Zurich University
of Applied Sciences

School of
Engineering

P12 – Selfplay for Tic Tac Toe

1. Background research

Amongst several striking ideas, Peter Abbeel prominently mentions "self play" as one important ingredient of future successful AI agents trained with reinforcement learning (RL) in his recent summary talk at EECS [1], where he places this topic in larger context of "learning to learn".

Make yourself acquainted with Prof. Abbeel's ideas from the abovementioned talk on "learning to learn" [4] with a specific focus on how the "self play" he proposes [2][3] works. Compare it with the work of David Silver et al. on the game of Go [5].

2. Learning to play Tic Tac Toe, WarGames-style

In their introductory book to RL, Sutton & Barto relate the example of the game Tic Tac Toe that due to its simplicity can be taught to an agent with a very simple RL technique [6, p. 7-10]. Please implement this example in a way where the agent learns to play Tic Tac Toe by repeated games against itself (i.e., against a non-static opponent that itself improves over time). Please use Python to implement the challenge and fulfill the API for the agent that is implicitly defined in `t3_controller.py` and `t3_engine.py`.

Check the progress ogf your agent while letting it play against humans (as first- and second mover). What do you observe, and why? Note down lessons learned from your development experience. Do you come to similar conclusions as the author of [7]?

References

[1] Peter Abbeel, "Learning to Learn Robotic Control", EECS colloquium, Oct 11, 2017, available online: https://www.youtube.com/watch?v=TERCdog1ddE&t=3575s (Feb 19, 2018)
[2] Maruan Al-Shedivat, Trapit Bansal, Yuri Burda, Ilya Sutskever, Igor Mordatch, Pieter Abbeel, "Continuous Adaptation via Meta-Learning in Nonstationary and Competitive Environments", Oct 10, 2017, available online: https://arxiv.org/abs/1710.03641 (Feb 19, 2018)
[3] Trapit Bansal, Igor Mordatch, Jakub Pachocki, Ilya Sutskever, Szymon Sidor, "Competitive Self-Play", OpenAI blog, Oct 11, 2017, available online: https://blog.openai.com/competitive-self-play/ (Feb 19, 2018)
[4] Chelsea Finn, "Learning to learn", BAIR blog, Jul 18, 2017, available onlinbe: http://bair.berkeley.edu/blog/2017/07/18/learning-to-learn/ (Feb 19, 2018)
[5] David Silver, Julian Schrittwieser, Karen Simonyan, et al., “Mastering the game of Go without human knowledge”, Nature, 550(7676), 354, 2017, available online: https://www.nature.com/articles/nature24270 (Sep 04, 2018)
[6] Richard S. Sutton, Andrew G. Barto, "Reinforcement Learning, 2nd Edition", MIT press (to appear), available online: http://incompleteideas.net/book/bookdraft2017nov5.pdf (Feb 19, 2018)
[7] Matthew Rahtz, "Lessons Learned Reproducing a Deep Reinforcement Learning Paper", Amid Fish Blog, April 06, 2018, available online: http://amid.fish/reproducing-deep-rl (Jul 17, 2018)