16
Bibliography

P. Abbeel and A. Y. Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the 21st International Conference on Machine Learning (ICML)*, 2004.

A. Agarwal, O. Chapelle, M. Dudík, and J. Langford. A reliable effective terascale linear learning system. *Journal of Machine Learning Research*, 15: 1111–1133, 2014. URL http://jmlr.org/papers/v15/agarwal14a.html.

Brian D. O. Anderson and John B. Moore. *Optimal Control: Linear Quadratic Methods*. Prentice-Hall, Inc., 1990.

B. D. Argall, S. Chernova, M. Veloso, and B. Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 2009.

S. Arora, E. Hazan, , and S. Kale. The multiplicative weights update method: A meta-algorithm and applications. *Theory of Computing*, 2012.

C. G. Atkeson. Using local trajectory optimizers to speed up global optimization in dynamic programming. In *Advances in Neural Information Processing Systems (NIPS)*, 1994.

C. G. Atkeson and J. Morimoto. Nonparametric representation of policies and value functions: A trajectory-based approach. In *Advances in Neural Information Processing Systems (NIPS)*, 2003.

C. G. Atkeson and S. Schaal. Robot learning from demonstration. In *Proceedings of the 14th International Conference on Machine Learning (ICML)*, 1997.

A. Bachrach, R. He, and N. Roy. Autonomous flight in unknown indoor environments. *International Journal of Micro Air Vehicles*, 1(4):217–228, 2009.

J. A. Bagnell. Robust supervised learning. In *Proceedings of the American Association for Artificial Intelligence (AAAI)*. AAAI Press / The MIT Press, 2005.

J. A. Bagnell and J. Schneider. Covariant policy search. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2003.

J. A. Bagnell, A. Y. Ng, S. Kakade, and J. Schneider. Policy search by dynamic programming. In *Advances in Neural Information Processing Systems*, 2003.

L. C. Baird. Residual algorithms: Reinforcement learning with function approximation. In *International Conference on Machine Learning*, 1995.

C. L Baker, R. Saxe, and J. B Tenenbaum. Action understanding as inverse planning. *Cognition*, 113(3):329–349, 2009.

Y. Bengio. Learning deep architectures for AI. *Foundations and Trends in Machine Learning*, 2009.

Justin A Boyan and Andrew W Moore. Generalization in reinforcement learning: Safely approximating the value function. In *Advances in neural information processing systems*, 1995.

S. P Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix inequalities in system and control theory*, volume 15. SIAM, 1994.

Massimo Caccia, Lucas Caccia, William Fedus, Hugo Larochelle, Joelle Pineau, and Laurent Charlin. Language gans falling short. *CoRR*, abs/1811.02549, 2018.

N. Cesa-Bianchi, Y. Freund, D. Haussler, D. P. Helmbold, R. E. Schapire, and M. K. Warmuth. How to use expert advice. *Journal of the ACM*, 1997.

A. Coates, P. Abbeel, and A. Y Ng. Apprenticeship learning for helicopter control. *Communications of the ACM*, 52(7):97–105, 2009.

H. Daumé III, J. Langford, and D. Marcu. Search-based structured prediction. *Machine Learning*, 2009.

A. Dragan and S. Srinivasa. Formalizing assistive teleoperation. In *Robotics: Science and Systems*, 2012.

Miroslav Dudik, Steven J. Phillips, and Robert E. Schapire. Performance guarantees for regularized maximum entropy density estimation. volume 3120, pages 472–486, 2004.

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

Peter D. Grunwald and A. Philip Dawid. Game theory, maximum entropy, minimum discrepancy and robust bayesian decision theory, 2004. URL http://arxiv.org/abs/math/0410076.

X. Guo, S. Singh, H. Lee, R. L Lewis, and X. Wang. Deep learning for real-time atari game play using offline monte-carlo tree search planning. In *Advances in Neural Information Processing Systems*, 2014.

H. He, H. Daume III, and J. Eisner. Imitation learning by coaching. In *NIPS*, 2012.

Martial H Hebert. *Intelligent unmanned ground vehicles: autonomous navigation research at Carnegie Mellon*. Kluwer Academic Publishers, 1997.

Jonathan Ho and Stefano Ermon. Generative adversarial imitation learning. 2016. URL http://arxiv.org/abs/1606.03476.

D.A. Huang, A.M. Farahmand, K. M. Kitani, and J. A. Bagnell. Approximate maxent inverse optimal control and its application for mental simulation of human interactions. In *AAAI Conference on Artificial Intelligence*, 2015.

L. D Jackel, E. Krotkov, M. Perschbacher, J. Pippine, and C. Sullivan. The DARPA LAGR program: Goals, challenges, methodology, and phase i results. *Journal of Field Robotics*, 23(11-12):945–973, 2006.

D. H. Jacobson and D. Q. Mayne. *Differential Dynamic Programming*. Elsevier, 1970.

Arthur Jacot-Guillarmod, Franck Gabriel, and Clement Hongler. Neural tangent kernel: Convergence and generalization in neural networks. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 8580–8589. Curran Associates, Inc., 2018. URL http://papers.nips.cc/paper/8076-neural-tangent-kernel-convergence-and-generalization-in-neural-networks.pdf.

E. T. Jaynes. *Probability theory: The logic of science*. Cambridge University Press, 2003.

S. Kakade and J. Langford. Approximately optimal approximate reinforcement learning. In *Proceedings of the 19th International Conference on Machine Learning (ICML)*, 2002.

M. Kalakrishnan, J. Buchli, P. Pastor, M. Mistry, and S. Schaal. Learning, planning, and control for quadruped locomotion over challenging terrain. *The International Journal of Robotics Research*, 30(2):236–258, 2011.

R. E. Kalman. When is a linear control system optimal? *Journal of Fluids Engineering*, 86(1):51–60, 1964.

B. Kim, A. M. Farahmand, J. Pineau, and D. Precup. Learning from limited demonstrations. In *Advances in Neural Information Processing Systems*, pages 2859–2867, 2013.

K. M Kitani, B. D. Ziebart, J. A. Bagnell, and M. Hebert. Activity forecasting. In *European Conference on Computer Vision*. Springer, October 2012.

J. Kober and J. Peters. Imitation and reinforcement learning - practical algorithms for motor primitive learning in robotics. 2010.

J Zico Kolter, Pieter Abbeel, and Andrew Y Ng. Hierarchical apprenticeship learning with application to quadruped locomotion. In *Advances in Neural Information Processing Systems*, pages 769–776, 2007.

H. Kretzschmar, M. Kuderer, and W. Burgard. Learning to predict trajectories of cooperatively navigating agents. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, China, 2014. DOI: 10.1109/ICRA.2014.6907442. URL http://ais.informatik.uni-freiburg.de/publications/papers/kretzschmar14icra.pdf.

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *IEEE*, 1998.

J. Leonard, J. How, S. Teller, M. Berger, S. Campbell, G. Fiore, L. Fletcher, E. Frazzoli, A. Huang, S. Karaman, et al. A perception-driven autonomous urban vehicle. *Journal of Field Robotics*, 25(10):727–774, 2008.

S. Levine and V. Koltun. Continuous inverse optimal control with locally optimal examples. In *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, pages 41–48, 2012.

Lihong Li. A worst-case comparison between temporal difference and residual gradient with linear function approximation. *Proceedings of the 25th international conference on Machine learning - ICML '08*, pages 560–567, 2008. DOI: 10.1145/1390156.1390227. URL http://portal.acm.org/citation.cfm?doid=1390156.1390227.

L. Ljung. Convergence analysis of parametric identification methods. *IEEE Transactions on Automatic Control*, 1978.

O. Miksik, D. Munoz, J. A. Bagnell, and M. Hebert. Efficient temporal consistency for streaming video scene analysis. In *ICRA*, 2013.

K. Mülling, J. Kober, O. Kroemer, and J. Peters. Learning to select and generalize striking movements in robot table tennis. *The International Journal of Robotics Research*, 32(3):263–279, 2013.

D. Munoz, J. A. Bagnell, and M. Hebert. Stacked hierarchical labeling. In *ECCV*, 2010.

M. Nechyba and J. A. Bagnell. Stabilizing human control strategies through reinforcement learning. In *Proc. IEEE Hong Kong Symp. on Robotics and Control*, volume 1, pages 39–44, April 1999.

A. Y Ng and S. J Russell. Algorithms for inverse reinforcement learning. In *Icml*, pages 663–670, 2000.

Takayuki Osa, Joni Pajarinen, and Gerhard Neumann. *An Algorithmic Perspective on Imitation Learning*. Now Publishers Inc., Hanover, MA, USA, 2018. ISBN 168083410X, 9781680834109.

D. Pomerleau. ALVINN: An Autonomous Land Vehicle in a Neural Network. In *Advances in Neural Information Processing Systems (NIPS)*, 1989.

M. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 1994.

V. Ramakrishna, D. Munoz, M. Hebert, J. A. (Drew) Bagnell, and Y. A. Sheikh. Pose machines: Articulated pose estimation via inference machines. In *European Conference on Computer Vision*, 2014.

N. Ratliff. *Learning to Search: Structured Prediction Techniques for Imitation Learning*. PhD thesis, Carnegie Mellon University, 2009.

N. Ratliff, J. A. Bagnell, and M. Zinkevich. (Online) subgradient methods for structured prediction. In *Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2007.

N. Ratliff, B. Ziebart, K. Peterson, J A. Bagnell, M. Hebert, A. K. Dey, and S. Srinivasa. Inverse optimal heuristic control for imitation learning. AISTATS, 2009a.

N. Ratliff, M. Zucker, J. A. Bagnell, and S. Srinivasa. Chomp: Gradient optimization techniques for efficient motion planning. In *ICRA*, 2009b.

Nathan D Ratliff, J Andrew Bagnell, and Martin A Zinkevich. Maximum margin planning. In *Proceedings of the 23rd international conference on Machine learning*, pages 729–736, 2006.

Nathan D Ratliff, David Silver, and J Andrew Bagnell. Learning to search: Functional gradient techniques for imitation learning. *Autonomous Robots*, 27(1):25–53, 2009c.

S. Ross. Comparison of imitation learning approaches on Super Tux Kart, 2010a. URL http://www.youtube.com/watch?v=V00npNnWzSU.

S. Ross. Comparison of imitation learning approaches on Super Mario Bros, 2010b. URL http://www.youtube.com/watch?v=anOI0xZ3kGM.

S. Ross and J. A. Bagnell. Efficient reductions for imitation learning. In *AISTATS*, 2010.

S. Ross and J. A. Bagnell. Reinforcement and imitation learning via interactive no-regret learning. *arXiv preprint arXiv:1406.5979*, 2014.

S. Ross, G. J. Gordon, and J. A. Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *AISTATS*, 2011a.

S. Ross, D. Munoz, M. Hebert, and J. A. Bagnell. Learning message-passing inference machines for structured prediction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011b.

S. Ross, N. Melik-Barkhudarov, K. S. Shankar, A. Wendel, D. Dey, J. A. Bagnell, and M. Hebert. Learning monocular reactive uav control in cluttered natural environments. In *ICRA*, 2013a.

S. Ross, N. Melik-Barkhudarov, K. S. Shankar, A. Wendel, D. Dey, J. A. Bagnell, and M. Hebert. Autonomous vision-based flight through a forest, 2013b. URL http://www.youtube.com/watch?v=o0_0hp1pHBw.

S. Ross, J. Zhou, Y. Yue, D. Dey, and J. A. Bagnell. Learning policies for contextual submodular prediction. In *ICML*, 2013c.

Stephane Ross and J. Andrew Bagnell. Stability Conditions for Online Learnability. *arXiv:1108.3154*, 2011. URL http://arxiv.org/abs/1108.3154.

J. Rust. *Do people behave according to Bellman's principle of optimality?* Hoover Institution, Stanford University, 1992.

John Rust. Structural estimation of markov decision processes. *Handbook of econometrics*, 4(4), 1994.

A. Safonova and J. K Hodgins. Construction and optimal search of interpolated motion graphs. In *ACM Transactions on Graphics (TOG)*, volume 26, page 106. ACM, 2007.

Ankan Saha, Prateek Jain, and Ambuj Tewari. The Interplay Between Stability and Regret in Online Learning. *arXiv preprint arXiv:1211.6158*, pages 1–19, 2012. URL http://arxiv.org/abs/1211.6158.

Robert E. Schapire and Manfred K. Warmuth. On the worst-case analysis of temporal-difference learning algorithms. *Machine Learning*, 22(1):95–121, 1996. ISSN 0885-6125. DOI: 10.1007/BF00114725.

S. Shalev-Shwartz and S. Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

D. Silver. *Learning Preference Models for Autonomous Mobile Robots in Complex Domains*. PhD thesis, Carnegie Mellon University, 2010.

Bharath K. Sriperumbudur, Arthur Gretton, Kenji Fukumizu, Gert R. G. Lanckriet, and Bernhard Schölkopf. A note on integral probability metrics and ϕ-divergences. 2009. URL http://arxiv.org/abs/0901.2698.

Wen Sun and J Andrew Bagnell. Online bellman residual algorithms with predictive error guarantees. *UAI*, 2015.

Wen Sun and J. Andrew (Drew) Bagnell. Online bellman residual and temporal difference algorithms with predictive error guarantees. In *Proceedings of The 25th International Joint Conference on Artificial Intelligence - IJCAI 2016*, April 2016.

Wen Sun, Geoffrey J Gordon, Byron Boots, and J Bagnell. Dual policy iteration. In *Advances in Neural Information Processing Systems*, 2018.

R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

U. Syed and R. E Schapire. A game-theoretic approach to apprenticeship learning. In *Advances in neural information processing systems*, pages 1449–1456, 2007.

Yuval Tassa, Tom Erez, and Emanuel Todorov. Synthesis and stabilization of complex behaviors through online trajectory optimization. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 4906–4913. IEEE, 2012.

C. Urmson, J. Anhalt, J. A. Bagnell, C. Baker, R. Bittner, MN Clark, J. Dolan, D. Duggins, T. Galatali, C. Geyer, et al. Autonomous driving in urban environments: Boss and the urban challenge. *Journal of Field Robotics*, 25(8): 425–466, 2008.

P. Vernaza and D. D. Lee. Efficient dynamic programming for high-dimensional, optimal motion planning by spectral learning of approximate value function symmetries. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2011.

K. Waugh, B. D. Ziebart, and J. A. Bagnell. Computational rationalization: The inverse equilibrium problem. In *Proceedings of the International Conference on Machine Learning*, June 2011.

Shih-En Wei, Varun Ramakrishna, Takeo Kanade, and Yaser Sheikh. Convolutional pose machines. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4724–4732, 2016.

C. Wellington and A. Stentz. Online adaptive rough-terrain navigation vegetation. In *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, volume 1, pages 96–101 Vol.1, 2004.

B. D. Ziebart, A. Maas, J. A. Bagnell, and A. K. Dey. Maximum entropy inverse reinforcement learning. In *Proceedings of the 23rd National Conference on Artificial Intelligence (AAAI)*, 2008a.

B. D. Ziebart, A. Maas, A. Dey, and J. A. Bagnell. Navigate like a cabbie: Probabilistic reasoning from observed context-aware behavior. In *UBI-COMP: Ubiquitious Computation*, 2008b.

B. D. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J. A. Bagnell, M. Hebert, A. Dey, and s. Srinivasa. Planning-based prediction for pedestrians. In *Proc. IROS 2009*, October 2009.

B. D. Ziebart, J. Andrew Bagnell, and A. K. Dey. Modeling interaction via the principle of maximum causal entropy. In *Proceedings of the 27th International Conference on Machine Learning*, 2010.

B. D. Ziebart, A. Dey, and J. A. Bagnell. Probabilistic pointing target prediction via inverse optimal control. In *International Conference on Intelligent User Interfaces (IUI 2012)*, 2012.

B. D. Ziebart, J. A. Bagnell, and A. K. Dey. The principle of maximum causal entropy for estimating interacting processes. *IEEE Transactions on Information Theory*, February 2013.

M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning (ICML)*, 2003.

M. Zucker, N. Ratliff, M. Stolle, J. Chestnutt, J Andrew Bagnell, C. G. Atkeson, and J. Kuffner. Optimization and learning for rough terrain legged locomotion. *The International Journal of Robotics Research*, 30(2):175–191, 2011.