

Etude de cas en Matlab

Enoncé du projet

Ecole Sup Galilée - Coursus Ingénieur - 1^{ère} année
2016-2017

John Chaussard
LAGA – Université Paris 13
chaussard@math.univ-paris13.fr



Apprendre à une machine à jouer au morpion

Le projet consistera à apprendre à un programme à jouer au morpion en lui apprenant uniquement les règles du jeu, et en le laissant apprendre la stratégie à adopter pour gagner.

Tout ce que vous devrez coder dans le programme seront

- . **Les conditions de victoire ou de défaite du jeu** : trois pions d'un même joueur alignés.
- . **Les mouvements autorisés** : il faut jouer dans l'une des neuf cases, pas le droit de jouer sur une case déjà occupée, chaque joueur joue à tour de rôle
- . **Le jeu s'arrête** en cas de victoire ou s'il n'y a plus de place sur le plateau.

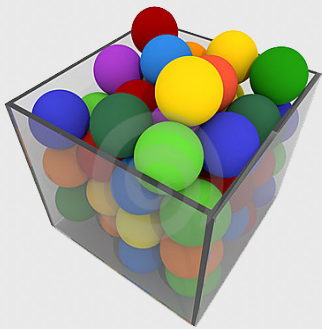


Apprendre à une machine à jouer au morpion

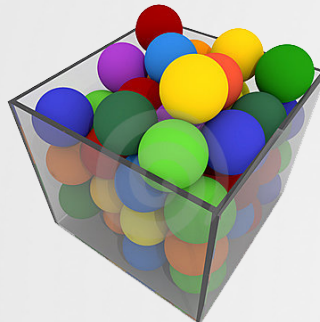
Nous **laisserons ensuite la machine jouer au hasard**, et petit à petit découvrir d'elle-même les stratégies gagnantes.

Pour ce faire, chaque « état » possible de la grille de morpions sera associé à une « boîte » dans laquelle seront disposées des balles numérotées de 1 à 9.

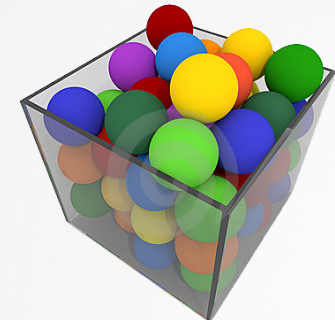
O		O
	X	
X		



X		O
	X	
X		O

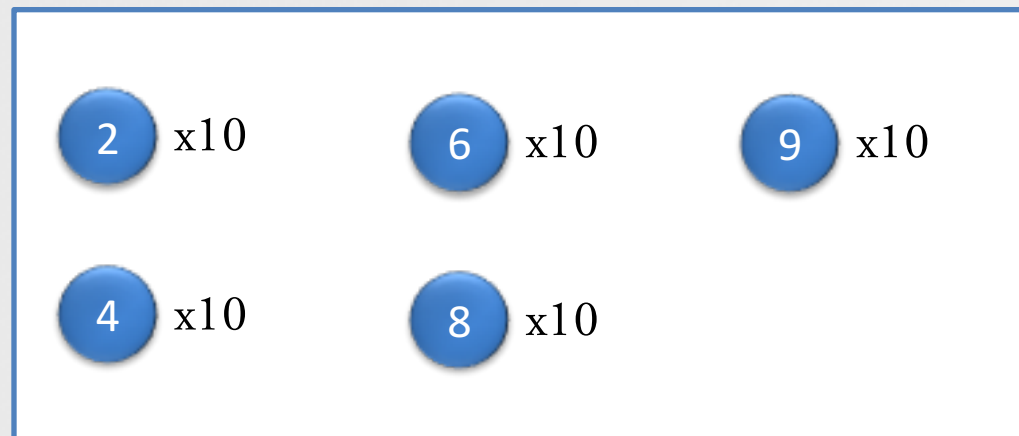
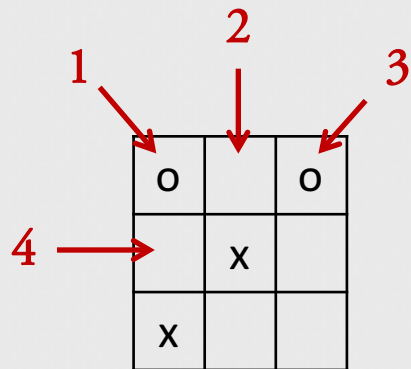


		O
X		



Apprendre à une machine à jouer au morpion

A chaque position de la grille, nous associons un chiffre de 1 à 9.



Dans la boîte associée à une grille, on dispose :

- . Dix balles avec le numéro **n** si la case **n** de la grille associée est libre
- . Aucune balle avec le numéro **n** si la case **n** de la grille n'est pas libre.

On réalise cette opération pour les $3^9=19683$ grilles possibles.



Apprendre à une machine à jouer au morpion

On commence le jeu contre la machine : imaginons que l'on a les croix, la machine a les ronds, et le joueur humain commence à jouer.

Imaginons que l'on joue ce coup :

x		

C'est au tour de la machine qui va regarder dans la boîte associée à cette grille :

1 x10	3 x10	5 x10	8 x10
2 x10	4 x10	6 x10	9 x10

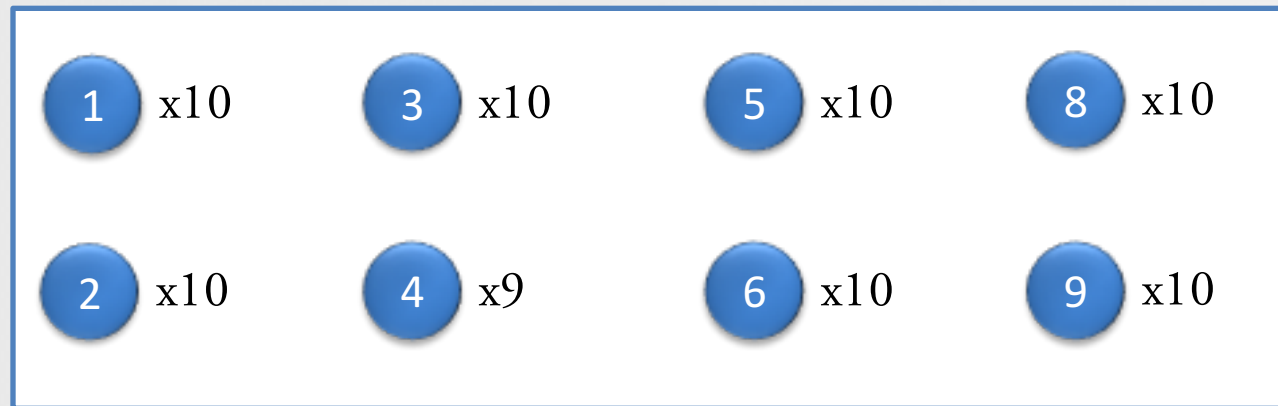


Apprendre à une machine à jouer au morpion

La machine **pioche au hasard une balle dans cette boîte**, et place son pion (les ronds) sur la case associée à ce chiffre.

Si la machine tire un 4

4



Alors elle jouera

O		
X		

On met de côté la boîte ainsi que la balle tirée, sans la replacer dans sa boîte.

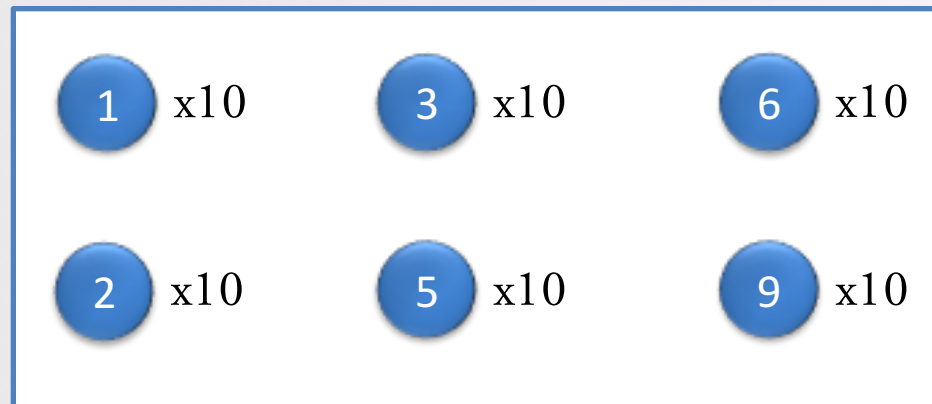


Apprendre à une machine à jouer au morpion

A nous de jouer... Imaginons que l'on joue ce coup :

o		
x	x	

C'est au tour de la machine qui va regarder dans la boîte associée à cette grille :



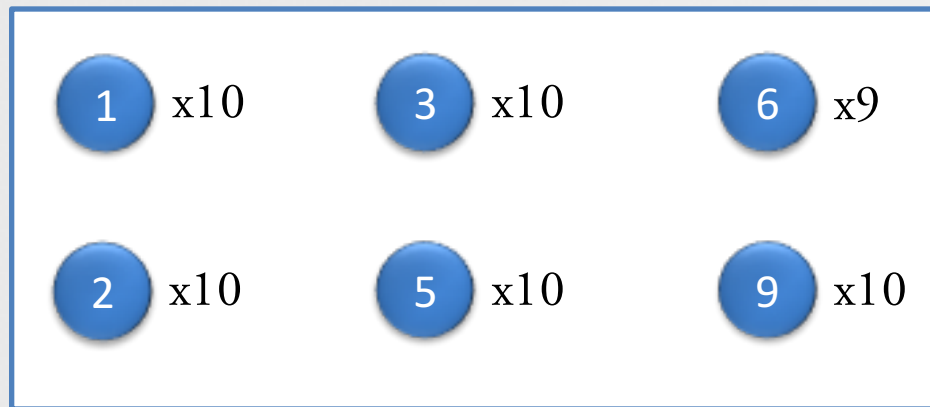


Apprendre à une machine à jouer au morpion

La machine **pioche au hasard une balle dans cette boîte**, et place son pion (les ronds) sur la case associée à ce chiffre.

Si la machine tire un 6

6



Alors elle jouera

o		o
x	x	

On met de côté la boîte ainsi que la balle tirée, sans la replacer dans sa boîte.

Apprendre à une machine à jouer au morpion

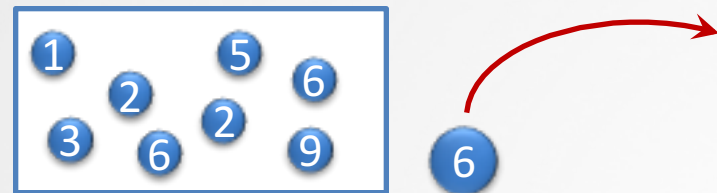
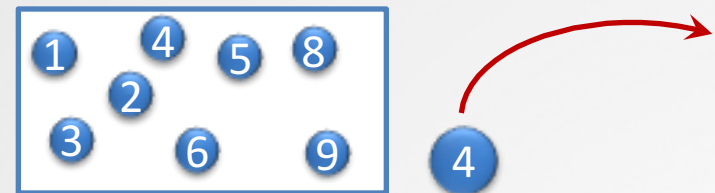
On continue à jouer, jusqu'à ce que la partie soit terminée...

Une fois la partie terminée, on **modifie la configuration de la machine afin de récompenser ou punir ses choix.**

On regarde toutes les boîtes dans lesquelles la machine a pioché pendant cette partie :

Si la machine a perdu la partie, on jette les balles qui ont été piochées.

Cela permettra, à l'avenir, de diminuer les chances de reproduire ces décisions qui nous ont fait perdre.



...

Apprendre à une machine à jouer au morpion

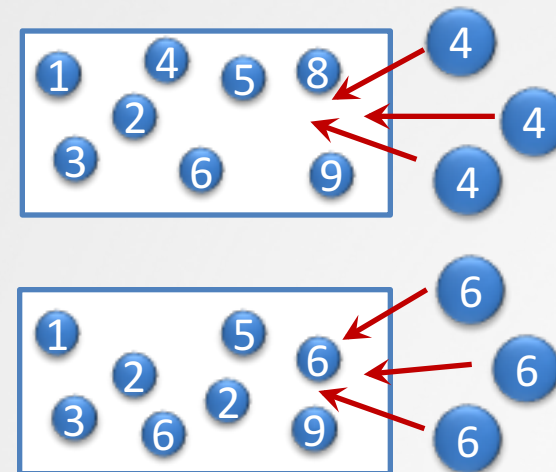
On continue à jouer, jusqu'à ce que la partie soit terminée...

Une fois la partie terminée, on modifie la configuration de la machine afin de récompenser ou punir ses choix.

On regarde toutes les boîtes dans lesquelles la machine a pioché pendant cette partie :

Si la machine a gagné la partie, on remet les balles qui ont été piochées dans leurs boîtes, et on ajoute deux balles avec le même numéro.

Cela permettra, à l'avenir, d'augmenter les chances de reproduire ces décisions qui nous ont fait gagner.



...



Apprendre à une machine à jouer au morpion

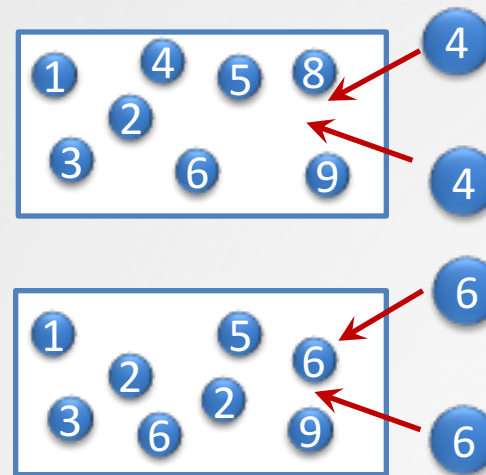
On continue à jouer, jusqu'à ce que la partie soit terminée...

Une fois la partie terminée, on modifie la configuration de la machine afin de récompenser ou punir ses choix.

On regarde toutes les boîtes dans lesquelles la machine a pioché pendant cette partie :

Si la partie s'est terminée en nul, on remet les balles qui ont été piochées dans leurs boîtes, et on ajoute une balle avec le même numéro.

Cela permet de récompenser une partie nulle, et éviter de « décourager » la machine lors des premières parties où elle aura du mal à gagner.



...



Apprendre à une machine à jouer au morpion

Petit à petit, au fur et à mesure des parties, **la machine « apprend » à jouer au morpion**, et à favoriser les décisions menant à la victoire.

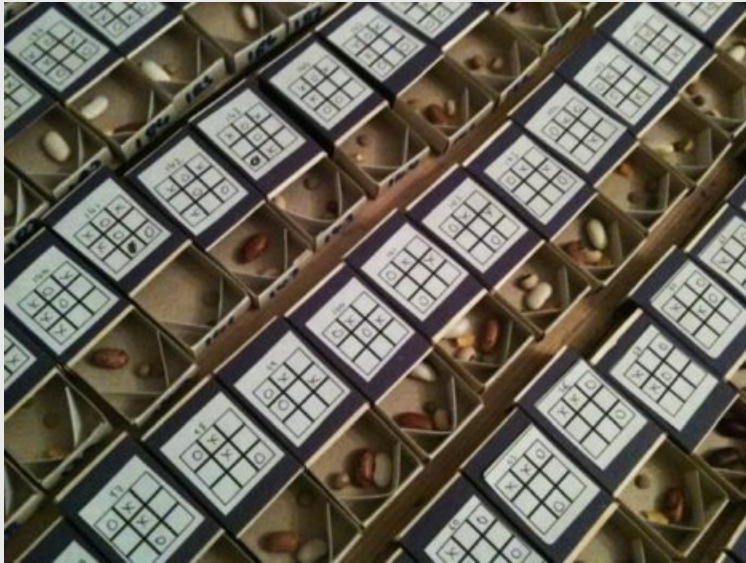
Elle n'a basé son apprentissage que sur des décisions effectuées au hasard, et sur un mécanisme de récompense /sanction à la fin d'une partie.

On peut apprendre à **la machine à jouer en jouant contre elle-même**, ou bien elle peut apprendre en jouant contre elle-même !

Apprendre à une machine à jouer au morpion

Ce dispositif a été inventé et mis en place « manuellement » par Donald Michie, en 1961.

Il utilisait des boîtes d'allumettes avec des billes de couleur pour effectuer les tirages aléatoires de balles. En prenant en compte les symétries des différents état de la grille, seulement 304 boîtes d'allumettes sont nécessaires.



Au bout de 220 parties, la machine ne perdait jamais...

Plus d'informations : <http://images.math.cnrs.fr/Une-machine-en-boites-d-allumettes-qui-apprend-a-jouer-au-Morpion.html>



Pour vous aider

Interrogez-vous sur différents points du projet :

- . Quelle structure de données pourrait représenter les boîtes de bille ?
- . Combien d'états différents une grille de morpion peut-elle avoir ?
- . Comment traduire l'état d'une grille en un simple nombre unique, tel que deux grilles avec des états différents soient traduites par deux nombres différents ?



Ce que vous devrez rendre

Pour ce projet, chaque équipe devra rendre un programme Matlab qui :

- . Possèdera une interface graphique, permettant à deux joueurs de jouer, un joueur de jouer contre l'ordinateur, ou l'ordinateur de jouer contre lui-même pour s'entraîner.
- . Devra être capable de sauvegarder les résultats de son « apprentissage » dans un fichier, ainsi que de les recharger.
- . L'interface devra clairement signifier, en cas de victoire, qui a gagné.

Votre programme sera noté en fonction de la facilité d'utilisation de votre interface, ainsi que de la qualité de votre code pour la phase d'apprentissage.



Ce que vous devrez rendre

Pour ce projet, chaque équipe devra rendre un rapport :

- . Contenant une introduction claire au problème et une conclusion, ainsi qu'une table des matières si nécessaire
- . Expliquant très clairement comment exécuter et utiliser l'interface graphique
- . Décrire votre solution, vos idées d'implémentations, vos choix, et les problèmes que vous avez pu rencontrer.

Votre rapport sera noté aussi bien sur le fond (le contenu) que sur la forme (la présentation) : ne bâclez pas votre travail !

Le rapport devra être uniquement en format PDF



Ce que vous devrez rendre

Vous devrez, à la fin de votre travail, envoyer tous vos fichiers sur l'ENT. A vous de tout mettre tout (sauf le rapport en PDF) dans un fichier zip.

Vous enverrez donc deux fichiers sur l'ENT : un fichier PDF contenant votre rapport, et un fichier ZIP contenant votre code source. A vous de vérifier, une fois l'envoi terminé sur l'ENT, que tout a été bien envoyé en re-téléchargeant votre travail et en vérifiant qu'aucun fichier n'est corrompu.

Date limite de rendu : le 27 Novembre à 23h59mn59s



Ce que vous devrez rendre

Attention, tout plagiat (de rapport ou de code) sera sanctionné par un zéro **pour les copieurs et pour les copiés.**

Ne donnez pas votre travail à un autre groupe : il est possible de les aider en leur expliquant où est leur problème ou en leur donnant des indications, mais vous ne devez pas donner votre rapport ou votre programme à autrui.

Le plagiat de code sur Internet sera aussi sanctionné par un zéro.