

Les programmes.

CHESS

Dans les années 80, Control Data était champion des programmes d'Echecs sur ordinateur avec Chess 4 développé sur CDC6600 par l'université de Chicago.

Sa force était de 2100 ELO, soit le niveau d'une première série.

Les GMI lui sont supérieurs avec 2600. Boris Spassky l'avait battu facilement à travers la liaison Minneapolis et Epernay. Toutefois Spassky avait décliné de jouer en partie rapide ou blitz de 5 minutes pour toute la partie. L'arbitre lui a accordé 15 mn.

Ensuite Chess a été confronté en simultané contre 10 joueurs. Parmi les invités de Moët et Chandon, était présent Dave Scott le commandant d'Apollo 15 qui a été le premier à rouler sur la Lune.

J'étais son coach, l'arbitre nous a accordé la nulle...

Chess 4 est construit à partir de l'élagage AlphaBeta pour balayer l'arborescence de façon optimale.

Des règles heuristiques sont prédéfinies, comme l'évaluation de la position : matériel, avance de développement, lignes ouvertes...

DEEPMIND

Le champion actuel est le programme développé par DEEPMIND.

Son niveau est considérable 3600 ELO, largement au-dessus des 2800 ELO du champion Magnus Carlsen, qui jusqu'ici a évité la confrontation en direct.

En fait sur 10 parties, Carlsen n'en gagnerait pas une seule. Même avec le handicap d'un cavalier, il ne partirait pas favori.

Le programme est basé sur un nouveau principe : l'auto-apprentissage à travers l'algorithme AlphaZero. Le Zero symbolise que le programme est nul au départ.

En quelques heures, à partir seulement de la notion des déplacements des pièces, le programme a développé sa propre stratégie en jouant des millions de parties avec la méthode de Monte-Carlo, c'est à dire par tirage au sort des coups !

Un exemple simpliste : mesurer la taille du bifteck dans votre assiette. Secouez fortement une salière, le nombre de grains de sel permet d'évaluer sa surface.

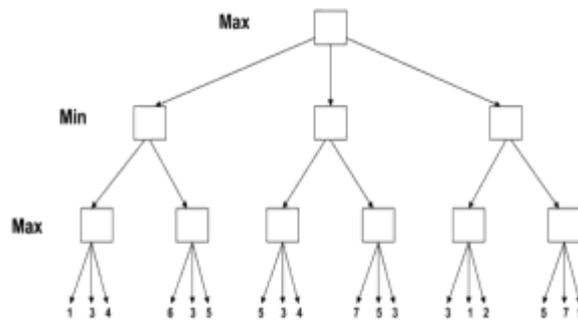
La méthode MC est utilisée dans de nombreux domaines.

Les Algorithmes.

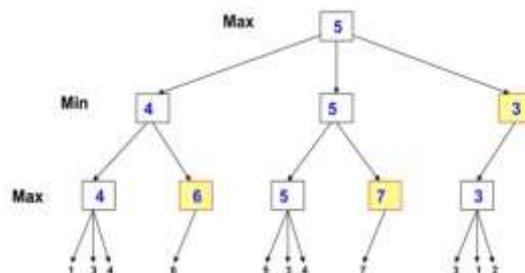
- **Elagage AlphaBeta CHESS**

Cette technique permet de réduire le nombre de nœuds évalués par l'algorithme MiniMax, qui effectue une exploration complète de l'arbre jusqu'à un niveau donné. Le maxi correspond à la stratégie de l'ordinateur et le mini à votre stratégie.

Exemple sur 2 niveaux



Les valeurs 1,3,4... correspondent à l'évaluation de la position après 2 coups. Remontée des valeurs en alternant mini et maxi.



Plusieurs branches sont éliminées, diminuant le nombre de tirages, car le résultat ne serait pas amélioré. La justification mathématique étant que la valeur Alpha devient supérieure à la valeur Beta.

A vous de vérifier ce point en rajoutant des étages avec des valeurs arbitraires.

- **AlphaZero DEEPMIND.**

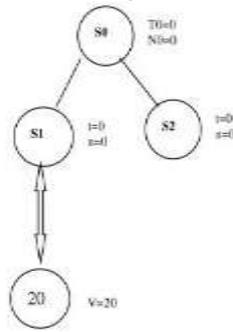
L'algorithme se déroule en 4 étapes :

- Choix du nœud selon la formule ci-dessous,
- Expansion des nœuds,
- Simulation par tirage,
- Mise à jour par remonté du résultat.

$UCB1(S_i) = V_i + 2 \times \text{Racine}(\log N/n_i)$

N précise le nombre total d'itérations n_i
 T cumule les résultats des parties t_i
 V est la valeur moyenne au nœud : T divisé par N .

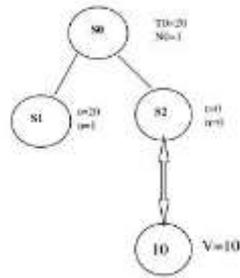
Le choix pour la première itération se porte sur S1 ou S2, UCB1 étant identique (infini)



Simulation d'une partie par tirage aléatoire- la double flèche.

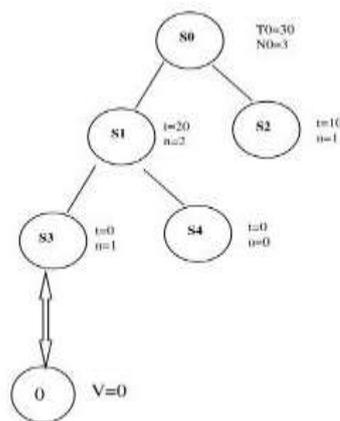
Remonté du résultat ici 20 (valeur arbitraire), 1ère itération.

Etape suivante, choix de l'autre nœud S2 et tirage d'une nouvelle partie :



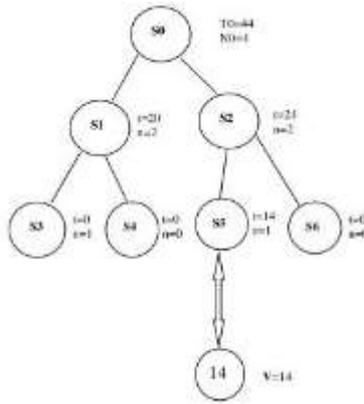
Remonté du résultat ici 10, itération 1 pour le nœud S2.

Choix du nœud suivant pour expansion S3 sur S1. UCB1 donne 12.5 pour S1, 10.5 pour S2



Remonté du résultat, ici zéro pour l'itération 2.

Choix du nœud suivant S2, son UCB1 étant maintenant supérieur à celui de S1.

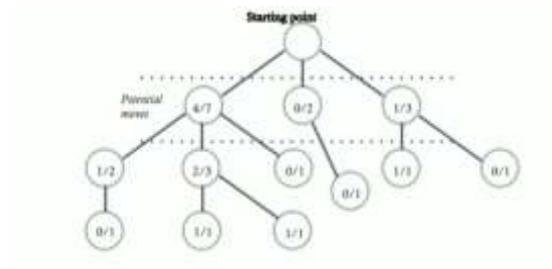


Nouveau tirage au sort et remonté des résultats ici 14.

A vous de continuer de nouvelles itérations avec des valeurs arbitraires.

Contrairement à l'exemple, AlphaZero appliqué aux Echecs simule une partie complète par tirage aléatoire de l'ensemble des coups. Le résultat est 1 gain ou 0 perte de la partie.

Sélection, Expansion
Simulation, Mise à jour



Plus de simulations sont lancées, plus l'arbre des nœuds s'élargit et plus leurs valeurs sont précises. Le rapport de parties gagnées par itération s'améliore.

En quelques heures, sur un puissant ordinateur de 5000 TDU, DEEPMIND a développé sa propre stratégie en jouant des millions de parties avec la méthode de Monte-Carlo.

Lors du déroulement d'une véritable partie, DEEPMIND avec sa stratégie supérieure se limite au calcul de 60 000 coups par minute. La puissance de calcul n'étant plus essentielle contrairement à sa phase d'apprentissage.

DEEPMIND a atteint le niveau de 3600 ELO nettement supérieur aux 2600 à 2800 ELO des GMI (Grands Maîtres Internationaux).