Reinforcement learning

Reinforcement learning is an area of ML in which an "agent" learns from the way that their environment responds to their actions. This topic area includes a wide range of RL techniques.

8a. Model-based reinforcement learning - SimPLe
   Citation: Lukasz Kaiser, Mohammad Babaeizadeh, Piotr Milos, Blazej Osinski, Roy H Campbell, Konrad Czechowski, Dumitru Erhan, Chelsea Finn, Piotr Kozakowski, Sergey Levine, Afroz Mohiuddin, Ryan Sepassi, George Tucker, Henryk Michalewski. Model Based Reinforcement Learning for Atari. In Proceedings of the 2020 International Conference on Learning Representations (ICLR '20). [Notebook on MBRL in general, solutions also available] [PDF][Open review] [Website][Blog post][Github (official SimPLe implementation)]

8b. Multi-armed bandits
   Here's a reference book: Aleksandrs Slivkins. Introduction to Multi-Armed Bandits. [PDF] [Blog post overview and associated Github] [Notebooks on multi-armed bandits and application to a recommendation system; solutions also available] [Notebook on Thompson sampling]

8c. Proximal Policy Optimization
   Citation: John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, Oleg Klimov. Proximal Policy Optimization Algorithms. arXiv:1707.06347. 2017. [PDF][Blog post][Notebook (tensor2tensor)][Also in OpenAI baselines] [Also in stable baselines and here's a tutorial notebook][More notebooks] [Also in SpinningUp][Blog post overview on policy gradient algorithms]

8d. Actor-Advantage Critic (A3C/A2C)
   Citation: Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Timothy P. Lillicrap, Tim Harley, David Silver, Koray Kavukcuoglu. Asynchronous Methods for Deep Reinforcement Learning. In Proceedings of the 2016 International Conference on Machine Learning (ICML '16). [PDF][Github (official, no notebook)] [Also in stable baselines, with notebook][Notebook; solutions also available] [Also available in Atari model zoo, with demo notebook - change method to A2C] [A3C notebooks][More notebooks] [Blog post on policy gradient algorithms]

8e. Soft actor-critic
   Citation: Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, Sergey Levine. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. In Proceedings of the 2018 International Conference on Machine Learning (ICML '18). [PDF] [Blog post] [Tutorial notebook (via Tensorflow)] [Also in stable baselines and here's a tutorial notebook (with HER)] [Also in SpinningUp][Blog post overview on policy gradient algorithms]

8f. TD3
Citation: Scott Fujimoto, Herke van Hoof, David Meger. Addressing Function Approximation Error in Actor-Critic Methods. In Proceedings of the 2018 International Conference on Machine Learning (ICML '18). [PDF][Github (official)] [Github (uses PyBullet - doesn't require MujoCo license)] [Notebook] [In stable baselines, and here's a tutorial notebook] [Notebook using stable baselines version][Also in SpinningUp] [Blog post overview on policy gradient algorithms]

8g. Dueling network
Citation: Ziyu Wang, Tom Schaul, Matteo Hessel, Hado van Hasselt, Marc Lanctot, Nando de Freitas. Dueling Network Architectures for Deep Reinforcement Learning. In Proceedings of the 2016 International Conference on Machine Learning (ICML '16). [PDF] [Notebook (not official)] [Another notebook, and another (not official)] [Also in stable baselines and here's a tutorial notebook]

8h. Rainbow
Citation: Matteo Hessel, Joseph Modayil, Hado van Hasselt, Tom Schaul, Georg Ostrovski, Will Dabney, Dan Horgan, Bilal Piot, Mohammad Azar, David Silver. Rainbow: Combining Improvements in Deep Reinforcement Learning. In Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI '18) [Tutorial with notebooks DQN to Rainbow] [Atari model zoo, including notebooks][Dopamine, including notebooks][PDF]

8i. Fast Task Inference with Variational Intrinsic Successor Features
Citation: Steven Hansen, Will Dabney, Andre Barreto, Tom Van de Wiele, David Warde-Farley, Volodymyr Mnih. Fast Task Inference with Variational Intrinsic Successor Features. In Proceedings of the 2020 International Conference on Learning Representations (ICLR '20). [PDF] [Notebook (official via authors)] (Warning: this is very new work.)

8j. MuZero
Citation: Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy Lillicrap, David Silver. Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model. arXiv:1911.08265. 2019. [PDF][Pseudocode (official via authors)][Github with notebook (not official)] [Other unofficial implementations]

Interesting applications (including AutoML!):

8k. Neural architecture search
Citation: Hieu Pham, Melody Y. Guan, Barret Zoph, Quoc V. Le, Jeff Dean. Efficient Neural Architecture Search via Parameter Sharing. In Proceedings of the 35th International Conference on Machine Learning, PMLR 80:4095-4104, 2018. [PDF] [Github (official implementation)][Github (language model implementation)] [Github (unofficial PyTorch implementation)] [Also in nni]

8l. Data augmentation
Citation: Ekin D. Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, Quoc V. Le. AutoAugment: Learning Augmentation Policies from Data. In Proceedings of the 2019 Conference on Computer Vision and Pattern Recognition (CVPR '19).
[PDF] [Blog post] [Github - official][Unofficial example]
[Unofficial implementation (includes notebook)]
[Also check out DeepAugment (with notebook), uses Bayesian optimization instead]
[or FastAutoAugment, for PyTorch]

8m. Tuning RNNs for music generation
Citation: Natasha Jaques, Shixiang Gu, Richard E. Turner, Douglas Eck. Tuning Recurrent Neural Networks With Reinforcement Learning. In 2017 International Conference on Learning Representations (ICLR '17) workshop [PDF]
Citation: Natasha Jaques, Shixiang Gu, Richard E. Turner, Douglas Eck. Generating Music by Fine-Tuning Recurrent Neural Networks with Reinforcement Learning. In Proceedings of the 2020 Deep Reinforcement Learning Workshop, NIPS '16. [PDF]
[Github][Blog post][Notebook]

8n. Reformulating questions
Citation: Christian Buck Jannis Bulian Massimiliano Ciaramita Wojciech Paweł Gajewski Andrea Gesmundo Neil Houlsby Wei Wang. Ask the Right Questions: Active Question Reformulation with Reinforcement Learning. In Proceedings of the Sixth International Conference on Learning Representations (ICRL '18). [PDF]
[Blog post][Github] (Warning: this project is not supported by the authors.)

8o. Text-based games
Citation: Marc-Alexandre Côté, Ákos Kádár, Xingdi Yuan, Ben Kybartas, Tavian Barnes, Emery Fine, James Moore, Ruo Yu Tao, Matthew Hausknecht, Layla El Asri, Mahmoud Adada, Wendy Tay, Adam Trischler. TextWorld: A Learning Environment for Text-based Games. In Proceedings of the Computer Games Workshop at IJCAI 2018. [PDF] [Website] [Github including notebooks]
[Notebook using TextWorld in Colab]