EVALUATING ROBUSTNESS OF DEEP REINFORCEMENT LEARNING FOR AUTONOMOUS DRIVING

By Bartu Mehmet Ortal

b.m.ortal@student.tudelft.nl

Supervisors

Matthijs Spaan, Moritz Zanger

TU Delft

1) Research Question

How does augmentation of the typical RL objective with a maximum policy entropy term affect training and the robustness of final policies under various testing conditions?

2) Background Information

- Off-policy algorithm
- Incorporates clipped double-Q trick.
- Enables target smoothing.
- Provides entropy regularization.
- Experience replay
- Promotes exploration and prevents converging to a bad local optimum

3) Methodology

Tune hyper-parameters on CarRacing environment

Use Carla environment to simulate traffic

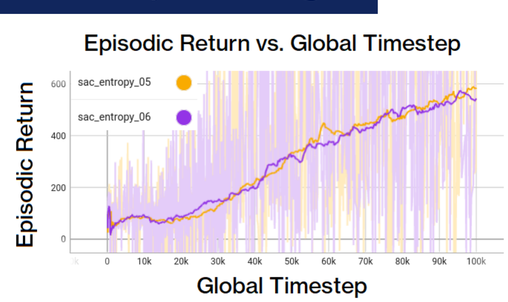
SAC - Implementation already exists in CleanRL repository. This is being used with some modification

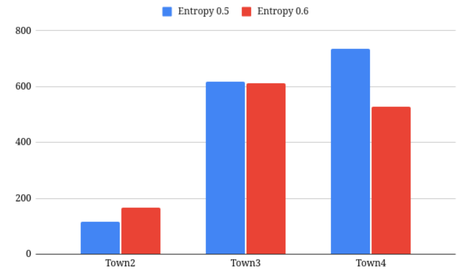
Using the implementation and Carla environment, train 100 thousand steps.

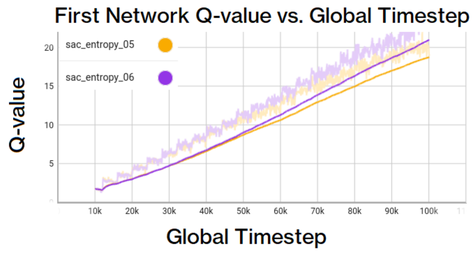
Run evaluation on 3 different maps 10 times each.

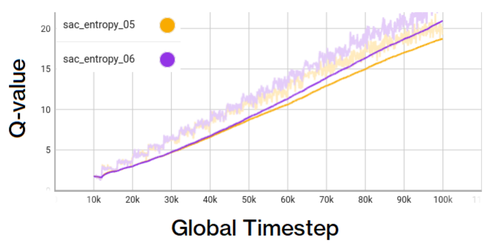
Change entropy value during training to train other agents and repeat the same steps

4) Results / Findings









5) Limitations / Future Work

- Hardware Limitations
- 100 thousand steps is not enough
- More than 10 evaluation per map
- Changing more parameters

6) Conclusion

Increasing entropy increases the robustness of the trained agent.

Decreasing entropy promotes exploitation and trains the agent better for the environment that it is in.

Related Literature

[1]Tuomas Haarnoja et al. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. 2018. arXiv: 1801.01290 [cs.LG].
[2]Wenjie Shi, Shiji Song, and Cheng Wu. Soft Policy Gradient Method for Maximum Entropy Deep Reinforcement Learning. 2019. arXiv: 1909.03198 [cs.LG].
[3]Tuomas Haarnoja, Sehoon Ha, Aurick Zhou, Jie Tan, George Tucker, and Sergey Levine. Learning to walk via deep reinforcement learning, 2019
[4]Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel, and Sergey Levine. Soft actor-critic algorithms and applications, 2019
[5]Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning, 2013