

# Étude des algorithmes UCB et $\epsilon$ -Glouton

Enseignant : Liva Ralaivola

Date : 13 Décembre 2017 à 23 :59

## 1 Modalités pratiques

- À rendre, dans une archive (format zip) :
  - un document (format pdf) de 6 pages maximum (contenu ci-dessous)
  - le code Python qui a permis de conduire toutes les expériences
- Date de rendu : 17 janvier 2017
- Moyen : par mail à {prenom.nom}@lif.univ-mrs.fr

## 2 Étude à mener

Les études à conduire visent à mesurer l'efficacité et la robustesse des algorithmes UCB et  $\epsilon$ -glouton. Les notations utilisées ici sont les mêmes que dans le document décrivant ces algorithmes.

### 2.1 Regret

Une première partie de l'étude consiste à mesurer l'évolution du regret au cours du temps — des définitions sont données dans le document décrivant les algorithmes UCB et  $\epsilon$ -glouton.

Si  $k^*$  est l'indice du bras optimal, c'est-à-dire :

$$k^* = \operatorname{argmax}_{k \in \{1, \dots, K\}} \mu_k,$$

le regret dépend non seulement de  $T$  (le nombre d'étapes de l'algorithme), de  $K$  (le nombre de bras), du nombre de fois que chaque bras a été joué et des récompenses collectées mais aussi de la quantité  $\Delta$  suivante :

$$\Delta = \mu_{k^*} - \max_{i \neq k^*} \mu_i,$$

qui est l'écart entre la valeur du bras optimal et la seconde valeur (on considère qu'il n'y a qu'un bras optimal et donc  $\Delta$  est une quantité strictement positive).

- Proposez une définition de regret.
- Proposez un protocole pour mesurer l'évolution du regret en fonction de  $\Delta$  (notamment).
- Donnez les résultats (tableaux, graphiques, etc.) des expériences menées, commentez-les. Mesurez notamment l'influence de  $\alpha$  (UCB) et  $\epsilon$  ( $\epsilon$ -glouton) sur le fonctionnement des algorithmes.

### 2.2 Robustesse

Une hypothèse classique des algorithmes de bandit est que les distributions associées aux bras sont fixes et donc que la récompense moyenne de chaque bras n'évolue pas au cours du temps.

La deuxième partie de l'étude que vous mènerez visera à analyser la robustesse de UCB et  $\epsilon$ -glouton à des changements des distributions des bras.

- Proposez et motivez un protocole d'analyse de la robustesse des algorithmes. Commentez les résultats observés (en particulier, vous regarderez la robustesse des algorithmes en fonction des paramètres d'exploration,  $\alpha$  ou  $\epsilon$ ).
- Proposez un algorithme de votre invention qui modifierait UCB ou  $\epsilon$ -glouton pour les rendre robustes à des changements de distribution. Illustrez cette robustesse par des graphiques et des tableaux.