# Bibliography and Review of RL literature

seancarver@datamachines.io

September 10, 2020

## 1 Review

I'd like to put annotations if not a full literature review here, but that will take time.

## References

[1] T. P. Le, N. A. Vien, and T. Chung, "A deep hierarchical reinforcement learning algorithm in partially observable markov decision processes," *Ieee Access*, vol. 6, pp. 49089–49102, 2018.

[2] A. Ma, M. Ouimet, and J. Cortés, "Hierarchical reinforcement learning via dynamic subspace search for multi-agent planning," *Autonomous Robots*, vol. 44, no. 3, pp. 485–503, 2020.

[3] H. Tang, J. Hao, T. Lv, Y. Chen, Z. Zhang, H. Jia, C. Ren, Y. Zheng, Z. Meng, C. Fan, *et al.*, "Hierarchical deep multiagent reinforcement learning with temporal abstraction," *arXiv preprint arXiv:1809.09332*, 2018.

[4] S. Omidshafiei, J. Pazis, C. Amato, J. P. How, and J. Vian, "Deep decentralized multi-task multi-agent reinforcement learning under partial observability," *arXiv preprint arXiv:1703.06182*, 2017.

[5] M. Lanctot, V. Zambaldi, A. Gruslys, A. Lazaridou, K. Tuyls, J. Pérolat, D. Silver, and T. Graepel, "A unified game-theoretic approach to multiagent reinforcement learning," in *Advances in neural information processing systems*, pp. 4190–4203, 2017.

[6] S. Iqbal and F. Sha, "Actor-attention-critic for multi-agent reinforcement learning," in *International Conference on Machine Learning*, pp. 2961–2970, PMLR, 2019.

[7] Y. Yang, R. Luo, M. Li, M. Zhou, W. Zhang, and J. Wang, "Mean field multi-agent reinforcement learning," *arXiv preprint arXiv:1802.05438*, 2018.

[8] J. D. Co-Reyes, Y. Liu, A. Gupta, B. Eysenbach, P. Abbeel, and S. Levine, "Self-consistent trajectory autoencoder: Hierarchical reinforcement learning with trajectory embeddings," *arXiv preprint arXiv:1806.02813*, 2018.

[9] T. Haarnoja, K. Hartikainen, P. Abbeel, and S. Levine, "Latent space policies for hierarchical reinforcement learning," *arXiv preprint arXiv:1804.02808*, 2018.

[10] C. Florensa, Y. Duan, and P. Abbeel, "Stochastic neural networks for hierarchical reinforcement learning," *arXiv preprint arXiv:1704.03012*, 2017.

[11] O. Nachum, S. Gu, H. Lee, and S. Levine, "Near-optimal representation learning for hierarchical reinforcement learning," *arXiv preprint arXiv:1810.01257*, 2018.

[12] R. T. Icarte, E. Waldie, T. Klassen, R. Valenzano, M. Castro, and S. McIlraith, "Learning reward machines for partially observable reinforcement learning," in *Advances in Neural Information Processing Systems*, pp. 15523–15534, 2019.

[13] M. Hausknecht and P. Stone, "Deep recurrent q-learning for partially observable mdps," *arXiv preprint arXiv:1507.06527*, 2015.

[14] J. Heinrich and D. Silver, "Deep reinforcement learning from self-play in imperfect-information games," *arXiv preprint arXiv:1603.01121*, 2016.

[15] V. Krishnamurthy, *Partially observed Markov decision processes*. Cambridge University Press, 2016.

[16] O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D. H. Choi, R. Powell, T. Ewalds, P. Georgiev, *et al.*, "Grandmaster level in starcraft ii using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350–354, 2019.