# Introduction

Ce projet a été réalisé par Raphaël LAPERTOT, Anthony LAFOURCADE et Thomas FAILLAUFAIX dans le cadre de l’unité PR-3001A « Résolution de problème et algorithme » de l’ESIEE Paris durant les mois de janvier et février 2019.

Le projet consiste à réaliser informatiquement un go-moku, ainsi qu’une intelligence artificielle (non-machine-learning) capable de jouer contre le joueur qui joue aussi.

*Mais enfin, qu’est-ce que le Go-moku ?*

Un go-moku est, pour faire court, un morpion où la grille n’est pas forcément 3 x 3, et le nombre de pions à aligner n’est pas forcément 3. Le joueur et l’ordinateur jouent tour par tour, l’un après l’autre, et le but est donc d’aligner un certain nombre de pions (*A*), diagonales comprises, dans une grille carrée (*N* x *N*).

Par exemple, voici une partie de go-moku, où la grille carrée est de dimension 4 x 4, et le nombre de pions à aligner est de 3. Ici, le joueur des croix (X) a gagné car il a aligné 3 pions.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| O | X | X | X |
|  | O |  |  |
| O | X |  |  |

Nous avons décidé de faire le projet en python pour avoir un haut niveau d’abstraction, en dépit des performances.

Notre programme est divisé en deux sous-fichiers :

* Le fichier fonction, qui comme son nom l’indique, regroupe toutes les fonctions du programme
* Le fichier play, qui permet de lancer le programme et de jouer contre l’ordinateur.

Nous avons mis à disposition toutes les versions de notre programme dans des dossiers pour permettre d’avoir une trace de l’avancement du code. Pour lancer le programme final, rendez-vous dans le dossier « Version finale » en ligne de commande, puis lancez « python play\_final.py ».

# MinMax

Nous allons maintenant implémenter un algorithme **MinMax**, permettant de prédire des futures positions (à un degré donné) pour évaluer la position actuelle, et donner le prochain meilleur coup, de façon un peu bourrine (mais on l’améliorera avec un élagage alpha-beta).

Brièvement, **MinMax** retourne le prochain coup de façon à minimiser la meilleure position de l’adversaire du joueur courant.

On crée donc deux fonctions :

* **minimax**, qui va s’appeler récursivement en diminuant son degré (sa profondeur) actuel jusqu’à atteindre le degré voulu, où il va utiliser **EvalPosition**. Cette fonction prend en paramètres :
  + La dimension de notre gomoku : *N*
  + La position actuelle dont on veut évaluer la position : *P*
  + Le nombre de pions à aligner : *A*
  + Le degré (la profondeur) actuel : on a déjà prédit combien de coups par rapport à la position à évaluer ? : *degre\_actuel*
  + Le degré voulu : combien de coups on veut prédire à l’avance : *degre\_voulu*
  + Quel joueur joue à ce tour ? : *joueur*

Cette fonction va retourner le score de la position qu’on lui a donné, ainsi que les coordonnées du coup optimal

* **MinMax**, qui va effectuer l’appel principal. Cette fonction a les mêmes paramètres que minimax à l’exception de *degre\_actuel*.

Remarque : Utiliser **MinMax** à un degré 0 revient à faire simplement appel à **EvalPosition** sur la position courante.

Si le degré voulu est strictement positif, **MinMax** vérifie si le degré voulu est supérieur au nombre de cases vides, puis appelle minimax (avec un *degré\_actuel* à 0), qui procède comme ci-décrit.

D’abord, on détermine toutes les prochaines positions possibles :

* Chaque case vide est un potentiel prochain coup à jouer

🡺 On récupère toutes les cases vides

* Pour chacune de ces cases vides
  + On crée une nouvelle position où le joueur courant (humain ou ordinateur) a joué dans cette case vide
  + On sauvegarde cette nouvelle position

De cette manière, on récupère toutes les prochaines positions.

Il nous faut maintenant évaluer chacune de ces positions. C’est ici qu’on fait apparaitre l’aspect récursif de **minimax** :

* On initialise le *score optimal* à –inf. si c’est le joueur qui joue, +inf. si c’est l’ordinateur, et le *coup optimal* à Null
* On diminue le degré actuel de 1
* Si le degré devient égal au degré voulu, pour chaque prochaine position *Next\_P*
  + On fait simplement **EvalPosition** sur cette position *Next\_P* pour récupérer son score
  + Sinon on fait **minimax** sur cette position pour récupérer le score de *Next\_P*, avec les mêmes paramètres (degré a déjà été diminué de 1), sauf *joueur* où on spécifie cette fois l’autre joueur, et *P* qui devient la prochaine position *Next\_P*
  + Pour chaque prochaine position, on vérifie si son score est meilleur (= inférieur si c’est l’ordi qui va jouer, =supérieur si c’est l’humain) que le *score optimal* (le meilleur score parmi les scores des positions qu’on a déjà parcouru)
    - Si c’est le cas, ce nouveau score devient le *score optimal*, et le *coup optimal* devient le coup joué pour arriver à la prochaine position
* Enfin, on retourne le *score optimal* et le *coup optimal* précédemment calculé

Et le tour est joué ! Faisons maintenant un exemple de déroulement sur une partie simple : avec un degré voulu de 2

|  |  |  |
| --- | --- | --- |
| X | O |  |
| O |  |  |
| X |  |  |

Ici, on admet que c’est à l’ordinateur (X) de jouer.

Les 5 coups possibles suivants sont les suivants :

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | O | X |  | X | O |  |  | X | O |  |  | X | O |  |  | X | O |  |
| O |  |  |  | O | X |  |  | O |  | X |  | O |  |  |  | O |  |  |
| X |  |  |  | X |  |  |  | X |  |  |  | X | X |  |  | X |  | X |

Ici, minimax va calculer le score de chacune des prochaines positions (les *sous-positions* où l’humain a joué un nouveau coup) de chacune de ces positions avec **EvalPosition**, et donc évaluer chacune de ces positions ci-dessus en gardant le score où l’humain est le plus avantagé (🡺 le score maximum) parmi les scores des *sous-positions*.

Enfin, on refait la même chose pour évaluer la position de base, sauf que cette fois on garde le score où l’ordinateur est le plus avantagé (🡺 le score minimum) parmi les scores des positions calculés. Au final, dans cet exemple, le 2ème cas est le meilleur car il aura deux lignes potentielles à compléter, et aura donc un score minimum (optimal pour l’ordinateur).

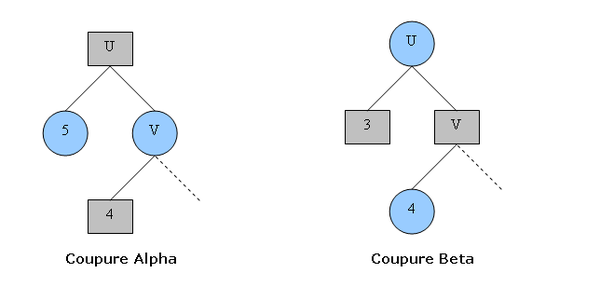
# Elagage alpha-beta :

MinMax permet d’anticiper tous les prochains coups, ce qui peut très vite devenir long. L’idée de cette implémentation est de retirer les parcours inutiles, d’éliminer les coups qu’on sait pas optimaux, de façon à accélérer l’algorithme. Il exploite le fait qu’un coup est forcément favorable (ou nul) au joueur qui l’a joué.

Pour cela, on ajoute deux paramètres à **minimax** : *alpha* et *beta*, qui représentent respectivement le *score* optimal du degré précédent (minimum pour *alpha*, maximum pour *beta*). On n’utilise qu’un des deux à la fois, dépendant de si le nœud courant est un nœud max (tour de l’humain : dans ce cas c’est *beta* qui est utilisé) ou si c’est un nœud min (tour de l’ordinateur : dans ce cas on utilise alpha).

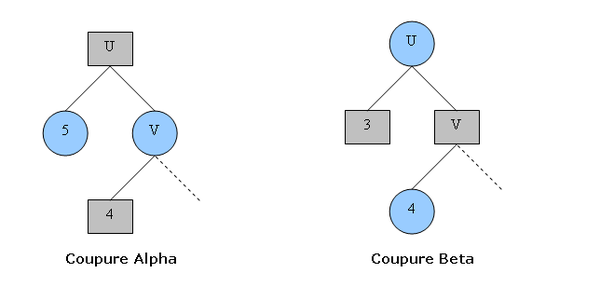
L’algorithme est le suivant :

* Si on est au tour du joueur,
  + On fait comme ce que faisait minimax
  + Durant le parcours des sous-positions, à la fin du parcours de la sous-position, on vérifie si notre *score optimal* (maximum) est supérieur à *beta*, le minimum de la position mère.
    - Si c’est le cas, ça veut dire que le score de la position courante (sous position de la position mère) sera forcément supérieur (car on cherche le max 🡺 il ne peut qu’augmenter) au *score optimal* de la position mère, qui cherche à être minimum. On arrête donc le parcours des sous-positions en retournant le *score optimal* actuel ainsi que le coup à jouer pour arriver dans cette position.



(Source : [https://fr.wikipedia.org/wiki/Élagage\_alpha-bêta](https://fr.wikipedia.org/wiki/%C3%89lagage_alpha-bêta) )

* + - Par exemple, ici, l’humain est en gris (et cherche à maximiser son score en U). Comme son *score optimal* ici vaut 5, et que le score de la prochaine position V est d’au plus 4 (minimisant le score pour l’ordinateur), on ne calcule pas le reste des sous-positions de V.
* Si on est au tour de l’ordinateur,
  + On fait comme ce que faisait minimax
  + Durant le parcours des sous-positions, à la fin du parcours de la sous-position, on vérifie si notre *score optimal* (minimum) est inférieur à *alpha*, le minimum de la position mère.
    - Si c’est le cas, ça veut dire que le score de la position courante (sous position de la position mère) sera forcément inférieur (car on cherche le min 🡺 il ne peut que diminuer) au *score optimal* de la position mère, qui cherche à être maximum. On arrête donc le parcours des sous-positions en retournant le *score optimal* actuel ainsi que le coup à jouer pour arriver dans cette position.



(Source : [https://fr.wikipedia.org/wiki/Élagage\_alpha-bêta](https://fr.wikipedia.org/wiki/%C3%89lagage_alpha-bêta) )

* + - Par exemple, ici, l’ordinateur est en bleu (et cherche à minimiser son score en U). Comme son *score optimal* ici vaut 3, et que le score de la prochaine position V est d’au moins 4 (maximisant le score pour l’humain), on ne calcule pas le reste des sous-positions de V.

Explications scoring :

Cette première version est très simple. Premièrement, on échelonne les valeurs entre -1000 et +1000, avec un pas de 1000/*A*. Par exemple, si on doit aligner 3 pions, le scoring va prendre les valeurs entre -1000, -750, -500, -250, 0, 250, 500, 750 et 1000.

Ensuite, on détermine le score de chaque joueur en comptant le nombre maximum de pions qu’il a réussi à aligner (pour l’instant, on ne considère pas le blocage, mais on le traitera), puis on le multiplie par ce pas de 1000/A. Par exemple, dans cette position (4 pions à aligner) :

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| O | X | X | X |
|  | O |  |  |
| O | X |  |  |

Le joueur O a maximum deux pions d’alignés : il a donc un score de 2\*(1000/A) = 500

Le joueur X, quant à lui, a 3 pions d’alignés : il a donc un score de 3\*(1000/A) = 750

Enfin, pour avoir un score global de la position, on soustrait les deux scores. Dans notre projet, nous avons choisi de soustraire le score du joueur X (l’ordinateur) au score du joueur O (l’humain). Dans notre exemple, le score serait donc de 500 – 750 = -250. L’ordinateur est considéré comme ayant un avantage (ce qui ne sera plus le cas après l’ajout de notre fonction blocage !).