

Sur l'utilisation de politiques non-stationnaires pour les processus de décision Markoviens à horizon infini [★]

Bruno Scherrer¹, Boris Lesner¹

Inria, Villers-lès-Nancy, F-54600, France

`bruno.scherrer@inria.fr, boris.lesner@inria.fr`

Résumé : Nous considérons les processus de décision Markoviens stationnaires et à horizon infini, pour lesquels on sait qu'il existe une politique optimale qui est stationnaire. En utilisant les algorithmes *itérations sur les valeurs* et *itérations sur les politiques* avec une erreur ϵ à chaque itération, il est connu que l'on peut calculer une politique stationnaire qui est $\frac{2\gamma}{(1-\gamma)^2} \epsilon$ -optimale. Après avoir montré que cette garantie est fine, nous développons des variations d'*itérations sur les valeurs* et *itérations sur les politiques* qui calculent des politiques non-stationnaires qui peuvent être $\frac{2\gamma}{1-\gamma} \epsilon$ -optimales, ce qui représente une amélioration significative dans la situation usuelle où γ est proche de 1. De manière surprenante, nous montrons ainsi que le problème de "calculer une politique non-stationnaire approximativement optimale" est beaucoup plus simple que celui de "calculer une politique stationnaire approximativement optimale".

★. L'article a été publié, en langue anglaise, à la conférence NIPS 2012.