

Projet SINF2275

« Data mining and decision making »

Exercice reinforcement learning

Année académique 2006-2007

Professeurs : Marco Saerens
Adresse : Université catholique de Louvain
Information Systems Research Unit (ISYS)
Institut d'Administration et de Gestion
Place des Doyens 1
B-1348 Louvain-la-Neuve
Belgique
Téléphone : 010 47.92.46.
Fax : 010 47.83.24.
Courriel : marco.saerens@ucLouvain.be

Objectif

L'objectif de ce travail est la mise en pratique concrète d'un certain nombre de techniques d'analyse de données quantitatives (et de prise de décision), à travers l'étude d'un cas pratique nécessitant l'utilisation de logiciels de traitement statistique de données (R, S-Plus, Matlab). L'application envisagée ici est de type « **renforcement learning** » ou de « **processus de décision de Markov** ».

Contexte

Ce projet a pour but l'étude des techniques de « **renforcement learning** » ou de « **processus de décision de Markov** ». L'objectif est de modéliser un problème de décision et de trouver sa solution à travers l'implémentation d'un algorithme. Les algorithmes qui devront être implémentés ont été présentés au cours et font l'objet de plusieurs articles et chapitres de livres qui vous ont été transmis.

Enoncé du problème

La réalisation du projet se fera par groupes de trois au maximum. Il vous est demandé de déterminer la solution du problème suivant :

Supposons que vous jouez à un jeu simple qui possède 7 cases. La case 1 est la case initiale (départ) et la case 7 est la case gagnante (arrivée). Si vous parvenez à la case 7, vous avez gagné.

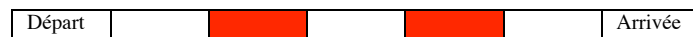
Pour avancer, vous disposez de deux dés :

- Le dé « risqué » dont les valeurs vont de 0 à 2, et qui vous permet d'avancer de ce nombre de cases, avec une probabilité 1/3.
- Le dé « sécurité » qui ne possède que deux valeurs, 0 ou 1. Il permet d'avancer de 0 ou 1 case, avec une probabilité 1/2.

Il y a également deux cases « piège » : la case 3 et la case 5. Si vous jouez avec le dé « risqué » et que vous tombez sur une de ces deux cases, vous devez retourner à la case départ (case 1). Par contre, si vous utilisez le dé « sécurité », vous ne rencontrez aucun problème : ce dé vous permet d'être immunisé contre le retour à la case départ.

Nous vous demandons de déterminer la stratégie optimale, à savoir, pour toute case, lequel des deux dés vous devez jouer pour arriver à la case 7 en jouant le moins de coups possible, en moyenne. Vous pouvez déterminer cette solution pour deux cas de figure :

- Il faut tomber **exactement** sur la case arrivée pour avoir gagné. Le plateau de jeu est un cercle, ce qui veut dire que lorsque vous dépassez la case arrivée, vous poursuivez à partir de la case départ.
- Vous avez gagné dès le moment où vous dépassez la case arrivée.



Pensez à écrire un code où on peut facilement modifier la valeur de :

- Le nombre de case total
- La position et le nombre des cases piège
- Le nombre de face des dés

afin de pouvoir les modifier et observer les répercutions que ça aura sur la solution finale.

Outils logiciels

Pour le projet le choix du logiciel est entièrement libre. Vous pouvez aussi bien le réaliser sous R, Matlab ou simplement le programmer en java.

Rapport

La solution de cet exercice se trouvera dans le rapport final.

Organisation du travail en groupe

Les étudiants travailleront par groupes de trois personnes au maximum.

Evaluation du projet

A travers le rapport.
