déjà construit) :

- a) on ajoute ce sommet à l'attracteur
- b) pour chaque prédécesseur s de ce sommet :
  - (i) on retranche 1 à son compteur de successeurs non gagnants ;
  - (ii) si s est dans S_a, il est gagnant, on déclenche donc un appel récursif dessus ;
  - (iii) sinon, s est dans S_b : si tous les successeurs de s n'a que des successeurs gagnants (le compteur associé est nul), il est gagnant, on déclenche donc un appel récursif dessus.

3) Jeux impartiaux

Dans le cas des jeux impartiaux, il est inutile de séparer les états en deux catégories. On représente donc le jeu par un graphe (S, E) et les états finaux se répartissent en positions gagnantes et positions perdantes (en général, on considère que toutes les positions finales sont perdantes).

Dans le cas où le graphe est acyclique, les sommets se répartissent alors en deux catégories :

- soit le joueur ayant la main a une stratégie gagnante
- soit l'autre joueur a une stratégie gagnante.

Dans le cas général, s'ajoutent les sommets correspondants à des parties nulles (infinies ou se terminant sur une position considérée comme nulle).

4) Jeux de gain à somme nulle

On suppose un jeu muni d'une fonction gain à somme nulle : autrement dit, le gain d'un joueur est égal à la perte de l'autre joueur. En considérant une perte comme un gain négatif, en additionnant les gains des deux joueurs, on obtient zéro. Chaque joueur veut choisir la meilleure stratégie afin de garantir un gain maximal dans le pire cas (c'est-à-dire dans le cas où l'autre joueur joue au mieux).

Il s'agit donc de suivre l'adage : $<$ Ce que tu peux espérer de mieux est d'éviter le pire $>$.

Algorithme Min-Max

L'arbre de jeu est un arbre étiqueté par les états et dont les arêtes correspondent aux coups possibles (alternativement par le joueur A et par le joueur B). La racine de l'arbre est étiqueté par l'état initial du jeu. Le nombre de fils de la racine est le nombre de coups possibles du premier joueur.

Le premier principe de l'algorithme min-max est donc de déployer l'arbre de jeu et d'attribuer une note aux situations atteintes à la fin de l'arbre : les feuilles de l'arbre. Cette note doit refléter la qualité de la situation. Elle doit donc être croissante avec la qualité de la situation : meilleure est la situation, plus haute est la note. On appelle fonction d'évaluation la fonction d'attribution de la note à une situation. La note attribuée correspond au gain espéré si l'on atteint cette situation.

Puisque seul le gain final nous intéresse, ce sont bien uniquement aux situations positionnées en une feuille de l'arbre qu'est appliquée la fonction d'évaluation. Le problème qui se pose maintenant est de déterminer comment utiliser ces valeurs des feuilles de l'arbre de jeu.

Le principe de l'algorithme min-max est de propager vers le haut de l'arbre les valeurs qui ont été attribuées aux situations feuilles afin d'attribuer, par propagation, une note aux autres situations présentes au sein de l'arbre de jeu jusqu'à la situation initiale située à la racine de l'arbre.

- Quand c'est au joueur de jouer, la note de la situation courante est donc le max des notes des situations suivantes.

- Mais une fois sur deux, c'est l'adversaire qui joue. Il va donc falloir "simuler" ce que joue l'adversaire. Pour ne pas prendre de risque dans notre stratégie, on va supposer que l'adversaire joue le mieux possible en défaveur du premier joueur. C'est là qu'intervient la propriété "jeu à somme nulle". Les intérêts des joueurs sont opposés. Ainsi, lorsque c'est à l'adversaire de jouer, la note de la situation courante est donc le min des notes des situations suivantes.

On obtient donc un arbre valué représentant les chemins possibles menant de l'état initial aux états finaux. Les feuilles sont étiquetées par un entier relatif correspondant au gain obtenu par Alice.
Les sommes de niveau pair correspondent aux états contrôlés par le premier joueur (par exemple Alice) et les sommets de niveau impair correspondent aux états contrôlés par le second joueur (par exemple Bob). A partir des feuilles de l'arbre, la valuation des sommets est donc définie par :

- $f(x) = \max\{f(y),\, y \text{ fils de } x\}$ si x est un état contrôlé par Alice (niveau pair)

- $f(x) = \min\{f(y),\, y \text{ fils de } x\}$ si x est un état contrôlé par Bob.

Alors $f(0)$, où 0 est la racine est le gain maximal que peut espérer Alice lorsque Bob joue au mieux.

Alice joue ⇂								**2**							
↿ max						↙				↘					
Bob joue ⇂				1								**2**			
↿ min			↙		↘						↙		↘		
Alice joue ⇂		5				1				3				**2**	
↿ max	↙		↘		↙		↘		↙		↘		↙		↘
Bob joue ⇂	1		5		1		1		3		0		**2**		1

L'objectif est soit de minimiser la perte potentielle maximale (min-max) soit de maximiser le gain potentiel minimal (max-min).

Utilisation d'une heuristique : En pratique, il est impossible de développer entièrement l'arbre du jeu et de dire si une feuille correspond à une position gagnante ou non. Dans ce cas, il convient de disposer d'une fonction d'évaluation (heuristique), capable d'estimer la qualité d'une position. On définit alors une profondeur de recherche (appelée "horizon"). La qualité d'une position est souvent évaluée en effectuant plusieurs essais aléatoires de parties issues de ce sommet et en considérant une valeur associée aux essais effectués.