

Jean-Baptiste Caillau^{*}, Christophe Cazanave^{*}, and Marc Monticelli^{*}

¹Université Côte d’Azur, CNRS, LJAD, France

Jeu de Nim : l’IA en boîte (d’allumettes)

- **Niveau** : 6ème à Terminale
- **Durée** : 1h30
- **Objectifs** : expliquer les idées sous-jacentes à l’*apprentissage par renforcement* utilisé en intelligence artificielle, illustrer le besoin de mathématiciens dans des applications désormais omniprésentes
- **Pré-requis** : aucun
- **Notions** : jeu, stratégie, algorithme
- **Matériel** (par binôme) : 8 allumettes, $50 \times 3 = 150$ billes (ou perles à repasser...) de 3 couleurs différentes, 8 boîtes d’allumettes

Introduction (5 min)

On parle beaucoup des avancées spectaculaires récentes en intelligence artificielle (DeepBlue, AlphaGo, chatGPT) : comment fonctionnent ces “machines” qui battent désormais systématiquement les grands maîtres d’échecs, de go, rédigent des programmes informatiques... ou des dissertations ? Le but de cet atelier est d’ouvrir ladite machine (la boîte d’allumettes !) et d’illustrer sur l’exemple du jeu de Nim un algorithme élémentaire d’intelligence artificielle qui apprend *seul* itérativement (par *renforcement*) la stratégie gagnante de ce jeu.

Manipulation no. 1 (20-30 minutes)

Donner les règles du Jeu de Nim (on peut par exemple illustrer avec un extrait vidéo de Fort Boyard) :

1. On dispose côte à côte 8 allumettes¹.

1. Le nombre 8 est différent de 1 modulo 4, ce qui assure que le premier joueur a une stratégie gagnante.

2. Les joueurs retirent chacun leur tour de 1 à 3 allumettes.
3. Le joueur qui prend la dernière allumette perd la partie.

Laisser les élèves s'approprier les règles en jouant par binômes pendant plusieurs parties. Une fois que les règles sont bien comprises, les amener à remarquer qu'ils ne peuvent jamais vous battre si vous commencez. Ils comprennent ainsi que le premier joueur a une stratégie gagnante qu'ils doivent découvrir. Cette découverte pourra être favorisée par un changement du nombre initial d'allumettes (partir de 4, 5...).² Débriefing et laisser au tableau un schéma illustrant le coup gagnant à jouer en fonction du nombre d'allumettes restantes. S'assurer que tous les élèves ont bien compris la stratégie et qu'ils savent "récupérer la main" si le premier joueur se trompe de coup.

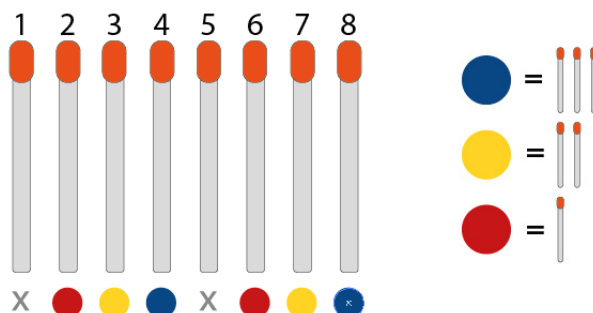


FIGURE 1 – Schéma des coups à jouer en fonction du nombre d'allumettes restantes

0.1 Intelligence artificielle (5-10 minutes)

Le but de la suite est de construire une machine qui apprend à jouer. Insister sur le fait qu'il ne s'agit pas de programmer la solution qu'ils viennent de trouver dans un ordinateur : ce serait alors le programmeur qui serait intelligent et non la machine. On propose donc un algorithme d'apprentissage automatique : la machine apprend à jouer *toute seule* ; on ne lui donne que les règles du jeu et une méthode d'apprentissage pour s'améliorer partie après partie. On peut illustrer avec l'histoire des échecs et comparer les programmes classiques (type *Stockfish*) et le programme révolutionnaire *AlphaZero*. Les élèves sont intéressés aussi par les machines qui jouent seules

². Pour les niveaux plus avancés, on peut utiliser le vocabulaire des congruences. Pour les autres, on remarque juste qu'il y a une répétition périodique des positions gagnantes et perdantes (et des coups à jouer).

aux jeux vidéos.³ La méthode d'apprentissage repose sur deux principes très naturels : l'exploration, en testant tous les coups possibles, et le renforcement, en privilégiant les coups qui font gagner et en oubliant ceux qui font perdre.

0.2 Manipulation "à la Menace" (20-30 minutes)

On suit l'idée de la machine MENACE (Matchbox Educable Noughts and Crosses Engine) de Donald Michie [2, 1] qui jouait au morpion, en l'adaptant au jeu de Nim. Par groupe de 4, les élèves ajoutent sous chaque allumette du jeu une boîte remplie avec 4 billes de chacune des 3 couleurs. Les boîtes servent à faire jouer la machine. Lorsque c'est au tour de la machine, les élèves tirent au hasard dans la boîte de la position courante une bille et la posent devant la boîte. Chaque couleur correspond à un coup possible (prendre les mêmes couleurs que dans la figure 1).

Les élèves commencent l'apprentissage de la machine en suivant le protocole des figures 2 et 3. Un binôme fait jouer la machine, l'autre binôme joue comme un expert. La machine joue en premier. À la fin de chaque partie, les élèves renforcent la machine. On échange les rôles des binômes après 4 ou 5 parties.

Lors de la première partie, s'assurer que les élèves ont bien compris la mise en place⁴ et comment faire jouer la machine⁵ et comment effectuer la phase de renforcement. On constate empiriquement qu'il faut environ 100 parties pour obtenir une machine experte. Pour une fête de la science, on doit pouvoir atteindre ce nombre de parties. Pour une activité en classe, c'est trop fastidieux, on peut passer alors à une simulation numérique après avoir joué une dizaine de parties et constaté que la machine commence parfois à gagner.

0.3 Simulation numérique

Une applet java est disponible sur le site de la MMI[4] à Lyon. Constaté que la machine obtient le motif périodique de la stratégie gagnante (voir Fig.1). Commenter le cas des boîtes correspondant aux positions perdantes (il y restera des billes de plusieurs couleurs). Montrer et commenter les dif-

3. On peut donner un rapide aperçu des différentes formes d'intelligence artificielle que l'on ne va pas illustrer (apprentissage supervisé et apprentissage non supervisé). Introduire le vocabulaire d'apprentissage *par renforcement*.

4. Un petit détail souvent difficile à intégrer est que dans les deux premières boîtes, on ne met pas toutes les couleurs car certains coups ne sont plus possibles.

5. Au bout d'un moment, il peut arriver qu'une boîte ne contienne plus de billes du tout. Dans ce cas, on réinitialise la boîte avec 4 billes de chaque couleur.

The diagram shows four boxes, labeled Boite 1, Boite 2, Boite 3, and Boite 4, each containing a different number of colored balls. A legend on the right indicates the value of each color: blue = 2, yellow = 1, and red = 0.5.

Boite	Red Balls	Yellow Balls	Blue Balls
Boite 1	4	0	0
Boite 2	4	4	0
Boite 3	4	4	4
Boite 4	4	4	4

The diagram illustrates the experimental setup. At the top, there are three horizontal bars with a red segment on the left and a grey segment on the right. A triangle points down from the middle bar. Below these are five boxes labeled Boite 1, Boite 2, Boite 6, Boite 7, and Boite 8. Each box contains a different combination of colored balls (red, yellow, and blue).

The diagram illustrates a multi-armed bandit problem with 8 boxes. Boite 1 contains 4 red balls. Boite 5 contains 10 balls (4 red, 3 yellow, 3 blue) and a lever with a triangle. Boite 7 contains 6 balls (2 red, 2 yellow, 2 blue). Boite 8 contains 6 balls (2 red, 2 yellow, 2 blue). A blue ball is shown being pulled from Boite 5.

férences de rapidité de convergence selon que le second joueur est naïf, expert ou que la machine joue contre elle-même. On peut conclure en insistant sur toutes les questions mathématiques liées à ce genre d'algorithme (par exemple, autour de la vitesse de convergence) et sur l'utilité de former de former des étudiants compétents en mathématiques et en informatique.

4

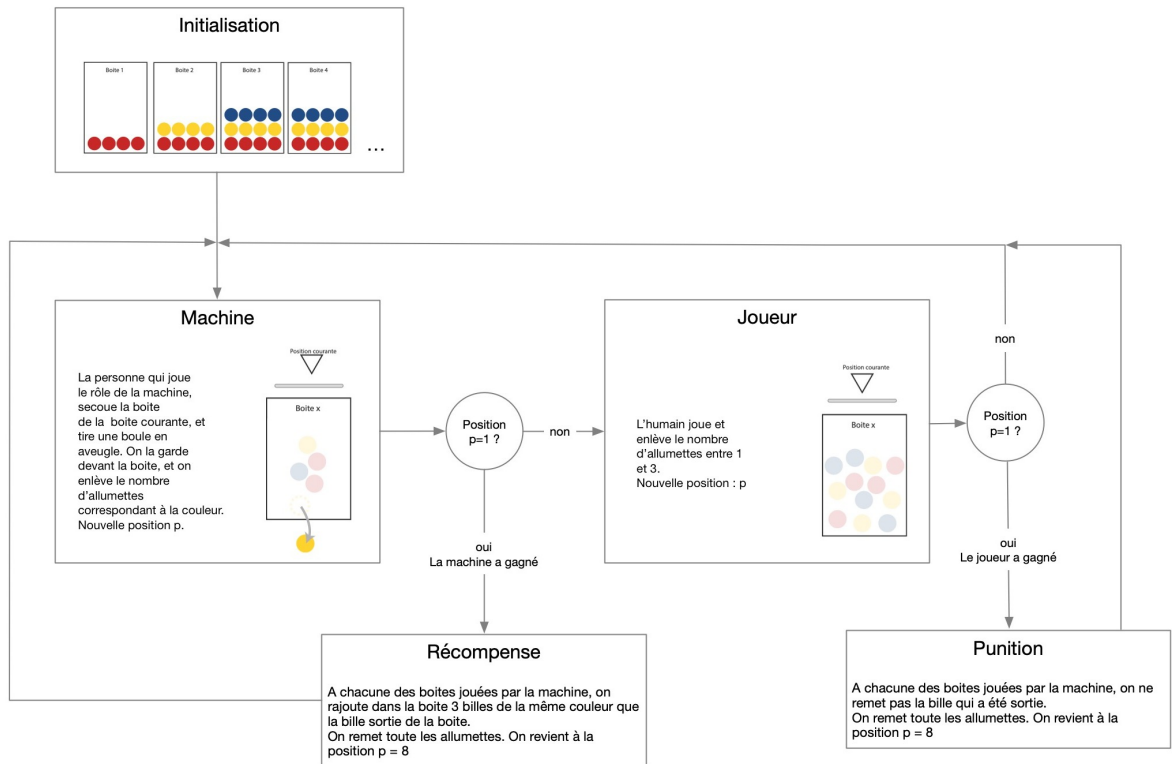


FIGURE 3 – Protocole d'apprentissage de la machine.

voire se reporter à l'ouvrage de référence sur le sujet [5].

Références

- [1] Michie, D. Experiments on the Mechanization of Game-Learning Part I. Characterization of the Model and its parameters. *Comput. J.* **6** (1963), 232-236.
- [2] Article Wikipédia *Matchbox Educable Noughts and Crosses Engine*
- [3] Franck, E. ; Privat, Y. Contrôler, optimiser, décider. *Gazette de la SMF* **174** (2022), 9-21.
- [4] Jouer au de NIM contre une machine, Maison des Mathématiques et de l'informatique de Lyon
- [5] Sutton, R. S. ; Barto, A. G. Reinforcement learning : an introduction, MIT press, 2015