

Un jeu pour comprendre l'apprentissage automatique

L'informatique débranchée permet de faire comprendre le principe de l'apprentissage automatique. L'activité présentée se base sur le jeu de Nim, dont les coups gagnants sont « mémorisés » à l'aide de pois chiches stockés dans des boîtes de lait...

* Cette activité a été élaborée à partir d'une idée d'Éric Duchêne et Aline Parreau.

C', est Alan Turing qui, le premier, a tenté de relever le défi consistant à créer un programme capable de rivaliser avec les humains pour un jeu de stratégie, dans son cas pour les échecs. Ce défi a passionné chercheurs et grand public pendant plus d'un demi siècle, et a culminé avec, à la fin du XXe siècle, la victoire du programme d'IBM Deep Blue sur le champion humain Garry Kasparov. Plus récemment, en 2017, le programme AlphaGo a vaincu Ke Jie, le champion du monde humain de go, l'un des jeux considérés comme les plus difficiles. En effet, le jeu de go offre beaucoup plus de parties possibles que les échecs. De nouvelles méthodes autour des réseaux de neurones et de l'apprentissage profond ont alors été employées.

Des robots à Marienbad

Destinée à illustrer par l'expérimentation une version très simplifiée des méthodes d'apprentissage automatique à des élèves de 8 à 13 ans, cette activité* se base sur un jeu facile d'accès : le jeu de Nim. Il s'agit d'un jeu classique (voir en encadré), dont de nombreuses variantes sont complètement résolues depuis très longtemps.

Les jeux de Nim

Le jeu de Nim, ou jeu de Marienbad, se joue classiquement à deux joueurs. Dans un tas d'allumettes, les joueurs enlèvent, à tour de rôle, une, deux ou trois allumettes (à leur choix). Le joueur qui prend la dernière allumette perd la partie. Ce jeu, qui possède de nombreuses variantes, a beaucoup été étudié car sa pratique permet la découverte de stratégies gagnantes.



Les élèves jouent avec des Carambar. Il sera également nécessaire de se munir de contenants opaques (ici des boîtes métalliques de lait pour nourrisson) pour mémoriser les coups. Des objets colorés, indiscernables au toucher, seront utilisés pour coder les actions à effectuer. Des pois chiches colorés avec de l'encre et séchés au four font parfaitement l'affaire. Enfin, des étiquettes repositionnables seront utilisées pour différencier les contenants. L'activité débute par la découverte du jeu de Nim à sept Carambar. Les élèves sont invités à se familiariser avec les règles du jeu. Ils découvrent assez rapidement l'existence de stratégies gagnantes. Une variante plus simple (quatre friandises au départ auxquels on ne peut retirer qu'un ou deux éléments à chaque coup) est ensuite utilisée pour le jeu avec les robots. Dans un premier temps, le robot, représenté par un unique contenant, n'a pas de mémoire. Il est rempli de deux graines de couleurs différentes qui servent, après un tirage aléatoire, à indiquer s'il supprime une ou deux sucreries. Cette étape permet d'insister sur le fait que les actions du robot ne sont pas programmées à l'avance. Le principe du programme repose sur des actions choisies aléatoirement mais dont la probabilité va évoluer au cours de l'apprentissage.

Apprendre aléatoirement

Dans un second temps, le robot mémorise les coups gagnants ; c'est ce qui va permettre de dégager une stratégie gagnante après un grand nombre de parties. Les trois « mémoires » de la machine sont trois boîtes étiquetées respectivement 2, 3 et 4. Le tirage des graines s'effectue dans celle portant le numéro égal au nombre de Carambar restants. Pour « apprendre », deux robots jouent l'un contre l'autre. On initialise les deux engins avec une graine de chaque couleur dans chacune de leurs trois



mémoires. Pendant la partie, chaque pois chiche tiré est laissé devant la boîte dont il sort. À la fin, on remet les graines dans les boîtes en doublant celles du robot gagnant. Après un certain nombre de parties, on observe que le robot qui commence finit par perdre de plus en plus souvent...

Le programme d'apprentissage s'adapte au jeu de Nim à sept Carambar. Il faudra alors utiliser trois couleurs de graines (on peut prendre une, deux ou trois friandises) et six « mémoires » étiquetées de 2 à 7, qui correspondent au nombre de sucreries qu'il est possible de rencontrer au cours de la partie. On enlèvera la couleur correspondant à l'action de retirer trois Carambar dans la boîte 2. On optimise le remplissage en supprimant les graines correspondant aux coups « idiots », comme, par exemple, prendre les deux bonbons restants, ce qui fait perdre automatiquement.

En fin d'activité, le regroupement du contenu des « mémoires » de tous les robots de la classe permet une analyse issue des données d'un grand nombre d'expériences. En comptant les graines de différentes couleurs dans les six boîtes, on observe :

- la mémoire 2 a enregistré une majorité de graines dont la couleur correspond à l'action « prendre un seul Carambar » puisque c'est le seul coup gagnant quand il n'en reste que deux ;
- la mémoire 3 a enregistré l'action « prendre deux Carambar » ;
- dans la mémoire 4, on retrouve la couleur correspondant à « prendre trois Carambar » ;
- Tous les coups possibles à partir du tirage dans la mémoire 5 mènent à une position gagnante pour l'adversaire, on ne devrait donc rien observer de particulier ;
- Pour la mémoire 6, la couleur correspondant à l'action « prendre un Carambar » domine ;
- Dans la mémoire 7 enfin, c'est l'action « prendre deux Carambar ».

À ce moment, il est utile de revenir à la notion de probabilité et à son rôle dans le mécanisme d'apprentissage automatique. Dans le cas du jeu à sept Carambar, le joueur qui joue le premier possède une stratégie gagnante, c'est-à-dire une façon de jouer qui lui permet de gagner n'importe quelle partie, quels que soient les coups choisis par son adversaire. Comme on l'a vu, ce n'est pas le cas dans la variante à quatre Carambar, puisque chacun des coups possibles pour le premier joueur aboutit à une configuration dans laquelle son adversaire peut choisir un coup gagnant. Pour effectuer cette analyse, il est possible de représenter le jeu sous la forme d'un graphe, dans lequel les états du jeu (le nombre de sucreries) sont reliés par les coups (prendre des bonbons) qui permettent de passer de l'un à l'autre.

F.M. & M.M.



Références :

Une approche de ce type a été mise en œuvre pour des jeux plus sophistiqués, par exemple le *tic-tac-toe*, voir à cette adresse :

- <http://images.math.cnrs.fr/Une-machine-en-boites-d-allumettes-qui-apprend-a-jouer-au-Morpion.html>
- *Les jeux de Nim*. Jacques Bouteilou, Association pour le développement de la culture scientifique, 1996.