**RL algorithms paper list**

**Nips**

[1] Z. Song, R. Parr, X. Liao, and L. Carin, “Linear Feature Encoding for Reinforcement Learning,” no. Nips, 2016.

[2] A. Krishnamurthy and J. Langford, “PAC Reinforcement Learning with Rich Observations,” no. Nips, pp. 1–9, 2016.

[3] A. Tamar, “Policy Gradient for Coherent Risk Measures,” *Adv. Neural Inf. Process. Syst. 28 (NIPS 2015)*, pp. 1–9, 2015.

[4] G. Comanici, “Basis Refinement Strategies for Linear Value Function Approximation in MDPs,” *Adv. Neural Inf. Process. Syst. 28 (NIPS 2015)*, pp. 1–9, 2015.

[5] C. Dann and E. Brunskill, “Sample Complexity of Episodic Fixed-Horizon Reinforcement Learning,” *Neural Inf. Process. Syst.*, pp. 1–9, 2015.

[6] M. Kobilarov, “Sample Complexity Bounds for Iterative Stochastic Policy Optimization,” *Adv. Neural Inf. Process. Syst.*, pp. 3096–3104, 2015.

[7] A. R. Mahmood, H. van Hasselt, and R. S. Sutton, “Weighted importance sampling for off-policy learning with linear function approximation,” *Proc. Adv. Neural Inf. Process. Syst.*, pp. 1–12, 2014.

[8] B. Szörényi, G. Kedenburg, and R. Munos, “Optimistic planning in Markov decision processes using a generative model,” *Nips*, pp. 1–9, 2014.

[9] I. Osband and B. Van Roy, “Near-optimal Reinforcement Learning in Factored MDPs,” *Adv. Neural Inf. Process. Syst.*, pp. 1–13, 2014.

[10] I. Osband and B. Van Roy, “Model-based Reinforcement Learning and the Eluder Dimension,” *Adv. Neural Inf. Process. Syst.*, pp. 1–9, 2014.

[11] M. Petrik, “RAAM : The Benefits of Robustness in Approximating Aggregated MDPs in Reinforcement Learning,” *Adv. Neural Inf. Process. Syst.*, pp. 1–9, 2014.

[12] B. Piot, M. Geist, and O. Pietquin, “Difference of Convex Functions Programming for Reinforcement Learning,” *Adv. Neural Inf. Process. Syst.*, vol. 2958, pp. 2519–2527, 2014.

[13] O. Maillard, T. Mann, and S. Mannor, “How hard is my MDP?‘ The distribution-norm to the rescue,’” *Adv. Neural Inf. …*, pp. 1–9, 2014.

**Icml**

[1] K. Lakshmanan, R. Ortner, and D. Ryabko, “Improved Regret Bounds for Undiscounted Continuous Reinforcement Learning.”

[2] B. Lesner and B. Scherrer BRUNOSCHERRER, “Non-Stationary Approximate Modified Policy Iteration.”

[3] H. Bou Ammar, R. Tutunov, and E. Eaton, “Safe Policy Search for Lifelong Reinforcement Learning with Sublinear Regret.”

[4] Z. Qin, W. Li, and F. Janoos, “Sparse Reinforcement Learning via Convex Optimization.”

[5] B. Scherrer BRUNOSCHERRER, “Approximate Policy Iteration Schemes: A Comparison.”

[6] T. A. Mann, “Scaling Up Approximate Value Iteration with Options: Better Policies with Fewer Iterations Shie Mannor.”

[7] R. C. Grande, T. J. Walsh, and J. P. How, “Sample Efficient Reinforcement Learning with Gaussian Processes.”

[8] A. Tamar, S. Mannor, and H. Xu, “Scaling Up Robust MDPs using Function Approximation.”

[9] Z. Zhang and D. Hsu, “Covering Number for Efficient Heuristic-Based POMDP Planning.”

[10] T. Dick, A. György, C. Szepesvári, and S. Ca, “Online Learning in Markov Decision Processes with Changing Cost Sequences.”

[11] P. S. Thomas and E. Brunskill, “Data-Efficient Off-Policy Policy Evaluation for Reinforcement Learning,” *arXiv1604.00923 [cs]*, vol. 48, 2016.

[12] J. Pérolat, B. Piot BILALPIOT, F. Matthieu Geist, C. Bruno Scherrer BRUNOSCHERRER, and I. Olivier Pietquin, “Softened Approximate Policy Iteration for Markov Games,” vol. 48, 2016.

[13] B. Balle, M. Gomrokchi, and D. Precup, “Differentially Private Policy Evaluation,” *arXiv*, vol. 48, pp. 1–27, 2016.

[14] J. Ho, J. K. Gupta, and S. Ermon, “Model-Free Imitation Learning with Policy Optimization,” *Icml 2016*, vol. 48, 2016.

[15] N. Jiang and L. Li, “Doubly Robust Off-policy Value Evaluation for Reinforcement Learning,” *Icml*, vol. 48, p. 14, 2015.

[16] L. A. Prashanth, J. Cheng, M. Fu, S. Marcus, and L. G. Jun, “Cumulative Prospect Theory Meets Reinforcement Learning : Estimation and Control,” *Work. Pap.*, vol. 48, pp. 1–27, 2015.

[17] N. Jiang and L. Li, “Doubly Robust Off-policy Value Evaluation for Reinforcement Learning,” *Icml*, vol. 48, p. 14, 2015.

[18] E. Brunskill and L. Li, “PAC-inspired Option Discovery in Lifelong Reinforcement Learning,” 2014.