Master Données, Apprentissage et Connaissances-DAC

RAPPORT PROJET MOGPL

DICE BATTLE

Realisé par :

Djeddal Hanane
Touzari Liticia
GROUPE 2

December 7, 2019

Contents

1 Introduction

La theorie des jeux est le domaine de mathématique qui permet la modélisation des interactions stratégiques. Dans ce contexte, où plusieurs agents, dit joueurs, cherchent à maximiser leurs gains, le but est de trouver une stratégie optimale pour chaque joueur et d'envisager les eventuelles relations liant l'interêt de chauqe joueur, et les possibilités d'un equilibre.
Ce domaine, ayant l'apparence d'un thème restreint, se conforme rapidement à des problèmes d'une grande complexité.
Le but de ce projet est d'etudier les stratégies possibles pour le jeu: Dic-Battle; de comparer expérimentalement ces stratégie et à la fin, d'introduire une interface permettant de jouer contre l'ordinateur ou un humain.

2 Description du jeu

Deux joueurs s'affrontent dans un jeu de dés. Le but est d'être le premier à atteindre au moins N points.Le nombre de points marqués est 1 si l'un des dés au moins tombe sur 1, dans le cas contraire c'est la somme des dés. On dit qu'un joueur a un gain égal à 1 s'il remporte la partie, un gain égal à 0 si la partie est nulle, un gain égal à -1 s'il perd.
Dans ce projet on va étudier deux variante:
- variante séquentielle : les joueurs jouent à tour de rôle;
- variante simultanée : les joueurs jouent simultanément à chaque tour. Dans ce cas si les deux joueurs atteignent N points ou plus lors du même tour, c'est le joueur qui dépasse le plus les N points qui l'emporte; si les deux joueurs obtiennent le même score, alors ils sont ex-aequo.

3 Probabilités

Soit P(d,k) la probabilité qu'un joueur qui lance d dés obtienne k points. Donc:
- P(d,1)=1-$(\frac{5}{6})^d$: c'est la probabilité qu'au moins un dé tombe sur 1.
- P(d, k) = 0 pour $2 \leq k \leq 2d$-1 et pour $k>6d$.
- P(d, k) =$(\frac{5}{6})^d Q(d,k)$ pour $2d \leq k \leq 6d$ avec Q(d,k) est la probabilité d'obtenir k points en jetant d dés sachant qu'aucun dé n'est tombé sur 1. Q(d,k) peut se calculer à l'aide de la formule de récurrence suivante pour $d \geq 2$ et $2d \leq k \leq 6d$: $Q(d,k) = \sum_{j=2}^{6} \frac{Q(d-1,k-j)}{5}$.

Explication : Si l'on obtient k points en lançant le dieme dè alors on a deja obtenue k' dés aprés avoir lancé d-1 dés telque k' ne peut prendre que des valeurs allant de k-2 à k-6, on a donc 5 cas probables avec chaqu'un une probabilité P(d-1,k') donc chaque cas à une probabilité $(\frac{5}{6})$P(d-1,k') . La probabilité d'avoir k point en lancant d dès est donc la somme des probabilité des états qui l'induisent. Donc on a bien $Q(d,k) = \sum_{j=2}^{6} \frac{Q(d-1,k-j)}{5}$.

Cas d'initialisation : $Q(1,k) = (\frac{1}{5})$ pour k=2,...,6

4 Variante séquentielle

4.1 Stratégie aveugle

La stratégie aveugle consiste à lancer d dès telque d représente le nombre de dès qui maximise l'espérance EP(d)=$4d(\frac{5}{6})^d + 1 - (\frac{5}{6})^d$ du nombre de points obtenus.

Cette stratégie n'est pas optimale.
Posons N=100, D=10, les deux joueurs étants dans un état (i,j)=(98,94) et c'est le tour du joueur 1 de jouer. D'aprés la stratégie aveugle le joueur devrait lancer 6 dés, sa probabilité de gagner serait donc $\sum_{k=2}^{36} P(6,k)$. D'aprés la table P construite précédement (question 1) cette somme de probabilité est égale à 0.335, or si le joueur lance un seul dé il a une chance $\frac{5}{6} = 0,833$ de gagner. On en conclu que cette stratégie n'est pas optimale.

4.2 Programmation dynamique

Dans cette section, on calcule une stratégie en se basant sur la programmation dynamique. Il s'agit de construire une table EG de dimensions (N-1+6D)x(N-1+6D) resprésentant l'éspérance de gain dans un état (i,j) tel que: état(i,j) : état où le premier joueur a cumulé i points, le deuxième joueur a cumulé j points, et c'est au joueur 1 de jouer.

Formule de récurrence: Le jeu étant à somme nulle, l'espérance de gain du joueur 1 est égale à (-) l'espérance de gain de son adversaire. Étant un état (i,j), le joueur 2 calcule son espérance de gain, pour un 'd' donné que lance le joueur 1, de la manière suivante:
EG2(j,i)=$\sum_{k=1}^{6d} P(d,k)EG1(j,i+k)$.
Les deux joeurs jouent de façon optimale, la même table est utilisée, et donc le joueur 1 est en mesure de calculer EG2 et de joueur de façon à minimiser cet espérance. On aura donc :
EG1(i,j) = - min (EG2(j,i)) = - min($\sum_{k=1}^{6d} P(d,k)EG(j,i+k)$)
<u>Initialisation:</u>
Si on est dans un état où $i \geq N$ alors soit il a gagné ($i \geq j$), soit il a perdu ($i \leq j$), soit la partie est nulle (i = j).

$$EG(i,j) = \begin{cases} 1 & \text{si } i > j, i \geq N \\ -1 & \text{si } i < j, j \geq N \\ 0 & \text{si } i = j, i \geq N \end{cases}$$

La stratégie optimale: Cette approche permet de trouver la stratégie optimale dans chaque état. Il suffit, lors du calcul de l'espérance de gain, de sauvegarder le nombre de dés 'd' permettant d'obtenir ce résultat.

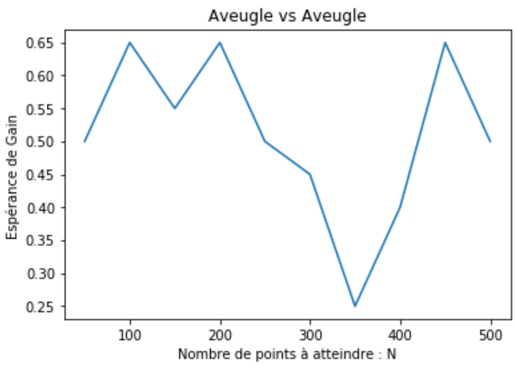
Difficulté possible: Si on supposait que le nombre de points marqués est 0 (et non plus 1) si on obtient au moins un 1, les points cumulés par les joueurs risquent à ne pas changer pendant plusieurs tours. C'est à dire, l'algorithme reste dans un état (i,j), et ça pour une durée indéterminée, voir à l'infinie. Donc l'algorithme risque à ne pas se terminer.

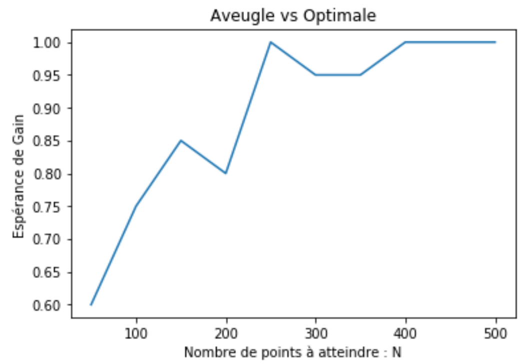
4.3 Mise en oeuvre

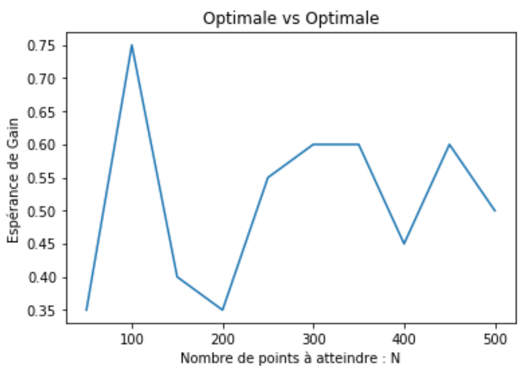
Implémentation: On implémente trois stratégies:
<u>Stratégie aveugle:</u>
Il s'agit de retourner le 'd' qui maximise la fonction de l'espérance.
<u>Stratégie optimale:</u>
En plus de la table EG, la fonction retourne une table 'strat' de même dimension qui donne le nombre de dés optimal à jouer dans chaque état.
<u>Stratégie aléatoire:</u>
On implémente de plus une stratégie aléatoire qui retoune aléatoirement un nombre entre 1 et D de dés à jouer.

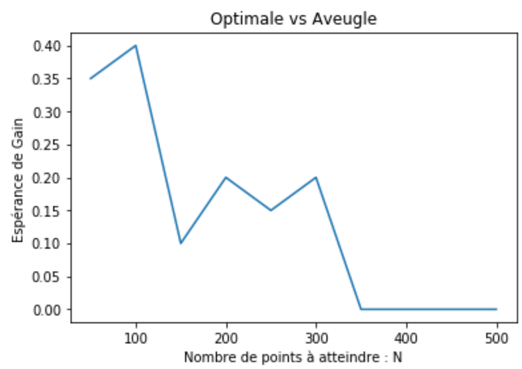
Evaluation: On simule plusieurs parties où le joueur 1 utilise une des trois stratégies contre une autre et on calcule son espérance de gain :
<u>En variant N:</u> on fixe D à une valeur passée en argument (ici D=10),et pour chaque valeur

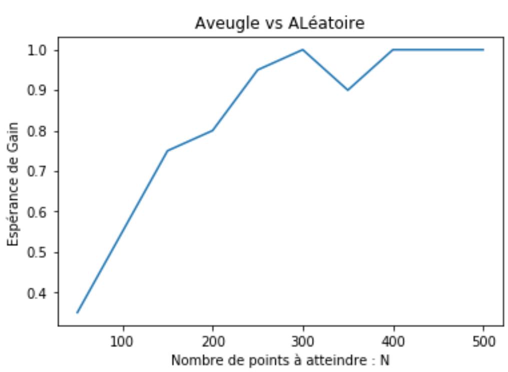
de N, on et on compte le gain du joueur 1.

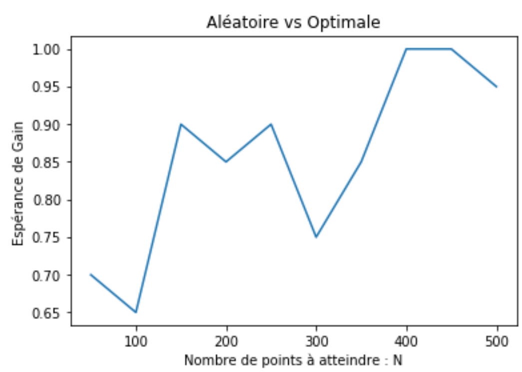






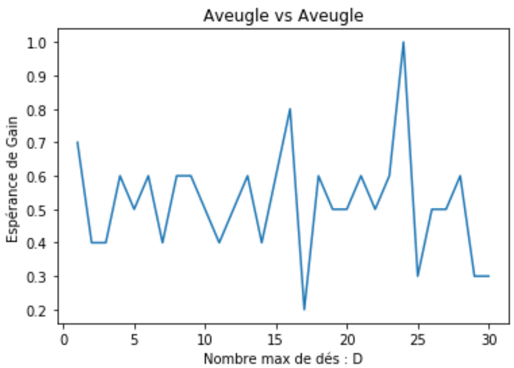


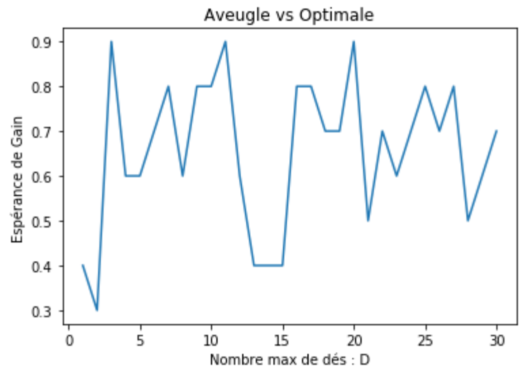


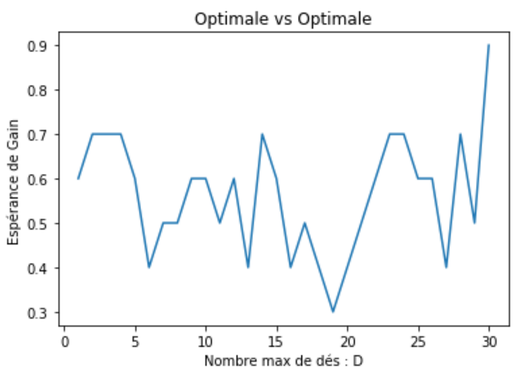


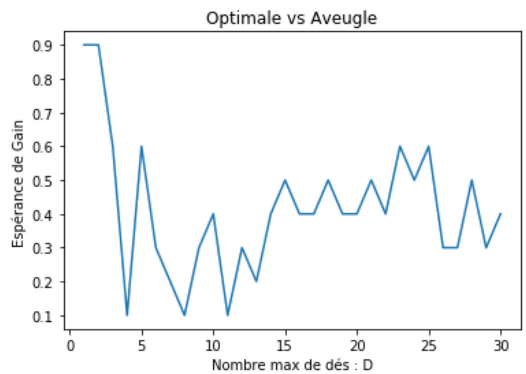
On générale, on constate que la stratégie aveugle donne des bons résultats contre les autres stratégies avec un espérance de gain plus en plus sûr en augmentant N.

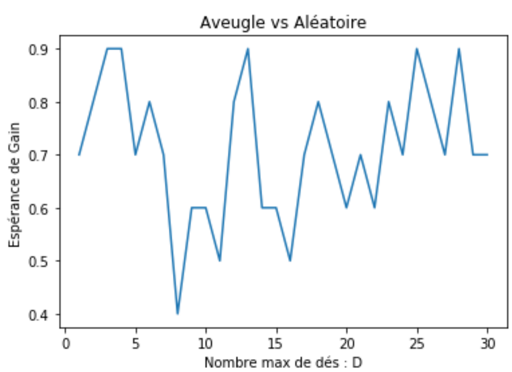
En variant D: on fixe N à une valeur passée en argument (ici N=50),et pour chaque valeur de D, et on compte le gain du joueur 1.

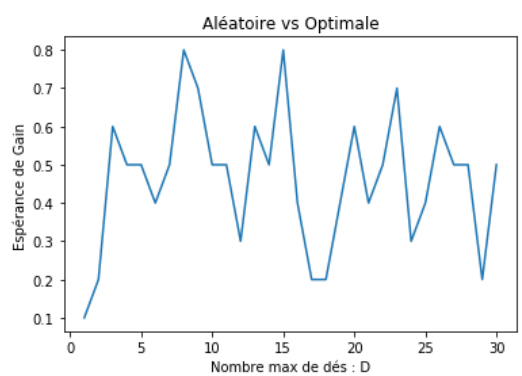






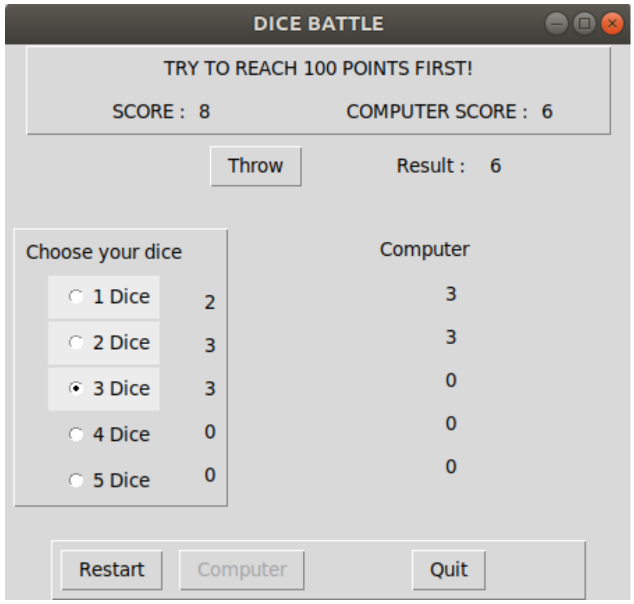






Pour les différentes stratégies, la variation de D entre une variation non-monotone de l'espérance de gain.

Interface de jeu: L'interface ci-desous permet de jouer contre l'ordinateur pour N=100 et D=5.



5 Variante simultanée

Si les deux joueurs jouent à chaque tour de façon simultanée, alors la stratégie optimale dans un état(i,j) donné est randomisée. Ainsi, nous considérons des stratégies randomisées, consistant à donner un vecteur p= [p(1), p(2), . . . , p(D)] où p(d) est la probabilité avec laquelle nous lançons d dés.

5.1 Jeu en un coup

On considère en première approche une version simplifiée où l'on ne lance qu'une fois les dés. Le gagnant est simplement celui qui remporte le plus de points.

Espérance de gain On note EG1(d1,d2) l'espérance de gain du joueur 1 s'il lance d1 dés alors que le joueur 2 en lance d2.
On a donc : EG1(d1,d2)= $\sum_{k1=1}^{6d1} kP(d1,k1) - \sum_{k2=1}^{6d2} kP(d2,k2)$, représentant le nombre de points gagnés (k1) fois la probabilité de gagner k1 points en lançant d1 dés pour le joueur 1 moins la perte (k2) fois la probabilité de perte qui n'est autre que la probabilité de gain du joueur 2 en lançant d2 dès puisque le jeu est à somme nulle.
Pour D=3, on a la matrice de gain suivante :

EG1(d1,d2)	1	2	3
1	0	-2.36111111	-3.86574074
2	2.36111111	0	-1.50462963
3	3.86574074	1.50462963	0

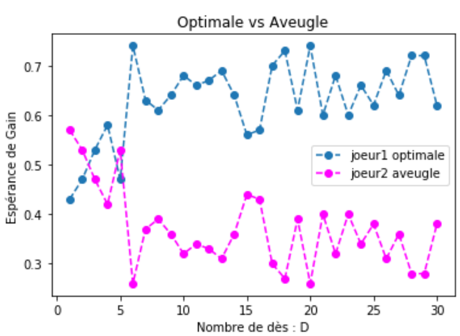
Stratégie optimale du joueur 2 Si le joueur 2 connaît la stratégie du joueur 1 : le vecteur de probabilités [p1(1), p1(2), . . . , p1(D)], sa stratégie optimale consisterait donc à minimiser les gains de son adversaire. $\min_{dj}(\sum_{i=1}^{D} p1(di)EG1(di,dj))$, i,j=1,...,D.

Maximiser l'éspérance de gain du joueur 1 Le joueur 1 cherche à maximiser son gain, sachant que le joeur 2 répond lui aussi de façon optimale, il chercherait donc à maximiser le sien en minimisant le gain du premier joueur comme vu precédement $\min_{dj}(\sum_{i=1}^{D} p1(di)EG1(di,dj))$ et donc le joueur 1 aurait intérêt à maximiser : $\max_{di}(\min_{dj}(\sum_{i=1}^{D} p1(di)EG1(di,dj)))$, i,j=1,...,D.

Programme Lineaire

$$\begin{array}{ll} \text{Max} & Z \\ \text{contraintes} & Z \leq \sum_{i=1}^{D} p1(di)EG1(di,dj), j = 1,...,D \\ & pi \geq 0 \qquad\qquad\qquad\qquad , i = 1,...,D \end{array}$$

Comparaison entre la stratégie Optimale et la stratégie deterministe Aveugle Pour D=1,...,30



On voit bien que l'espérance de gain pour la stratégie optimale varie entre 0,75 et 0,6 contre 0,3 à 0,4 pour la stratégie aveugle.

5.2 Cas général

Dans cette partie on va étendre notre stratégie randomisée au cas général de notre jeu où il s'agit de dépasser N points.

Esperance de gain du joeur 1 L'espérance de gain EG1(i,j) du joueur 1 lorsque $i \geq N$ ou $j \geq N$ est calculé comme suit : $EG1(i,j) = \frac{i-j}{|i-j|}$ si $i \neq j$, 0 sinon.

Espérance de gain $E^{d1,d2}$(i,j) du joueur 1 Soit EG1(k,l) l'espérance de gain du joueur 1 lorsque $k{>}i$ et $l{>}j$ (lorsque les joueurs jouent de manière optimale).
Supposons que le joueur 1 lance d1 dés et que le joueur 2 lance d2 dés. L'espérance de gain du joeur 1 est donnée comme suit :
$E^{d1,d2}$(i,j)=$\sum_{k1=1}^{6d1}\sum_{k2=1}^{6d2} P(d1,k1)P(d2,k2)EG(i+k1,j+k2)$.

Espérance de gain EG(i,j) du joueur 1 Par la suite le calcule de EG se fait en appliquant la résolution linéaire à $E^{d1,d2}$, pour chaque case EG(i,j).
La stratégie optimale retournée pour l'état (i,j) étant le vecteur p retourné par la résolution linéaire.

6 Conclusion

On a pu étudier tout au long du projet trois stratégies pour le jeu Dice Battle. En sequentiel on a vu une premiere strategie Aveugle qui n'est pas optimale néaumoins facile à implémenter, la stratégie dynamique quant à elle fournit de meilleurs résultats mais nécéssite le cacul des probabilités de gains à partir de chaque état possible. En simultané, notre stratégie randomisée fournie de bons résultats contre une stratégie aléatoire ou aveugle, mais pour un adversaire utilisant la même approche on obtient un gain nul puisque le jeu est à somme nulle.