

## References

- [1] Brenna D. Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, 2009.
- [2] J. L. Austerweil, S. Brawner, A. Greenwald, E. Hilliard, M. Ho, M. L. Littman, J. MacGlashan, and C. Trimbach. The impact of outcome preferences in a collection of non-zero-sum grid games. AAAI Spring Symposium 2016 on Challenges and Opportunities in Multiagent Learning for the Real World, 2016.
- [3] Monica Babes, Vukosi N. Marivate, Michael L. Littman, and Kaushik Subramanian. Apprenticeship learning about multiple intentions. In *International Conference on Machine Learning*, pages 897–904, 2011.
- [4] Michael Bacharach. Interactive team reasoning: A contribution to the theory of co-operation. *Research in Economics*, 53(2):117–147, 1999.
- [5] Michael Bacharach, Natalie Gold, and Robert Sugden. *Beyond individual choice: teams and frames in game theory*. Princeton University Press, 2006.
- [6] C. L. Baker and J. B. Tenenbaum. Modeling human plan recognition using Bayesian theory of mind. In G. Sukthankar, R. P. Goldman, C. Geib, D. Pynadath, and H. Bui, editors, *Plan, Activity and Intent Recognition: Theory and Practice*, pages 177–204. Morgan Kaufmann, 2014.
- [7] Andrew G. Barto, S. J. Bradtke, and Satinder P. Singh. Learning to act using real-time dynamic programming. *Artificial Intelligence*, 72(1):81–138, 1995.
- [8] Gary S. Becker. A theory of social interactions. Working Paper 42, National Bureau of Economic Research, June 1974.
- [9] Richard Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, 1957.
- [10] Joyce Berg, John Dickhaut, and Kevin McCabe. Trust, reciprocity, and social history. *Games and economic behavior*, 10(1):122–142, 1995.
- [11] Dimitri P. Bertsekas. *Dynamic Programming: Deterministic and Stochastic Models*. Prentice-Hall, Englewood Cliffs, NJ, 1987.
- [12] A. Boularias, J. Kober, and J. Peters. Relative entropy inverse reinforcement learning. In *JMLR Workshop and Conference Proceedings Volume 15: AISTATS 2011*, pages 182–189, Cambridge, MA, USA, April 2011. MIT Press.
- [13] Craig Boutilier. Planning, learning and coordination in multiagent decision processes. In *Proceedings of the 6th conference on Theoretical aspects of rationality and knowledge*, pages 195–210. Morgan Kaufmann Publishers Inc., 1996.
- [14] Craig Boutilier, Thomas Dean, and Steve Hanks. Decision-theoretic planning: Structural assumptions and computational leverage. *Journal of Artificial Intelligence Research*, 11:1–94, 1999.
- [15] Benjamin Burchfiel, Carlo Tomasi, and Ronald Parr. Distance minimization for reward learning from scored trajectories. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.
- [16] C. F. Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, Princeton, NJ, 2003.
- [17] Mike Cassidy. Centaur chess shows power of teaming human and machine. [http://www.huffingtonpost.com/mike-cassidy/centaur-chess-shows-power\\_b\\_6383606.html](http://www.huffingtonpost.com/mike-cassidy/centaur-chess-shows-power_b_6383606.html), 2014.
- [18] Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *Quarterly journal of Economics*, pages 817–869, 2002.

- [19] Jaedeug Choi and Kee-Eung Kim. Nonparametric Bayesian inverse reinforcement learning for multiple reward functions. In *Proceedings of the 25th International Conference on Neural Information Processing Systems, NIPS'12*, pages 305–313, USA, 2012. Curran Associates Inc.
- [20] Gregg Collins and Louise Pryor. Planning under uncertainty: Some key issues. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1567–1573, 1995.
- [21] V. Conitzer and T. Sandholm. AWESOME: A general multiagent learning algorithm that converges in self-play and learns a best response against stationary opponents. *Machine Learning*, 67:23–43, 2007.
- [22] A. Cournot. *Recherches sur les Principes Mathématiques de la Théorie la Richesse*. Hachette, 1838.
- [23] Alexander G Cunningham, Enric Galceran, Ryan M Eustice, and Edwin Olson. MPDM: Multipolicy decision-making in dynamic, uncertain environments for autonomous driving. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1670–1677. IEEE, 2015.
- [24] Constantinos Daskalakis, Paul W Goldberg, and Christos H Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.
- [25] Marie desJardins and Michael L. Littman. AAAI-13 preface. In *AAAI 2013*, 2013.
- [26] Carlos Diuk, Lihong Li, and Bethany R Leffler. The adaptive k-meteorologists problem and its application to structure learning and feature selection in reinforcement learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 249–256. ACM, 2009.
- [27] Dmitri Dolgov, Sebastian Thrun, Michael Montemerlo, and James Diebel. Path planning for autonomous vehicles in unknown semi-structured environments. *The International Journal of Robotics Research*, 29(5):485–501, 2010.
- [28] Ernst Fehr and Klaus M Schmidt. A theory of fairness, competition, and cooperation. *Quarterly journal of Economics*, pages 817–868, 1999.
- [29] A.M. Fink. Equilibrium in a stochastic  $n$ -person game. *Journal of Science in Hiroshima University*, 28:89–93, 1964.
- [30] D. Fudenberg and D. K. Levine. *The Theory of Learning in Games*. MIT Press, Cambridge, 1998.
- [31] Kobi Gal, Avi Pfeffer, Francesca Marzo, and Barbara J. Grosz. Learning social preferences in games. In *AAAI-04*, pages 226–231, 2004.
- [32] T. Goff, A. Greenwald, E. Hilliard, W. Ketter, A. Loomis, and E. Sodomka. JACK: A Java auction configuration kit. In *Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets.*, volume 136 of *Lecture Notes in Business Information Processing*, pages 45–60. Springer, 2013.
- [33] Nakul Gopalan and Stefanie Tellex. Modeling and solving human-robot collaborative tasks using POMDPs. In *Robotics: Science and Systems 2015: Workshop on Model Learning for Human–Robot Communication*, 2015.
- [34] A. Greenwald and K. Hall. Correlated  $Q$ -learning. In *Proceedings of the Twentieth International Conference on Machine Learning*, pages 242–249, 2003.
- [35] A. Greenwald, J. Li, and E. Sodomka. Approximating equilibria in sequential auctions with incomplete information and multi-unit demand. In *NIPS '12: Advances in Neural Information Processing Systems*, volume 25, pages 2330–2338, December 2012.
- [36] A. Greenwald, J. Li, E. Sodomka, and M. L. Littman. Solving for best-responses in extensive-form games using reinforcement learning methods. In *RLDM '13: Proceedings of 1st Multidisciplinary Conference Reinforcement Learning and Decision Making*, pages 116–120, October 2013.

- [37] A. Greenwald, E. Sodomka, E. Stix, J. Stix, and D. Storch. An empirical analysis of auctioneer profitability in Quibids penny auctions. In *Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets.*, volume 197 of *Lecture Notes in Business Information Processing*, pages 56–68. Springer, 2014.
- [38] Werner Guth, Rolf Schmittberger, and Bernd Schwarze. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4):367–388, 1982.
- [39] J. Hu and M. Wellman. Nash  $Q$ -learning for general-sum stochastic games. *Machine Learning Research*, 4:1039–1069, 2003.
- [40] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2):99–134, 1998.
- [41] D. Kahneman, J. L. Knetsch, and R. Thaler. Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review*, 76:728–741, 1986.
- [42] D. Kahneman, P. Slovic, and A. Tversky, editors. *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press, 1982.
- [43] Michihiro Kandori, George J. Mailath, and Rafael Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56, January 1993.
- [44] Michael Kearns, Yishay Mansour, and Andrew Y. Ng. A sparse sampling algorithm for near-optimal planning in large Markov decision processes. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99)*, pages 1324–1331, 1999.
- [45] Beomjoon Kim and Joelle Pineau. Socially adaptive path planning in human environments using inverse reinforcement learning. *International Journal of Social Robotics*, 8(1):51–66, 2016.
- [46] Max Kleiman-Weiner, Mark K. Ho, Joseph L. Austerweil, Michael L. Littman, and Joshua B. Tenenbaum. Coordinate to cooperate or compete: Abstract goals and joint intentions in social interaction. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, 2016.
- [47] Levente Kocsis and Csaba Szepesvári. Bandit based Monte-carlo planning. In *Machine Learning: ECML 2006*, pages 282–293, 2006.
- [48] J Zico Kolter and Andrew Y Ng. Regularization and feature selection in least-squares temporal difference learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 521–528. ACM, 2009.
- [49] Lihong Li, Jason D Williams, and Suhrid Balakrishnan. Reinforcement learning for dialog management using least-squares policy iteration and fast feature selection. In *INTERSPEECH*, pages 2475–2478, 2009.
- [50] M. Littman. Friend or foe  $Q$ -learning in general-sum Markov games. In *Proceedings of Eighteenth International Conference on Machine Learning*, pages 322–328, June 2001.
- [51] Michael L Littman. Markov games as a framework for multi-agent reinforcement learning. In *Proceedings of the eleventh international conference on machine learning*, volume 157, pages 157–163, 1994.
- [52] Michael L. Littman. Friend-or-foe  $Q$ -learning in general-sum games. In *Proceedings of the Eighteenth International Conference on Machine Learning*, pages 322–328. Morgan Kaufmann, 2001.
- [53] Robert Loftin, James MacGlashan, Michael L. Littman, Matthew E. Taylor, David L. Roberts, and Jeff Huang. A strategy-aware technique for learning behaviors from discrete human feedback. In *Proceedings of the Twenty-Eighth Association for the Advancement of Artificial Intelligence Conference (AAAI-14)*, 2014.

- [54] Manuel Lopes, Francisco Melo, and Luis Montesano. Active learning for reward estimation in inverse reinforcement learning. In *Machine Learning and Knowledge Discovery in Databases*, pages 31–46. Springer, 2009.
- [55] J. MacGlashan, M. Babes-Vroman, M. desJardins, M. Littman, S. Muresan, S. Squire, S. Tellex, D. Arumugam, and L. Yang. Grounding english commands to reward functions. In *Proceedings of Robotics: Science and Systems*, 2015.
- [56] James MacGlashan. BURLAP. <http://burlap.cs.brown.edu>, 2016.
- [57] James MacGlashan, Michael Littman, and Fiery Cushman. Flexible theft and resolute punishment: Evolutionary dynamics of social behavior among reinforcement-learning agents. In *Proceedings of the Annual Conference of the Cognitive Science Society*, 2014.
- [58] James MacGlashan and Michael L. Littman. Between imitation and intention learning. In *Proceedings of IJCAI*, 2015.
- [59] Pattie Maes and Robyn Kozierok. Learning interface agents. In *AAAI*, volume 93, pages 459–465, 1993.
- [60] B. Mayer, E. Sodomka, and A. Greenwald. The price of independence in simultaneous auctions. In *AA-MAS '13: Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, pages 1227–1228, May 2013.
- [61] B. Mayer, E. Sodomka, A. Greenwald, and M.P. Wellman. Accounting for price dependencies in simultaneous sealed-bid auctions. In *EC '13: Proceedings of the 14th ACM Conference on Electronic Commerce*, pages 679–696, June 2013.
- [62] Bernard Michini and Jonathan P. How. Bayesian nonparametric inverse reinforcement learning. In Peter A. Flach, Tijl De Bie, and Nello Cristianini, editors, *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2012, Bristol, UK, September 24-28, 2012. Proceedings, Part II*, pages 148–163. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
- [63] Enrique Munoz de Cote and Michael L. Littman. A polynomial-time Nash equilibrium algorithm for repeated stochastic games. In *24th Conference on Uncertainty in Artificial Intelligence (UAI'08)*, 2008.
- [64] Andrew Y. Ng, Daishi Harada, and Stuart J. Russell. Policy invariance under reward transformations: Theory and application to reward shaping. In *Proceedings of the Sixteenth International Conference on Machine Learning, ICML '99*, pages 278–287, San Francisco, CA, USA, 1999. Morgan Kaufmann Publishers Inc.
- [65] Andrew Y. Ng and Stuart Russell. Algorithms for inverse reinforcement learning. In *International Conference on Machine Learning*, pages 663–670, 2000.
- [66] Ronald Parr, Lihong Li, Gavin Taylor, Christopher Painter-Wakefield, and Michael L Littman. An analysis of linear models, linear value-function approximation, and feature selection for reinforcement learning. In *Proceedings of the 25th international conference on Machine learning*, pages 752–759. ACM, 2008.
- [67] Liam Paull. What is Duckietown? <http://duckietown.mit.edu/>, 2016.
- [68] Dean A. Pomerleau. *Neural Network Perception for Mobile Robot Guidance*. Kluwer Academic Publishing, 1993.
- [69] Qifeng Qiao and Peter A. Beling. Inverse reinforcement learning with Gaussian process. *CoRR*, abs/1208.2112, 2012.
- [70] Matthew Rabin. Incorporating fairness into game theory and economics. *American Economic Review*, 83(5):1281–1302, 1993.

- [71] Lawrence R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, February 1989.
- [72] Deepak Ramachandran and Eyal Amir. Bayesian Inverse Reinforcement Learning. *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, 51:2586–2591, 2007.
- [73] Tummalapalli Sudhamsh Reddy, Vamsikrishna Gopikrishna, Gergely Zaruba, and Manfred Huber. Inverse reinforcement learning for decentralized non-cooperative multiagent systems. In *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*, pages 1930–1935. IEEE, 2012.
- [74] Peter C. Reiss and Frank A. Wolak. Structural econometric modeling: Rationales and examples from industrial organization. In J.J. Heckman and E.E. Leamer, editors, *Handbook of Econometrics*, volume 6A, chapter 64. Elsevier, 1 edition, 2007.
- [75] D. E. Rumelhart and J. L. McClelland, editors. *Parallel Distributed Processing: Explorations in the Microstructures of Cognition. Volume 1: Foundations*. The MIT Press, Cambridge, MA, 1986.
- [76] T.C. Schelling. *The Strategy of Conflict*. Harvard University Press, 1980.
- [77] John R. Searle. *The Construction of Social Reality*. Simon and Schuster, 1995.
- [78] L.S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America*, 39:1095–1100, 1953.
- [79] Kyriacos Shiarlis, Joao Messias, and Shimon Whiteson. Inverse reinforcement learning from failure. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 1060–1068. International Foundation for Autonomous Agents and Multiagent Systems, 2016.
- [80] Satinder Singh, Richard L Lewis, and Andrew G Barto. Where do rewards come from. In *Proceedings of the Annual Conference of the Cognitive Science Society*, pages 2601–2606, 2009.
- [81] Eric Sodomka, Elizabeth Hilliard, Michael Littman, and Amy Greenwald. Coco-Q: Learning in stochastic games with side payments. *JMLR Workshop and Conference Proceedings: Proceedings of the 30th International Conference on Machine Learning*, 28(3):1471–1479, 2013.
- [82] Richard S. Sutton, Doina Precup, and Satinder Singh. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, 112(1–2):181–211, 1999.
- [83] Michael Tomasello, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(05):675–691, 2005.
- [84] Blase Ur, Elyse McManus, Melwyn Pak Yong Ho, and Michael L. Littman. Practical trigger-action programming in the smart home. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2014.
- [85] Luis Von Ahn. Human computation. In *Design Automation Conference, 2009. DAC’09. 46th ACM/IEEE*, pages 418–419. IEEE, 2009.
- [86] Payne John W., Bettman James R., and Johnson Eric J. *The adaptive decision maker*. Cambridge University Press, 1993.
- [87] Christopher J. C. H. Watkins and Peter Dayan. Q-learning. *Machine Learning*, 8(3):279–292, 1992.
- [88] M. Wellman, A. Greenwald, and P. Stone. *Autonomous Bidding Agents: Strategies and Lessons from the Trading Agent Competition*. Intelligent Robotics and Autonomous Agents. MIT Press, Cambridge, 2007.
- [89] M. P. Wellman, E. Sodomka, and A. Greenwald. Self-confirming price prediction strategies for simultaneous one-shot auctions. In *UAI ’12: Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, pages 893–902, August 2012.

- [90] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. In *AAAI Conference on Artificial Intelligence*, pages 1433–1438, 2008.
- [91] M. Zinkevich, A. Greenwald, and M. Littman. Cyclic equilibria in Markov games. In *Advances in Neural Information Processing Systems 18*. MIT Press, 2006.