WEEKLY RESEARCH PROGRESS REPORT: 36

Shangshang Wang
School of Information Science and Technology
ShanghaiTech University
wangshsh2@shanghaitech.edu.cn

1 QUOTE OF LAST WEEK'S PLAN

- Complete beamer presentation on *Toward Information Theory blessed Deep Reinforcement Learning*.
- ICML 2019 workshop on self-supervised learning.
- Graph Machine Learning survey (Wu et al., 2019).
- Inverse RL survey (Arora & Doshi, 2018).
- Causal Inference survey (Yao et al., 2020).
- Machine Teaching survey (Zhu et al., 2018).

2 PLANNED ACCOMPLISHMENTS

- Complete beamer presentation on *Toward Information Theory blessed Deep Reinforcement Learning*.
- ICML 2019 workshop on self-supervised learning.
- Graph Machine Learning survey (Wu et al., 2019).
- Inverse RL survey (Arora & Doshi, 2018).
- Causal Inference survey (Yao et al., 2020).
- Machine Teaching survey (Zhu et al., 2018).
- Boltzmann operater in exploration
- Difference between population-based method and multi-agent framework

3 OTHER ACCOMPLISHMENTS

Paper Reading

- QXplore: Q-learning Exploration by Maximizing Temporal Difference Error (Simmons-Edler et al., 2019)
- The Effects of Memory Replay in Reinforcement Learning (Liu & Zou, 2018)
- A Survey of Learning Causality with Data: Problems and Methods (Liu & Zou, 2018)
- Algorithms for inverse reinforcement learning

Topic learning

- Self-supervised Learning: Autoencoder (Weng, 2018)
- Attention and Transformer: All you need is Attention (Vaswani et al., 2017)
- Capsule Networks (Hinton et al., 2011) and (Sabour et al., 2017)
- Wasserstein Distance and optimal transport: WGAN (Arjovsky et al., 2017)
- Multi-agent Reinforcement Learning: 2018 SJTU MARL tutorial and hd Thesis: Deep Multi-agent Reinforcement Learning

4 Issues and Problem to Solve

- Self-consistency in path-consistency RL
- Take a deeper look: The role of Replay Buffer

5 Next Week's Plan

- Systematically review of probablity and oprimization for shcool courses
- Five kinds of Exploration methods in RL
- Survey the way of reusing data (increase sample efficiency) in RL

References

Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein gan, 2017.

Saurabh Arora and Prashant Doshi. A survey of inverse reinforcement learning: Challenges, methods and progress. *arXiv preprint arXiv:1806.06877*, 2018.

Geoffrey E Hinton, Alex Krizhevsky, and Sida D Wang. Transforming auto-encoders. In *International conference on artificial neural networks*, pp. 44–51. Springer, 2011.

Ruishan Liu and James Zou. The effects of memory replay in reinforcement learning. In *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 478–485. IEEE, 2018.

Sara Sabour, Nicholas Frosst, and Geoffrey E Hinton. Dynamic routing between capsules. In *Advances in neural information processing systems*, pp. 3856–3866, 2017.

Riley Simmons-Edler, Ben Eisner, Daniel Yang, Anthony Bisulco, Eric Mitchell, Sebastian Seung, and Daniel Lee. Qxplore: Q-learning exploration by maximizing temporal difference error, 2019.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pp. 5998–6008, 2017.

Lilian Weng. From autoencoder to beta-vae. *lilianweng.github.io/lil-log*, 2018. URL `http://lilianweng.github.io/lil-log/2018/08/12/from-autoencoder-to-beta-vae.html`.

Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S Yu. A comprehensive survey on graph neural networks. *arXiv preprint arXiv:1901.00596*, 2019.

Liuyi Yao, Zhixuan Chu, Sheng Li, Yaliang Li, Jing Gao, and Aidong Zhang. A survey on causal inference, 2020.

Xiaojin Zhu, Adish Singla, Sandra Zilles, and Anna N. Rafferty. An overview of machine teaching, 2018.