

**References for the ICGI 2012 tutorial
“Learning Probabilistic Finite-State Machines”
by J. Castro and R. Gavaldà**

References

- [Abe-Warmuth92] Naoki Abe and Manfred K. Warmuth. On the computational complexity of approximating distributions by probabilistic automata. *Machine Learning*, 9:205–260, 1992.
- [Beimel+00] Amos Beimel, Francesco Bergadano, Nader H. Bshouty, Eyal Kushilevitz, and Stefano Varricchio. Learning functions represented as multiplicity automata. *J. ACM*, 47(3):506–530, 2000.
- [Balle-Castro-G12] Borja Balle, Jorge Castro, and Ricard Gavaldà. Bootstrapping and learning pdfa in data streams. In *11th Intl. Conf. on Grammatical Inference (ICGI)*, volume 21 of *JMLR Workshop and Conference Proceedings*, pages 34–48, 2012.
- [Bailly+09] Raphaël Bailly, François Denis, and Liva Ralaivola. Grammatical inference as a principal component analysis problem. In *26th Annual International Conference on Machine Learning (ICML)*, page 5. ACM, 2009.
- [Baum+70] L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Statist.*, 41(1):164–171, 1970.
- [Balle+11] Borja Balle, Ariadna Quattoni, and Xavier Carreras. A spectral learning algorithm for finite state transducers. In *Machine Learning and Knowledge Discovery in Databases - European Conference (ECML PKDD)*, volume 6911 of *Lecture Notes in Computer Science*, pages 156–171. Springer, 2011.
- [Balle+12] Borja Balle, Ariadna Quattoni, and Xavier Carreras. Local loss optimization in operator models: A new insight into spectral learning. In *International Conference on Machine Learning (ICML)*, 2012.
- [Castro-G08] Jorge Castro and Ricard Gavaldà. Towards feasible PAC-learning of probabilistic deterministic finite automata. In *9th International Colloquium on Grammatical Inference: Algorithms and Applications (ICGI)*, volume 5278 of *Lecture Notes in Computer Science*, pages 163–174. Springer, 2008.
- [Carrasco-Oncina94] Rafael C. Carrasco and José Oncina. Learning stochastic regular grammars by means of a state merging method. In

Second International Colloquium on Grammatical Inference and Applications (ICGI), volume 862 of *Lecture Notes in Computer Science*, pages 139–152. Springer, 1994.

- [Carrasco-Oncina99] Rafael C. Carrasco and José Oncina. Learning deterministic regular grammars from stochastic samples in polynomial time. *ITA*, 33(1):1–20, 1999.
- [Clark-Thollard04] Alexander Clark and Franck Thollard. PAC-learnability of probabilistic deterministic finite state automata. *Journal of Machine Learning Research*, 5:473–497, 2004.
- [Dupont+05] Pierre Dupont, François Denis, and Yann Esposito. Links between probabilistic automata and hidden Markov models: probability distributions, learning models and induction algorithms. *Pattern Recognition*, 38(9):1349–1371, 2005.
- [Denis-Esposito04] François Denis and Yann Esposito. Learning classes of probabilistic automata. In *17th Annual Conference on Learning Theory (COLT)*, volume 3120 of *Lecture Notes in Computer Science*, pages 124–139. Springer, 2004.
- [Denis+06] François Denis, Yann Esposito, and Amaury Habrard. Learning rational stochastic languages. In *19th Annual Conference on Learning Theory (COLT)*, volume 4005 of *Lecture Notes in Computer Science*, pages 274–288. Springer, 2006.
- [delaHiguera10] Colin de la Higuera. *Grammatical Inference Learning Automata and Grammars*. Cambridge University Press, 2010.
- [dlHT00] Colin de la Higuera and Franck Thollard. Identification in the limit with probability one of stochastic deterministic finite automata. In *5th International Colloquium on Grammatical Inference: Algorithms and Applications (ICGI)*, volume 1891 of *Lecture Notes in Computer Science*, pages 141–156. Springer, 2000.
- [Guttman+05] Omri Guttman, S. V. N. Vishwanathan, and Robert C. Williamson. Learnability of probabilistic automata via oracles. In *16th International Conference on Algorithmic Learning Theory (ALT)*, pages 171–182, 2005.
- [Hsu+09] Daniel Hsu, Sham M. Kakade, and Tong Zhang. A spectral algorithm for learning hidden Markov models. In *22nd ACM Conference on Learning Theory (COLT)*. ACM, 2009.
- [Kermorvant-Dupont02] Christopher Kermorvant and Pierre Dupont. Stochastic grammatical inference with multinomial tests. In *6th International Colloquium on Grammatical Inference: Algorithms and Applications (ICGI)*, volume 2484 of *Lecture Notes in Computer Science*, pages 149–160. Springer, 2002.

- [Kearns+94] Michael J. Kearns, Yishay Mansour, Dana Ron, Ronitt Rubinfeld, Robert E. Schapire, and Linda Sellie. On the learnability of discrete distributions. In *Twenty-Sixth Annual ACM Symposium on Theory of Computing (STOC)*, pages 273–282. ACM, 1994.
- [Luque+12] Franco M. Luque, Ariadna Quattoni, Borja Balle, and Xavier Carreras. Spectral learning for non-deterministic dependency parsing. In *13th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 409–419. The Association for Computer Linguistics, 2012.
- [Mossel-Roch05] Elchanan Mossel and Sébastien Roch. Learning nonsingular phylogenies and hidden Markov models. In *37th Annual ACM Symposium on Theory of Computing (STOC)*, pages 366–375. ACM, 2005.
- [Palmer-Goldberg07] Nick Palmer and Paul W. Goldberg. PAC-learnability of probabilistic deterministic finite state automata in terms of variation distance. *Theor. Comput. Sci.*, 387(1):18–31, 2007.
- [Ron+96] Dana Ron, Yoram Singer, and Naftali Tishby. The power of amnesia: Learning probabilistic automata with variable memory length. *Machine Learning*, 25(2-3):117–149, 1996.
- [Rudich85] Steven Rudich. Inferring the structure of a Markov chain from its output. In *26th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 321–326, 1985.
- [Shalizi-Shalizi04] Cosma Rohilla Shalizi and Kristina Lisa Shalizi. Blind construction of optimal nonlinear recursive predictors for discrete sequences. In *20th Conference in Uncertainty in Artificial Intelligence (UAI)*, pages 504–511, 2004.
- [Thollard+00] Franck Thollard, Pierre Dupont, and Colin de la Higuera. Probabilistic dfa inference using Kullback-Leibler divergence and minimality. In *Seventeenth International Conference (ICML)*, pages 975–982. Morgan Kaufmann, 2000.
- [Vidal+05a] Enrique Vidal, Franck Thollard, Colin de la Higuera, Francisco Casacuberta, and Rafael C. Carrasco. Probabilistic finite-state machines - part i. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(7):1013–1025, 2005.
- [Vidal+05b] Enrique Vidal, Franck Thollard, Colin de la Higuera, Francisco Casacuberta, and Rafael C. Carrasco. Probabilistic finite-state machines - part ii. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(7):1026–1039, 2005.