# Rémi EYRAUD, PhD

Laboratoire d'Informatique et Systèmes
Centre de Mathématiques et d'Informatique
39 rue Joliot-Curie
13453 Marseille Cedex 13
remi.eyraud@gmail.com
http://pageperso.lis-lab.fr/~remi.eyraud/


## Current Situation

Ph.D in Computer Science

Maitre de Conférences at the Aix-Marseilles University (France) – tenure

Chair of the steering committee of the International Community in Grammatical Inference

## Academic Degrees

**Doctoral Dissertation,** "Grammatical Inference of Context-Free Languages", *University Jean Monnet of Saint-Étienne, 2*006. Under the co-supervision of Pr. Colin de la Higuera (University of Nantes) and Pr. Jean-Christophe Janodet (University of Evry).

**MSc. in Computer Science,** *Ecole Nationale des Mines de Saint-Etienne*, 2003. Concentration: multi-agent systems, machine learning, grammatical inference. Final thesis: "Learning automata with a splitting-state process". First year thesis: "Simulations of social networks using cellular automata".

**Licence in Computer Science,** *"University Jean-Monnet, Saint-Etienne"*, 2001. Major in Computer Science, minor in Mathematics.

## Publications

All publications can be found here.

### Book Chapter

"Efficiency in the identification in the limit paradigm", Rémi Eyraud, Jeffrey Heinz, Ryo Yoshinaka, in *Topics in Grammatical Inference*, Springer, 2016

### International Journals:

- "Designing and Learning Substitutable Plane Graph Grammars", Rémi Eyraud, Jean-Christophe Janodet, Tim Oates, Frédéric Papadopoulos, ***Fundamenta Informaticae***, 2016
- "Learning Strictly Local Subsequential Functions", Jane Chandlee, Rémi Eyraud, Jeffrey Heinz, ***Transactions of the Association for Computational Linguistics***, 2014.
- "PAutomaC: a probabilistic automata and hidden Markov models learning competition", Sicco Verwer, Rémi Eyraud, Colin de la Higuera, ***Machine Learning Journal***, Springer, 2014
- "Using Contextual Representations to Efficiently Learn Context-Free Languages", Alexander Clark, Rémi Eyraud, Amaury Habrard, ***Journal of Machine Learning Research***, 2010
- "Polynomial Identification of Substitutable Context-Free Languages", Alexander Clark, Rémi Eyraud, ***Journal of Machine Learning Research***, 2007
- "LARS : a Learning Algorithm for Rewriting Systems", Rémi Eyraud, Colin de la Higuera, Jean-Christophe Janodet, ***Machine Learning Journal***, Springer, 2007

## Peer-Reviewed International Conferences:

- "Explaining Black Boxes on Sequential Data using Weighted Automata", Stéphane Ayache, Rémi Eyraud, and Noé Goudian, proceedings of the 14th ***International Conference on Grammatical Inference,*** PMLR, vol. 88, *to appear,* 2018
- "Sp2Learn: A Toolbox for the spectral learning of weighted automata", Denis Arrivault, Dominique Benielli, François Denis and Rémi Eyraud, proceedings of the 13th ***International Conference on Grammatical Inference,*** pages 3 to 17, JMLR W&C, vol. 57, 2016
- "Output Strictly Local Functions", Jane Chandlee, Rémi Eyraud, Jeffrey Heinz, proceedings of the 14th Meeting on the ***Mathematics of Language***, pages 112 to 125, Association for Computational Linguistics, 2015
- "Very efficient learning of structured classes of subsequential functions from positive data", Adam Jardine, Jane Chandlee, Rémi Eyraud, Jeffrey Heinz , proceedings of the 12th ***International Conference on Grammatical Inference,*** pages 94 to 108, JMLR W&C, vol. 34, 2014
- "Learning Substitutable Binary Plane Graph Grammars", Rémi Eyraud, Jean-Christophe Janodet, Tim Oates, proceedings of the 11th ***International Conference on Grammatical Inference***, pages 114 to 128, JMLR W&C, vol. 21, 2012
- "A Polynomial Algorithm for the Inference of Context-Free Languages", Alexander Clark, Rémi Eyraud, Amaury Habrard, proceedings of the 9th ***International Colloquium on Grammatical Inference***, pages 29 to 42, LNAI 5278, Springer, 2008
- "Learning Auxiliary Fronting with Grammatical Inference", Alexander Clark, Rémi Eyraud, proceedings of the 10th Conference on ***Computational Natural Language Learning***, pages 125 to 132, 2006
- "Identification in the Limit of Substitutable Context-Free Languages", Alexander Clark, Rémi Eyraud, proceedings of the 16ith International Conference on ***Algorithmic Learning Theory***, pages 283 to 296, LNAI 3734, Springer, 2005
- "Representing Languages by Learnable Rewriting Systems", Rémi Eyraud, Colin de la Higuera, Jean-Christophe Janodet, proceedings of the 7th ***International Colloquium on Grammatical Inference***, pages 139 to 150, LNAI 3264, Springer, 2004

## Short papers, posters, workshops:

- "Scikit-SpLearn: A Toolbox for the spectral learning of weighted automata compatible with scikit-learn", Denis Arrivault, Dominique Benielli, François Denis and Rémi Eyraud, proceedings of the ***Conference francophone en Apprentissage*** (French Machine Learning Conference)***,*** 2017
- "Results of the Sequence PredIction ChallengE (SPiCe): a Competition on Learning the Next Symbol in a Sequence ", Borja Balle, Rémi Eyraud, Franco M. Luque, Ariadna Quattoni, Sicco Verwer, proceedings of the 11th ***International Conference on Grammatical Inference***, pages 243 to 248, JMLR proceedings, 2016
- "New Polynomial Bounds for the Identification in the Limit Paradigm using Generative Grammars", Rémi Eyraud, Jeffrey Heinz, ***ICALP Satellite Workshop on Learning Theory and Complexity***, 2013, Riga, Latvia.
- "Results of the PAutomaC Probabilistic Automaton Learning Competition ", Sicco Verwer, Rémi Eyraud, Colin de la Higuera, proceedings of the 11th ***International Conference on Grammatical Inference***, pages 243 to 248, JMLR proceedings, vol. 21, 2012.
- "A note on contextual binary feature grammars", Alexander Clark, Rémi Eyraud, Amaury Habrard, ***EACL 2009 workshop on Computational Linguistic Aspects of Grammatical Inference***, Athens, Greece, 2009.
- "Spring School in Machine Learning. Teaching experiences", Cécile Capponi, François Denis, Rémi Eyraud, Amaury Habrard, Liva Ralaivola, ***PASCAL Workshop: Teaching Machine Learning***, Saint-Etienne, France, 2008.

- "Two methods to learn context-free languages", Rémi Eyraud, ***workshop ML4NLP*****,** Amsterdam, 2007.
- "Learning Auxiliary Fronting with Grammatical Inference", Alexander Clark, Rémi Eyraud, proceedings of the 28th Annual Conference of the ***Cognitive Science Society***, pages 1127 to 1132, Vancouver, Canada, 2006.
- "Deux Techniques d'Apprentissage de Langages Hors-Contextes", Rémi Eyraud, proceedings of the 8th ***Conférence francophone d'Apprentissage Automatique***, pages 186 and 187, Trégastel, France, 2006.

## Career

**September 2007 – present:** Maitre de conférences (Junior Professor – tenure), Aix-Marseille University, LIS UMR CNRS 7020, France

**March 2018:** Invited lecturer at the Stony Brook University, USA

**July 2014 – December 2014:** Invited researcher at the Delaware University, USA

**July 2012 – December 2013:** Sabbatical founded by the CNRS; Invited researcher at the University of Maryland, Baltimore County, USA

**January 2007 – September 2007:** non-permanent researcher at the University of Amsterdam, The Netherlands

**September 2006 – January 2007:** non-permanent full-time lecturer (ATER) at University of Saint-Etienne, France

**September 2003 – August 2006:** Ministry of Research founded Ph. D program with teaching duties (Allocataire-Moniteur)

## Research Activities

**Principal collaborations**

- ✓ Alexander Clark, King's College, London, UK (7 co-authored papers, long visit)
- ✓ Jeffrey Heinz, University of Delaware, USA (5 co-authored papers, long stay)
- ✓ Colin de la Higuera, Nantes University, France (4 co-authored papers, co-organised events)
- ✓ Jean-Christophe Janodet, Evry University, France (4 co-authored papers)
- ✓ Sicco Verwer, Delft University, The Netherlands (3 co-authored papers, co-organised international challenges)
- ✓ Tim Oates, University of Maryland, USA (2 co-authored papers, co-organised event)
- ✓ Jane Chandlee, Haverford College, USA (2 co-authored papers, on-going work)
- ✓ François Denis, Aix-Marseille University, France (2 co-authored papers, ongoing project)
- ✓ Amaury Habrard, Saint-Etienne University, France (2 co-authored papers)
- ✓ Borja Balle, Lancaster University, UK, (1 co-authored paper, co-organised events)
- ✓ Ryo Yoshinaka, Kyoto University, Japan (1 co-authored paper, short visits)

**Invited talks**

- ✓ Invited speaker at the *Journées annuelles de Vérification, GdR IM, Grenoble*, 2018. Title: Spectral Learning of Weighted Automata, from theory to a toolbox.
- ✓ Invited speaker at the *Linguistic departement of the Stony Brook University*, USA, 2018. Title: Recent advances in Grammatical Inference.
- ✓ Invited speaker at the *ELC Workshop on Learning Theory and Complexity*, Kyoto, 2014. Title: Efficiency in the identification in the limit paradigm

- ✓ Invited seminar at *the Computer and Information Science department of the University of Delaware*, 2013. Title: Recent advances in grammatical inference of non-regular languages
- ✓ Invited seminar at the *Computer Science department of the University of Maryland, Baltimore County*, 2012. Title: Learning context-free grammars

**Editing & Reviewing:**

- ✓ Co-editor of a special issue on grammar learning for Fundamenta Informaticae, 2016
- ✓ Member of the program committee of the international conferences EMNLP-CoNLL'07, ICGI'08, ICGI'12, ICGI'14, ICGI'16, ICGI'18
- ✓ Reviewer for the Machine Learning Journal (2014, 2015), Theoretical Computer Science (2013, 2014, 2015, 2016), Journal of Machine Learning Research (2010, 2012, 2015), Information Processing Letters (2007), Advances in Complex Systems (2006)
- ✓ Reviewer for the international conferences ECML'05, ECML'10, ICML'15, NIPS'16
- ✓ Reviewer for the francophone conference *CAp'05, CAp'06, CAp'16, CAp'18*

**Tutoring:**

- ✓ **PhD**
  - ○ Member of the PhD proposal and defense committees of Adam Jardine (University of Delaware, USA), 2013-2016
  - ○ Co-tutoring with Tim Oates of John Clemens (University of Maryland, Baltimore County, USA), 2012-2015
- ✓ **MSc.**
  - ○ Final internship and thesis of a MSc student, subject: *Inside the black-box: extracting WA from RNN*, 2018
  - ○ Last year project of a MSc. Student, subject: *Automatic detection of typewriter model,* 2017
  - ○ Final internship of an applied MSc. student, subject: *Image classification via graph features*, 2016
  - ○ Last year project of 4 applied MSc. students, subject: *Image indexation by content*, 2011
  - ○ Last year project of 4 applied MSc. students, subject: *Development of a Firefox plug-in for film recommendation (allocine.fr*), 2010
  - ○ First year thesis of a MSc. student subject: *learning formal languages: theory and practice*, 2009
  - ○ Final internship and thesis of a MSc. student, subject: *statistical grammatical inference: from theory to implementation*, 2008
  - ○ First year project of 3 students, subject: *Implementation of a grammatical inference algorithm*, 2008
  - ○ First year thesis of a student, Subject: *Learning Languages defined with the use of String Rewriting Systems*, 2006
  - ○ First year thesis of a MS. student, Subject: *Implementation of a learning algorithm: LARS*, 2005
- ✓ **Undergraduate:** Two months internships of third year students. Subjects: *Development of a toolbox for spectral learning* & *Image classification via graph feature extraction & Machine Learning: teaching via examples*, 2014, 2015, 2018

**Past projects:**

- ✓ ANR DECODA, 2010-2013, *Speech Analytics in recorded call-center conversations*
- ✓ ANR LAMPADA, 2009-2014, *Learning Algorithms, Models and sPArse representations for structured DAta*
- ✓ ANR SEQUOIA, 2009-2012, *Probabilistic syntactic analysis with large coverage of French*.
- ✓ ANR MARMOTA, 2005-2008, *MAchine learning for pRobabilistic MOdels Tree lAnguages*
- ✓ Member of the European networks of excellence PASCAL: *Pattern Analysis, Statistical modelling and ComputAtional Learning* (2003-2007) and PASCAL2 (2008-2013)

**Software:**

Implementation of 2 toolboxes for spectral learning, Sp2Learn and Scikit-SpLearn, during a 1 year LabEx archimède project (both in production), Free BSD licensed

## Teaching

This teaching was done during a 3 years formation (*monitorat*) between 2003 and 2006, then during a 6 month full-time lecturer contract at the university Jean-Monnet (2006) and finally during my current position (≈ 220 hours per year).

| | |
|---|---|
| ✓**2nd year MSc. Computer Science**: | XML* (2 years) |
| | Classification, Learning, Decision (3 years) |
| | Advanced Classification |
| | Data Mining* (4 years) |
| | Research Methodology* |
| | Machine Learning* (3 years) |
| ✓**1st year MSc. Computer Science**: | Theoretical Computer Science* |
| | Introduction to Data Science* |
| | Data Science |
| ✓**3rd year License Computer Science**: | Algebraic Languages (2 years) |
| | Data Bases (10 years) |
| | Compilation (6 years) |
| | Translation & Semantic (3 years) |
| | Advanced Algorithmic (3 years) |
| | Introduction to Machine Learning* |
| ✓**2nd year License Computer Science:** | Rational Languages* (2 years) |
| ✓**2nd year License Physics:** | Imperative Programming* (2 years) |
| ✓**1st year License Computer Science:** | Functional Programming: Caml (3 years) |
| | Introduction to Algorithms |
| | Programming in C (2 years) |

*Note : In charge of the courses "Advanced Algorithmic", "Algebraic Languages", "XML: principles and tools", "Programming in C", "Data mining", "Compilation", "Data Bases", "Translation & Semantic", "Classification, Learning, Decision" and "Theoretical Computer Science". All courses correspond roughly to 60-70 hours, except those noted with a star, most of which are half unit of about 30 hours.*

A 4 days professional formation titled "Machine Learning: from Bases to Deep Learning" was given at Leroy Merlin Headquater in April 2018, in the context of CNRS Formation Entreprise.

## Collective Responsibilities

- ✓ Co-chair of the LearnAut 2018 workshop at the Federated Logic Conference (FloC 2018), Oxford, UK, 2018
- ✓ Co-chair of the LearnAut workshop at the Logic In Computer Science (LICS 2017) conference, Reykjavik, Iceland, 2017
- ✓ Co-organiser of the SPiCe on-line competition about learning the next symbol in sequences, 2016
- ✓ Co-organiser of the PAutomaC on-line competition about learning probabilistic finite state machines, 2012
- ✓ In charge of the Master program in computer science "Advanced Data Bases" 2009-2012
- ✓ In charge of the Master program in computer science "Reliability and Security in Computer Science" 2010-2012
- ✓ Head of the organisation committee and member of the scientific committee of the PASCAL2 Bootcamp 2010, Marseilles, France

- ✓ Elected at the council of the UFR (≃ French faculty) Mathematics - Computer Science - Mechanics of the University of Provence (2008-2012). Member of the research and pedagogic council of this UFR.
- ✓ Elected at the executive committee of the national association for computer science SPECIF, 2009-2011
- ✓ Member of the organisation committee of the first French summer school on Machine Learning, EPIT 2008, in charge of communication
- ✓ Member of the organisation committee of the *SPECIF* congress 2006 on Ph. D cursus, Saint-Etienne
- ✓ Member of the organisation committee of the grammatical inference workshop 2006*,* Saint-Étienne
- ✓ Elected at the council of the doctoral school of Saint-Étienne, 2002-2007
- ✓ President of the Association Stéphanoise des jEunes Chercheurs, 2003-2004
  Vice-president of the Association Stéphanoise des jEunes Chercheurs, 2004-2005
- ✓ Member of the executive committee of the Confédération des Jeunes Chercheurs, 2005-2007

## Science popularisation

- ✓ Member of the organization committee of the Treize Minutes Marseille, a science popularization event, 2014, 2015, 2016, 2017, 2018
- ✓ Co-creator and Animator of the network Communauté Française des Docteurs in Marseilles, 2015
- ✓ Animation of a debate (“Causerie”) at the Fêtes de la science 2015, Villa mediterranée, Marseilles
- ✓ Talk titled « Computer science : education, jobs and dangers » at the French high school *La Condamine* of Quito (Ecuador), January 2007
- ✓ Animation of a debate with high school students, titled "*To talk and to discuss with a computer*", during the Vogue du Net organized by the town of Saint-Étienne in may 2006
- ✓ Animation of a debate, “the new dangers of internet”, at the occasion of the Science Party 2006, at La Rotonde