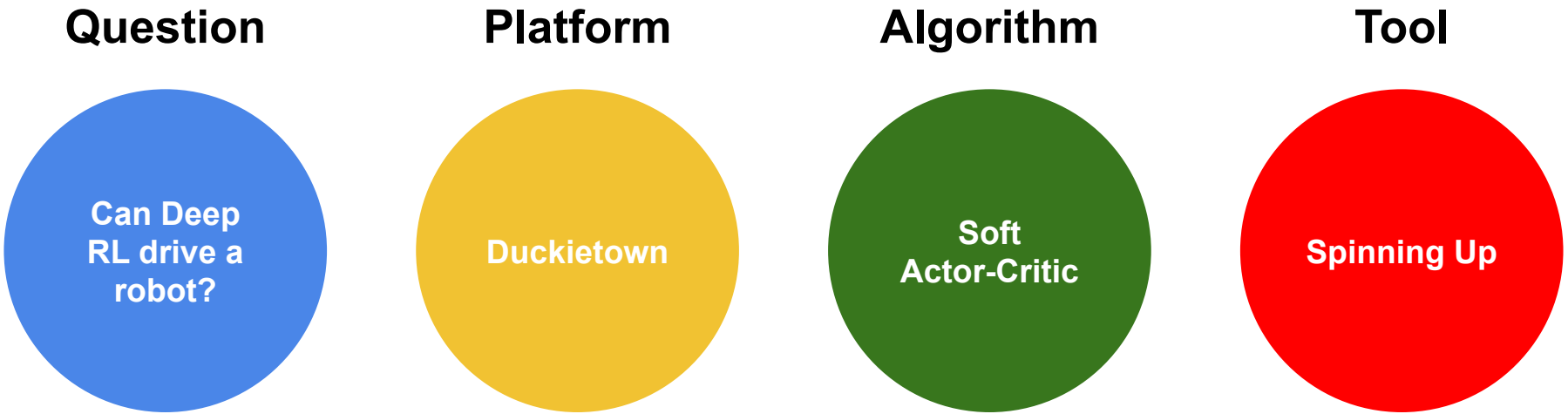
Taxicab: Duckietown with Soft Actor-Critic

UCI CS 175, Roy Fox | 8 March 2019

Team 8: Julian Freedberg, Bryon Tjanaka, Carleton Zhao

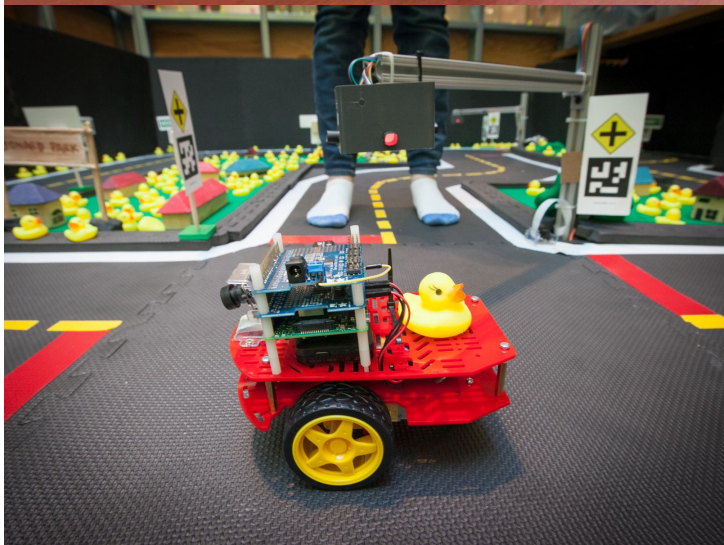
taxicab175.github.io

Project Overview



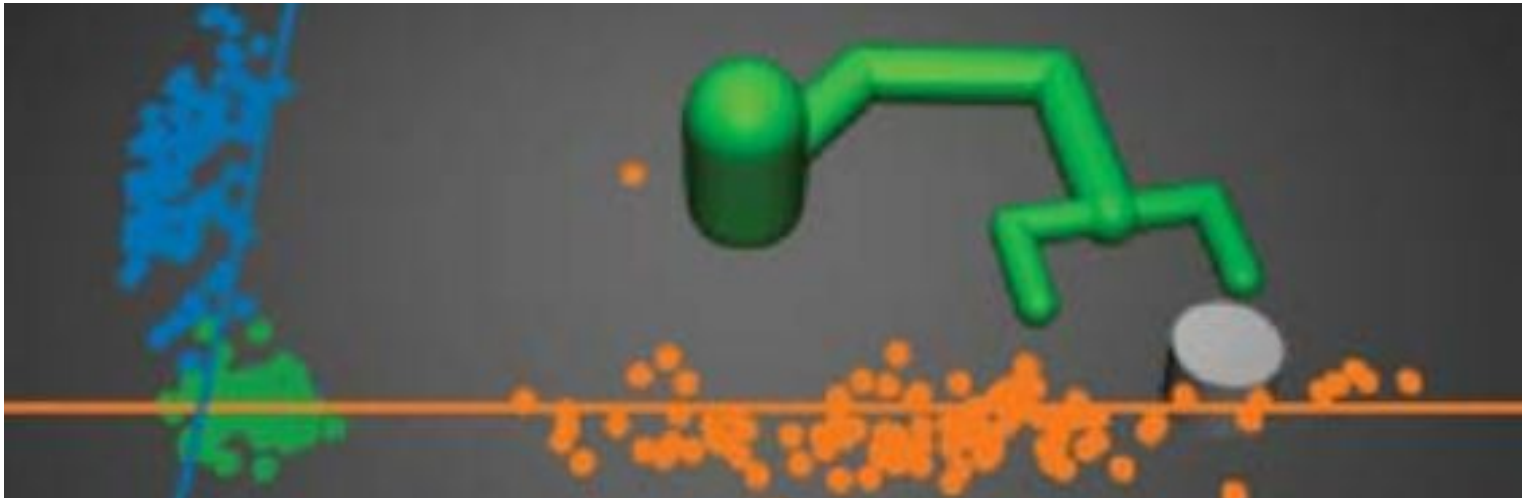
Duckietown

- Self-driving car platform
- AI Driving Olympics
  - Lane Following Challenge
- Gym Duckietown
  - **Observations:** Color image
  - **Actions:** [-1, 1] per wheel
  - **Rewards:** Follow the road



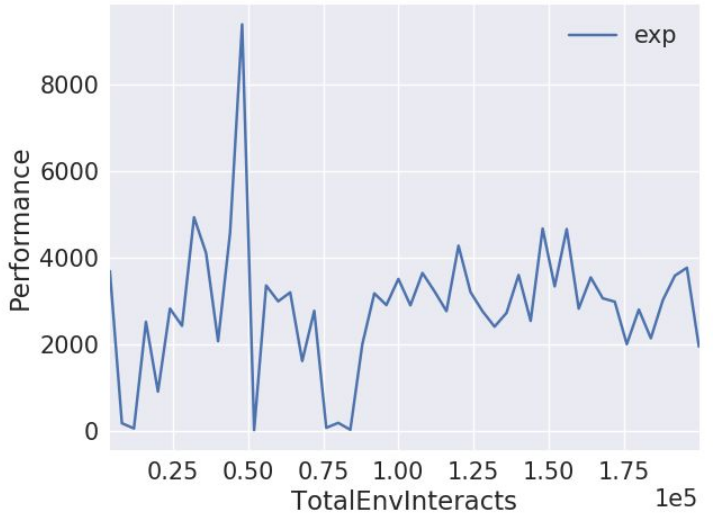
Soft Actor Critic (SAC)

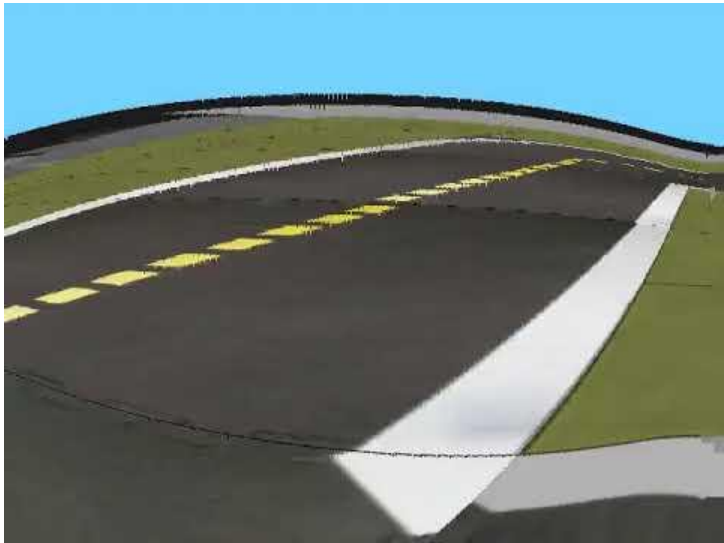
- State-of-the-art
- **Characteristics**
  - Maximum-Entropy
  - Off-policy
- **Networks:** 2 Q-networks + 1 policy network



Milestone 1: Implement and Evaluate SAC

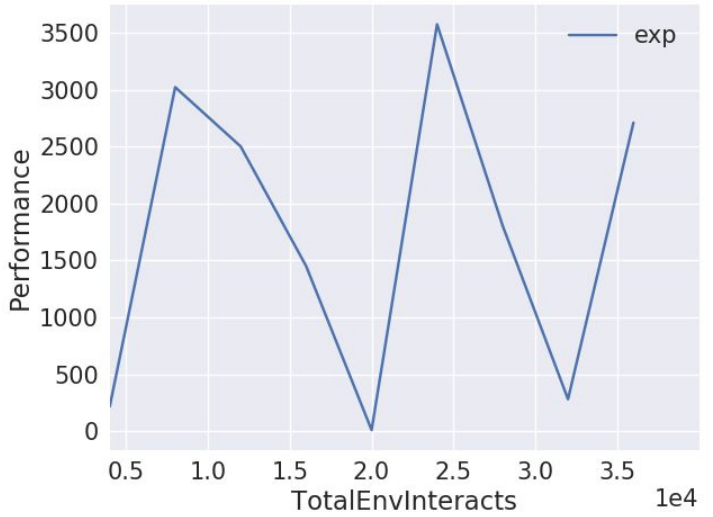
- Objective
  - Run the simulator and evaluate performance
- Methods
  - Train Spinning Up SAC for 200,000 timesteps
- Results
  - Spinning!

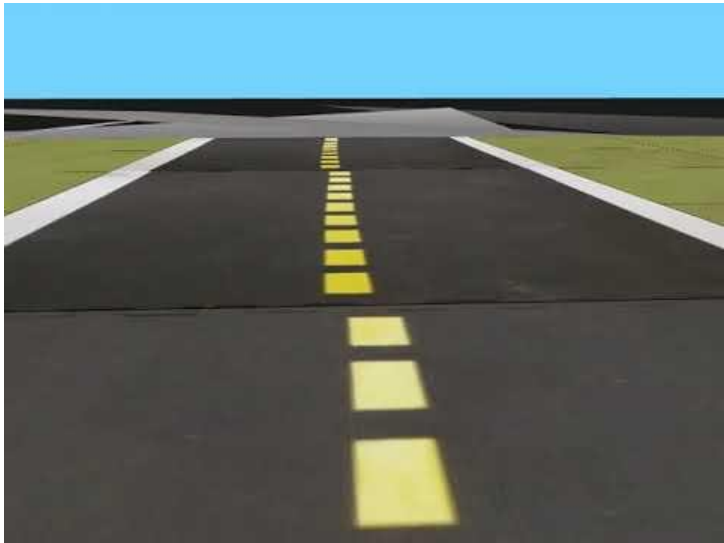




Milestone 2: Short Straight Track

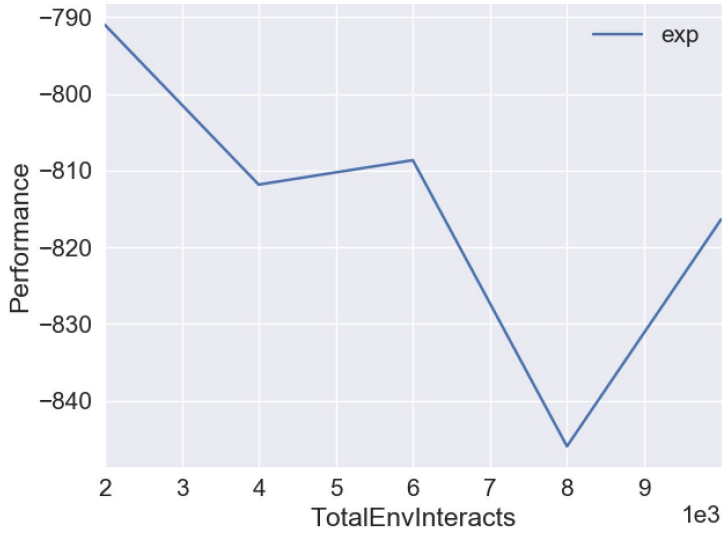
- Objective
  - Drive straight
- Methods
  - Convolutional Layers
  - Straight Track
  - 40,000 training steps
- Results
  - Somewhat Straight





Milestone 3: Longer Straight Track

- Objective
  - Drive straight… reliably
- Revisions
  - Reward function
  - Longer track
- Results
  - More Spinning :(
  - Return: 3000



Challenges

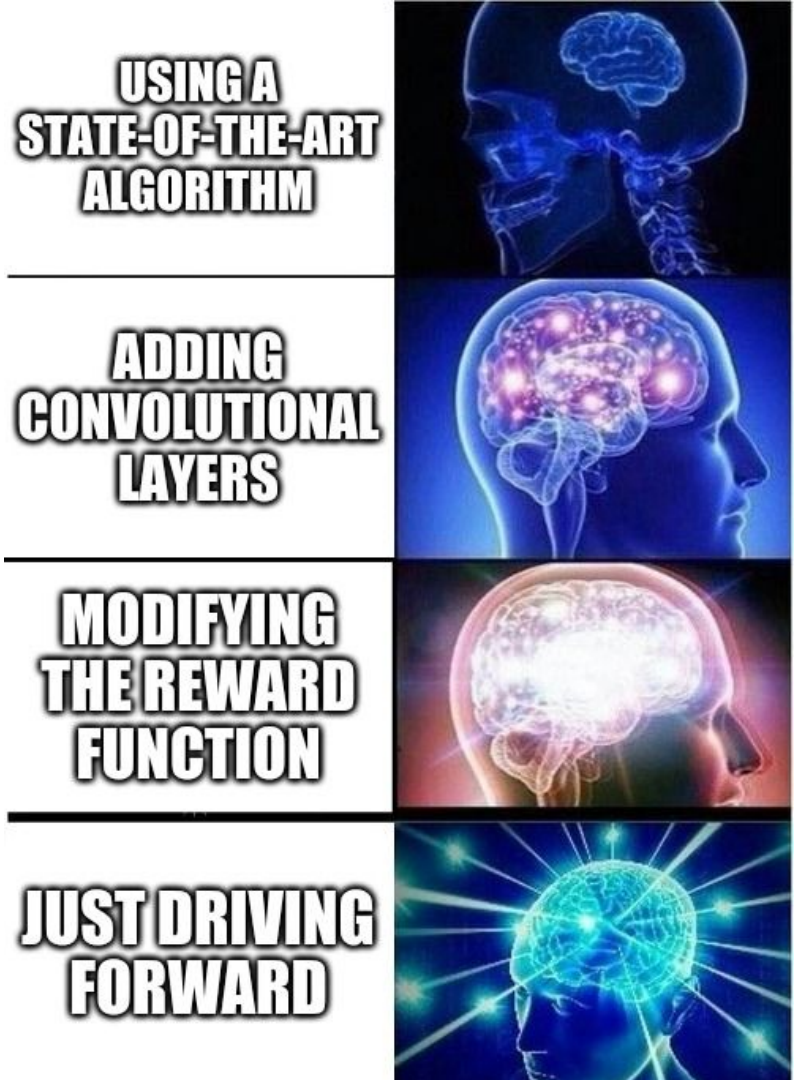
Deep RL

Training

Documentation

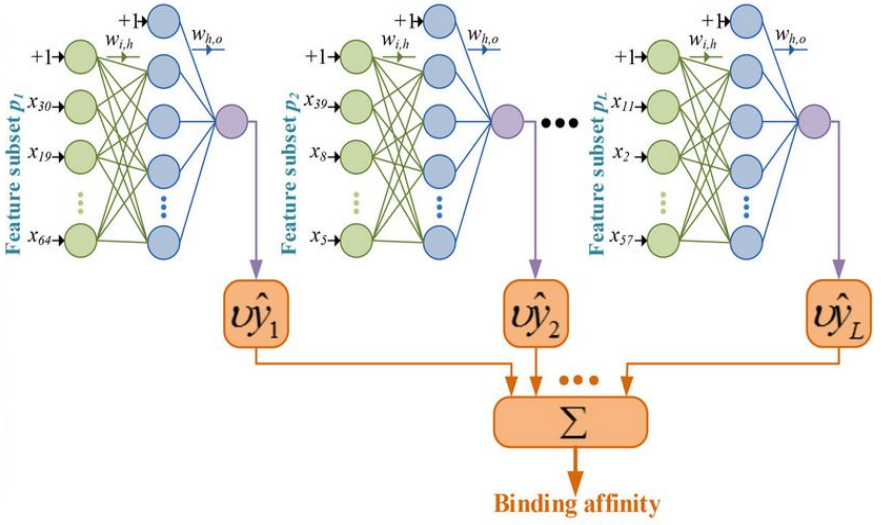
Conclusion

Return: 6819 ¯_(ツ)_/¯ ⟶



Future Work

- Infrastructure
  - Training on HPC or AWS
- Environment
  - Complex Tracks
- Algorithms
  - Imitation Learning
  - Ensembles



Acknowledgements

This research was performed as part of the Winter 2020 offering of CS 175: Project in Artificial Intelligence taught by Roy Fox at the University of California, Irvine. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the instructor.

In addition to Professor Fox, the authors would like to thank the members of the Duckietown Slack for providing technical help throughout the course of the project.

References

1. V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, “Human-level control through deep reinforcement learning,” Nature, vol. 518, no. 7540, pp. 529–533, 2015.
2. N. Heess, D. TB, S. Sriram, J. Lemmon, J. Merel, G. Wayne, Y. Tassa, T. Erez, Z. Wang, S. M. A. Eslami, M. A. Riedmiller, and D. Silver, “Emergence of locomotion behaviours in rich environments,” CoRR, vol. abs/1707.02286, 2017.
3. L. Paull, J. Tani, H. Ahn, J. Alonso-Mora, L. Carlone, M. Cap, Y. F. Chen, C. Choi, J. Dusek, Y. Fang, D. Hoehener, S. Liu, M. Novitzky, I. F. Okuyama, J. Pazis, G. Rosman, V. Varricchio, H. Wang, D. Yershov, H. Zhao, M. Benjamin, C. Carr, M. Zuber, S. Karaman, E. Frazzoli, D. Del Vecchio, D. Rus, J. How, J. Leonard, and A. Censi, “Duckietown: An open, inexpensive and flexible platform for autonomy education and research,” in 2017 IEEE International Conference on Robotics and Automation (ICRA), pp. 1497–1504, May 2017.
4. T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, “Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor,” in Proceedings of the 35th International Conference on Machine Learning (J. Dy and A. Krause, eds.), vol. 80 of Proceedings of Machine Learning Research, (Stockholmsmässan, Stockholm Sweden), pp. 1861–1870, PMLR, 10–15 Jul 2018.
5. S. Faußer and F. Schwenker, “Neural network ensembles in reinforcement learning,” Neural Processing Letters, vol. 41, pp. 55–69, Feb 2015.
6. M. Chevalier-Boisvert, F. Golemo, Y. Cao, B. Mehta, and L. Paull, “Duckietown environments for openai gym.” `https://github.com/duckietown/gym-duckietown`, 2018.
7. J. Achiam, “Spinning Up in Deep Reinforcement Learning,” 2018.
8. T. Hester, M. Vecerik, O. Pietquin, M. Lanctot, T. Schaul, B. Piot, D. Horgan, J. Quan, A. Sendonaris, I. Osband, G. Dulac-Arnold, J. Agapiou, J. Leibo, and A. Gruslys, “Deep q-learning from demonstrations,” 2018.
9. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016. `http://www.deeplearningbook.org`
10. J. Zilly, J. Tani, B. Considine, B. Mehta, A. F. Daniele, M. Diaz, G. Bernasconi, C. Ruch, J. Hakenberg, F. Golemo, A. K. Bowser, M. R. Walter, R. Hristov, S. Mallya, E. Frazzoli, A. Censi, and L. Paull, “The ai driving olympics at neurips 2018,” arXiv preprint arXiv:1903.02503, 2019.

Questions?