LES BOÎTES QUI APPRENNENT

ADRIENNE TUYNMAN

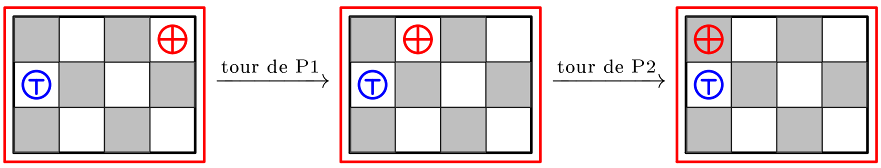
Niveau de l'activité : à partir du lycée

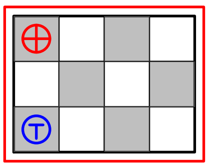
Durée de l'activité : Autour d'une heure et demie, peut être raccourci à une heure et quart avec une préparation préalable

L'objectif de l'activité est de comprendre comment une machine peut apprendre à jouer à un jeu relativement simple. Il s'agit d'abord de se familiariser avec le jeu, puis de comprendre comment un algorithme peut apprendre à gagner à ce jeu.

1. Le jeu

Le principe est le suivant : sur un échiquier, on dispose un certain nombre de tours. Chaque joueur doit bouger une pièce, chacun à son tour, uniquement vers le bas ou vers la gauche. Les tours peuvent se superposer sur le plateau. Le joueur qui pose la dernière tour sur le coin en bas à gauche a gagné. Voici un exemple de partie entre deux joueurs, P1 et P2.





Le joueur P2 peut gagner la partie après ce coup de P1.

2. L'activité

Dans un premier temps, les élèves vont commencer par se familiariser avec le jeu, en y jouant par deux ou en petits groupes, selon leur effectif. Cette phase peut éventuellement être réalisée préalablement avec l'enseignant afin de gagner du temps.

Dans un second temps, ils vont pouvoir observer comment une « machine » apprend. La machine sera matérialisée de la manière suivante : on dispose d'une collection de boîtes, chacune correspondant à une configuration de la partie. Chaque boîte contient des papiers représentant les mouvements possibles pour la configuration.

Le reste de l'activité consiste pour les élèves à comprendre le fonctionnement d'un algorithme d'intelligence artificielle connu sous le nom d'*apprentissage par renforcement*. A l'état initial, tous les mouvements possibles sont disposés dans les boîtes. Les élèves vont alors jouer contre la machine, animé par l'un.e des élèves ou par l'animateur qui piochera au hasard son prochain coup. En fonction du résultat de la partie, en particulier selon que la machine ait gagné ou perdu, le contenu des boîtes sera modifié, matérialisant ce que la machine a appris de la partie. Partie après partie, les élèves verront alors que l'ensemble des mouvements possibles se raréfie, jusqu'à atteindre une situation où seuls les mouvements gagnants sont disponibles.

Enfin, en fonction du déroulement, du temps restant, et du niveau des élèves, il est envisageable d'exposer une stratégie gagnante pour le jeu. Cette dernière repose sur une démonstration assez complexe portant sur l'opération logique « ou exclusif », mais néanmoins très intéressante. On pourra éventuellement se demander si la machine implémente cette stratégie, et pourquoi.

3. Notions d'informatique abordées

La première notion informatique étudiée est l'algorithmique : les élèves devront comprendre et pouvoir suivre l'algorithme permettant à la machine d'apprendre à bien jouer. Ils réfléchiront aussi sur la notion d'apprentissage en informatique : la machine n'est pas consciente, ne réfléchit pas, mais va s'améliorer en « apprenant » de ses erreurs. Enfin, la stratégie générale repose sur des portes logiques, ce qui peut leur permettre de comprendre l'utilité de tels outils, qui ne sont pas qu'abstraits.

LES BOÎTES QUI APPRENNENT - FICHE PROFESSEUR

ADRIENNE TUYNMAN

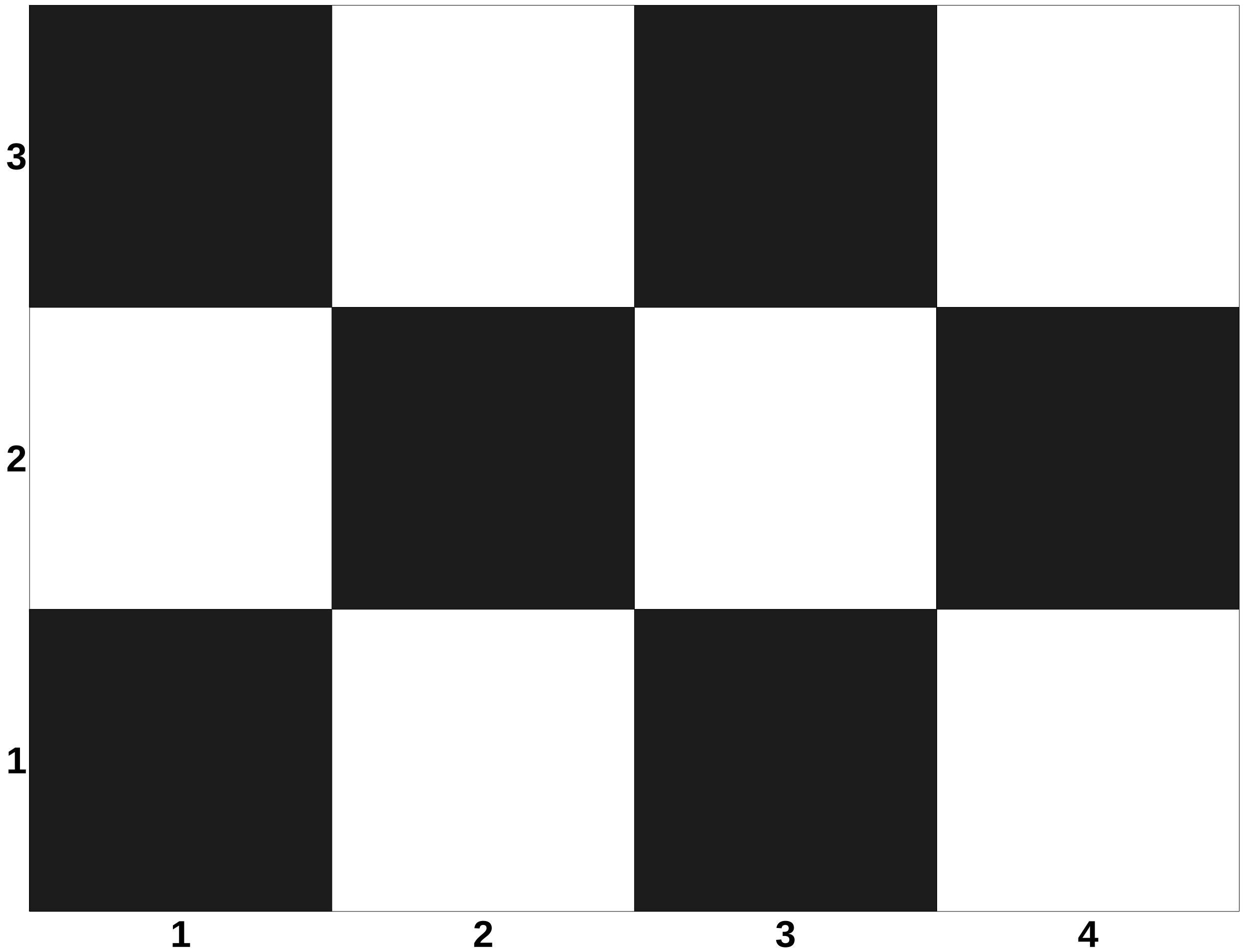
Durée : 1h35
Matériel :
- 48 gobelets
- 2 sets de papiers "Mouvements"
- Plateaux 4×3 (un par groupe de deux ou trois élèves)
- Jetons (deux par échiquier)

Public :
1ère - Terminale
Déroulement :

Durée	**Phase**	**Activités et consignes**	**Matériel**
5"	Introduction de la séance	*On entend beaucoup parler aujourd'hui de l'intelligence artificielle, et de l'apprentissage informatique. Pourtant, une machine n'est pas consciente, et ne peut donc pas réfléchir comme nous. Nous allons voir comment une machine peut apprendre à jouer à un jeu.*	
5"	Apprentissage du jeu	Dessiner un petit échiquier au tableau avec au moins deux tours de couleurs différentes, et modéliser une partie. Répondre aux questions des élèves portant sur les règles du jeu.	Tableau noir ou velleda
15"	Jeu en groupe	Les élèves sont répartis en groupes de deux ou trois et jouent au jeu avec deux tours : une tour est placée dans le coin en haut à droite, et la deuxième tour est sur la case au dessus du coin opposé. Pendant ce temps, les responsables construisent les machines en disposant les gobelets selon deux échiquiers 4×3 et en mettant les papiers de mouvement dans les bons gobelets. Les papiers intitulés 4,3,1 vont par exemple dans le gobelet de coordonnées 4,3 du premier échiquier.	Plateaux de jeu, gobelets, papiers Mouvements
5"	Explication des machines	*Nous avons construit deux machines avec des gobelets. Chaque gobelet représente une certaine configuration du plateau. Ces machines peuvent jouer au jeu : si on pioche dans le gobelet correspondant à la situation, le papier pioché va donner le coup de la machine. A la fin de la partie, si la machine a perdu, on lui enlève le papier correspondant au dernier choix qu'elle a fait.*	

Durée	**Phase**	**Activités et consignes**	**Matériel**
30"	Apprentissage des machines	*Vous allez faire deux équipes, et chaque équipe va avoir une machine à faire apprendre. Après, on va faire jouer les machines l'une contre l'autre, et on verra laquelle va gagner.* Chaque machine est opérée par un animateur, qui va jouer pour la machine. Tour à tour, les élèves jouent une partie contre la machine. C'est eux qui choisissent la configuration de départ : la première tour peut être n'importe où sur le plateau, et la deuxième est soit dans le coin de fin, soit juste au dessus. On veut au moins une bonne quinzaine de parties par machine.	
15"	Tournoi des machines	On fait quatre parties, avec la même position de départ, et en alternant la première machine qui joue.	
15"	Débriefing	Faire participer les élèves : - que s'est-il passé ? - lors de l'entraînement, vaut-il mieux perdre ou gagner contre la machine ? - y a-t-il d'autres méthodes pour l'enseignement que d'enlever les coups perdants ? - combien de temps ça prendrait pour que la machine soit parfaite ? Présenter la stratégie gagnante. Cette explication est bien évidemment accompagnée d'une simulation au tableau. Les machines implémentent-elles cette stratégie ? Pourquoi, ou pourquoi pas ? Souligner que, puisque les élèves ne jouent pas nécessairement de manière optimale, peut-être reste-t-il des coups faisables par la machine qui la font perdre.	
5"	Ouverture	Les problèmes de l'apprentissage : si les coups possibles sont trop nombreux, ou les configurations trop compliquées à expliquer à la machine, il y a des problèmes d'implémentation ; vitesse de convergence. Proposer aux élèves de programmer un algorithme qui met en place cette machine, ou d'apprendre à une machine à jouer au jeu avec plus de tours et un plus grand plateau, pour voir en combien de parties la machine est optimale.	

Plateau



Mouvements

4,3,1	**1 vers le bas : 1,1**
4,3,1	**1 vers le bas : 4,2**
4,3,1	**2 vers le bas : 4,1**
4,3,1	**1 vers la gauche : 3,3**
4,3,1	**2 vers la gauche : 2,3**
4,3,1	**3 vers la gauche : 1,3**
4,3,2	**1 vers le bas : 4,2**
4,3,2	**2 vers le bas : 4,1**
4,3,2	**1 vers la gauche : 3,3**
4,3,2	**2 vers la gauche : 2,3**
4,3,2	**3 vers la gauche : 1,3**
3,3,1	**1 vers le bas : 3,2**
3,3,1	**2 vers le bas : 3,1**
3,3,1	**1 vers la gauche : 2,3**
3,3,1	**2 vers la gauche : 1,3**
3,3,1	**1 vers le bas : 1,1**
3,3,2	**1 vers le bas : 3,2**
3,3,2	**2 vers le bas : 3,1**
3,3,2	**1 vers la gauche : 2,3**
3,3,2	**2 vers la gauche : 1,3**
2,3,1	**1 vers le bas : 2,2**
2,3,1	**2 vers le bas : 2,1**
2,3,1	**1 vers la gauche : 1,3**
2,3,1	**1 vers le bas : 1,1**
2,3,2	**1 vers le bas : 2,2**
2,3,2	**2 vers le bas : 2,1**
2,3,2	**1 vers la gauche : 1,3**
1,3,1	**1 vers le bas : 1,2**
1,3,1	**2 vers le bas : 1,1**
1,3,1	**1 vers le bas : 1,1**
1,3,2	**1 vers le bas : 1,2**
1,3,2	**2 vers le bas : Victoire de la machine**
2,1,1	**1 vers la gauche : 1,1**
2,1,1	**1 vers le bas : 1,1**
2,1,2	**1 vers la gauche : Victoire de la machine**
1,1,1	**1 vers le bas : Victoire de la machine**

4,2,1	**1 vers le bas : 4,1**
4,2,1	**1 vers la gauche : 3,2**
4,2,1	**2 vers la gauche : 2,2**
4,2,1	**3 vers la gauche : 1,2**
4,2,1	**1 vers le bas : 1,1**
4,2,2	**1 vers le bas : 4,1**
4,2,2	**1 vers la gauche : 3,2**
4,2,2	**2 vers la gauche : 2,2**
4,2,2	**3 vers la gauche : 1,2**
3,2,1	**1 vers la gauche : 2,2**
3,2,1	**2 vers la gauche : 1,2**
3,2,1	**1 vers le bas : 3,1**
3,2,1	**1 vers le bas : 1,1**
3,2,2	**1 vers la gauche : 2,2**
3,2,2	**2 vers la gauche : 1,2**
3,2,2	**1 vers le bas : 3,1**
2,2,1	**1 vers le bas : 2,1**
2,2,1	**1 vers la gauche : 1,2**
2,2,1	**1 vers le bas : 1,1**
2,2,2	**1 vers le bas : 2,1**
2,2,2	**1 vers la gauche : 1,2**
1,2,1	**1 vers le bas : 1,1**
1,2,1	**1 vers le bas : 1,1**
1,2,2	**1 vers le bas : Victoire de la machine**
4,1,1	**1 vers la gauche : 3,1**
4,1,1	**2 vers la gauche : 2,1**
4,1,1	**3 vers la gauche : 1,1**
4,1,1	**1 vers le bas : 1,1**
4,1,2	**1 vers la gauche : 3,1**
4,1,2	**2 vers la gauche : 2,1**
4,1,2	**3 vers la gauche : Victoire de la machine**
3,1,1	**1 vers la gauche : 2,1**
3,1,1	**2 vers la gauche : 1,1**
3,1,1	**1 vers le bas : 1,1**
3,1,2	**1 vers la gauche : 2,1**
3,1,2	**2 vers la gauche : Victoire de la machine**

LES BOÎTES QUI APPRENNENT, ANNEXE : MODE D'EMPLOI DE LA MACHINE

ADRIENNE TUYNMAN

1. Le matériel pour une machine

Un plateau en échiquier 4×3
Deux jetons, un rouge et un bleu
24 gobelets
Les papiers d'une fiche "mouvements" découpés

2. Préparation

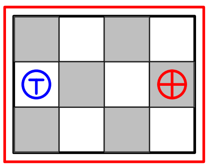
Disposer le plateau sur la table, avec le jeton rouge dans n'importe quelle case, et le jeton bleu soit en $(1, 1)$, soir en $(1, 2)$

Répartir les papiers Mouvements dans les gobelets : disposer les gobelets en deux grilles de même forme que l'échiquier, puis déposer les papiers numérotés $i, j, 1$ dans le gobelet de coordonnées i, j du premier échiquier, et les papiers numérotés $i, j, 2$ dans le gobelet de coordonnées i, j du deuxième échiquier. Le nombre de papiers dans les gobelets n'est pas toujours le même ; c'est normal.

C'est la machine qui joue le premier coup.

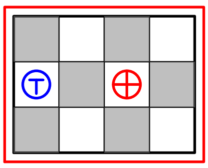
3. Le jeu

Il va falloir tirer un papier dans le bon gobelet pour pouvoir exécuter le coup suivant. Pour cela, on regard d'abord la position du jeton blanc. Si le jeton blanc est dans le coin, on va considérer la deuxième grille de gobelets ; si il est dans la case au dessus du coin, on va considérer la première. Ensuite, les coordonnées du jeton noir vont donner le gobelet dans lequel on doit tirer le papier.



Dans cette situation, on pioche un papier dans le gobelet $(4, 2)$ de la première grille, donc un papier numéroté $4, 2, 1$.

L'opérateur effectue l'opération inscrite sur le papier pioché : il s'agit, sauf dans le cas où c'est indiqué, de faire bouger le jeton rouge. Par exemple, dans cette situation, si l'opérateur tire le papier intitulé "4,2,1 : 1 vers la gauche, 3,2", le plateau devient :



Ensuite, si le papier tiré par l'opérateur était le dernier du gobelet, on le replace dans le gobelet approprié (on ne doit pas avoir de gobelet vide). Sinon, il garde le papier devant lui, en prenant garde d'ordonner les papiers par ordre chronologique.

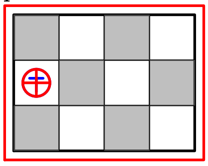
4. Fin de partie

Si la machine a gagné et a posé le dernier jeton sur la case de coin, l'opérateur remet tous les papiers dans les gobelets correspondants.

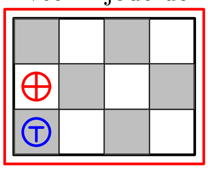
Si la machine a perdu, l'opérateur ne remet pas dans le gobelet le dernier papier tiré, qui est éliminé, mais remet les autres. Il ne devrait y avoir aucun gobelet vide, car l'opérateur n'a gardé que les papiers qui ne vidaient pas leur gobelet. Si chaque papier tiré durant la partie était le dernier de son gobelet, il ne se passe rien.

On peut recommencer la partie, en laissant les papiers éliminés au fil des parties passées de côté. Ce sont ces papiers éliminés qui permettent le processus d'apprentissage.

Continuons l'exemple ci-dessus. Supposons que le joueur choisisse de jouer le coup :



Supposons que l'opérateur pioche le papier 1,2,1 : Tour blanche sur le coin d'arrivée. Il joue donc le coup :



Et le joueur gagne alors.

L'opérateur doit alors remettre le papier 4,2,1 dans le gobelet correspondant, et se débarasser du papier 1,2,1, puisque ce papier était le dernier de la machine à être joué. Toutefois, si le papier 1,2,1 était le dernier de son gobelet, alors le papier 4,2,1 est le dernier tiré, et c'est celui-là qu'on enlève du jeu.

5. Explication

Les papiers éliminés représentent les "mauvais choix", en quelque sorte. Une fois qu'on élimine tous les choix qui font perdre la machine, il ne reste que les bons choix, et la machine est optimale.

LES BOÎTES QUI APPRENNENT, ANNEXE : COMPLÉMENTS THÉORIQUES

ADRIENNE TUYNMAN

1. Stratégie optimale

1.1. **Introduction.** La stratégie présentée ici devrait pouvoir être mise en oeuvre par toute personne sachant décomposer un nombre sous écriture binaire, mais la démonstration de son optimalité nécessite de bonnes bases en raisonnement mathématique. Elle ne doit par conséquent être abordée que si le niveau des élèves le permet.

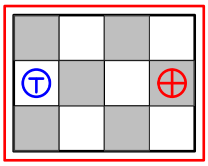
Il s'agit ici de la retranscription en français du cas $n = 3$ de la stratégie optimale du jeu de Nim à n tas telle qu'elle est donnée sur le site https ://brilliant.org/wiki/nim/ .

1.2. **La stratégie gagnante.** On a donc deux jetons sur un échiquier, le premier de coordonnées $(i+1, j+1)$ et le deuxième de coordonnées $(1, k+1)$ où i, j, k sont des entiers positifs ou nuls.

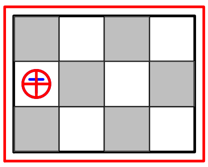
La stratégie optimale consiste à utiliser la XOR somme, qu'on notera ici $\oplus$. On rappelle que la XOR somme consiste à décomposer un nombre en son écriture en binaire, puis pour chaque bit, renvoyer $1 \oplus 1 = 0 \oplus 0 = 0$ et $0 \oplus 1 = 1 \oplus 0 = 1$. Par exemple, $5 \oplus 3 = (4+1) \oplus (2+1) = (1 \oplus 0) \times 4 + (0 \oplus 1) \times 2 + (1 \oplus 1) \times 1 = 4 + 2 = 6$.

On choisit alors le plus grand entier p tel que 2^p soit inférieure ou égal à $i \oplus j \oplus k$. On prend alors n parmi i, j, k tel que le p^e bit soit 1. On calcule alors $n' = n \oplus (i \oplus j \oplus k)$, puis on utilise cette nouvelle valeur pour bouger un pion sur le plateau. Si on avait pris $n = i$, on bouge le pion noir pour le mettre aux coordonnées $(n'+1, j+1)$.

Par exemple, dans le cas suivant :



On a $(i, j, k) = (3, 1, 1)$. On a donc $i \oplus j \oplus k = 3$. On prend $p = 1$ et $n = i = 3$. D'où $i' = 3 \oplus 3 = 0$. Le coup à jouer est donc :



Nous allons à présent montrer que cette stratégie permet de gagner toutes les parties gagnables.

1.3. **Preuve.** Commençons par remarquer que, lorsqu'un joueur peut gagner la partie en un coup, c'est à dire quand exactement deux coordonnées parmi i, j, k

sont nulles, on a nécessairement $i \oplus j \oplus k \neq 0$. En effet, pour tout entier x, on a $x \oplus 0 = 0 \oplus x = x$.

Généralisons, et montrons que, tant que la XOR somme des coordonnées est non-nulle au début du tour du joueur, il peut gagner la partie, et que dans le cas où la XOR somme est nulle au début du tour du joueur, il ne peut pas gagner la partie si son adversaire applique la stratégie ci-dessus, ce qui montre bien l'optimalité de cette stratégie. Il suffit de montrer que, si la XOR somme est nulle, un mouvement ne peut que la rendre non nulle, et que, si elle est non-nulle, la stratégie la rend nulle.

Supposons donc que $i \oplus j \oplus k = 0$, et que, sans perte de généralités, le joueur change la première coordonnée du jeton noir de i à i'. On veut donc calculer $i' \oplus j \oplus k$.

Supposons par l'absurde que $i \oplus i' = 0$. On a alors $i = i \oplus 0 = i \oplus (i \oplus i')$ puis, en utilisant l'associativité de la XOR somme (c'est à dire, le fait qu'on peut effectuer les calculs dans n'importe quel ordre et bouger les parenthèses) : $i = (i \oplus i) \oplus i'$ puis $i = 0 \oplus i' = i'$, ce qui est absurde. Par conséquent, on a $i \oplus i' \neq 0$.

$i' \oplus j \oplus k = i' \oplus 0 \oplus j \oplus k = i' \oplus (i \oplus i) \oplus j \oplus k = (i' \oplus i) \oplus (i \oplus j \oplus k)$

$i' \oplus j \oplus k = (i' \oplus i) \oplus 0$ par hypothèse.

Donc $i' \oplus j \oplus k \neq 0$.

Dans le cas où ce n'est pas i qui est modifié, mais j ou k, on utilise la commutativité de la XOR somme : $a \oplus b = b \oplus a$, et on applique le même raisonnement.

Ainsi, tout joueur commençant son tour avec une XOR somme des coordonnées nulle n'a pas d'autre choix que de la rendre non-nulle, et ne peut donc pas gagner, puisque une partie gagnée vérifie $i = j = k = 0$.

Maintenant, supposons que la XOR somme des coordonnées est non nulle, montrons que la stratégie présentée est correcte et permet de rendre la XOR somme nulle.

Tout d'abord, on a choisi le plus grand p tel que $2^p \leq i \oplus j \oplus k$. Ainsi, on a nécessairement qu'il existe n parmi i, j, k tel que le p^e bit de n est un 1. Par conséquent, on a que le p^e bit de $n \oplus 2^p$ est nul, et que les bits supérieurs sont inchangés, puisque que $i \oplus j \oplus k < 2^{p+1}$. Par conséquent, si $n' = n \oplus (i \oplus j \oplus k)$, on a bien $n' < 2^p \leq n$ et on peut bien bouger le jeton vers ses nouvelles coordonnées dans le respect des règles du jeu.

De plus, on a alors, en supposant $i = n$:

$i' \oplus j \oplus k = i' \oplus i \oplus i \oplus j \oplus k = (i \oplus j \oplus k) \oplus (i \oplus i')$.

$i' \oplus j \oplus k = (i \oplus j \oplus k) \oplus (i \oplus i \oplus (i \oplus j \oplus k)) = (i \oplus j \oplus k) \oplus (i \oplus j \oplus k)$

Donc $i' \oplus j \oplus k = 0$.

Par conséquent, en utilisant cette stratégie, un joueur qui a au début de son tour une XOR somme non nulle peut la rendre nulle, et par conséquent empêcher son adversaire de gagner et retrouver une XOR somme non nulle au début de son tour suivant.

On a ainsi montré l'optimalité de la stratégie. On en déduit de plus que certaines parties ne sont pas gagnables, même en jouant au mieux.

Cette stratégie étant optimale, on en déduit que toute stratégie développée par la machine va suivre le même schéma, et, dans toutes les configurations où c'est possible, va rendre la XOR-somme nulle.

2. Durée de convergence

Intuitivement, il suffit de compter le nombre de papiers qu'on enlève de la machine. Il y a 72 papiers répartis dans 23 gobelets ; par conséquent, en enlevant 49 papiers, on déterminise la machine, elle n'a qu'un choix possible par situation.

Comme on a enlevé à chaque fois le dernier choix pris, on garantit qu'on enlève uniquement des "mauvais" choix, soit parce qu'ils sont dans une situation irrémédiablement perdante (et dans ce cas, aucun choix n'est bon), soit parce que le choix est un qui rend la partie irrémédiablement perdante, étant donné que les coups suivants étaient déjà fixés (n'avaient qu'un papier dans le gobelet) donc étaient, par le même raisonnement, les meilleurs choix possibles.

Par conséquent, la machine a besoin d'au moins 49 parties perdues pour être optimale. On en déduit également qu'il vaut mieux perdre des parties contre la machine plutôt que d'en gagner, puisque gagner contre la machine ne lui apprend rien.

3. Autres méthodes de renforcement

On pourrait, par exemple, doubler le nombre de papiers pour un certain mouvement si ce mouvement amène à la victoire. Finalement, le plus important est, au fil du temps, de diminuer la probabilité des coups qui ont tendance à faire perdre, et augmenter la probabilité des coups qui ont tendance à faire gagner.

4. Addendum : sur la forme du plateau

La forme du plateau a été choisie ainsi afin de ne pas poser de problèmes de lecture de coordonnées (on peut voir du premier coup d'oeil où sont les abscisses et où sont les ordonnées) ; quant à sa taille, elle a été choisie pour avoir une durée de convergence raisonnable. Un plateau 2x3 ne nécessite que 12 défaites, et un plateau 5x4 en nécessiterait une bonne soixantaine ; une cinquantaine de défaites est un bon entre-deux.