



# OpenDP a free Reinforcement Learning toolbox for discrete time control problems

Sylvain Gelly, Olivier Teytaud

► **To cite this version:**

Sylvain Gelly, Olivier Teytaud. OpenDP a free Reinforcement Learning toolbox for discrete time control problems. NIPS Workshop on Machine Learning Open Source Software, Dec 2006, Whistler (B.C.), 2006. <inria-00117392>

**HAL Id: inria-00117392**

**<https://hal.inria.fr/inria-00117392>**

Submitted on 1 Dec 2006

**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.

# OpenDP a free Reinforcement Learning toolbox for discrete time control problems

OpenDP (<http://opendp.sourceforge.net>) is a free software (under GPL) written in C++ for control problems: (i) with discrete time steps (either finite horizon, or approximation of infinite horizon by temporal-ring for stationary problems), (ii) where the transition function  $newState = T(state, decision)$  can be encoded in C++, (iii) possibly stochastic (with the function  $T$  above depending on some random process).

OpenDP is designed to be a general framework for all control problems of the previous form. One can easily add new solving methods to the framework and can compare the results to already existing methods. OpenDP includes some benchmark problems, with parameterizable dimension and difficulty. Some benchmarks are stochastic and some are deterministic. Other benchmark problems can easily be added.

OpenDP includes policy search methods in their simplest way, but its core solving method is stochastic dynamic programming. In this method, the Bellman value is computed backward in time by iterating: (i) sampling points in the state domain, (ii) for each point, compute the best expected value given the Bellman function, (iii) the examples and values given (i) and (ii) to learn a regression which gives the next Bellman function.

For each of the three steps considered above, a toolbox is provided including classical methods. For the sampling step, we provide classes generating some classical quasi random sequences (low discrepancy sequences with or without scrambling, low dispersion, ...). For the continuous-optimisation steps, OpenDP is interfaced with famous optimisation frameworks as Opt++, Coin DFO, EO, Beagle, which include for example quasi-Newton algorithms, Hooke&Jeeves, CMAES, SAES and some other algorithms. For the learning step, OpenDP is interfaced with the Weka library and so we include all the regressions implemented in Weka. OpenDP is also interfaced with the Torch library (including SVM and Neural Networks) and the FANN library. The algorithms included in the interfaced toolboxes are available as plugins and so OpenDP can be used even if the libraries are not present.

The download area of OpenDP can be found at <http://opendp.sourceforge.net/>. OpenDP is in beta stage of developpement. Some typical command-lines provide extensive experiments (see the webpage above), providing a simple way of testing modifications of the source code. OpenDP also comes with a graphical user interface written in Qt, with curve visualizations in 2D, and Bellman function visualization in 3D.

The OpenDP source code is cross-platform. OpenDP actually works under Linux and MS Windows and binary-packages can be downloaded for theses platforms. The building process and some mecanisms have to be ported to Mac OS X.

## Acknowledgment

Authors are Sylvain Gelly<sup>1</sup>, Olivier Teytaud<sup>1</sup>, Jérémie Mary<sup>2</sup>. This work was supported in part by the Pascal Network of Excellence.

---

<sup>1</sup>TAO (INRIA), LRI, UMR CNRS, Univ. Paris-Sud; email [firstname.lastname@lri.fr](mailto:firstname.lastname@lri.fr).

<sup>2</sup>Grappa, Univ. Lille UMR CNRS, [Jeremie.Mary@grappa.univ-lille3.fr](mailto:Jeremie.Mary@grappa.univ-lille3.fr)