Théorie algorithmique de l'information
et
Apprentissage

Laurent Signac

Février 1999

1 Introduction.

De nombreux jeux vidéos actuels reposent sur l'exploration de mondes virtuels dont il faut découvrir les règles. Nous nous intéressons à la possibilité pour une machine de réussir face à ce type de situations alors que ni son programme ni son concepteur ne savent exactement ce qu'il convient de faire. L'exemple choisi consiste à diriger une «poule» à la recherche de sa nourriture qu'un fermier sadique s'évertue à lui cacher au milieu de grains empoisonnés.

2 Le problème de reconnaissance.

2.1 Formalisation.

Le joueur voit son «énergie» augmenter lorsqu'il «mange» de bons grains et diminuer s'il en prend de mauvais ou ne fait rien. Ces grains sont représentés par des matrices booléennes carrées d'ordre 100 et les comestibles par un sous-ensemble $\mathcal{C}$ de ces matrices. La «règle» à découvrir est donc $\mathcal{C}$ et, pour éviter toute indication à priori, ce sous-ensemble sera défini interactivement de manière à induire le système en erreur. Le jeu est supposé se dérouler sans limitation de durée jusqu'à épuisement de l'énergie.

Proposition 1 *S'il dispose de ressources non bornées, le joueur se maintient en vie indéfiniment en mémorisant le graphe de la fonction caractéristique de $\mathcal{C}$.*

Le problème ne se situe donc pas au niveau de la calculabilité mais, dès lors que la taille mémoire est limitée à une valeur raisonnable que nous fixons à 10^8 bits, à celui de la complexité. Les résultats de théorie de l'information vont nous permettre de l'aborder sous cet angle.

2.2 Rappels sur la Théorie Algorithmique de l'information.

Définition 1 *Un mot m est dit aléatoire si sa longueur est inférieure à la taille du plus petit programme sans saisie capable de l'écrire.*

Théorème 1 *La probabilité pour un mot de ne pas être aléatoire décroît exponentiellement avec sa longueur.*

Ce résultat conforte notre incrédulité face à quelqu'un qui prétendrait avoir obtenu la suite 1,2,3,4,5,6,1,2,3,4,5,6 en jouant aux dés. Cette forte probabilité pour $\mathcal{C}$ d'être aléatoire est inquiétante car, par définition dans ce cas, aucune fonction booléenne codée sur 10^8 bits ne pourrait le reconnaître. Pire encore, ce caractère aléatoire est le plus souvent indécidable :

Théorème 2 *Dans tout système formel, il existe un plus long mot dont le caractère aléatoire soit décidable.*

Si la probabilité de construire une fonction booléenne reconnaissant parfaitement $\mathcal{C}$ est négligeable, il reste l'espoir d'en trouver une ne commettant qu'un taux acceptable d'erreurs. Malheureusement :

Théorème 3 *Si l'ensemble est aléatoire, le taux d'erreurs d'une fonction booléenne codée sur 10^8 bits est compris entre $\frac{1}{3}$ et $\frac{2}{3}$.*

2.3 Une nouvelle hiérarchie.

En fait, la situation n'est pas toujours aussi désespérée qu'il n'y paraît. Selon la puissance de calcul utilisée par le «fermier», trois classes de jeux peuvent être définies :

Définition 2 *Un jeu est dit figé s'il existe une fonction booléenne codée sur* 10^8 *bits qui assure la pérennité.*

L'existence d'une telle fonction ne prouve évidemment pas celle d'un algorithme d'apprentissage mais elle est alors possible.

Définition 3 *Un jeu est dit chaotique si aucun algorithme ne peut assurer la pérennité sur un espace mémoire limité à* 10^8 *bits.*

Définition 4 *Un jeu est dit agité s'il n'est ni figé ni chaotique.*

Proposition 2 *Aucune des trois classes de jeux n'est vide.*

3 Algorithmes adaptatifs.

3.1 Conditions de survie.

Face à un jeu agité, un système ne peut survivre qu'avec une correction rapide de ses erreurs. Cette correction accroît ses connaissances et n'est possible qu'au prix d'un oubli des souvenirs les moins utiles. Pour modéliser l'alternance de jour et de nuit, nous interromprons les séquences de jeu par des périodes de repos. Examinons alors les qualités et défauts des méthodes classiques :

1. En l'absence de procédures d'acquisition et surtout d'élimination de règles, les systèmes experts ne peuvent être utilisés que pour des jeux figés.
2. Malgré leurs facultés d'oubli, les réseaux de neurones présentent l'inconvénient de la lenteur de convergence de leurs algorithmes d'apprentissage. Et le contrôle de l'oubli y reste très hasardeux.
3. En dépit de son côté séduisant, la programmation génétique (qui pourrait être effectuée durant le repos) n'est pas très utile dans ce cadre car l'algorithme sélectionné ne pourra faire mieux que la fonction d'évaluation utilisée.

3.2 Le modèle d'apprentissage diurne et d'oubli nocturne.

1. L'algorithme d'apprentissage suivant, qui construit sa propre base d'exemples et contre-exemples, offre l'avantage d'une correction rapide des erreurs qu'elle que soit son heuristique de comparaison :

   ```
   Matrice:=donnée ;
   Si (Comestible(PPV(Matrice))) ou (Jeûne trop long) alors manger;
   Si Comestible(Matrice) XOR Comestible(PPV(matrice)) alors
           ajouter matrice à la base;
   {Remarque: PPV = Plus Proche Voisin}
   ```

   Par hypothèse, il entraîne une saturation de la mémoire.
2. L'algorithme d'oubli suivant offre l'avantage d'un coût nul et d'un risque prévisible minimal. Notons que son implémentation impose l'interruption des liens avec les capteurs et actionneurs.

   ```
   Pour chaque souvenir,
           le retirer de la base;
           lui appliquer l'algorithme d'apprentissage;
   ```

   Théorème 4 *L'alternance des deux algorithmes précédents n'évite pas la saturation de la mémoire.*
3. L'algorithme d'oubli est complété par une heuristique de marquage-démarquage.

4 Intégration du comportement.

4.1 Les fonctions vitales.

La multiplicité des décisions à prendre dans un environnement un peu plus riche est traitée en divisant le problème de survie en fonctions booléennes :«Dois-je m'approcher de cet objet ?» ou «Dois-je m'écarter de celui-ci ?» et en gérant les inévitables contradictions entre ces fonctions.

4.2 Modèle par juxtaposition.

L'arbitrage des conflits est réalisé en associant à chaque fonction de décision une importance préprogrammée. Par exemple, la priorité maximale sera accordée à la fuite devant un prédateur, l'urgence d'écarter un rival augmentera avec le jeûne et les graines les plus proches seront examinées en premier. Nous retrouvons ainsi, mais par un cheminement différent, le modèle de Brooks.

5 Conclusion.

En raison de son caractère indécidable, la notion d'aléatoire est souvent perçue négativement. Elle offre pourtant une base théorique solide pour l'étude des algorithmes d'apprentissage et démontre, avec une forte probabilité, le caractère optimal de méthodes pourtant bien étranges à priori.