



HAL
open science

Model-Free Monte Carlo-like Policy Evaluation

Raphael Fonteneau, Susan Murphy, Louis Wehenkel, Damien Ernst

► **To cite this version:**

Raphael Fonteneau, Susan Murphy, Louis Wehenkel, Damien Ernst. Model-Free Monte Carlo-like Policy Evaluation. Journées MAS et Journée en l'honneur de Jacques Neveu, Aug 2010, Talence, France. inria-00510325

HAL Id: inria-00510325

<https://hal.inria.fr/inria-00510325>

Submitted on 18 Aug 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Journées MAS 2010, Bordeaux

Session : Apprentissage par renforcement

Model-Free Monte Carlo-like Policy Evaluation

par Raphael Fonteneau, Susan A. Murphy, Louis Wehenkel et **Damien Ernst**

We propose an algorithm for estimating the finite-horizon expected return of a closed loop control policy from an a priori given (off-policy) sample of one-step transitions. It averages cumulated rewards along a set of “broken trajectories” made of one-step transitions selected from the sample on the basis of the control policy. Under some Lipschitz continuity assumptions on the system dynamics, reward function and control policy, we provide bounds on the bias and variance of the estimator that depend only on the Lipschitz constants, on the number of broken trajectories used in the estimator, and on the sparsity of the sample of one-step transitions.

Adresses :

Raphael FONTENEAU

Université de Liège

Institut Montefiore, Sart-Tilman B28

B-4000 Liège BELGIQUE

E-mail : raphael.fonteneau@ulg.ac.be

<<http://sites.google.com/site/raphaelfonteneau/home>>

Susan A. MURPHY

University of Michigan

439 West Hall, 1085 S. Univ. University of Michigan Ann Arbor, MI 48109-1107

E-mail : samurphy@umich.edu

<<http://www.stat.lsa.umich.edu/~samurphy/>>

Louis WEHENKEL

Université de Liège

Institut Montefiore, Sart-Tilman B28

B-4000 Liège BELGIQUE

E-mail : L.Wehenkel@ulg.ac.be

<<http://www.montefiore.ulg.ac.be/~lwh/>>

Damien ERNST

Université de Liège

Institut Montefiore, Sart-Tilman B28

B-4000 Liège BELGIQUE

E-mail : dernst@ulg.ac.be

<<http://www.montefiore.ulg.ac.be/~ernst/>>

Session : Apprentissage par renforcement