Exercices sur les algorithmes probabilités
Nom du rédacteur : Johanne Cohen
`https://www.lri.fr/~jcohen/pages/enseignement.html`

# Jeu du gobelet

Le jeu consiste à trouver une balle sous $N$ gobelets (notons $[N]$ : ensemble des gobelets). Deux joueurs participent à ce jeu : un joueur que l'on notera par $J$ et un adversaire. A chaque partie $t$ :

1. L'adversaire va mettre une balle dans un gobelet, ou une balle dans plusieurs gobelets. A cet instant de la partie, le joueur $J$ ne sait pas où se trouve la ou les balles.
2. Le joueur $J$ choisit un gobelet $x^t$ ***en fonction d'une politique*** $\mathcal{A}$.
3. L'adversaire montre le contenu de tous les gobelets.
4. Si le joueur $J$ a choisi le gobelet où se trouve une balle, alors il gagne en recevant 1 euro sinon il perd (en recevant 0 euro)

Le joueur $J$ va faire $T$ parties. Son objectif est bien sur de gagner le plus de parties. Par la suite de l'exercice, nous introduisons la notion de ***perte*** (ou *loss*) pour chacun des gobelets $i$ à chaque partie $t$ :

$$\ell_i^t = \begin{cases} 0 & si \quad \text{une balle est sous le gobelet } i \text{ à la partie } t \\ 1 & sinon \end{cases}$$

Nous pouvons étendre cette définition de perte sur l'ensemble de $t$ parties.

1. La ***perte du gobelet*** $i$ correspond au nombre de parties où le gobelet $i$ n'a pas contenu de balle (pendant les $t$ premières parties).

$$L_i^t = \sum_{k=1}^{t} \ell_i^k$$

2. La ***perte du joueur*** $J$ ***utilisant la politique*** $\mathcal{A}$ correspond au nombre de parties où le gobelet $i$ n'a pas contenu de balle (pendant les $t$ premières parties).

$$L_{\mathcal{A}}^t = \sum_{k=1}^{t} \ell_{x^k}^k$$

### Premier algorithme : Algorithme glouton

Voici l'algorithme que le joueur $J$ va utiliser.

1. A la partie 1 : le joueur $J$ choisit le gobelet 1 ($=x^1$).
2. A chaque partie $t = 2, \cdots, T$
   (a) $L_{min}^{t-1} := \min_{i\in[N]} L_i^{t-1}$ ;

(b) $S^{t-1} := \{i : L_i^{t-1} = L_{min}^{t-1}\}$ ;

(c) $x^t := \min S^{t-1}$ ;

(d) Le joueur $J$ choisit le gobelet $x^t$.

**Question 1**. Quel est l'algorithme de placement des balles du pire adversaire s'il sait que le joueur $J$ fait ses choix selon cette stratégie ?

1. A la partie 1 mettre une balle dans le gobelet 2
2. A chaque partie $t$
   (a) Si $S^{t-1} = [N]$, alors mettre une balle dans le gobelet 2
   (b) Sinon, choisir un gobelet qui n'est pas dans $S^{t-1}$

**Question 2**. Combien de parties le joueur $J$ gagne-t-il dans ce cas ? Quelle est la perte du joueur $J$ ($L_{glouton}^T$) ?

Aucune, $L_{glouton}^T = T$

Dans ce contexte, nous n'allons pas mesurer la performance de l'algorithme qu'utilise le joueur $J$ par rapport au nombre de parties gagnées. Nous allons la mesurer par rapport à la notion de ***regret***[1] (problème du bandit manchot).

Pour cela, nous allons comparer la perte totale $L_{glouton}^T$ du joueur $J$ par rapport à la perte totale $\min_{i \in [N]}\{L_i^T\}$ du "meilleur" gobelet que l'on notera $L_{min}^T$.

**Question 3**. Que peut-on dire si à l'étape $t$, le joueur $J$ perd et que $L_{min}^t$ n'augmente pas ?

Au moins un gobelet est enlevé de $S_t$ car le joueur $J$ utilise un gobelet qui était dans $S_t$. Comme il a perdu, la perte totale de ce gobelet aura augmentée $x^t \notin S_{t+1}$.

**Question 4**. Montrer que $L_{glouton}^T \leq N \cdot L_{min}^T + (N-1)$

Le cas précédent ne peut se réaliser au plus $N$ fois avant que $L_{min}^t$ augmente de 1. $L_{glouton}^t \leq N \cdot L_{min}^T + (N - |S^t|)$

### Mauvaise performance pour une politique déterministe

Soit $\mathcal{D}$ un algorithme déterministe.

**Question 5**. Supposons que l'adversaire connaît l'algorithme $\mathcal{D}$ joué par le joueur $J$. Montrer que l'adversaire peut construire un algorithme de placement des balles de l'adversaire telle que $L_{\mathcal{D}}^T = T$ et $L_{min}^T \leq \lfloor \frac{T}{N} \rfloor$.

L'adversaire génère la suite de la façon suivante

1. A chaque partie $t$ :

1. Le regret d'un algorithme $A$ est $regret_A = L_A^T - \min_{i \in [N]}\{L_i^T\}$

(a) mettre la perte $\ell^t_{x_t} = 1$

(b) et les autres à 0

Cela garantie que $L^t_{\mathcal{D}} = t$ (puisque le joueur perd tout le temps).

Comme il y a $N$ gobelets, il existe des gobelets que l'algorithme joue au plus $\frac{T}{N}$ fois. La perte d'un gobelet correspond au nombre de fois que l'algorithme le choisit.

**Deuxième algorithme : Algorithme glouton probabiliste (GP)**

1. A la partie 1 :
   (a) $p_i^1 = \frac{1}{|N|} \forall i \in [N]$
   (b) le joueur $J$ choisit le gobelet $x^1$ en fonction de la loi de la probabilité $p^1$.
2. A chaque partie $t = 2, \cdots, T$
   (a) $L_{min}^{t-1} := \min_{i \in [N]} L_i^{t-1}$ ;
   (b) $S^{t-1} := \{i : L_i^{t-1} = L_{min}^{t-1}\}$ ;
   (c) $p_i^t = \frac{1}{|S^{t-1}|} \forall i \in S^{t-1}$ (sinon $p_i = 0$ $\forall i \notin S^{t-1}$)
   (d) Le joueur $J$ choisit le gobelet $x^t$ en fonction de la loi de la probabilité $p^t$.

**Question 6**. Soit $t < T$ un entier tel que $L_{min}^{t+1} = L_{min}^t$.

1. Montrer que $S^{t+1} \subseteq S^t$.
2. Montrer que la perte moyenne de l'algorithme vaut : $\frac{k}{|S^t|}$, c'est-à-dire

$$E[\ell_{x_t}^t] = \tfrac{k}{|S^t|} \text{ avec } k = |S^t| - |S^{t+1}|$$

Les éléments de $S^{t+1}$ sont ceux de $S_t$ qui ont reçu une balle : leur perte n'augmente pas.
L'algorithme choisit de fac con uniforme le gobelet sur $S_t$. Donc $E[\ell_{x_t}^t] = 1 * \frac{1}{|S^t|} * k + 0$.

**Question 7**. Notons $t_j$ l'étape où $L_{min}^t$ vaut $j$ pour la première fois.

1. Quelle est la valeur de $t_0$ ?
2. Montrer que la perte moyenne de l'algorithme entre l'étape $t_j$ et $t_{j+1} - 1$ est majorée par : $1 + \ln N$, c'est-à-dire

$$\sum_{z=t_j}^{t_{j+1}-1} E[\ell_{x_t}^t] \leq \sum_{z=1}^{N} \frac{1}{z} \leq 1 + \ln N$$

$t_0 = 0$
Remarquons que $\frac{k}{|S^t|} \leq \frac{1}{|S^t|} + \frac{1}{|S^t|-1} + \cdots + \frac{1}{|S^t|-k+1}$

$$\sum_{z=t_j}^{t_{j+1}-1} E[\ell_{x_t}^t] = \sum_{z=t_j}^{t_{j+1}-1} \frac{|S^t| - |S^{t+1}|}{|S^t|}$$

$$\sum_{z=t_j}^{t_{j+1}-1} E[\ell_{x_t}^t] \leq \sum_{z=t_j}^{t_{j+1}-1} \frac{1}{|S^t|} + \frac{1}{|S^t|-1} + \cdots + \frac{1}{|S^t| - |S^t| + |S^{t+1}| + 1}$$

$$\sum_{z=t_j}^{t_{j+1}-1} E[\ell_{x_t}^t] \leq \sum_{z=1}^{N} \frac{1}{z} \leq 1 + \ln N$$

**Question 8**. Montrer que $L_{GR}^T \leq (1+\ln N)\cdot L_{min}^T + \ln N$ avec $L_{GR}^T$ correspondant à la perte totale du joueur $J$ si il choisit son gobelet en fonction de l'algorithme glouton probabiliste.

$L_{GR}^t \leq 1/N + 1/(N-1) + \cdots + 1/(|S_t|+1) + (1+\ln N)L_{min}^t$