





## REFERENCES

- [1] Raunak Bhattacharyya, Derek J. Phillips, et al. 2019. Simulating emergent properties of human driving behavior using multi-agent reward augmented imitation learning. In *International Conference on Robotics and Automation (ICRA)*.
- [2] L. Busoniu, R. Babuska, and B. De Schutter. 2008. A Comprehensive Survey of Multiagent Reinforcement Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 38, 2 (March 2008), 156–172. <https://doi.org/10.1109/TSMCC.2007.913919>
- [3] Laurent Dinh, David Krueger, and Yoshua Bengio. 2014. NICE: Non-linear independent components estimation. *arXiv preprint arXiv:1410.8516* (2014).
- [4] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. 2016. Density estimation using Real NVP. *arXiv preprint arXiv:1605.08803* (2016).
- [5] Mathieu Germain, Karol Gregor, Iain Murray, and Hugo Larochelle. 2015. Made: Masked autoencoder for distribution estimation. In *International Conference on Machine Learning (ICML)*. 881–889.
- [6] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in Neural Information Processing Systems (NeurIPS)*. 2672–2680.
- [7] Karol Gregor, Ivo Danihelka, Andriy Mnih, Charles Blundell, and Daan Wierstra. 2014. Deep AutoRegressive Networks. In *International Conference on Machine Learning (ICML)*. 1242–1250.
- [8] Jonathan Ho, Xi Chen, Aravind Srinivas, Yan Duan, and Pieter Abbeel. 2019. Flow++: Improving Flow-Based Generative Models with Variational Dequantization and Architecture Design. In *International Conference on Machine Learning (ICML)*, Vol. 97. PMLR, 2722–2730.
- [9] Max Jaderberg, Wojciech M Czarnecki, Iain Dunning, Luke Marris, Guy Lever, Antonio Garcia Castañeda, Charles Beattie, Neil C Rabinowitz, Ari S Morcos, Avraham Ruderman, Nicolas Sonnerat, Tim Green, Louise Deason, Joel Z Leibo, David Silver, Demis Hassabis, Koray Kavukcuoglu, and Thore Graepel. 2019. Human-level performance in 3D multiplayer games with population-based reinforcement learning. *Science* 364, 6443 (May 2019), 859–865.
- [10] Diederik P. Kingma and Prafulla Dhariwal. 2018. Glow: Generative flow with invertible 1x1 convolutions. In *Advances in Neural Information Processing Systems (NeurIPS)*. 10215–10224.
- [11] Diederik P. Kingma, Tim Salimans, Rafal Jozefowicz, Xi Chen, Ilya Sutskever, and Max Welling. 2016. Improved variational inference with inverse autoregressive flow. In *Advances in Neural Information Processing Systems (NeurIPS)*. 4743–4751.
- [12] Diederik P. Kingma and Max Welling. 2013. Auto-Encoding Variational Bayes. In *International Conference on Learning Representations (ICLR)*.
- [13] Marc Lanctot, Vinicius Zambaldi, Audrunas Gruslys, et al. 2017. A unified game-theoretic approach to multiagent reinforcement learning. In *Advances in Neural Information Processing Systems (NeurIPS)*. 4190–4203.
- [14] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. 2015. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971* (2015).
- [15] Gabriel Peyré, Marco Cuturi, et al. 2019. Computational optimal transport. *Foundations and Trends® in Machine Learning* 11, 5-6 (2019), 355–607.
- [16] Dean Pomerleau. 1991. Efficient training of artificial neural networks for autonomous navigation. *Neural Computation* 3 (1991), 88–97.
- [17] Danilo Rezende and Shakir Mohamed. 2015. Variational Inference with Normalizing Flows. In *International Conference on Machine Learning (ICML)*. 1530–1538.
- [18] Edward Schmerling, Karen Leung, Wolf Vollprecht, and Marco Pavone. 2017. Multimodal Probabilistic Model-Based Planning for Human-Robot Interaction. In *IEEE International Conference on Robotics and Automation (ICRA)*. 1–9.
- [19] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).
- [20] Sriram Srinivasan, Marc Lanctot, Vinicius Zambaldi, Julien Pérolat, Karl Tuyls, Rémi Munos, and Michael Bowling. 2018. Actor-critic policy optimization in partially observable multiagent environments. In *Advances in Neural Information Processing Systems (NeurIPS)*. 3422–3435.