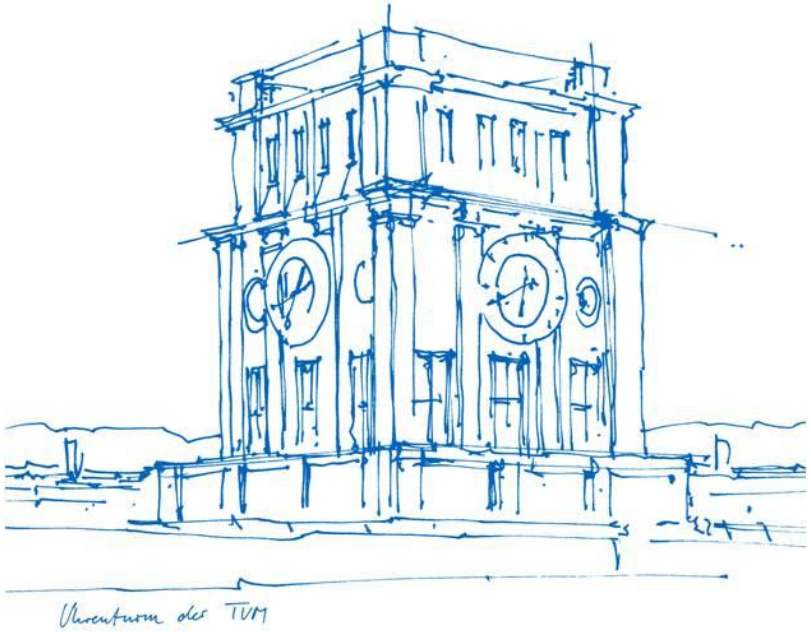
Munchausen RL with Continuous Action Space

Marcel Brucker

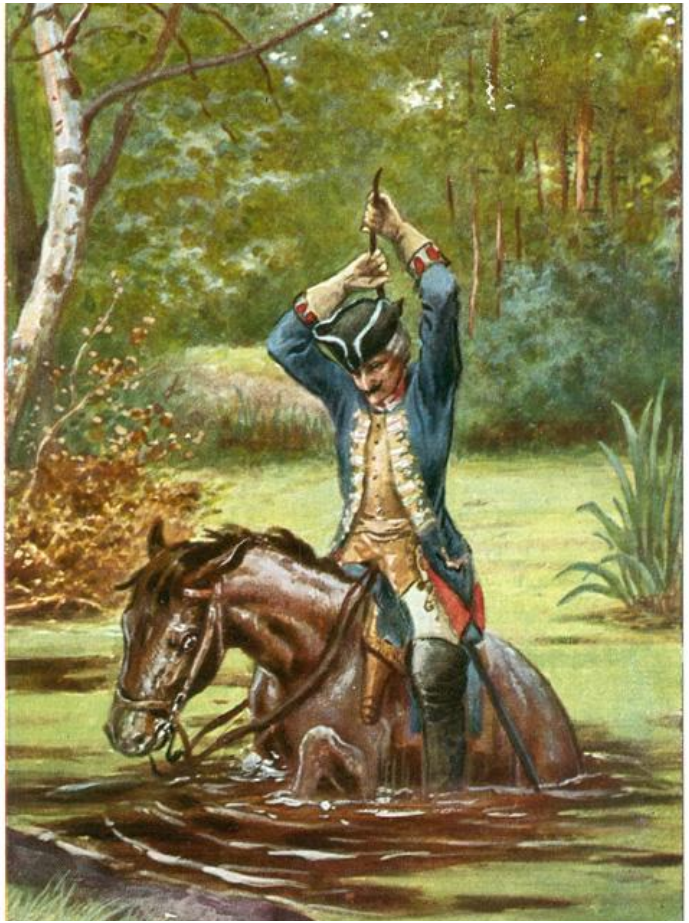
Finn Süberkrüb

Technische Universität München

München, 10.06.2021

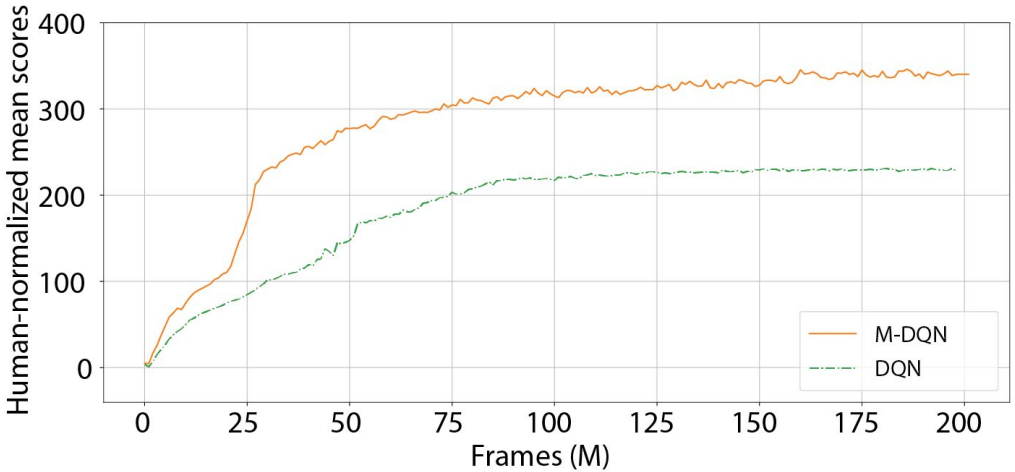


Baron Munchausen pulls himself out of the swamp by his own hair.



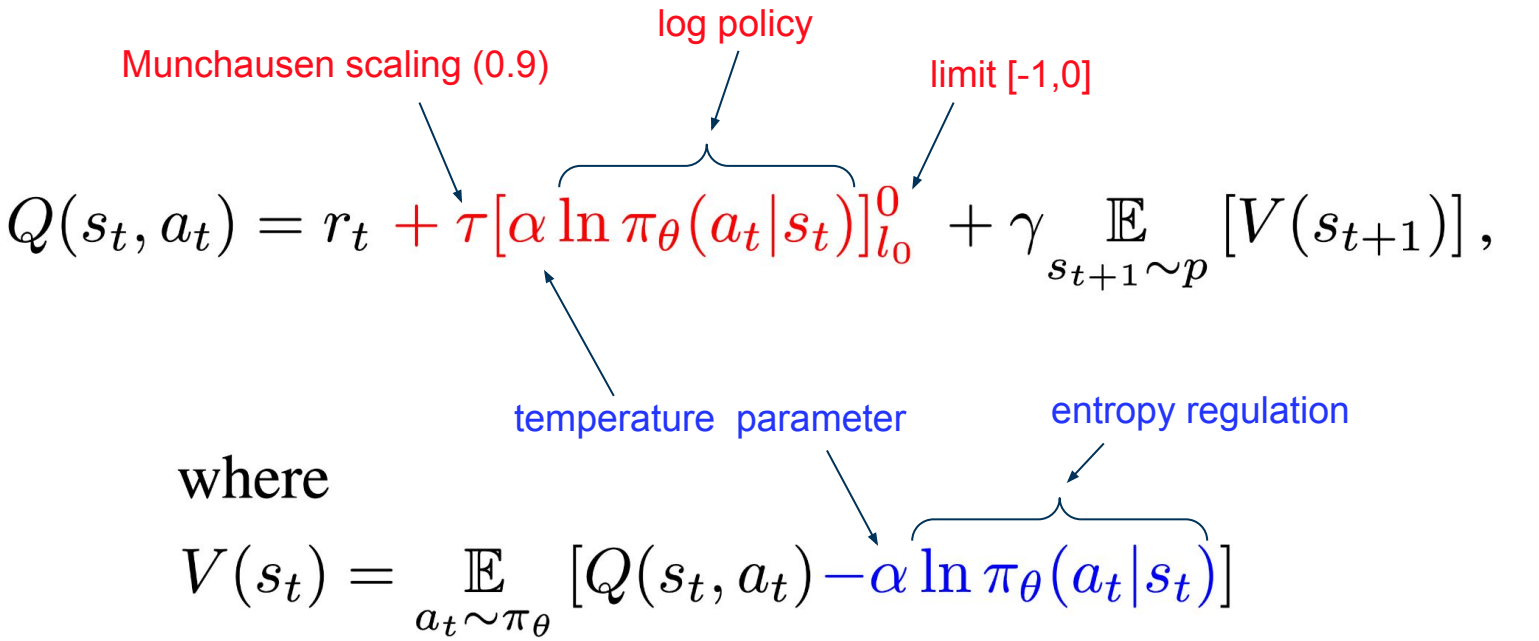
Author: Oskar Herrfurth (Wikimedia Commons)

M-DQN results on 60 Atari games[1]



Human-normalized mean score over the 60 Atari games from the dopamine baselines.

Munchausen RL[1]

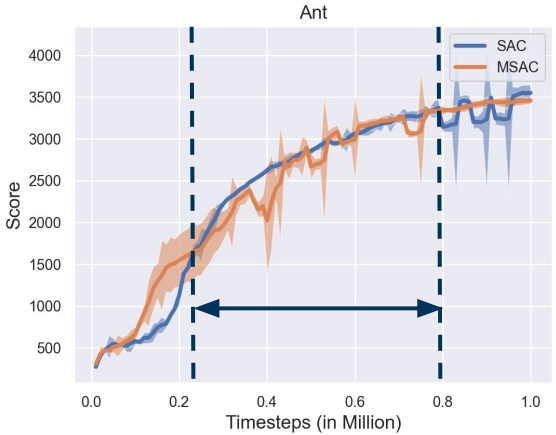


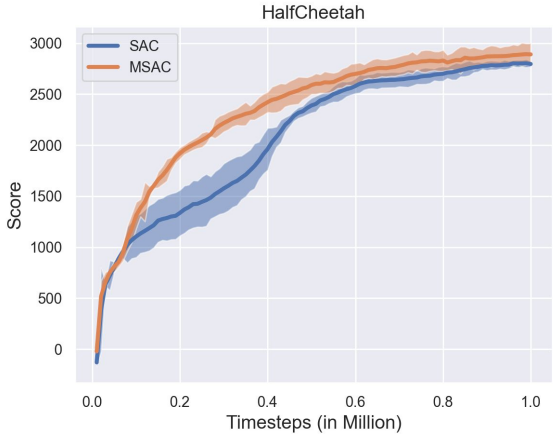
Used Tools

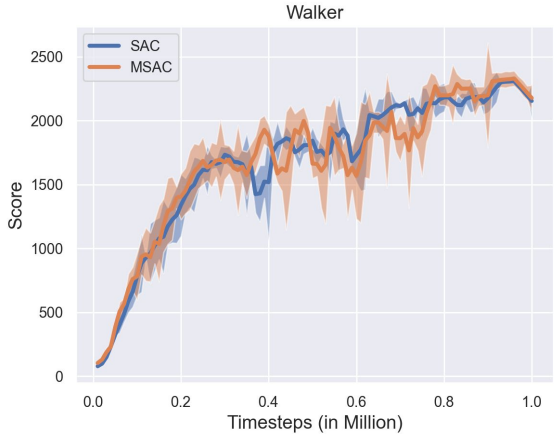
- Stable Baselines3 (SB3)
- RL Baselines3 Zoo

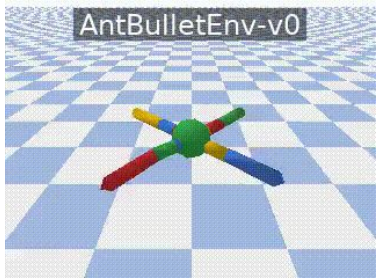
- OpenAI Gym
- PyBullet Gymperium

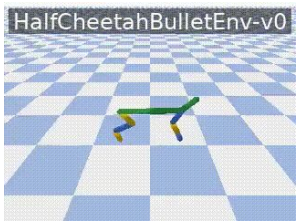
Some Preliminary Results







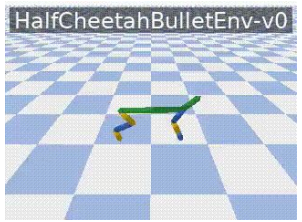


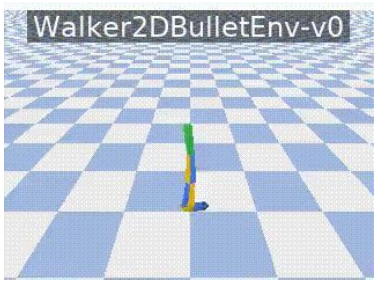


SAC

vs.

MSAC





Next Steps

- More experiments in other environments
- Munchausen parameter optimization (random search)
  - clipping parameters
  - munchhausen scaling
- Can different probability distributions of the selected actions bring improvements?
- Can we find a reason why Munchausen RL does not yet bring so much improvement for our continuous action spaces?

References

[1] Nino Vieillard, Olivier Pietquin and Matthieu Geist (2020). Munchausen Reinforcement Learning. arXiv:2007.14430.

[2] Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel and Sergey Levine (2019). Soft Actor-Critic Algorithms and Applications. arXiv:1812.05905.

[3] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, Sergey Levine (2018). Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. arXiv:1801.01290.

[4] Make a four-legged creature walk forward as fast as possible. https://gym.openai.com/envs/Ant-v2/.

[5] Antonin Raffin, Ashley Hill, Maximilian Ernestus, Adam Gleave, Anssi Kanervisto and Noah Dormann (2019). Stable Baselines3. GitHub repository, https://github.com/DLR-RM/stable-baselines3

[6] Antonin Raffin (2020). RL Baselines3 Zoo. GitHub repository, https://github.com/DLR-RM/rl-baselines3-zoo

[7] Benjamin Ellenberger (2018-2019). PyBullet Gymperium. GitHub repository, https://github.com/benelot/pybullet-gym

[8] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang and Wojciech Zaremba (2016). OpenAI Gym. arXiv:1606.01540.