IN PARTNERSHIP WITH:
**CNRS**

**Université Paris-Sud (Paris 11)**

# Activity Report 2015

# Project-Team TAO

## Machine Learning and Optimisation

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)

RESEARCH CENTER
**Saclay - Île-de-France**

THEME
**Optimization, machine learning and statistical methods**

# Table of contents

# Project-Team TAO

**Keywords:**

**Computer Science and Digital Science:** information Coming Later Format Csv

**Other Research Topics and Application Domains:** information Coming Later Format Csv

*Creation of the Project-Team: 2004 November 04*

# 1. Members

**Research Scientists**

Marc Schoenauer [Team leader, Inria, Senior Researcher, HdR]
Anne Auger [Inria, Researcher]
Guillaume Charpiat [Inria, Researcher]
Cyril Furtlehner [Inria, Researcher]
Gregory Grefenstette [Inria, Advanced Research position, HdR]
Nikolaus Hansen [Inria, Senior Researcher, HdR]
Odalric Maillard [Inria, Researcher]
Yann Ollivier [CNRS, Researcher, HdR]
Michèle Sebag [CNRS, Senior Researcher, HdR]
Olivier Teytaud [Inria, Researcher, HdR]
Balász Kégl [CNRS, Senior Researcher (Associate Member), HdR]

**Faculty Members**

Philippe Caillou [Univ. Paris XI, Associate Professor]
Aurélien Decelle [Univ. Paris XI, Associate Professor]
Cécile Germain [Univ. Paris XI, Professor, HdR]
Isabelle Guyon [Univ. Paris XI, Professor, HdR]
Nicolas Spyratos [Univ. Paris XI, Professor Emeritus (Associate Member), HdR]
Véronique Ventos [Univ. Paris XI, Associate Professor (Associate Member)]

**Engineers**

Emmanuel Benazera [Inria, until Mar 2015]
Mohamed Bouatira [Inria, from Feb 2015]
Jean-Baptiste Hoock [Inria, from Oct 2015]
Lovro Ilijasic [Inria, until Sep 2015, granted by Conseil Régional d'Ile-de-France]
Luis Marti Orosa [Inria, from Nov 2015]

**PhD Students**

Chaouki Boufenar [UNIV étrangère]
Ouassim Ait Elhara [Univ. Paris XI, STIC]
Sandra Cecilia Astete Morales [Inria]
Asma Atamna [Inria]
Nacim Belkhir [AUT ENTREPRISE PRIV, granted by Thalès (CIFRE)]
Jeremy Bensadon [Univ. Paris XI, Maths]
Vincent Berthier [Univ. Paris XI]
Marie Liesse Cauwet [Inria]
Alexandre Chotard [Inria, until Nov 2015]
Jeremie Decock [Inria, until Feb 2015, granted by ADEME]
Benjamin Donnot [AUT ENTREPRISE PRIV, from Oct 2015]
Guillaume Doquet [Univ. Paris XI, STIC,from Sept. 2015]

Nicolas Galichet [Inria, until Sept. 2015]
François Gonard [Inst. de Recherche Technologique SystemX]
Jialin Liu [Inria]
Sourava Prasad Mishra [AUT ENTREPRISE PUBL]
Karima Rafes [AUT ENTREPRISE PRIV]
Thomas Schmitt [Inria]
Guohua Zhang [UNIV étrangère, until Aug 2015]

**Post-Doctoral Fellows**

Nicolas Galichet [Inria, from Oct. 2015]
Edgar Galvan Lopez [UNIV étrangère, from April 2015]
Lawrence Githiari [Inria, from Jul 2015]
Gaetan Marceau Caron [Inria, granted by Conseil Régional d'Ile-de-France]
Duc Manh Nguyen [Inria, from Nov 2015]
Phillipe Rodrigues Sampaio [Inria, from Sep 2015]
Lawrence Muchemi [Inria, from July 2015]
Marco Bressan [AUT ENTREPRISE PRIV, until Jun 2015]

**Visiting Scientists**

Yohei Akimoto [UNIV étrangère, until Sep 2015]
Aditya Gopalan [UNIV étrangère, from Apr 2015 until May 2015]
Dimitri Klimenko [UNIV étrangère, from Nov 2015]
Luigi Malago [UNIV étrangère, Assistant Professor, until Sep 2015]

**Administrative Assistant**

Olga Mwana Mobulakani [Inria]

**Others**

Arturo Castellanos [Inria, Internship, from Jul 2015 until Sep 2015]
Victor Estrade [Univ. Paris XI, Internship, until Apr 2015]
Robyn Ffrancon [Inria, Internship, until Jul 2015]
Ching-Nung Lin [Inria, Internship, from May 2015 until Oct 2015]
Corentin Tallec [Ecole Polytechnique, Internship, from Mar 2015 until Jul 2015]
Wen Wei [SUPELEC, Internship, from Apr 2015 until Sep 2015]
Bouatira Mohamed [Inria, Young Research Engineer, from February 2015 until January 2017]

# 2. Overall Objectives

## 2.1. Presentation

Data Mining (DM), acknowledged to be one of the main ten challenges of the 21st century [1], aims at building (partial) phenomenological models from the massive amounts of data produced in scientific labs, industrial plants, banks, hospitals or supermarkets. Machine Learning (ML) likewise aims at modeling the complex systems underlying the available data; the main difference between DM and ML disciplines is the emphasis put on the acquisition, storage and management of large-scale data.

DM and ML problems can be set as optimization problems, thus leading to two possible approaches. Note that this alternative has been characterized by H. Simon (1982) as follows. *In complex real-world situations, optimization becomes approximate optimization since the description of the real-world is radically simplified until reduced to a degree of complication that the decision maker can handle. Satisficing seeks simplification in a somewhat different direction, retaining more of the detail of the real-world situation, but settling for a satisfactory, rather than approximate-best, decision.*

[1] MIT Technological Review, feb. 2001.

The first approach is to simplify the learning problem to make it tractable by standard statistical or optimization methods. The alternative approach is to preserve as much as possible the genuine complexity of the goals (yielding “interesting” models, accounting for prior knowledge): more flexible optimization approaches are therefore required, such as those offered by Evolutionary Computation.

Symmetrically, optimization techniques are increasingly used in all scientific and technological fields, from optimum design to risk assessment. Evolutionary Computation (EC) techniques, mimicking the Darwinian paradigm of natural evolution, are stochastic population-based dynamical systems that are now widely known for their robustness and flexibility, handling complex search spaces (e.g., mixed, structured, constrained representations) and non-standard optimization goals (e.g., multi-modal, multi-objective, context-sensitive), beyond the reach of standard optimization methods.

The price to pay for such properties of robustness and flexibility is twofold. On one hand, EC is tuned, mostly by trials and errors, using quite a few parameters. On the other hand, EC generates massive amounts of intermediate solutions. It is suggested that the principled exploitation of preliminary runs and intermediate solutions, through Machine Learning and Data Mining techniques, can offer sound ways of adjusting the parameters and finding shortcuts in the trajectories in the search space of the dynamical system.

## 2.2. Context and overall goal of the project

The overall goals of the project are to model, predict, understand, and control physical or artificial systems. The central claim is that Learning and Optimization approaches must be used, adapted and integrated in a seamless framework, in order to bridge the gap between the system under study on the one hand, and the expert’s goal as to the ideal state/functionality of the system on the other hand.

Specifically, our research context is based on the following assumptions:

1. The systems under study range from large-scale engineering systems to physical or chemical phenomenons, including robotics and games. Such systems, sometimes referred to as *complex systems*, can hardly be modeled based on first principles due to their size, their heterogeneity and the incomplete information aspects involved in their behavior.
2. Such systems can be observed; indeed selecting the relevant observations and providing a reasonably appropriate description thereof are part of the problem to be solved. A further assumption is that these observations are sufficient to build a reasonably accurate model of the system under study.
3. The available expertise is sufficient to assess the system state, and any modification thereof, with respect to the desired states/functionalities. The assessment function is usually not a well-behaved function (differentiable, convex, defined on a continuous domain, etc.), barring the use of standard optimization approaches and making Evolutionary Computation a better suited alternative.

In this context, the objectives of TAO are threefold:

1. Investigating how specific prior knowledge and requirements can be accommodated in Machine Learning thanks to evolutionary computation (EC) and more generally Stochastic Optimization;
2. Investigating how statistical Machine Learning can be used to interpret, study and enhance evolutionary computation;
3. Facing diversified and real-world applications, requiring and suggesting new integrated ML/EC approaches.

# 3. Research Program

## 3.1. The Four Pillars of TAO

This Section describes TAO main research directions at the crossroad of Machine Learning and Evolutionary Computation. Since 2008, TAO has been structured in several special interest groups (SIGs) to enable the agile investigation of long-term or emerging theoretical and applicative issues. The comparatively small size of TAO SIGs enables in-depth and lively discussions; the fact that all TAO members belong to several SIGs, on the basis of their personal interests, enforces the strong and informal collaboration of the groups, and the fast information dissemination.

The first two SIGs consolidate the key TAO scientific pillars, while the others evolve and adapt to new topics.

The **Stochastic Continuous Optimization** SIG (OPT-SIG) takes advantage of the fact that TAO is acknowledged the best French research group and one of the top international groups in evolutionary computation from a theoretical and algorithmic standpoint. A main priority on the OPT-SIG research agenda is to provide theoretical and algorithmic guarantees for the current world state-of-the-art continuous stochastic optimizer, CMA-ES, ranging from convergence analysis (Youhei Akimoto's post-docs) to a rigorous benchmarking methodology. Incidentally, this benchmark platform COCO has been acknowledged since 2009 as "the" international continuous optimization benchmark, and its extension is at the core of the ANR project NumBBO (started end 2012). Another priority is to address the current limitations of CMA-ES in terms of high-dimensional or expensive optimization and constraint handling (respectively Ouassim Ait El Hara's, Ilya Loshchilov's PhDs and Asma Atamna's).

The **Optimal Decision Making under Uncertainty** SIG (UCT-SIG) benefits from the MoGo expertise and the team previous activity reports) and its past and present world records in the domain of computer-Go, establishing the international visibility of TAO in sequential decision making. Since 2010, UCT-SIG resolutely moves to address the problems of **energy management** from a fundamental and applied perspective. On the one hand, energy management offers a host of challenging issues, ranging from long-horizon policy optimization to the combinatorial nature of the search space, from the modeling of prior knowledge to non-stationary environment to name a few. On the other hand, the energy management issue can hardly be tackled in a pure academic perspective: tight collaborations with industrial partners are needed to access the true operational constraints. Such international and national collaborations have been started by Olivier Teytaud during his three stays (1 year, 6 months, 6 months) in Taiwan, and witnessed by the FP7 STREP Citines, the ADEME Post contract, and the METIS I-lab with SME Artelys.

The **Data Science** SIG (DS-SIG) now includes the activities related to the CDS and ISN Lidex in Saclay. On the one hand, it replaces and extends the former *Distributed systems* SIG, that was devoted to the modeling and optimization of (large scale) distributed systems, and itself was extending the goals of the original *Autonomic Computing* SIG, initiated by Cécile Germain-Renaud and investigating the use of statistical Machine Learning for large scale computational architectures, from data acquisition (the Grid Observatory in the European Grid Initiative) to grid management and fault detection. But these activities have become more and more application-driven, from High Energy Physics for the highly distributed computation to the Social Sciences for the multi-agents approaches – hence the change of focus of this SIG. A major result of this theme has been the creation 2 years ago of the Paris-Saclay Center for Data Science, co-chaired by Balázs Kégl, and the organization of the Higgs-ML challenge (http://higgsml.lal.in2p3.fr/), most popular challenge ever on the Kaggle platform.

On the other hand, several activities around Digital Humanities involving Gregory Grefenstette, Cécile Germain-Renaud, Michèle Sebag and Philippe Caillou, have widely extended previous work around the modeling of multi-agent systems and the exploitation of simulation results in the SimTools RNSC network frame. Digital Humanities involves adding semantics to underspecified collections of societal information: in an historical perspective (as in the new TAO H2020 project, EHRI-II on holocaust archives, or in the Gregorius project on church history); or an economical and societal perspective (as in the Cartolabe and AMIQAP projects); or an individual perspective (as in the ongoing Personal Semantics project). The key challenge here is to use learning algorithms to find structure and extract knwoledge from poorly structured or unstructured information, and to provide intelligible results and/or means to interact with the user.

The **Designing Criteria** SIG (CRI-SIG) focuses on the design of learning and optimization criteria. It elaborates on the lessons learned from the former *Complex Systems* SIG, showing that the key issue in challenging applications often is to design the objective itself. Such targeted criteria are pervasive in the study and building of autonomous cognitive systems, ranging from intrinsic rewards in robotics to the notion of saliency in vision and image understanding, and that of automatic algorithm selection and parameterization. The desired criteria can also result from fundamental requirements, such as scale invariance in a statistical physics perspective, and guide the algorithmic design. Additionally, the criteria can also be domain-driven and reflect the expert priors concerning the structure of the sought solution (e.g., spatio-temporal consistency); the challenge is to formulate such criteria in a mixed non convex/non differentiable objective function, nevertheless amenable to tractable optimization.

The **Deep Learning and Information Theory** SIG (DEEP-SIG) involves Yann Ollivier, Guillaume Charpiat, Michèle Sebag. This SIG originated from some extensions of the wwork done in the *Distributed Systems* SIG that have been developped in the context of the TIMCO FUI project (started end 2012 and just ended); the challenge was not only to port ML algorithms on massively distributed architectures, but to see how these architectures can inspire new ML criteria and methodologies. The coincidence of this project with the arrival of Yann Ollivier in TAO gradualy lead this work toward Deep Networks. This year, in addition to studying various theoretical and practical aspects of deep learning, we provide information-theoretic perspectives on the design and optimization of deep learning models, such as using the Fisher information matrix to optimize the parameters, or using minimum description length criteria to choose the right model structure (topology of the neural graph, addition or removal of parameters...) and to provide regularization and model selection.

# 4. Application Domains

## 4.1. Energy Management

Energy management, our prioritary application field, involves sequential decision making with:

- stochastic uncertainties (typically weather);
- both high scale combinatorial problems (as induced by nuclear power plants) and non-linear effects;
- high dimension (including hundreds of hydroelectric stocks);
- multiple time scales:
  - minutes (dispatching, ensuring the stability of the grid), essentially beyond the scope of our work, but introducing constraints for our time scales;
  - days (unit commitment, taking care of compromises between various power plants);
  - years, for evaluating marginal costs of long term stocks (typically hydroelectric stocks);
  - decades, for investments.

Significant challenges also include:

- spatial distribution of problems; due to capacity limits we can not consider a power grid like Europe + North Africa as a single “production = demand” constraint; with extra connections we can equilibrate excess production by renewables for remote areas, but not in an unlimited manner.
- other uncertainties, which might be modelized by adversarial or stochastic frameworks (e.g. technological breakthroughs, decisions about ecological penalization).

We have had several related projects (Citines, a European (FP7) project; IOMCA, a ANR project), and we now work on the POST project, a ADEME BIA about investments in power systems. Our collaboration with company Artelys (working on optimization in general, and in particular on energy management) is formalized as an Inria ILAB.

**Technical challenges**: Our work focuses on the combination of reinforcement learning tools, with their anytime behavior and asymptotic guarantees, with existing fast approximate algorithms. Our goal is to extend the state of the art by taking into account non-linearities which are often neglected in power systems due to the huge computational cost. We study various modelling errors, such as bias due to finite samples, linearization, and propose corrections.

**Related Activities**:

- Joint team with Taiwan, namely the Indema associate team (see Section 9.4.2.1).
- Ilab METIS, in progress with Artelys (see Section 6.6) for industrialization of our work. In particular, the Crystal tool is adopted by the European Community (http://www.artelys.com/news/120/90/Energy-The-European-Commission-Chooses-Artelys-Crystal)
- Organization of various forums and meetings around Energy Management
- Visit of Edgar Galvan Lopez also includes applications to energy management, more precisely Demand-Side Management systems. In [40], Differential Evolution is used to generate optimal plans to use the accumulators of electrical vehicles in order to reduce the peak household consumption loads.

## 4.2. TAO & Humanities

Several projects related to research for Humanities and/or research transfer have started in 2015:

- Personal semantics (Gregory Grefenstette). In the current digital world, individuals generate increasing amount of personal data. Our work involves discovering semantic axes for organizing and exploiting this data for personal use.
- Gregorius (Cécile Germain & Gregory Grefenstette). An application of semantic structuring and automatic enrichment of existing digital humanities archives.
- Cartolabe (Ph. Caillou, Gregory Grefenstette, Jean-Daniel Fekete - AVIZ, Michèle Sebag). The Cartolabe project applies machine learning techniques to determine comprehensible structures in unstructured data. The goal is to use raw textual data, and underspecified ontologies, to provide intuitive access to pertinent research activities in a large research organisation.
- AmiQap (Ph. Caillou, Michèle Sebag). The multivariate analysis of questionaire data relative to the quality of life at work, in relation with the socio-economical indicators of firms, aims at investigating the relationship between quality of life and economic performances (depending on the activity domain).
- Collaborative Hiring (Ph. Caillou, Michèle Sebag). Thomas Schmitt's PhD, started in 2014, aims at handling job offers and demands matching as a collaborative filtering problem.
- IODS (Wikidata for Science).

Significant challenges include some Big Data problems:

- learning interpretable clusters from bottom-up treatment of mixed text and numerical data
- aligning bottom-up clusters with existing manually created top-down structures
- building a unified system integrating the "dire d'experts".

**Partners**:

- Amiqap is funded by ISN, with Telecom SES, RITM and La Fabrique as partners.
- The collaborative hiring study is conducted in cooperation with J.P. Nadal from EHESS.

# 5. Highlights of the Year

## 5.1. Highlights of the Year

**DataScienc@LHC** First of a series of workshops officially organized at CERN - TAO leader on the ML side.

### 5.1.1. Awards

- Best Paper Award in the Genetic Programming track at GECCO 2015 (Madrid, July 2015) for the paper .
- First place in the Taxonomy Induction task of SemEval 2015 (Denver, June 2015)[55]

BEST PAPER AWARD:

[39]
R. FFRANCON, M. SCHOENAUER. *Memetic Semantic Genetic Programming*, in "Genetic and Evolutionary Computation COnference (GECCO 2015)", Madrid, Spain, A. ESPARCIA, S. SILVA (editors), ACM, July 2015, pp. 1023-1030, Best Paper Award in the GP track - see http://www.sigevo.org/gecco-2015/papers.html, https://hal.inria.fr/hal-01169074

# 6. New Software and Platforms

## 6.1. CMA-ES

Covariance Matrix Adaptation Evolution Strategy
KEYWORDS: Numerical optimization - Black-box optimization - Stochastic optimization
SCIENTIFIC DESCRIPTION

The CMA-ES is considered as state-of-the-art in evolutionary computation and has been adopted as one of the standard tools for continuous optimisation in many (probably hundreds of) research labs and industrial environments around the world. The CMA-ES is typically applied to unconstrained or bounded constraint optimization problems, and search space dimensions between three and a hundred. The method should be applied, if derivative based methods, e.g. quasi-Newton BFGS or conjugate gradient, (supposedly) fail due to a rugged search landscape (e.g. discontinuities, sharp bends or ridges, noise, local optima, outliers). If second order derivative based methods are successful, they are usually faster than the CMA-ES: on purely convex-quadratic functions, f(x)=xTHx, BFGS (Matlabs function fminunc) is typically faster by a factor of about ten (in terms of number of objective function evaluations needed to reach a target function value, assuming that gradients are not available). On the most simple quadratic function f(x)=||x||2=xTx BFGS is faster by a factor of about 30.
FUNCTIONAL DESCRIPTION

The CMA-ES is an evolutionary algorithm for difficult non-linear non-convex black-box optimisation problems in continuous domain.

- Participants: Nikolaus Hansen and Emmanuel Benazera
- Contact: Nikolaus Hansen
- URL: https://www.lri.fr/~hansen/cmaesintro.html

## 6.2. COCO

COmparing Continuous Optimizers
KEYWORDS: Benchmarking - Numerical optimization - Black-box optimization - Stochastic optimization
SCIENTIFIC DESCRIPTION

COmparing Continuous Optimisers (COCO) is a tool for benchmarking algorithms for black-box optimisation. COCO facilitates systematic experimentation in the field of continuous optimization. COCO provides: (1) an experimental framework for testing the algorithms, (2) post-processing facilities for generating publication quality figures and tables, (3) LaTeX templates of articles which present the figures and tables in a single document.

The COCO software is composed of two parts: (i) an interface available in different programming languages (C/C++, Java, Matlab/Octave, R, Python) which allows to run and log experiments on multiple test functions testbeds of functions (noisy and noiseless) are provided (ii) a Python tool for generating figures and tables that can be used in the LaTeX templates.

FUNCTIONAL DESCRIPTION

The Coco Platform provides the functionality to automatically benchmark optimization algorithms for unbounded, unconstrained optimization problems in continuous domains. Benchmarking is a vital part of algorithm engineering and a necessary path to recommend algorithms for practical applications. The Coco platform releases algorithm developers and practitioners alike from (re-)writing test functions, logging, and plotting facilities by providing an easy-to-handle interface in several programming languages. The Coco platform has been developed since 2007 and has been used extensively within the "Blackbox Optimization Benchmarking (BBOB)" workshop series since 2009. Overall, 123 algorithms and algorithm variants by contributors from all over the world have been benchmarked with the platform so far and all data is publicly available for the research community for the submissions to BBOB-2013).

- Participants: Dimo Brockhoff, Tea Tulsar, Dejan Tulsar, Thanh-Do Tran, Nikolaus Hansen, Anne Auger, Marc Schoenauer, Ouassim Ait Elhara and Asma Atamna
- Partners: Université technique de Dortmund - Université technique de Prague
- Contact: Dimo Brockhoff
- URL: http://coco.gforge.inria.fr/doku.php

## 6.3. PTraces

Personal digital Traces

KEYWORDS: Information retrieval - Taxonomy induction - Personal Big Data - Ontology

SCIENTIFIC DESCRIPTION Personal digital Traces (PTraces) is a platform for fetching and annotating personal data. Ptraces provides us a demonstration platform for personal semantics and ontology induction. Ptraces includes: (1) modules for securely fetching personal data from external applictions (2) an annotation component that annotates personal information with user-chosen taxonomies (facets). These facets arean experimental framework for testing the algorithms, (3) an information retrieval component

The PTraces software is composed of two parts: (i) A Java backbone, built on Elastic Search (ii) a web-browser user interface for configuring the system (choosing data sources, choosing ontologies to activate) and for query retrieved data.

FUNCTIONAL DESCRIPTION

The system accepts taxonomies in the SKOS format. These taxonomies are presented to the user, along with a number of connectors to otuside data sources (gmail, twitter, facebook, fitbit, ...). The user chooses which sources to index, and which taxonomies to apply. After this initialisation, the system accesses the sources using the indentification tokens that the user has supplied, fetches the user data, annotates the data using the activated taxonomies, and sets up a local server that that user can access to search in their own data, using a browser interface.

- Participants: Gregory Grefenstette, Lawrence Muchemi, Mohamed Bouatira
- Contact: Mohamed Bouatira
- URL: https://gforge.inria.fr/scm/browser.php?group_id=7217

## 6.4. Cartolabe

FUNCTIONAL DESCRIPTION

Scientific cartography from articles

- Contact: Philippe Caillou
- URL: (project starting in 2016 - no url yet)

## 6.5. GO

Grid Observatory
KEYWORDS: Green computing - Autonomic computing
FUNCTIONAL DESCRIPTION

The Grid Observatory (GO) software suite collects and publishes traces of the EGI (European Grid Initiative) grid usage. With the release and extensions of its portal, the Grid Observatory has made a database of grid usage traces available to the wider computer science community since 2008. These data are stored on the grid, and made accessible through a web portal without the need of grid credentials. The GO is fully integrated with the evolution of EGI monitoring.

- Participants: Cécile Germain-Renaud, Julien Nauroy and Martine Sebag
- Contact: Cécile Germain-Renaud
- URL: http://grid-observatory.org/

## 6.6. METIS

KEYWORDS: Optimization - Energy
FUNCTIONAL DESCRIPTION

Many works in Energy Optimization, in particular in the case of high-scale sequential decision making, are based on one software per application, because optimizing the software eventually implies losing generality. Our goal is to develop with Artelys a platform, METIS, which can be used for several applications. In 2012 we interfaced existing codes in Artelys and codes developed in the TAO team, experiments have been performed and test cases have been designed. Several codes have been developped, tested on real world problems, and are (depending on which code) under the process of (i) open source diffusion (ii) code protection.

A big work is the development of a new, independent, open source, simulator for the French power grid, currently under extension to other European countries.

- Participants: Olivier Teytaud, Jérémie Decock, Jean-Joseph Christophe, Vincent Berthier, Marie-Liesse Cauwet and Sandra Cecilia Astete Morales
- Partner: Artelys
- Contact: Olivier Teytaud
- URL: http://www.lri.fr/~teytaud

## 6.7. Game Test Bed

KEYWORDS: An open source game test bed.
FUNCTIONAL DESCRIPTION

GTB is an open source library of games, including solvers. It includes an interface to noisy continuous optimization of parametric policies, and noise-free continuous optimization, leading to a preliminary continuous optimization platform with real-world test cases.

- Participants: Olivier Teytaud is the only developer, some feedback from Sandra Astete-Morales, and an interfacing with works by Marcus Gallagher at Univ. Queensland in under discussion.
- URL: https://gforge.inria.fr/projects/gametestbed/

## 6.8. MOGO

KEYWORDS: Computer Go - Monte-Clarlo - UCT
FUNCTIONAL DESCRIPTION

MoGo and its Franco-Taiwanese counterpart MoGoTW is a Monte-Carlo Tree Search program for the game of Go, which made several milestones of computer-Go in the past (first wins against professional players in 19x19, first win with disadvantageous side in 9x9 Go). Recent results include 7 wins out of 12 against professional players (in Brisbane, 2012) in 7x7, and recently an optimization of the random seed which brings a significant improvement in Go and (unpublished) on the difficult case of phantom-Go. However, the work in the UCT-SIG has now shifted to energy management.

- Participants: Sylvain Gelly, Rémi Munos, Olivier Teytaud, Yizao Wang and Jean-Baptiste Hoock
- Contact: Olivier Teytaud
- URL: https://www.lri.fr/~teytaud/taiwanopen2009.html

## 6.9. MultiBoost

multi-purpose boosting package
KEYWORDS: Multi-class - Multi-label classification
FUNCTIONAL DESCRIPTION

The MultiBoost package provides a fast C++ implementation of multi-class/multi-label/multi-task boosting algorithms. It is based on AdaBoost.MH but it also implements popular cascade classifiers, Arc-GV , and FilterBoost . The package contains common multi-class base learners (stumps, trees, products, Haar filters). Further base learners and strong learners following the boosting paradigm can be easily implemented in a flexible framework.

- Participants: Balasz Kegl, Robert Busa-Fekete and Djalel Benbouzid
- Contact: Balasz Kegl
- URL: http://www.multiboost.org/

## 6.10. io.datascience

FUNCTIONAL DESCRIPTION

This Data as a Service (DaaS) platform [54] is developed in the context of the Center for Data Science and the TIMCO project. Its overall goals is to exploit the advances in semantic web techniques for efficient sharing and usage of scientific data. A related specific software is the Tester for Triplestore (TFT) software suite [49], which benchmarks the compliance of sparql databases wrt the RDF standard and publishes the results through the SparqlScore service.

- Contact: Cécile Germain
- URL: https://io.datascience-paris-saclay.fr/

# 7. New Results

## 7.1. Optimal Decision Making under Uncertainty

The Tao-uct-sig is working mainly on mathematical programming tools useful for power systems. In particular, we advocate a data science approach, in order to reduce the model error - which is much more critical than the optimization error, in most cases. Real data are the best way for handling uncertainties. Our main works are as follows:

**Noisy optimization** In the context of stochastic uncertainties, noisy optimization handles the model error by simulation-based optimization. Our results include:

- A formalization of noisy optimization in continuous domains, often poorly modeled in the evolutionary computation community [64], [6]. We also proposed heuristic rules for reaching slope -1/2 in log-log scale [34]. We also show that in some settings the slope -1 (classical in mathematical programming) can be recoved in evolution strategies (unpublished: http://www.lri.fr/~teytaud/mca.pdf), and we provided complexity bounds [20].
- An extension of portfolio algorithms for noisy optimization. Portfolio methods are already usual in combinatorial optimization, some works exists in the continuous case, this is the first work in the noisy case[8].
- Pragmatic approaches of noisy optimization, for improving robustness and for taking into account human expertise, including: Applying sieves methods in noisy optimization [27], paired optimization [35], and combining various policies [25].
- Upper bounds on noisy optimization in discrete domains [5].

**Quasi-random numbers** Continuous optimization is a key component of our works, hence we improve evolution strategies (which have simplicity and convenience qualities) by quasi-random numbers (showing that even in simple cases it is beneficial[52], and provides great improvements in highly multimodal cases (unpublished, http://www.lri.fr/~teytaud/qrr.pdf)). We also developped criteria and testbeds, pointing out some key points not widely studied in the optimization literature[26]. We also extended our earlier results in parallel optimization to additional algorithms[30], and used cutting planes as in the ellipsoid method, hence combining the best of both worlds, i.e. fast rates from cutting planes methodologies and parallel behavior as in evolution strategies[36].

**Dynamical problems** The dynamical nature of power systems is critical, as transient regimes, ramping constraints are ubiquitous in unit commitment and dispatch. Optimizing policies, with their temporal components, is a challenge when the high dimension and the nonlinearities are taken into account. Games provide a nice testbed for experiments and are used in several of our works. We provided:

- An original algorithm for learning opening books, by an unexpected use of random seeds[32]. The principle is to randomly sample policies, by modifying the random seed. This can be used for any stochastic policy: we generate thousands of deterministic policies (by setting the random seed to arbitrary values) and select the best ones. This can be applied for games (always the most convincing application for a proof of concept), and any control problem where stochastic policies are available.
- An extension of the previous work for dynamically adapting the probability distribution for specific positions[51]. This work provides a MCTS without the scalability limitations of MCTS. This work might give birth to many future works.

## 7.2. Continuous Optimization

**Markov Chain Analysis of Evolution Strategies** The theory of Markov chains with discrete time and continuous state space turns out to be very useful to analyze the convergence of adaptive evolution strategies (including simplified versions of the state-of-the are CMA-ES). Exploiting invariance of the algorithms, we can indeed construct homogeneous Markov chains underlying the algorithms whose stability implies the linear convergence of the algorithm [65]. We have also shown how the convergence on constrained problems can be analyzed with Markov chains theory [10].

However the stability can be very difficult to prove; even the irreducibility can be very challenging to prove with current Markov chain theory. We have hence been developing new theoretical tools exploiting deterministic control models to prove more easily the irreducibility and T-chain property of general Markov chains [67]. Those theoretical tools can be applied to the optimization algorithms we are interested in, and trivialize some stability proofs [1], [10].

**Benchmarking of continuous optimizers** We have been pursuing our effort towards improving the standards in benchmarking of continuous optimizers. We tackled the benchmarking of bi-objective problems and transferred and adapted standard benchmarking techniques from the single-objective optimization and classical derivative free optimization community to the field of EMO [28]. In addition, we have been rewritting part of the COCO platform to improve its modularity and make it less error prone and started its extension to multiobjective optimization.

**Concentration inequalities for sampling without replacement** We studied the concentration of measure phenomenon in the case of sampling without replacement, which is directly relevant for a recent MCMC techique for large data sets, see [7] accepted to the Bernoulli journal.

**Random projections for confident MCMC** In the paper [66] accepted at the NIPS "Bayesian Optimization Workshop", we study the benefit of replacing uniform subsampling by random projections in recent MCMC techniques for linear regression of tall datasets.

**Automatic step size adaptation** We have derived a new, low-cost strategy for online adaptation of the step size in stochastic gradient descent and related algorithms [72]. This problem is of crucial importance in many machine learning algorithms, as current approaches often rely on exploring a grid of step sizes and performing a full optimization for each of them, a lengthy process.

## 7.3. Data Science

**High Energy Physics** The success of the 2014 HiggsML challenge has created a willingness for structured collaboration from the High Energy Physics experiments. A working group has been set up and new challenges are currently explored. A yearly workshop has been decided, with a first edition at CERN 9-13 Nov. 2015, *DataScience@LHC*.

The challenge exemplifies a new machine learning task [58][56]: *learning to discover* evaluating the significance of a scientific discovery. It can be formally casted into a two-class classification problem, but with two major departures from a regular setting. 1) Discovery: labeled training examples of the positive class (signals) are not available and must be obtained from simulation. The learning machine can then address the “inverse problem” of predicting which events are signals in real data. 2) Evaluation: because the classes are enormously imbalanced and overlapping, the objective function of the classifier is a metric of a statistical test.

**Personal Semantics** Our algorithm for inducing a taxonomy from a set of domain terms placed first in the international Taxonomy Induction task, part of the SemEval 2015 conference in Denver. Since then, we have developed a robust technique for discovering the domain vocabulary for a new topic using a directed crawler we created. We are currently creating hundreds of taxonomy for personal themes (hobbies, illnesses) that can be integrated into our Personal Semantics platform PTraces. The challenges for the coming year will be deploying and evaluating the taxonomies, and introducing newer machine learning methods, such as Latent Dirichlet Allocation, for better recognizing domain vocabularies.

**Distributed system observation** The work on distributed system automated analaysis and description [59][60], has been persued thru the continued development of the GAMA multi-agent framework https://github.com/gama-platform/gama/wiki. The simulation framework has been applied to the study of a new anytime reverse auctions protocol [53]. Philippe Caillou is associated to the young researcher ANR ACTEUR, coordinated by Patrick Taillandier (IDEES, Rouen university). With this project, a new BDI cognitive agent model, designed to be easy to use for non computer scientist, has been proposed [29] and applied to Rouen traffic simulation [57]. Finaly, agent behavior has been extracted from human player logs to study the perception of emotive behaviors in board games [37].

**Digital humanities** Amiqap and Cartolabe projects both start in 2016. The Cartolabe project applies machine learning techniques to determine comprehensible structures in unstructured data. The goal is to use raw textual data, and underspecified ontologies, to provide intuitive access to pertinent research activities in a large research organisation. Amiqap studies the relation between worker well-being and company performance, in collaboration with Mines ParisTech sociology department and La Fabrique de l’Industrie for research, Secafi and DARES for the data.

These activities will benefit from Paola Tubaro's arrival (researcher CNRS in sociology and economy) in 2016.

## 7.4. Designing criteria

**Criterion design and optimization methods for computer vision** On the topic of large-scale image segmentation with multiple object detection, targetting as an application the analysis of high-resolution multispectral satellite images covering the Earth, challenges are numerous: scalable complexity, finding good features to distinguish objects, designing shape statistics as well as an optimization method able to incorporate them. We propose a solution [42], [43] based on the construction of binary partition trees and on their optimization, whose cost is alleviated thanks to theoretical results reducing the search space. Concerning video segmentation, we have extended previous work, on the inclusion of shape growth constraints into classical MRF settings (graph cuts with globally optimal segmentation), to the case of multimodal sequences of medical 3D scans [19]. We also studied a new family of metrics in [9], together with a redefinition of the associated gradient and practical ways to compute it. This allows the consideration of new types of priors on planar curve evolution, such as piecewise-rigid motions. Surprisingly, the problem of finding the best piecewise-rigid approximation of a motion turns out to be convex, and to be linked to sparsity approaches.

**Algorithm selection and configuration** Two PhD theses are still related to the former *Crossing the Chasm* SIG: Nacim Belkhir has worked on inline parameter tuning for the CMA-ES algorithm in the context of a large number of cores [21], and is now using surrogate models to compute the features of expensive continuous optimization (submitted). François Gonard's PhD is dedicated to algorithm selection. The original application domain is that of expensive car industry simulations (within the IRT-ROM project). Initial results concern combinatorial optimization, and François obtained a "Honorable mention from the jury" for his submission to the ICON Challenge (http://iconchallenge.insight-centre.org/), for its original approach coupling a pre-scheduler and an algorithm selector. A paper describing the algorithms and analyzing the results has been submitted.

**A statistical physics perspective** In the topic of MRF design, with motivating applications in large scale inference problems like traffic congestions, we have finalized in [13] an approach based on the disordered Ising model relying on approximate solutions to the Inverse Ising problem. To this specific problem we also propose new approximate solutions, compliant with the generalized belief-propagation algorithm in the static [63] and a new $l_0$ regularized method based on a maximum likelihood maximization for the dynamical case[11]. In fact in [63] we have developped a method adapted to the generalized belief propagation framework, aiming at adressing directly and systematically the loop corrections without loss of scalability, offering new possibilities in the context of inference by MRF models. In parallel, a better understanding of the so-called mean-field approximation when the phase space is clustered has been derived [68] giving a direct method to solve static inverse problem in the weak coupling limit. Apart from the method point of view, some consideration over what can be said on the data has been considered, still in the topic of MRF design. In this sense, it is shown in [69] that the reconstruction of the MRF model depends strongly on how the data are gathered, and how to remove redundant data and keep a good reconstruction.

**Multi-objective AI Planning** This activity had almost stopped since the end of the DESCARWIN ANR project. However, a productive intership resulted in some new benchmarks in the ZenoTravel domain together with an exact solver ensuring the knowledge of the true Pareto front [48], [47].

## 7.5. Deep Learning and Information Theory

**Natural Gradients for Deep Learning** Deep learning is now established as a state-of-the-art technology for performing different tasks such as image or sequence processing. Nevertheless, much of the computational burden is spent on tuning the hyper-parameters. On-going work, started during the TIMCO project, is proposing, in the framework of Riemannian gradient descents, invariant algorithms for training neural networks that effectively reduce the number of arbitrary choices, e.g., affine transformations of the activation functions or shuffling of the inputs. Moreover, the Riemannian gradient descent algorithms perform as well as the state-of-the-art optimizers for neural networks, and are even faster for training complex models. The proposed approach is based on Amari's theory of information geometry and consists in practical and well-grounded approximations for computing the Fisher metric. The scope of this framework is larger than Deep Learning and encompasses any class of statistical models.

**Training dynamical systems online without backtracking** with application to recurrent neural networks [73]. We propose an algorithm to learn the parameters of a dynamical system in an online, memoryless setting, thus requiring no backpropagation through time, and consequently scalable, avoiding the large computational and memory cost of maintaining the full gradient of the current state with respect to the parameters. The algorithm essentially maintains, at each time, a single search direction in parameter space. The evolution of this search direction is partly stochastic and is constructed in such a way to provide, at every time, an unbiased random estimate of the gradient of the loss function with respect to the parameters.

**Approximating Bayesian predictors thanks to Laplace's rule of succesion** Laplace's "add-one" rule of succession modifies the observed frequencies in a sequence of heads and tails by adding one to the observed counts. This improves prediction by avoiding zero probabilities and corresponds to a uniform Bayesian prior on the parameter. We prove that, for any exponential family of distributions, arbitrary Bayesian predictors can be approximated by taking the average of the maximum likelihood predictor and the sequential normalized maximum likelihood predictor from information theory, which generalizes Laplace's rule. The proof heavily involves the geometry provided by the Fisher information matrix. Thus it is possible to approximate Bayesian predictors without the cost of integrating or sampling in parameter space[46].

# 8. Bilateral Contracts and Grants with Industry

## 8.1. Bilateral Contracts with Industry

**Thales Research & Technology** 2014-2017 (30 kEuros), related to Nacim Belkhir's CIFRE PhD
Coordinator: Marc Schoenauer
Participants: Johann Dréo, Pierre Savéant, Nacim Belkhir

**Orange** 2013-2016 (30 kEuros), related to Robin Allesiardo's CIFRE PhD
Coordinator: Michèle Sebag
Participants: Raphael Feraud, Robin Allesiardo

**Réseau Transport d'Electricité** 2015-2018 (30 kEuros), related to Benjamin Donnot's CIFRE PhD
Coordinator: Olivier Teytaud
Participants: Benjamin Donnot, Antoine Marot

**Augure (SME)** 2013-2015 (150 kEuros). MoDyRum (Modélisation Dynamique d'un Réseau Médiatique), related to Marco Bressan's postdoc SME Augure)
Coordinator: Michèle Sebag
Participants: Philippe Caillou, Cyril Furtlehner, Marco Bressan

# 9. Partnerships and Cooperations

## 9.1. Regional Initiatives

**TIMCO** 2012-2015 (432 kEuros)
Coordinator: Bull SAS
Participants: Cécile Germain-Renaud, Julien Nauroy, Karima Rafes, Lovro Ilisajic, Gaetan Marceau Caron

**ROM** *Model Reduction and Multiphysics Optimization* 2014-2016 (50 Keuros)
Coordinator: IRT System X
Participants: Marc Schoenauer, Michèle Sebag, François Gonard (PhD)

**ISN** *A Collaborative Filtering Approach to Matching Job Openings and Job Seekers*, 2013-2016 (105 kEuros)

Related to Thomas Schmitt's PhD (funded by ISN).
Participants: Michèle Sebag, Thomas Schmitt

**AutoML** *An empirical approach to Machine Learning* 2014-2017 (104 kEuros)
Related to Sourava Mishra's PhD
Participants: Michèle Sebag, Balazs Kégl, Sourava Mishra

**ReMODeL** *Rewarded Multimodal Online Deep Learning* 2015-2016 (31,5 kEuros)
This project lies at the junction of reinforcement learning, deep learning, computational neuroscience and developmental robotic fields. It is closely related to the transversal DIGITEO robotic theme, Roboteo.
Participants: Michèle Sebag, Mathieu Lefort, Alexander Gepperth

**AMIQAP** 2015-2016 (12 months of Postdoctoral fellow). Project funded by ISN
Participants: Philippe Caillou, Michèle Sebag

## 9.2. National Initiatives

**SIMINOLE** 2010-2015 (1180kEuros, 250kEuros for TAO). Large-scale simulation-based probabilistic inference, optimization, and discriminative learning with applications in experimental physics, ANR project, Coordinator B. Kégl (CNRS LAL).

Participants: Emmanuel Benazera, Nikolaus Hansen, Marc Schoenauer, Cécile Germain-Renaud

**NUMBBO** 2012-2016 (290kEuros for TAO). Analysis, Improvement and Evaluation of Numerical Blackbox Optimizers, ANR project, Coordinator Anne Auger, Inria. Other partners: Dolphin, Inria Lille, Ecole des Mines de Saint-Etienne, TU Dortmund
Participants: Anne Auger, Nikolaus Hansen, Marc Schoenauer, Ouassim Ait ElHara

**ACTEUR** 2014-2018 (236kEuros). Cognitive agent development for urban simulations, ANR project, Coordinator P. Taillandier (IDEES, Univ Rouen).

Participants: Philippe Caillou

### *9.2.1. Other*

**POST** 2014-2018 (1,220 MEuros, including 500 kEuros for a 'private' cluster). Platform for the optimization and simulation of trans-continental grids
ADEME (Agence de l'Environnement et de la Maîtrise de l'Energie)
Coordinator: ARTELYS
Participants: Olivier Teytaud, Marie-Liesse Cauwet, Jérémie Decock, Sandra Cecilia Astete Morales, David L. Saint-Pierre, J. Decock

**E-LUCID** 2014-2017 (194 kEuros)
Coordinator: Thales Communications & Security S.A.S
Participants: Marc Schoenauer, Cyril Furtlehner

**FSN ADAMME** 2015-2018 (258 kEuros)
Coordinator: Bull SAS
Participants: Marc Schoenauer, Yann Ollivier, Gaetan Marceau Caron, Guillaume Charpiat, Cécile Germain-Renaud, Michèle Sebag

**CNES contract** 2015-2017 (70 kEuros)
Coordinator: Manuel Grizonnet (CNES) & Yuliya Tarabalka (Inria Sophia-Antipolis, Titane team)
Participants: Guillaume Charpiat

## 9.3. European Initiatives

### *9.3.1. FP7 & H2020 Projects*

**EHRI-II** 2015-2019 (7 969 kEuros). European Holcaust Research Infrastructure, H2020, Coordinator NIOD, Amsterdam. Digital Humanities.

Participants: Gregory Grefenstette

#### *9.3.1.1. Collaborations in European Programs, except FP7 & H2020*

#### *9.3.1.2. Collaborations with Major European Organizations*

**MLSpaceWeather** 2015-2019. Coupling physics-based simulations with Artificial Intelligence.
Coordinator: CWI
Participants: Michèle Sebag, Aurélien Decelle, Cyril Furtlehner.

## 9.4. International Initiatives

### *9.4.1. Inria International Labs*

### *9.4.2. Inria Associate Teams not involved in an Inria International Labs*

#### *9.4.2.1. CIADM*

Title: Computational intelligence and Decision making
International Partner (Institution - Laboratory - Researcher):
NUTN (Taiwan) - Multimedia Informatics Lab - Chang-Shing Lee
Start year: 2015
See also: http://www.lri.fr/~teytaud/indema.html
The associate team works on computation intelligence for decision making, with different application fields for the various partners: - power systems (Tao) - eLearning (Oase) - games (Ailab)

#### *9.4.2.2. s3-bbo*

Title: Threefold Scalability in Any-objective Black-Box Optimization
International Partner (Institution - Laboratory - Researcher):
Shinshu (Japan) - Tanaka-Hernan-Akimoto Laboratory - Hernan Aguirre
Start year: 2015
See also: http://francejapan.gforge.inria.fr/doku.php?id=associateteam
This associate team brings together researchers from the TAO and Dolphin Inria teams with researchers from Shinshu university in Japan. Additionally, researchers from the University of Calais are external collaborators to the team. The common interest is on black-box single and multi-objective optimization with complementary expertises ranging from theoretical and fundamental aspects over algorithm design to solving industrial applications. The work that we want to pursue in the context of the associate team is focused on black-box optimization of problems with a large number of decision variables and one or several functions to evaluate solutions, employing distributed and parallel computing resources. The objective is to theoretically derive, analyze, design, and develop scalable black-box stochastic algorithms including evolutionary algorithms for large-scale optimization considering three different axes of scalability: (i) decision space, (ii) objective space, and (iii) availability of distributed and parallel computing resources.

We foresee that the associate team will make easier the collaboration already existing through a proposal funded by Japan and open-up a long term fruitful collaboration between Inria and Shinshu university. The collaboration will be through exchanging researchers and Ph.D. students and co-organization of workshops.

### *9.4.3. Inria International Partners*

#### *9.4.3.1. Declared Inria International Partners*

*9.4.3.2. Informal International Partners*

Marc Schoenauer , partner of the ARC-DP (Australian Research Council Discovery Project) *bio-inspired computing methods for dynamically changing environments*. Coordinator: University of Adelaide (Frank Neumann), 5 years, 400 k$-AUS.

9.4.3.2.1. Participation In other International Programs

## 9.5. International Research Visitors

### *9.5.1. Visits of International Scientists*

- Holger Hoos, University of British Columbia, Canada, 3 weeks in February 2015, follow-up of his 3-months visit at Fall 2014, funded my MSR-Inria joint lab.
- Isabelle Guyon, Chalearn. April-July 2015, 1 month by University Paris Sud, 3 months with TIMCO.
- Youhei Akimoto, Shinshu University, September 2015, a month funded by Digiteo.
- Aditya Gopalan, Indian Institute of Science Bangalore, April 2015, three weeks funded by Digiteo.
- Edgar Galvan Lopez, University College Dublin, April 2015 - March 2016, funded with the ELEVATE Fellowship, the Irish Research Council's Career Development Fellowship co-funded by Marie Curie Actions.

*9.5.1.1. Internships*

Lin Ching-Nung

Date: Apr - Oct 2015

Institution: NDHU (Taiwan)

Supervisor: Olivier Teytaud

### *9.5.2. Visits to International Teams*

*9.5.2.1. Sabbatical programme*

*9.5.2.2. Explorer programme*

*9.5.2.3. Research stays abroad*

# 10. Dissemination

## 10.1. Promoting Scientific Activities

### *10.1.1. Scientific events organisation*

*10.1.1.1. General chair, scientific chair*

##### 10.1.1.2. Member of the organizing committees

- Cécile Germain, DataScience@LHC Workshop http://indico.cern.ch/event/395374/, 9-13 November 2015, CERN
- Marc Schoenauer, Steering Committee, Parallel Problem Solving from Nature (PPSN); Steering Committee, Learning and Intelligent OptimizatioN (LION).
- Michele Sebag, Steering Committee, Eur. Conf. on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD).

### 10.1.2. Scientific events selection

##### 10.1.2.1. Chair of conference program committees

##### 10.1.2.2. Member of the conference program committees

All TAO members are members of the Program Committees of the main conferences in the fields of Machine Learning, Evolutionary Computation, and Information Processing.

##### 10.1.2.3. Reviewer

All TAO member review papers for the most prestigious journals in the fields of Machine Learning and Evolutionary Computation.

### 10.1.3. Journal

##### 10.1.3.1. Member of the editorial boards

- Anne Auger, member of Editorial Board, *Evolutionary Computation Journal*, MIT Press.
- Gregory Grefenstette, member of Editorial Board, *Journal of Natural Language Engineering*, Cambridge University Press.
- Nikolaus Hansen, member of Editorial Board, *Evolutionary Computation Journal*, MIT Press.
- Marc Schoenauer, member of Advisory Board, *Evolutionary Computation Journal*, MIT Press; member of Editorial Board, *Genetic Programming and Evolutionary Machines*, Springer Verlag; action editor, *Journal of Machine Learning Research*(JMLR).
- Michèle Sebag: Editorial Board, Machine Learning, Springer Verlag.
- Olivier Teytaud, action editor, *Journal of Machine Learning Research* (JMLR).

##### 10.1.3.2. Reviewer - Reviewing activities

All members of the team reviewed numerous articles for international conferences and journals.

### 10.1.4. Invited talks

- Aurélien Decelle, workshop of Machine Learning at the Yukawa institute.
- Cyril Furtlehner, International Workshop on Innovative Algorithms for Big Data (Kyoto).
- Gregory Grefenstette, CLEF 2015, Toulouse, Sept. 8, 2015
- Gregory Grefenstette, Journée Recherche - les Big Data, Ecole Contrale Marseille, Nov 18, 2015
- Philippe Caillou, Workshop PICS at Hanoi, May 2015.
- Odalric Maillard, Journées YSP at IHP, Paris, Jan. 31, 2015.
- Odalric Maillard, Workshop CIMI at Toulouse, November 2015.
- Yann Ollivier, Deep Learning Workshop at ICML 2015.
- Yann Ollivier, Deep Learning Scoping Workshop of the Alan Turing Institute (Edinburgh).
- Yann Ollivier, American Mathematical Society - Mathematical Association of America Joint Math Meetings (San Antonio).

- Michèle Sebag, DATA 2015
- Michèle Sebag, AutoML Wshop, @ ICML 2015.
- Michèle Sebag, Constructive Machine Learning Wshop, @ICML 2015
- Michèle Sebag, Intelligence Artificielle et Interfaces Hommes-Machines 2015
- Michèle Sebag and Marc Schoenauer, Big Data Business Convention, *Seeded Influencer Ranking*, Dec. 2015.
- Marc Schoenauer, *Numbers don't Count*, Conférence en l'honneur de Laurence Halpern, January 2015.
- Marc Schoenauer*Rank-SVM for flexible optimization*, ISEGI seminar, Lisbon, Jan. 2015.
- Marc Schoenauer and Michèle Sebag, *More is different: When, and What For?*, JS Inria, Nancy, June 2015.
- Marc Schoenauer, DataScience@LHC, Nov. 9-13, 2015.

### *10.1.5. Leadership within the scientific community*

- Michèle Sebag, Chair of Steering Committee, ECML-PKDD.
- Marc Schoenauer, elected Chair of ACM-SIGEVO (Special Interest Group on Evolutionary Computation), July 2015 (2-years term).
- Marc Schoenauer, founding President of SPECIES (Society for the Promotion of Evolutionary Computation In Europe and Surroundings), that organizes the yearly series of conferences *EvoStar*.

### *10.1.6. Scientific expertise*

- Cécile Germain, evaluator for the H2020 calls: *ICT-2015 Topic ICT-16 – Big Data - research*
- Gregory Grefenstette, evaluator for FU21, Cap Digital, Digiteo (IASI) review board
- Gregory Grefenstette, project reviewer for the H2020 (SemCare): *Information and Communication Technologies ICT*
- Michele Sebag, evaluator for the Swedish Foundation for Strategic Research; CNRS PEPS Fascido, FNRS Belgium
- Michele Sebag, president of a CNRS hiring jury (EHESS, research engineer); member of hiring juries for Inria, U. Paris-1, U. Lille.

### *10.1.7. Research administration*

- Cécile Germain, elected member of the Scientific Council and of its board. University officer for scientific computing. Deputy head of the computer science departement, in charge of research.
- Philippe Caillou, elected member of the Scientific Council and Academic Council.
- Marc Schoenauer, *Délégué Scientifique* (aka VP-Research) for the Inria Saclay Île-de-France branch.
- Michele Sebag, deputy director of LRI, CNRS UMR 8623; elected member of the Research Council of Univ. Paris-Saclay; head of the Research Committee of Labex Digicosme; member of the Board of the Lidexes *Institut de la Société Numérique* and *Center for Data Science*.

## 10.2. Teaching - Supervision - Juries

### *10.2.1. Teaching*

Licence : Philippe Caillou, Computer Science for students in Accounting and Management, 192h, L1, IUT Sceaux, Univ. Paris Sud

Licence : Aurélien Decelle, Computer Architecture, 50h, L3, Univ. Paris-Sud

Licence : Aurélien Decelle, Machine Learning and Artificial Life, 27h, L2, Univ. Paris-Sud

Licence : Aurélien Decelle, Databases, 30h, L2, Univ. Paris-Sud

Master : Aurélien Decelle, Machine Learning, 6h, M2 Recherche, Univ. Paris-Sud

Master : Anne Auger, Optimisation, 12h, M2 Recherche, U. Paris-sud.

Master : Philippe Caillou, Multi-Agents Systems, 27h, M2 Recherche, U. Paris-sud.

Master : Philippe Caillou, Multi-Agent Based Simulation, 3h, M2 Recherche, U. Paris-Dauphine.

Master : Guillaume Charpiat (et Gaétan Marceau), Machine Learning, 18h, M2 Recherche, Centrale-Supelec.

Master : Odalric-Ambrym Maillard, Machine Learning, 6h, M2 Recherche, Univ. Paris-Sud

Master : Yann Ollivier, Deep learning, 4h, M2 Recherche, Telecom/Polytech.

Master : Michèle Sebag, Machine Learning, 12h; Deep Learning, 6h; Reinforcement Learning, 6h; M2 Recherche, U. Paris-sud.

**E-learning**

Mooc, SPOC, etc. : Enseignant ou auteur, titre du cours, durée en nombre de semaine, plate-forme, établissement porteur du cours, public ciblé, formation initiale ou continue, nombre d'inscrits

Pedagogical resources : enseignant, titre, type (video, pdf, exercice, ou autre), niveau, url

### *10.2.2. Supervision*

PhD: Yoann ISAAC, *Représentations redondantes pour les signaux d'électroencéphalographie*, 29/05/2015, Jamal Atif and Michèl`e Sebag.

Alexandre Chotard, *Markov Chain Analysis of Evolution Strategies*, Université Paris-Sud, 24/09/2015, Anne Auger and Nikolaus Hansen

PhD: Nicolas Galichet, Contributions to Multi-Armed Bandits: Risk-Awareness and Sub-Sampling for Linear Contextual Bandits, Université Paris-Sud, 28/09/2015, Michèle Sebag.

PhD: Guohua Zhang, *Exploratory Robotic Controllers: An Evolution and Information Theory Driven Approach*, Université Paris-Sud, 24/09/2015, Michèle Sebag and Jingzhong Zhang.

PhD: Jialin Liu, *Portfolio methods in uncertain contexts*, Université Paris-Sud, 11/12/2015, Olivier Teytaud and Marc Schoenauer.

PhD in progress: Ouassim AIT ELHARA, *Large-scale optimization and Evolution Strategies*, 1/09/2012, Anne Auger and Nikolaus Hansen.

PhD in progress: Sandra ASTETE-MORALES, *Noisy optimization, with applications to power systems*, 1/09/2013, Olivier Teytaud.

PhD in progress: Asma ATAMNA, *Evolution Strategies and Constrained Optimization*, 1/10/2013, Anne Auger and Nikolaus Hansen.

PhD in progress: Jérémy BENSADON, *Applications of Information Theory to Statistical Learning*, Yann Ollivier, defense planned for 02/02/2016.

PhD in progress: Nacim BELKHIR, *On-line parameter tuning*, 1/5/2014, Marc Schoenauer and Johann Dréo (Thalès), CIFRE Thalès.

PhD in progress: Vincent BERTHIER, *Large scale parallel optimization, with application to power systems*, 1/09/2013, Michèle Sebag et Olivier Teytaud.

PhD in progress: Marie-Liesse CAUWET, *Artificial intelligence with uncertainties, application to power systems*, 1/09/2013, Olivier Teytaud.

PhD in progress: François GONARD, *Automatic optimization algorithm selection and configuration*, 1/10/2014, Marc Schoenauer and Michèle Sebag, thèse IRT SystemX.

PhD in progress: Pierre-Yves MASSÉ, *Gradient Methods for Statistical Learning*, 1/10/2014, Yann Ollivier

PhD in progress: Sourava MISHRA, *AutoML: An empirical approach to Machine Learning*, 1/10/2014, Balazs Kégl and Michèle Sebag

PhD in progress: Thomas SCHMITT, *A Collaborative Filtering Approach to Matching Job Openings and Job Seekers*, 1/11/2014, Michèle Sebag.

PhD in progress : Emmanuel Maggiori, *Large-Scale Remote Sensing Image Classification*, 1/1/2015, Guillaume Charpiat (with Yuliya Tarabalka and Pierre Alliez, Inria Sophia-Antipolis)

PhD in progress: Mehdi Cherti *Learning to discover: supervised discrimination and unsupervised representation learning with applications in particle physics.* 01/10/2014, Balazs Kegl and Cécile Germain.

PhD in progress: Karima Rafes *Gestion et sécurité des données personnelles dans le web des objets*. 01/10/2014, Serge Abiteboul and Cécile Germain.

#### *10.2.3. Juries*

All senior TAO members participated to many PhD or HDR committees.

## 10.3. Popularization

**Odalric-Ambrym Maillard** May, participation to day "Petits débrouillards", during which secondary schools students come and visit a lab.

**Aurélien Decelle** May, participation to day "Petits débrouillards", during which secondary schools students come and visit a lab.

**Yann Ollivier** co-organizes the European Union Contest for Young Scientists (science fair for high school students from 30+ countries organized by the European Commission).

**Michele Sebag** , participation to IHP debate on AI and society; panel at Technion Research Day.

# 11. Bibliography

## Publications of the year

### Doctoral Dissertations and Habilitation Theses

[1] A. CHOTARD. *Markov chain Analysis of Evolution Strategies*, Université Paris Sud - Paris XI, September 2015, https://tel.archives-ouvertes.fr/tel-01252128

[2] N. GALICHET. *Contributions to Multi-Armed Bandits: Risk-Awareness and Sub-Sampling for Linear Contextual Bandits*, Université Paris-Sud, September 2015, https://hal.inria.fr/tel-01251046

[3] J. LIU. *Portfolio methods in uncertain contexts*, Inria, December 2015, https://hal.inria.fr/tel-01250552

[4] G. ZHANG. *Exploratory Robotic Controllers: An Evolution and Information Theory Driven Approach*, Université Paris-Sud, September 2015, https://hal.inria.fr/tel-01251238

### Articles in International Peer-Reviewed Journals

[5] Y. AKIMOTO, S. ASTETE-MORALES, O. TEYTAUD. *Analysis of runtime of optimization algorithms for noisy functions over discrete codomains*, in "Journal of Theoretical Computer Science (TCS)", 2015, vol. 605, 42:50 p. [*DOI :* 10.1016/J.TCS.2015.04.008], https://hal.inria.fr/hal-01194556

[6] S. Astete-Morales, M.-L. Cauwet, J. Liu, O. Teytaud. *Simple and Cumulative Regret for Continuous Noisy Optimization*, in "Journal of Theoretical Computer Science (TCS)", 2015, vol. 617, pp. 12–27, https://hal.inria.fr/hal-01194564

[7] R. Bardenet, O.-A. Maillard. *Concentration inequalities for sampling without replacement*, in "Bernoulli", 2015, vol. 21, n$^{o}$ 3, pp. 1361-1385 [*DOI :* 10.3150/14-BEJ605], https://hal.archives-ouvertes.fr/hal-01216652

[8] M.-L. Cauwet, J. Liu, R. Baptiste, O. Teytaud. *Algorithm Portfolios for Noisy Optimization*, in "Annals of Mathematics and Artificial Intelligence", November 2015, pp. 1-30 [*DOI :* 10.1007/s10472-015-9486-2], https://hal.archives-ouvertes.fr/hal-01223113

[9] G. Charpiat, G. Nardi, G. Peyré, F.-X. Vialard. *Piecewise rigid curve deformation via a Finsler steepest descent*, in "Interfaces and Free Boundaries", December 2015, https://hal.archives-ouvertes.fr/hal-00849885

[10] A. Chotard, A. Auger, N. Hansen. *Markov Chain Analysis of Cumulative Step-size Adaptation on a Linear Constrained Problem*, in "Evolutionary Computation", 2015, vol. 23, n$^{o}$ 4, pp. 611-640 [*DOI :* 10.1109/4235.873238], https://hal.inria.fr/hal-01215727

[11] A. Decelle, P. Zhang. *Inference of the sparse kinetic Ising model using the decimation method*, in "Physical Review E : Statistical, Nonlinear, and Soft Matter Physics", May 2015, vol. 91, n$^{o}$ 5 [*DOI :* 10.1103/PhysRevE.91.052136], https://hal.archives-ouvertes.fr/hal-01250830

[12] D. Feng, C. Germain. *Fault monitoring with sequential matrix factorization*, in "ACM Transactions on Autonomous and Adaptive Systems", 2015, vol. 10, n$^{o}$ 3, pp. 20:1–20:25 [*DOI :* 10.1145/2797141], https://hal.inria.fr/hal-01176013

[13] V. Martin, J.-M. Lasgouttes, C. Furtlehner. *Latent binary MRF for online reconstruction of large scale systems*, in "Annals of Mathematics and Artificial Intelligence", 2015, pp. 1-32 [*DOI :* 10.1007/s10472-015-9470-x], https://hal.inria.fr/hal-01186220

[14] Y. Ollivier. *Riemannian metrics for neural networks I: Feedforward networks*, in "Information and Inference", 2015, vol. 4, n$^{o}$ 2, pp. 108–153 [*DOI :* 10.1093/imaiai/iav006], https://hal.archives-ouvertes.fr/hal-00857982

[15] Y. Ollivier. *Riemannian metrics for neural networks II: Recurrent networks and learning symbolic data sequences*, in "Information and Inference", 2015, vol. 4, n$^{o}$ 2, pp. 154–193 [*DOI :* 10.1093/imaiai/iav007], https://hal.archives-ouvertes.fr/hal-00857980

[16] Y. Shogo, M. Ohzeki, A. Decelle. *Detection of Cheating by Decimation Algorithm*, in "Journal of the Physical Society of Japan", January 2015, vol. 84, 024801 p. [*DOI :* 10.7566/JPSJ.84.024801], https://hal.archives-ouvertes.fr/hal-01105415

[17] N. Sokolovska, O. Teytaud, S. Rizkalla, K. Clément, J.-D. Zucker. *Sparse Zero-Sum Games as Stable Functional Feature Selection*, in "PLoS ONE", September 2015, vol. 10, n$^{o}$ 9, e0134683 p. , MicroObese consortium [*DOI :* 10.1371/journal.pone.0134683], http://hal.upmc.fr/hal-01223887

## Invited Conferences

[18] F. HUTTER, B. KÉGL, R. CARUANA, I. GUYON, H. LAROCHELLE, E. VIEGAS. *Automatic Machine Learning (AutoML)*, in "ICML 2015 Workshop on Resource-Efficient Machine Learning, 32nd International Conference on Machine Learning", Lille, France, July 2015, http://hal.in2p3.fr/in2p3-01171463

## International Conferences with Proceedings

[19] E. ALBERTS, G. CHARPIAT, Y. TARABALKA, T. HUBER, M.-A. WEBER, J. BAUER, C. ZIMMER, B. H. MENZE. *A Nonparametric model for Brain Tumor Segmentation and Volumetry in Longitudinal MR Sequences*, in "MICCAI Brain Lesion Workshop", Munich, Germany, October 2015, https://hal.inria.fr/hal-01205916

[20] S. ASTETE-MORALES, M.-L. CAUWET, O. TEYTAUD. *Evolution Strategies with Additive Noise: A Convergence Rate Lower Bound*, in "Foundations of Genetic Algorithms", Aberythswyth, United Kingdom, Foundations of Genetic Algorithms, ACM, 2015, pp. 76–84, https://hal.inria.fr/hal-01077625

[21] N. BELKHIR, J. DRÉO, P. SAVÉANT, M. SCHOENAUER. *Parameter Setting for Multicore CMA-ES with Large Populations*, in "Artificial Evolution (EA 2015)", Lyon, France, S. BONNEVAY, P. LEGRAND, N. MONTMARCHÉ, E. LUTTON, M. SCHOENAUER (editors), Springer Verlag, October 2015, forthcoming, https://hal.inria.fr/hal-01236025

[22] V. BERTHIER. *Comparing optimizers on a unit commitment problem*, in "Artificial Evolution (EA2015)", Lyon, France, S. BONNEVAY, P. LEGRAND, N. MONTMARCHÉ, E. LUTTON, M. SCHOENAUER (editors), Springer Verlag, October 2015, to appear, https://hal.inria.fr/hal-01215804

[23] V. BERTHIER. *Experiments on the CEC 2015 expensive optimization testbed*, in "CEC 2015 - IEEE Congress on Evolutionary Computation", Sendai, Japan, IEEE Press, May 2015, pp. 1059 - 1066 [*DOI :* 10.1109/CEC.2015.7257007], https://hal.inria.fr/hal-01215805

[24] V. BERTHIER. *Progressive Differential Evolution on Clustering Real World Problems*, in "Artifical Evolution 2015", Lyon, France, EA 2015 - International Conference on Artificial Evolution., Springer, October 2015, https://hal.inria.fr/hal-01215803

[25] V. BERTHIER, A. COUËTOUX, O. TEYTAUD. *Combining policies: the best of human expertise and neurocontrol*, in "Artificial Evolution 2015", Lyon, France, Proceedings of Artificial Evolution 2015 (EA2015), 2015, forthcoming, https://hal.inria.fr/hal-01194516

[26] V. BERTHIER, O. TEYTAUD. *On the codimension of the set of optima: large scale optimisation with few relevant variables*, in "Artificial Evolution 2015", Lyon, France, Proceedings of Artificial Evolution 2015 (EA2015), 2015, forthcoming, https://hal.inria.fr/hal-01194519

[27] V. BERTHIER, O. TEYTAUD. *Sieves method in fuzzy control: logarithmically increase the number of rules*, in "IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)", Istanbul, Turkey, IEEE Press, August 2015, pp. 1 - 9, https://hal.inria.fr/hal-01215806

[28] D. BROCKHOFF, T.-D. TRAN, N. HANSEN. *Benchmarking Numerical Multiobjective Optimizers Revisited*, in "Genetic and Evolutionary Computation Conference (GECCO 2015)", Madrid, Spain, A. ESPARCIA,

S. SILVA (editors), July 2015, pp. 639-646 [*DOI :* 10.1145/2739480.2754777], https://hal.inria.fr/hal-01146741

[29] P. CAILLOU, B. GAUDOU, A. GRIGNARD, C. Q. TRUONG, P. TAILLANDIER. *A Simple-to-use BDI architecture for Agent-based Modeling and Simulation*, in "The Eleventh Conference of the European Social Simulation Association (ESSA 2015)", Groningen, Netherlands, September 2015, https://hal.inria.fr/hal-01216165

[30] M.-L. CAUWET, O. TEYTAUD, S.-Y. CHIU, K.-M. LIN, S.-J. YEN, D. L. SAINT-PIERRE, F. TEYTAUD. *Parallel Evolutionary Algorithms Performing Pairwise Comparisons*, in "Foundations of Genetic Algorithms", Aberythswyth, United Kingdom, J. HE, T. JANSEN, G. OCHOA, C. ZARGES (editors), Foundations of Genetic Algorithms, ACM, 2015, pp. 99-113, https://hal.inria.fr/hal-01077626

[31] M.-L. CAUWET, O. TEYTAUD, H.-M. LIANG, S.-J. YEN, H.-H. LIN, I.-C. WU, T. CAZENAVE, A. SAFFIDINE. *Depth, balancing, and limits of the Elo model*, in "IEEE Conference on Computational Intelligence and Games 2015", Tainan, Taiwan, I, August 2015, https://hal.archives-ouvertes.fr/hal-01223116

[32] T. CAZENAVE, J. LIU, O. TEYTAUD. *The Rectangular Seeds of Domineering, Atari-Go and Breakthrough*, in "2015 IEEE Conference on Computational Intelligence and Games (CIG)", Tainan, Taiwan, August 2015, pp. 530 - 531 [*DOI :* 10.1109/CIG.2015.7317904], https://hal.inria.fr/hal-01245531

[33] S.-Y. CHIU, C.-N. LIN, J. LIU, T.-C. SU, F. TEYTAUD, O. TEYTAUD, S.-J. YEN. *Differential Evolution for Strongly Noisy Optimization: Use $1.01^n$ Resamplings at Iteration n and Reach the -1/2 Slope*, in "2015 IEEE Congress on Evolutionary Computation (IEEE CEC)", Sendai, Japan, May 2015, https://hal.inria.fr/hal-01120892

[34] S.-Y. CHIU, C.-N. LIN, J. LIU, T.-C. SU, F. TEYTAUD, O. TEYTAUD, S.-J. YEN. *Differential evolution for strongly noisy optimization: Use 1.01n resamplings at iteration n and reach the - 1/2 slope*, in "2015 IEEE Congress on Evolutionary Computation (IEEE CEC 2015)", Sendai, Japan, May 2015, pp. 338 - 345 [*DOI :* 10.1109/CEC.2015.7256911], https://hal.inria.fr/hal-01245526

[35] J.-J. CHRISTOPHE, J. DECOCK, J. LIU, O. TEYTAUD. *Variance Reduction in Population-Based Optimization: Application to Unit Commitment*, in "Artificial Evolution (EA2015)", Lyon, France, S. BONNEVAY, P. LEGRAND, N. MONTMARCHÉ, E. LUTTON, M. SCHOENAUER (editors), 2015, forthcoming p. , https://hal.inria.fr/hal-01194510

[36] J. DECOCK, D. L. SAINT-PIERRE, O. TEYTAUD. *Evolutionary Cutting Planes*, in "Artificial Evolution (EA2015)", Lyon, France, S. BONNEVAY, P. LEGRAND, N. MONTMARCHÉ, E. LUTTON, M. SCHOENAUER (editors), 2015, forthcoming p. , https://hal.inria.fr/hal-01194540

[37] C. FAUR, P. CAILLOU, J.-C. MARTIN, C. CLAVEL. *A Socio-cognitive Approach to Personality: Machine-learned Game Strategies as Cues of Regulatory Focus*, in "Affective Computing and Intelligent Interaction (ACII 2015)", Xi'an, China, September 2015, https://hal.inria.fr/hal-01216540

[38] R. FFRANCON, M. SCHOENAUER. *Greedy Semantic Local Search for Small Solutions*, in "Companion Proceedings (workshops) of the Genetic and Evolutionary Computation COnference", TAO, Inria Saclay, France, July 2015, pp. 1293-1300, https://hal.inria.fr/hal-01169075

[39] *Best Paper*
R. FFRANCON, M. SCHOENAUER. *Memetic Semantic Genetic Programming*, in "Genetic and Evolutionary Computation COnference (GECCO 2015)", Madrid, Spain, A. ESPARCIA, S. SILVA (editors), ACM, July 2015, pp. 1023-1030, Best Paper Award in the GP track - see http://www.sigevo.org/gecco-2015/papers.html, https://hal.inria.fr/hal-01169074.

[40] E. GALVÁN-LÓPEZ, M. SCHOENAUER, C. PATSAKIS. *Design of an Autonomous Intelligent Demand-Side Management System by using Electric Vehicles as Mobile Energy Storage Units by Means of Evolutionary Algorithms*, in "7th International Joint Conference on Computational Intelligence (IJCCI 2015)", Lisbon, Portugal, A. C. ROSA, J. J. M. GUERVÓS, A. DOURADO, J. M. CADENAS, K. MADANI, A. E. RUANO, J. FILIPE (editors), SciTePress, November 2015, vol. 1, pp. 106–115 [*DOI :* 10.5220/0005607401060115], https://hal.inria.fr/hal-01254912

[41] G. GREFENSTETTE, L. MUCHEMI. *Extracting Hierarchical Topic Models from the Web for Improving Digital Archive Access*, in "Expert Workshop on Topic Models and Corpus Analysis", Dublin, Ireland, DARIAH Text & Data Analytics Working Group, December 2015, https://hal.inria.fr/hal-01251326

[42] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT. *Improved Partition Trees for Multi-Class Segmentation of Remote Sensing Images*, in "2015 IEEE International Geoscience and Remote Sensing Symposium - IGARSS 2015", Milan, Italy, IEEE, July 2015, https://hal.inria.fr/hal-01182772

[43] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT. *Optimizing Partition Trees for Multi-Object Segmentation with Shape Prior*, in "26th British Machine Vision Conference", Swansea, United Kingdom, September 2015, https://hal.inria.fr/hal-01182776

[44] L. MALAGÒ, G. PISTONE. *Information Geometry of Gaussian Distributions in View of Stochastic Optimization*, in "Foundations of Genetic Algorithms XIII", Aberystwyth, United Kingdom, ACM, January 2015, pp. 150-162, https://hal.inria.fr/hal-01108986

[45] W. MEI-HUI, C.-S. WANG, C.-S. LEE, O. TEYTAUD, J. LIU, S.-W. LIN, P.-H. HUNG. *Item response theory with fuzzy markup language for parameter estimation and validation*, in "2015 IEEE Conference on Fuzzy Systems (FUZZ-IEEE)", Istanbul, Turkey, IEEE Press, August 2015, pp. 1 - 7 [*DOI :* 10.1109/FUZZ-IEEE.2015.7337884], https://hal.inria.fr/hal-01245695

[46] Y. OLLIVIER. *Laplace's rule of succession in information geometry*, in "Geometric science of information", Palaiseau, France, F. NIELSEN, F. BARBARESCRO (editors), Geometric science of information, LNCS 9389, Springer Verlag, October 2015, pp. 311-319, https://hal.archives-ouvertes.fr/hal-01228952

[47] A. QUEMY, M. SCHOENAUER. *True Pareto Fronts for Multi-Objective AI Planning Instances*, in "European Conference on Combinatorial Optimization - EvoCOP", Copenhague, Denmark, F. CHICANO, G. OCHOA (editors), LNCS 9026, Springer Verlag, April 2015, pp. 197-208, https://hal.archives-ouvertes.fr/hal-01109777

[48] A. QUEMY, M. SCHOENAUER, V. VIDAL, J. DRÉO, P. SAVÉANT. *Solving Large MultiZenoTravel Benchmarks with Divide-and-Evolve*, in "Learning and Intelligent OptimizatioN - LION 9", Lille, France, C. DHAENENS, L. JOURDAN, M.-E. MARMION (editors), LNCS 8994, Springer Verlag, January 2015, pp. 262-267, https://hal.archives-ouvertes.fr/hal-01109776

[49] K. RAFES, J. NAUROY, C. GERMAIN. *Certifying the interoperability of RDF database systems*, in "LDQ 2015 - 2nd Workshop on Linked Data Quality", Portorož, Slovenia, 2nd Workshop on Linked Data Quality (LDQ), Springer, May 2015, n$^{o}$ 2, https://hal.inria.fr/hal-01147765

[50] A. SAFFIDINE, O. TEYTAUD, S.-J. YEN. *Go Complexities*, in "Advances in Computer Games", Leiden, Netherlands, Advances in Computer Games, 2015 [*DOI :* 10.1007/978-3-319-27992-3_8], https://hal.inria.fr/hal-01256660

[51] D. L. ST-PIERRE, J. LIU, O. TEYTAUD. *Nash Reweighting of Monte Carlo Simulations: Tsumego*, in "2015 IEEE Congress on Evolutionary Computation (IEEE CEC 2015)", Sendai, Japan, IEEE Press, May 2015, pp. 1458 - 1465 [*DOI :* 10.1109/CEC.2015.7257060], https://hal.inria.fr/hal-01245520

[52] O. TEYTAUD. *Quasi-random numbers improve the CMA-ES on the BBOB testbed*, in "Artificial Evolution (EA2015)", Lyon, France, S. BONNEVAY, P. LEGRAND, N. MONTMARCHÉ, E. LUTTON, M. SCHOENAUER (editors), Springer Verlag, 2015, 13 p. , to appear, https://hal.inria.fr/hal-01194542

### National Conferences with Proceedings

[53] I. BRIGUI-CHTIOUI, P. CAILLOU, S. PINSON. *Approche Anytime pour l'ajustement de l'incrément dans les enchères multicritères automatisées*, in "23es Journées Francophones sur les Systèmes Multi-Agents (JFSMA'15)", Rennes, France, L. VERCOUTER, G. PICARD (editors), Cépaduès, June 2015, pp. 153-162, https://hal.archives-ouvertes.fr/hal-01163019

[54] K. RAFES, C. GERMAIN. *A platform for scientific data sharing*, in "BDA2015 - Bases de Données Avancées", Île de Porquerolles, France, September 2015, https://hal.inria.fr/hal-01168496

### Conferences without Proceedings

[55] G. GREFENSTETTE. *InriaSAC: Simple Hypernym Extraction Methods*, in "SemEval 2015", Denver, United States, June 2015, https://hal.inria.fr/hal-01112844

[56] D. ROUSSEAU, G. COWAN, C. A. BOURDARIOS, B. KÉGL, C. GERMAIN-RENAUD, I. GUYON. *The ATLAS Higgs Machine Learning Challenge*, in "21st International Conference on Computing in High Energy and Nuclear Physics – CHEP2015", Okinawa, Japan, April 2015, Proceedings to appear in Journal of Physics : Conference Series, http://hal.in2p3.fr/in2p3-01141742

[57] P. TAILLANDIER, P. CAILLOU, B. GAUDOU, A. GRIGNARD, B. MATHIEU. *Providing Modelers with a Simple Yet Rich Tool to Define Cognitive Agents with the GAMA Platform*, in "Conference on Complex Systems (CCS 2015)", Tempe, United States, September 2015, https://hal.inria.fr/hal-01216183

### Scientific Books (or Scientific Book chapters)

[58] *NIPS 2014 Workshop on High-energy Physics and Machine Learning*, JMLR Workshop and Conference Proceedings, June 2015, vol. 42, 134 p. , https://hal.inria.fr/hal-01208543

[59] A. BANOS, P. CAILLOU, B. GAUDOU, N. MARILLEAU. *Agent-Based Model Exploration*, in "Agent-Based Spatial Simulation with NetLogo", Volume 1: Introduction and Bases, Elsevier, September 2015, 57 p. , https://hal.inria.fr/hal-01216192

[60] A. BANOS, P. CAILLOU, B. GAUDOU, N. MARILLEAU. *Exploration de modèles agent*, in "Simulation spatiale à base d'agents avec NetLogo 1", ISTE, February 2015, 49 p. , https://hal.inria.fr/hal-01216207

[61] N. HANSEN, D. V. ARNOLD, A. AUGER. *Evolution Strategies*, in "Handbook of Computational Intelligence", J. KACPRZYK, W. PEDRYCZ (editors), 2015, https://hal.inria.fr/hal-01155533

## Research Reports

[62] R. AKROUR, B. MAYEUR, M. SEBAG. *Direct Value Learning: Reinforcement Learning and Anti-Imitation*, Inria ; CNRS ; Université Paris-Sud 11, December 2015, n<sup>o</sup> RR-8836, 18 p. , https://hal.inria.fr/hal-01249377

[63] C. FURTLEHNER, A. DECELLE. *Cycle-based Cluster Variational Method for Direct and Inverse Inference*, Inria Saclay, équipe TAO, October 2015, n<sup>o</sup> RR-8788, 40 p p. , https://hal.inria.fr/hal-01214155

## Other Publications

[64] S. ASTETE-MORALES, M.-L. CAUWET, O. TEYTAUD. *Criteria and Convergence Rates in Noisy Optimization*, July 2015, Genetic and Evolutionary Computation Conference (GECCO 2015), Poster [*DOI :* 10.1145/2739482.2764722], https://hal.inria.fr/hal-01217128

[65] A. AUGER, N. HANSEN. *Linear Convergence of Comparison-based Step-size Adaptive Randomized Search via Stability of Markov Chains*, December 2015, working paper or preprint, https://hal.inria.fr/hal-00877160

[66] R. BARDENET, O.-A. MAILLARD. *A note on replacing uniform subsampling by random projections in MCMC for linear regression of tall datasets*, December 2015, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01248841

[67] A. CHOTARD, A. AUGER. *Verifiable Conditions for Irreducibility, Aperiodicity and T-chain Property of a General Markov Chain*, October 2015, working paper or preprint, https://hal.inria.fr/hal-01222222

[68] A. DECELLE, F. RICCI-TERSENGHI. *Solving the inverse Ising problem by mean-field methods in a clustered phase space with many states*, January 2015, 6 pages, 5 figures, https://hal.archives-ouvertes.fr/hal-01250824

[69] A. DECELLE, F. RICCI-TERSENGHI, P. ZHANG. *Data quality for the inverse Ising problem*, November 2015, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01250822

[70] C. GERMAIN, J. NAUROY, K. RAFES. *The Grid Observatory 3.0 - Towards reproducible research and open collaborations using semantic technologies*, May 2015, EGI Community Forum 2014, https://hal.inria.fr/hal-01104235

[71] G. GREFENSTETTE, K. RAFES. *Transforming Wikipedia into a Search Engine for Local Experts*, November 2015, working paper or preprint, https://hal.inria.fr/hal-01224114

[72] P.-Y. MASSÉ, Y. OLLIVIER. *Speed learning on the fly*, November 2015, preprint, https://hal.archives-ouvertes.fr/hal-01228955

[73] Y. OLLIVIER, G. CHARPIAT. *Training recurrent networks online without backtracking*, 2015, Preliminary version, https://hal.archives-ouvertes.fr/hal-01228954