[1] Pieter Abbeel and John Schulman. ,Deep Reinforcement Learning through Policy Optimization, Tutorial at NIPS 2016.  
[2] Kai Arulkumaran‚ Nat Dilokthanakul‚ Murray Shanahan‚ and Anil~Anthony Bharath. ,Classifying Options for Deep Reinforcement Learning, IJCAI Workshop on Deep Reinforcement Learning: Frontiers and Challenges, 2016  
[3] Pierre-Luc Bacon‚ Jean Harb‚ and Doina Precup. ,The Option-Critic Architecture, AAAI, 2017  
[4] J~Andrew Bagnell and Jeff Schneider. ,Covariant Policy Search, IJCAI, 2003  
[5] Dzmitry Bahdanau‚ Philemon Brakel‚ Kelvin Xu‚ Anirudh Goyal‚ Ryan Lowe‚ Joelle Pineau‚ Aaron Courville‚ and Yoshua Bengio. ,An Actor-Critic Algorithm for Sequence Prediction, ICLR, 2017  
[6] Leemon~C Baird~III. ,Advantage Updating, Technical report, DTIC, 1993.  
[7] Nir Baram‚ Oron Anschel‚ and Shie Mannor. ,Model-Based Adversarial Imitation Learning, NIPS Workshop on Deep Reinforcement Learning, 2016  
[8] Andrew~G Barto‚ Richard~S Sutton‚ and Charles~W Anderson. ,Neuronlike Adaptive Elements That Can Solve Difficult Learning Control Problems, IEEE Trans. on Systems, Man, and Cybernetics, iss.5, pp.834-846, 1983  
[9] Marc~G Bellemare‚ Yavar Naddaf‚ Joel Veness‚ and Michael Bowling. ,The Arcade Learning Environment: An Evaluation Platform for General Agents, IJCAI, 2015  
[10] Marc~G Bellemare‚ Sriram Srinivasan‚ Georg Ostrovski‚ Tom Schaul‚ David Saxton‚ and Rémi Munos. ,Unifying Count-Based Exploration and Intrinsic Motivation, NIPS, 2016  
[11] Marc~G Bellemare‚ Will Dabney‚ and Rémi Munos. ,A Distributional Perspective on Reinforcement Learning, ICML, 2017  
[12] Richard Bellman. ,On the Theory of Dynamic Programming, PNAS, vol.38, iss.8, pp.716-719, 1952  
[13] Yoshua Bengio‚ Jérôme Louradour‚ Ronan Collobert‚ and Jason Weston. ,Curriculum Learning, ICML, 2009  
[14] Yoshua Bengio‚ Aaron Courville‚ and Pascal Vincent. ,Representation Learning: A Review and New Perspectives, IEEE Trans. on Pattern Analysis and Machine intelligence, vol.35, iss.8, pp.1798-1828, 2013  
[15] Dimitri~P Bertsekas. ,Dynamic Programming and Suboptimal Control: A Survey from ADP to MPC, European Journal of Control, vol.11, iss.4-5, pp. 310-334, 2005  
[16] Greg Brockman‚ Vicki Cheung‚ Ludwig Pettersson‚ Jonas Schneider‚ John Schulman‚ Jie Tang‚ and Wojciech Zaremba. ,OpenAI Gym, arXiv:1606.01540, 2016  
[17] Lucian Busoniu‚ Robert Babuska‚ and Bart De~Schutter. ,A Comprehensive survey of Multiagent Reinforcement Learning, IEEE Trans. on Systems, Man, And Cybernetics, 2008  
[18] Murray Campbell‚ A~Joseph Hoane‚ and Feng-hsiung Hsu. ,Deep Blue, Artificial intelligence, vol.134, iss.1-2, pp.57-83, 2002  
[19] Rich Caruana. ,Multitask Learning, Machine Learning, vol.28, iss.1, pp.41-75, 1997  
[20] Silvia Chiappa‚ Sébastien Racaniere‚ Daan Wierstra‚ and Shakir Mohamed. ,Recurrent Environment Simulators, ICLR, 2017  
[21] Paul Christiano‚ Zain Shah‚ Igor Mordatch‚ Jonas Schneider‚ Trevor Blackwell‚ Joshua Tobin‚ Pieter Abbeel‚ and Wojciech Zaremba. ,Transfer from Simulation to Real World through Learning Deep Inverse Dynamics Model, arXiv:1610.03518, 2016  
[22] Giuseppe Cuccu‚ Matthew Luciw‚ Jürgen Schmidhuber‚ and Faustino Gomez. ,Intrinsically Motivated Neuroevolution for Vision-Based Reinforcement Learning, ICDL, 2011  
[23] Peter Dayan. ,Improving Generalization for Temporal Difference Learning: The Successor Representation, Neural Computation, vol.5, iss.4, pp.613-624, 1993  
[24] Jeffrey Dean‚ Greg Corrado‚ Rajat Monga‚ Kai Chen‚ Matthieu Devin‚ Mark Mao‚ Andrew Senior‚ Paul Tucker‚ Ke~Yang‚ Quoc~V Le‚ et~al. ,Large Scale Distributed Deep Networks, NIPS, 2012  
[25] Marc~P Deisenroth‚ Gerhard Neumann‚ and Jan Peters. ,A Survey on Policy Search for Robotics, Foundations and Trends{\textregistered, vol.2, iss.1--2, 2013  
[26] Misha Denil‚ Pulkit Agrawal‚ Tejas~D Kulkarni‚ Tom Erez‚ Peter Battaglia‚ and Nando de~Freitas. ,Learning to Perform Physics Experiments via Deep Reinforcement Learning, ICLR, 2017  
[27] Yan Duan‚ Xi~Chen‚ Rein Houthooft‚ John Schulman‚ and Pieter Abbeel. ,Benchmarking Deep Reinforcement Learning for Continuous Control, ICML, 2016  
[28] Yan Duan‚ John Schulman‚ Xi~Chen‚ Peter~L Bartlett‚ Ilya Sutskever‚ and Pieter Abbeel. ,RL$^2$: Fast Reinforcement Learning via Slow Reinforcement Learning, NIPS Workshop on Deep Reinforcement Learning, 2016  
[29] Gabriel Dulac-Arnold‚ Richard Evans‚ Hado van Hasselt‚ Peter Sunehag‚ Timothy Lillicrap‚ Jonathan Hunt‚ Timothy Mann‚ Theophane Weber‚ Thomas Degris‚ and Ben Coppin. ,Deep Reinforcement Learning in Large Discrete Action Spaces, arXiv:1512.07679, 2015  
[30] David Ferrucci‚ Eric Brown‚ Jennifer Chu-Carroll‚ James Fan‚ David Gondek‚ Aditya~A Kalyanpur‚ Adam Lally‚ J~William Murdock‚ Eric Nyberg‚ John Prager‚ et~al. ,Building Watson: An Overview of the DeepQA Project, AI Magazine, vol.31, iss.3, pp.59-79, 2010  
[31] Chelsea Finn‚ Xin~Yu Tan‚ Yan Duan‚ Trevor Darrell‚ Sergey Levine‚ and Pieter Abbeel. ,Deep Spatial Autoencoders for Visuomotor Learning, ICRA, 2016  
[32] Jakob Foerster‚ Yannis~M Assael‚ Nando de~Freitas‚ and Shimon Whiteson. ,Learning to Communicate with Deep Multi-Agent Reinforcement Learning, NIPS, 2016  
[33] Michael~C Fu. ,Gradient Estimation, Handbooks in Operations Research and Management Science, vol.13:, pp.575-616, 2006  
[34] Marta Garnelo‚ Kai Arulkumaran‚ and Murray Shanahan. ,Towards Deep Symbolic Reinforcement Learning, NIPS Workshop on Deep Reinforcement Learning, 2016  
[35] Peter~W Glynn. ,Likelihood Ratio Gradient Estimation for Stochastic Systems, Communications of the ACM, vol.33, iss.10, pp.75-84, 1990  
[36] Faustino Gomez and Jürgen Schmidhuber. ,Evolving Modular Fast-Weight Networks for Control, ICANN, 2005  
[37] Ian Goodfellow‚ Jean Pouget-Abadie‚ Mehdi Mirza‚ Bing Xu‚ David Warde-Farley‚ Sherjil Ozair‚ Aaron Courville‚ and Yoshua Bengio. ,Generative Adversarial Nets, NIPS, 2014  
[38] Audrunas Gruslys‚ Mohammad~Gheshlaghi Azar‚ Marc~G Bellemare‚ and Rémi Munos. ,The Reactor: A Sample-Efficient Actor-Critic Architecture, arXiv:1704.04651, 2017  
[39] Shixiang Gu‚ Timothy Lillicrap‚ Ilya Sutskever‚ and Sergey Levine. ,Continuous Deep Q-Learning with Model-Based Acceleration, ICLR, 2016  
[40] Shixiang Gu‚ Timothy Lillicrap‚ Zoubin Ghahramani‚ Richard~E Turner‚ and Sergey Levine. ,Q-Prop: Sample-Efficient Policy Gradient with an Off-Policy Critic, ICLR, 2017  
[41] Mance~E Harmon and Leemon~C Baird~III. ,Multi-Player Residual Advantage Learning with General Function Approximation, Technical report, DTIC, 1996.  
[42] Matthew Hausknecht and Peter Stone. ,Deep Recurrent Q-Learning for Partially Observable MDPs, AAAI Fall Symposium Series, 2015  
[43] Nicolas Heess‚ Jonathan~J Hunt‚ Timothy~P Lillicrap‚ and David Silver. ,Memory-Based Control with Recurrent Neural Networks, NIPS Workshop on Deep Reinforcement Learning, 2015  
[44] Nicolas Heess‚ Gregory Wayne‚ David Silver‚ Tim Lillicrap‚ Tom Erez‚ and Yuval Tassa. ,Learning Continuous Control Policies by Stochastic Value Gradients, NIPS, 2015  
[45] Nicolas Heess‚ Srinivasan Sriram‚ Jay Lemmon‚ Josh Merel‚ Greg Wayne‚ Yuval Tassa‚ Tom Erez‚ Ziyu Wang‚ Ali Eslami‚ Martin Riedmiller‚ et~al. ,Emergence of Locomotion Behaviours in Rich Environments, arXiv:1707.02286, 2017  
[46] Johannes Heinrich and David Silver. ,Deep Reinforcement Learning from Self-Play in Imperfect-Information Games, 2016.  
[47] Todd Hester‚ Matej Vecerik‚ Olivier Pietquin‚ Marc Lanctot‚ Tom Schaul‚ Bilal Piot‚ Andrew Sendonaris‚ Gabriel Dulac-Arnold‚ Ian Osband‚ John Agapiou‚ et~al. ,Learning from Demonstrations for Real World Reinforcement Learning, arXiv:1704.03732, 2017  
[48] Geoffrey Hinton‚ Oriol Vinyals‚ and Jeff Dean. ,Distilling the Knowledge in a Neural Network, 2014.  
[49] Jonathan Ho and Stefano Ermon. ,Generative Adversarial Imitation Learning, NIPS, 2016  
[50] Rein Houthooft‚ Xi~Chen‚ Yan Duan‚ John Schulman‚ Filip De~Turck‚ and Pieter Abbeel. ,VIME: Variational Information Maximizing Exploration, NIPS, 2016  
[51] Ahmed Hussein‚ Mohamed~Medhat Gaber‚ and Eyad Elyan. ,Deep Active Learning for Autonomous Navigation, EANN, 2016  
[52] Max Jaderberg‚ Volodymyr Mnih‚ Wojciech~Marian Czarnecki‚ Tom Schaul‚ Joel~Z Leibo‚ David Silver‚ and Koray Kavukcuoglu. ,Reinforcement Learning with Unsupervised Auxiliary Tasks, ICLR, 2017  
[53] Matthew Johnson‚ Katja Hofmann‚ Tim Hutton‚ and David Bignell. ,The Malmo Platform for Artificial Intelligence Experimentation, IJCAI, 2016  
[54] Leslie~P Kaelbling‚ Michael~L Littman‚ and Anthony~R Cassandra. ,Planning and Acting in Partially Observable Stochastic Domains, Artificial intelligence, vol.101, iss.1, pp.99-134, 1998  
[55] Sham~M Kakade. ,A Natural Policy Gradient, NIPS, 2002  
[56] Ken Kansky‚ Tom Silver‚ David~A Mély‚ Mohamed Eldawy‚ Miguel Lázaro-Gredilla‚ Xinghua Lou‚ Nimrod Dorfman‚ Szymon Sidor‚ Scott Phoenix‚ and Dileep George. ,Schema Networks: Zero-Shot Transfer with a Generative Causal Model of Intuitive Physics, ICML, 2017  
[57] Hilbert~J Kappen. ,Path Integrals and Symmetry Breaking for Optimal Control Theory, JSTAT, vol.2005, iss.11, 2005  
[58] Diederik~P Kingma and Max Welling. ,Auto-Encoding Variational Bayes, ICLR, 2014  
[59] Nate Kohl and Peter Stone. ,Policy Gradient Reinforcement Learning for Fast Quadrupedal Locomotion, ICRA, 2004  
[60] Vijay~R Konda and John~N Tsitsiklis. ,On Actor-Critic Algorithms, SICON, vol.42, iss.4, pp.1143-1166, 2003  
[61] Jan Koutník‚ Giuseppe Cuccu‚ Jürgen Schmidhuber‚ and Faustino Gomez. ,Evolving Large-Scale Neural Networks for Vision-Based Reinforcement Learning, GECCO, 2013  
[62] Tejas~D Kulkarni‚ Karthik Narasimhan‚ Ardavan Saeedi‚ and Josh Tenenbaum. ,Hierarchical Deep Reinforcement Learning: Integrating Temporal Abstraction and Intrinsic Motivation, NIPS, 2016  
[63] Tejas~D Kulkarni‚ Ardavan Saeedi‚ Simanta Gautam‚ and Samuel~J Gershman. ,Deep Successor Reinforcement Learning, NIPS Workshop on Deep Reinforcement Learning, 2016  
[64] Tze~Leung Lai and Herbert Robbins. ,Asymptotically Efficient Adaptive Allocation Rules, Advances in Applied Mathematics, vol.6, iss.1, pp. 4-22, 1985  
[65] Brenden~M Lake‚ Tomer~D Ullman‚ Joshua~B Tenenbaum‚ and Samuel~J Gershman. ,Building Machines That Learn and Think Like People, The Behavioral and Brain Sciences, 2016  
[66] Sascha Lange‚ Martin Riedmiller‚ and Arne Voigtlander. ,Autonomous Reinforcement Learning on Raw Visual Input Data in a Real World Application, IJCNN, 2012  
[67] Yann LeCun‚ Yoshua Bengio‚ and Geoffrey Hinton. ,Deep Learning, Nature, vol.521, iss.7553, pp.436-444, 2015  
[68] Joel~Z Leibo‚ Vinicius Zambaldi‚ Marc Lanctot‚ Janusz Marecki‚ and Thore Graepel. ,Multi-Agent Reinforcement Learning in Sequential Social Dilemmas, AAMAS, 2017  
[69] Sergey Levine and Pieter Abbeel. ,Learning Neural Network Policies with Guided Policy Search under Unknown Dynamics, NIPS, 2014  
[70] Sergey Levine and Vladlen Koltun. ,Guided Policy Search, ICLR, 2013  
[71] Sergey Levine‚ Chelsea Finn‚ Trevor Darrell‚ and Pieter Abbeel. ,End-to-End Training of Deep Visuomotor Policies, JMLR, vol.17, iss.39, pp.1-40, 2016  
[72] Sergey Levine‚ Peter Pastor‚ Alex Krizhevsky‚ and Deirdre Quillen. ,Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection, ISER, 2016  
[73] Ke~Li and Jitendra Malik. ,Learning to Optimize, 2017.  
[74] Xiujun Li‚ Lihong Li‚ Jianfeng Gao‚ Xiaodong He‚ Jianshu Chen‚ Li~Deng‚ and Ji~He. ,Recurrent Reinforcement Learning: A Hybrid Approach, arXiv:1509.03044, 2015  
[75] Yuxi Li. ,Deep Reinforcement Learning: An Overview, arXiv:1701.07274, 2017  
[76] Timothy~P Lillicrap‚ Jonathan~J Hunt‚ Alexander Pritzel‚ Nicolas Heess‚ Tom Erez‚ Yuval Tassa‚ David Silver‚ and Daan Wierstra. ,Continuous Control with Deep Reinforcement Learning, ICLR, 2016  
[77] Long-Ji Lin. ,Self-Improving Reactive Agents Based on Reinforcement Learning, Planning and Teaching, Machine Learning, vol.8, iss.3--4, pp.293-321, 1992  
[78] Luke Metz‚ Julian Ibarz‚ Navdeep Jaitly‚ and James Davidson. ,Discrete Sequential Prediction of Continuous Actions for Deep RL, arXiv:1705.05035, 2017  
[79] Piotr Mirowski‚ Razvan Pascanu‚ Fabio Viola‚ Hubert Soyer‚ Andy Ballard‚ Andrea Banino‚ Misha Denil‚ Ross Goroshin‚ Laurent Sifre‚ Koray Kavukcuoglu‚ et~al. ,Learning to Navigate in Complex Environments, ICLR, 2017  
[80] Volodymyr Mnih‚ Nicolas Heess‚ Alex Graves‚ and Koray Kavukcuoglu. ,Recurrent Models of Visual Attention, NIPS, 2014  
[81] Volodymyr Mnih‚ Koray Kavukcuoglu‚ David Silver‚ Andrei~A Rusu‚ Joel Veness‚ Marc~G Bellemare‚ Alex Graves‚ Martin Riedmiller‚ Andreas~K Fidjeland‚ Georg Ostrovski‚ et~al. ,Human-Level Control through Deep Reinforcement Learning, Nature, vol.518, iss.7540, pp.529-533, 2015  
[82] Volodymyr Mnih‚ Adria~Puigdomenech Badia‚ Mehdi Mirza‚ Alex Graves‚ Timothy~P Lillicrap‚ Tim Harley‚ David Silver‚ and Koray Kavukcuoglu. ,Asynchronous Methods for Deep Reinforcement Learning, ICLR, 2016  
[83] Shakir Mohamed and Danilo Jimenez~Rezende. ,Variational Information Maximisation for Intrinsically Motivated Reinforcement Learning, NIPS, 2015  
[84] Andrew~William Moore. ,Efficient Memory-Based Learning for Robot Control, Technical report, University of Cambridge, Computer Laboratory, 1990.  
[85] Rémi Munos‚ Tom Stepleton‚ Anna Harutyunyan‚ and Marc~G Bellemare. ,Safe and Efficient Off-Policy Reinforcement Learning, NIPS, 2016  
[86] Ofir Nachum‚ Mohammad Norouzi‚ Kelvin Xu‚ and Dale Schuurmans. ,Bridging the Gap Between Value and Policy Based Reinforcement Learning, arXiv:1702.08892, 2017  
[87] Anusha Nagabandi‚ Gregory Kahn‚ Ronald~S Fearing‚ and Sergey Levine. ,Neural Network Dynamics for Model-Based Deep Reinforcement Learning with Model-Free Fine-Tuning, arXiv:1708.02596, 2017  
[88] Arun Nair‚ Praveen Srinivasan‚ Sam Blackwell‚ Cagdas Alcicek‚ Rory Fearon‚ Alessandro De~Maria‚ Vedavyas Panneershelvam‚ Mustafa Suleyman‚ Charles Beattie‚ Stig Petersen‚ et~al. ,Massively Parallel Methods for Deep Reinforcement Learning, ICML Workshop on Deep Learning, 2015  
[89] Andrew~Y Ng and Stuart~J Russell. ,Algorithms for Inverse Reinforcement Learning, ICML, 2000  
[90] Andrew~Y Ng‚ Adam Coates‚ Mark Diel‚ Varun Ganapathi‚ Jamie Schulte‚ Ben Tse‚ Eric Berger‚ and Eric Liang. ,Autonomous Inverted Helicopter Flight via Reinforcement Learning, Experimental Robotics, 2006  
[91] Brendan O’Donoghue‚ Rémi Munos‚ Koray Kavukcuoglu‚ and Volodymyr Mnih. ,PGQ: Combining Policy Gradient and Q-Learning, ICLR, 2017  
[92] Junhyuk Oh‚ Xiaoxiao Guo‚ Honglak Lee‚ Richard~L Lewis‚ and Satinder Singh. ,Action-Conditional Video Prediction using Deep Networks in Atari Games, NIPS, 2015  
[93] Junhyuk Oh‚ Valliappa Chockalingam‚ Satinder Singh‚ and Honglak Lee. ,Control of Memory, Active Perception, and Action in Minecraft, ICLR, 2016  
[94] Ian Osband‚ Charles Blundell‚ Alexander Pritzel‚ and Benjamin Van~Roy. ,Deep Exploration via Bootstrapped DQN, NIPS, 2016  
[95] Emilio Parisotto and Ruslan Salakhutdinov. ,Neural Map: Structured Memory for Deep Reinforcement Learning, arXiv:1702.08360, 2017  
[96] Emilio Parisotto‚ Jimmy~L Ba‚ and Ruslan Salakhutdinov. ,Actor-Mimic: Deep Multitask and Transfer Reinforcement Learning, ICLR, 2016  
[97] Deepak Pathak‚ Pulkit Agrawal‚ Alexei~A Efros‚ and Trevor Darrell. ,Curiosity-Driven Exploration by Self-supervised Prediction, ICML, 2017  
[98] Peng Peng‚ Ying Wen‚ Yaodong Yang‚ Quan Yuan‚ Zhenkun Tang‚ Haitao Long‚ and Jun Wang. ,Multiagent Bidirectionally-Coordinated Nets: Emergence of Human-level Coordination in Learning to Play StarCraft Combat Games, arXiv:1703.10069, 2017  
[99] Dean~A Pomerleau. ,ALVINN, an Autonomous Land Vehicle in a Neural Network, Technical report, Carnegie Mellon University, Computer Science Department, 1989.  
[100] Alexander Pritzel‚ Benigno Uria‚ Sriram Srinivasan‚ Adrià Puigdomènech‚ Oriol Vinyals‚ Demis Hassabis‚ Daan Wierstra‚ and Charles Blundell. ,Neural Episodic Control, ICML, 2017  
[101] Marc’Aurelio Ranzato‚ Sumit Chopra‚ Michael Auli‚ and Wojciech Zaremba. ,Sequence Level Training with Recurrent Neural Networks, ICLR, 2016  
[102] Benjamin Recht‚ Christopher Re‚ Stephen Wright‚ and Feng Niu. ,Hogwild: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent, NIPS, 2011  
[103] Danilo~Jimenez Rezende‚ Shakir Mohamed‚ and Daan Wierstra. ,Stochastic Backpropagation and Approximate Inference in Deep Generative Models, ICML, 2014  
[104] Martin Riedmiller. ,Neural Fitted Q Iteration---First Experiences with a Data Efficient Neural Reinforcement Learning Method, ECML, 2005  
[105] Stéphane Ross‚ Geoffrey~J Gordon‚ and Drew Bagnell. ,A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning, AISTATS, 2011  
[106] David~E Rumelhart‚ Geoffrey~E Hinton‚ and Ronald~J Williams. ,Learning Representations by Back-Propagating Errors, Cognitive Modeling, vol.5, iss.3, 1988  
[107] Gavin~A Rummery and Mahesan Niranjan. ,{On-line Q-learning using Connectionist Systems, University of Cambridge, Department of Engineering, 1994.  
[108] Andrei~A Rusu‚ Sergio~Gomez Colmenarejo‚ Caglar Gulcehre‚ Guillaume Desjardins‚ James Kirkpatrick‚ Razvan Pascanu‚ Volodymyr Mnih‚ Koray Kavukcuoglu‚ and Raia Hadsell. ,Policy Distillation, ICLR, 2016  
[109] Andrei~A Rusu‚ Neil~C Rabinowitz‚ Guillaume Desjardins‚ Hubert Soyer‚ James Kirkpatrick‚ Koray Kavukcuoglu‚ Razvan Pascanu‚ and Raia Hadsell. ,Progressive Neural Networks, arXiv:1606.04671, 2016  
[110] Andrei~A Rusu‚ Matej Vecerik‚ Thomas Rothörl‚ Nicolas Heess‚ Razvan Pascanu‚ and Raia Hadsell. ,Sim-to-Real Robot Learning from Pixels with Progressive Nets, CoRL, 2017  
[111] Tim Salimans‚ Jonathan Ho‚ Xi~Chen‚ and Ilya Sutskever. ,Evolution Strategies as a Scalable Alternative to Reinforcement Learning, arXiv:1703.03864, 2017  
[112] Tom Schaul‚ Daniel Horgan‚ Karol Gregor‚ and David Silver. ,Universal Value Function Approximators, ICML, 2015  
[113] Tom Schaul‚ John Quan‚ Ioannis Antonoglou‚ and David Silver. ,Prioritized Experience Replay, ICLR, 2016  
[114] Jürgen Schmidhuber. ,A Possibility for Implementing Curiosity and Boredom in Model-Building Neural Controllers, SAB, 1991  
[115] Jürgen Schmidhuber and Rudolf Huber. ,Learning to Generate Artificial Fovea Trajectories for Target Detection, IJNS, vol.2, iss.01n02, pp.125-134, 1991  
[116] John Schulman‚ Nicolas Heess‚ Theophane Weber‚ and Pieter Abbeel. ,Gradient Estimation using Stochastic Computation Graphs, NIPS, 2015  
[117] John Schulman‚ Sergey Levine‚ Pieter Abbeel‚ Michael Jordan‚ and Philipp Moritz. ,Trust Region Policy Optimization, ICML, 2015  
[118] John Schulman‚ Philipp Moritz‚ Sergey Levine‚ Michael Jordan‚ and Pieter Abbeel. ,High-Dimensional Continuous Control using Generalized Advantage Estimation, ICLR, 2016  
[119] John Schulman‚ Pieter Abbeel‚ and Xi~Chen. ,Equivalence Between Policy Gradients and Soft Q-Learning, arXiv:1704.06440, 2017  
[120] John Schulman‚ Filip Wolski‚ Prafulla Dhariwal‚ Alec Radford‚ and Oleg Klimov. ,Proximal Policy Optimization Algorithms, arXiv:1707.06347, 2017  
[121] Bobak Shahriari‚ Kevin Swersky‚ Ziyu Wang‚ Ryan~P Adams‚ and Nando de~Freitas. ,Taking the Human out of the Loop: A Review of Bayesian Optimization, Proc. of the IEEE, vol.104, iss.1, pp.148-175, 2016  
[122] David Silver‚ Guy Lever‚ Nicolas Heess‚ Thomas Degris‚ Daan Wierstra‚ and Martin Riedmiller. ,Deterministic Policy Gradient Algorithms, ICML, 2014  
[123] David Silver‚ Aja Huang‚ Chris~J Maddison‚ Arthur Guez‚ Laurent Sifre‚ George van~den Driessche‚ Julian Schrittwieser‚ Ioannis Antonoglou‚ Veda Panneershelvam‚ Marc Lanctot‚ et~al. ,Mastering the Game of Go with Deep Neural Networks and Tree Search, Nature, vol.529, iss.7587, pp.484-489, 2016  
[124] Satinder Singh‚ Diane Litman‚ Michael Kearns‚ and Marilyn Walker. ,Optimizing Dialogue Management with Reinforcement Learning: Experiments with the NJFun System, JAIR, vol.16:, pp.105-133, 2002  
[125] Ivan Sorokin‚ Alexey Seleznev‚ Mikhail Pavlov‚ Aleksandr Fedorov‚ and Anastasiia Ignateva. ,Deep Attention Recurrent Q-Network, NIPS Workshop on Deep Reinforcement Learning, 2015  
[126] Bradley~C Stadie‚ Pieter Abbeel‚ and Ilya Sutskever. ,Third Person Imitation Learning, ICLR, 2017  
[127] Bradly~C Stadie‚ Sergey Levine‚ and Pieter Abbeel. ,Incentivizing Exploration in Reinforcement Learning with Deep Predictive Models, NIPS Workshop on Deep Reinforcement Learning, 2015  
[128] Alexander~L Strehl‚ Lihong Li‚ Eric Wiewiora‚ John Langford‚ and Michael~L Littman. ,PAC Model-Free Reinforcement Learning, ICML, 2006  
[129] Sainbayar Sukhbaatar‚ Arthur Szlam‚ and Rob Fergus. ,Learning Multiagent Communication with Backpropagation, NIPS, 2016  
[130] Richard~S Sutton and Andrew~G Barto. ,{Reinforcement Learning: An Introduction, MIT Press, 1998.  
[131] Richard~S Sutton‚ Doina Precup‚ and Satinder Singh. ,Between MDPs and Semi-MDPs: A Framework for Temporal Abstraction in Reinforcement Learning, Artificial intelligence, vol.112, iss.1--2, pp. 181-211, 1999  
[132] Gabriel Synnaeve‚ Nantas Nardelli‚ Alex Auvolat‚ Soumith Chintala‚ Timothée Lacroix‚ Zeming Lin‚ Florian Richoux‚ and Nicolas Usunier. ,TorchCraft: A Library for Machine Learning Research on Real-Time Strategy Games, arXiv:1611.00625, 2016  
[133] Aviv Tamar‚ Yi~Wu‚ Garrett Thomas‚ Sergey Levine‚ and Pieter Abbeel. ,Value Iteration Networks, NIPS, 2016  
[134] Haoran Tang‚ Rein Houthooft‚ Davis Foote‚ Adam Stooke‚ Xi~Chen‚ Yan Duan‚ John Schulman‚ Filip De~Turck‚ and Pieter Abbeel. ,\#Exploration: A Study of Count-Based Exploration for Deep Reinforcement Learning, NIPS, 2017  
[135] Yee~Whye Teh‚ Victor Bapst‚ Wojciech~Marian Czarnecki‚ John Quan‚ James Kirkpatrick‚ Raia Hadsell‚ Nicolas Heess‚ and Razvan Pascanu. ,Distral: Robust Multitask Reinforcement Learning, NIPS, 2017  
[136] Gerald Tesauro. ,Temporal Difference Learning and TD-Gammon, Communications of the ACM, vol.38, iss.3, pp.58-68, 1995  
[137] Gerald Tesauro‚ Rajarshi Das‚ Hoi Chan‚ Jeffrey Kephart‚ David Levine‚ Freeman Rawson‚ and Charles Lefurgy. ,Managing Power Consumption and Performance of Computing Systems using Reinforcement Learning, NIPS, 2008  
[138] Chen Tessler‚ Shahar Givony‚ Tom Zahavy‚ Daniel~J Mankowitz‚ and Shie Mannor. ,A Deep Hierarchical Approach to Lifelong Learning in Minecraft, AAAI, 2017  
[139] Emanuel Todorov‚ Tom Erez‚ and Yuval Tassa. ,MuJoCo: A Physics Engine for Model-Based Control, IROS, 2012  
[140] John~N Tsitsiklis and Benjamin Van~Roy. ,Analysis of Temporal-Difference Learning with Function Approximation, NIPS, 1997  
[141] Eric Tzeng‚ Coline Devin‚ Judy Hoffman‚ Chelsea Finn‚ Xingchao Peng‚ Sergey Levine‚ Kate Saenko‚ and Trevor Darrell. ,Towards Adapting Deep Visuomotor Representations from Simulated to Real Environments, WAFR, 2016  
[142] Nicolas Usunier‚ Gabriel Synnaeve‚ Zeming Lin‚ and Soumith Chintala. ,Episodic Exploration for Deep Deterministic Policies: An Application to StarCraft Micromanagement Tasks, ICLR, 2017  
[143] Hado van Hasselt. ,Double Q-Learning, NIPS, 2010  
[144] Hado van Hasselt‚ Arthur Guez‚ and David Silver. ,Deep Reinforcement Learning with Double Q-Learning, AAAI, 2016  
[145] Harm Vanseijen and Rich Sutton. ,A Deeper Look at Planning as Learning from Replay, ICML, 2015  
[146] Alexander Vezhnevets‚ Volodymyr Mnih‚ Simon Osindero‚ Alex Graves‚ Oriol Vinyals‚ John Agapiou‚ and Koray Kavukcuoglu. ,Strategic Attentive Writer for Learning Macro-Actions, NIPS, 2016  
[147] Alexander~Sasha Vezhnevets‚ Simon Osindero‚ Tom Schaul‚ Nicolas Heess‚ Max Jaderberg‚ David Silver‚ and Koray Kavukcuoglu. ,FeUdal Networks for Hierarchical Reinforcement Learning, ICML, 2017  
[148] Oriol Vinyals‚ Timo Ewalds‚ Sergey Bartunov‚ Petko Georgiev‚ Alexander~Sasha Vezhnevets‚ Michelle Yeo‚ Alireza Makhzani‚ Heinrich Küttler‚ John Agapiou‚ Julian Schrittwieser‚ et~al. ,StarCraft II: A New Challenge for Reinforcement Learning, arXiv:1708.04782, 2017  
[149] Niklas Wahlström‚ Thomas~B Schön‚ and Marc~P Deisenroth. ,Learning Deep Dynamical Models from Image Pixels, IFAC SYSID, vol.48, iss.28, 2015  
[150] Niklas Wahlström‚ Thomas~B Schön‚ and Marc~P Deisenroth. ,From Pixels to Torques: Policy Learning with Deep Dynamical Models, ICML Workshop on Deep Learning, 2015  
[151] Jane~X Wang‚ Zeb Kurth-Nelson‚ Dhruva Tirumala‚ Hubert Soyer‚ Joel~Z Leibo‚ Rémi Munos‚ Charles Blundell‚ Dharshan Kumaran‚ and Matt Botvinick. ,Learning to Reinforcement Learn, CogSci, 2017  
[152] Ziyu Wang‚ Nando de~Freitas‚ and Marc Lanctot. ,Dueling Network Architectures for Deep Reinforcement Learning, ICLR, 2016  
[153] Ziyu Wang‚ Victor Bapst‚ Nicolas Heess‚ Volodymyr Mnih‚ Rémi Munos‚ Koray Kavukcuoglu‚ and Nando de~Freitas. ,Sample Efficient Actor-Critic with Experience Replay, ICLR, 2017  
[154] Christopher~JCH Watkins and Peter Dayan. ,Q-Learning, Machine Learning, vol.8, iss.3-4, pp.279-292, 1992  
[155] Manuel Watter‚ Jost Springenberg‚ Joschka Boedecker‚ and Martin Riedmiller. ,Embed to Control: A Locally Linear Latent Dynamics Model for Control from Raw Images, NIPS, 2015  
[156] Théophane Weber‚ Sébastien Racanière‚ David~P Reichert‚ Lars Buesing‚ Arthur Guez‚ Danilo~Jimenez Rezende‚ Adria~Puigdomènech Badia‚ Oriol Vinyals‚ Nicolas Heess‚ Yujia Li‚ et~al. ,Imagination-Augmented Agents for Deep Reinforcement Learning, NIPS, 2017  
[157] Paul~John Werbos. ,Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences, Technical report, Harvard University, Applied Mathematics, 1974.  
[158] Daan Wierstra‚ Alexander Förster‚ Jan Peters‚ and Jürgen Schmidhuber. ,Recurrent Policy Gradients, Logic Journal of the IGPL, vol.18, iss.5, pp.620-634, 2010  
[159] Ronald~J Williams. ,Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning, Machine Learning, vol.8, iss.3-4, pp.229-256, 1992  
[160] Markus Wulfmeier‚ Peter Ondruska‚ and Ingmar Posner. ,Maximum Entropy Deep Inverse Reinforcement Learning, NIPS Workshop on Deep Reinforcement Learning, 2015  
[161] Kelvin Xu‚ Jimmy Ba‚ Ryan Kiros‚ Kyunghyun Cho‚ Aaron~C Courville‚ Ruslan Salakhutdinov‚ Richard~S Zemel‚ and Yoshua Bengio. ,Show, Attend and Tell: Neural Image Caption Generation with Visual Attention, ICML, 2015  
[162] Yuke Zhu‚ Roozbeh Mottaghi‚ Eric Kolve‚ Joseph~J Lim‚ Abhinav Gupta‚ Li~Fei-Fei‚ and Ali Farhadi. ,Target-Driven Visual Navigation in Indoor Scenes using Deep Reinforcement Learning, ICRA, 2017  
[163] Barret Zoph and Quoc~V Le. ,Neural Architecture Search with Reinforcement Learning, ICLR, 2017