Modelling of Human Behaviour in Traffic Interactions Using Inverse Reinforcement Learning

Santiago Amaya Corredor[1]

Chalmers University of technology, Sweden
santiago.amaya@vub.be

Abstract. For a long time, the idea of imbuing human-like intelligence into machines and technologies has been in the collective consciousness of human kind. Although this goal is still far from becoming a reality, the recent advances in artificial intelligence is making it possible for systems to move one step closer to this kind of achievements. Under the simplifying assumption that human decisions are taken according to a Markov Decision Process, in this thesis we learn a representation of the motivations or "reasoning" behind some specific human actions. This is a different paradigm from standard machine learning algorithms, where an extra layer of intelligence is added. The system is able to "mimic" human behavior (as standard supervised learning can) but it also obtains a sense of why it is acting in such a way. This could appear like an inconsequential step, but models trained with this approach could (potentially) navigate strange environments as well as familiar ones, only by applying the "common sense" obtained in previous training.

In this thesis, we are looking to extract through the use of Inverse Reinforcement Learning, the "reward function" behind the actions of human road users, and obtain a policy that allows an AI agent to maximize its reward as it navigates through the environment. We aim to create a system capable of reproducing human behaviour in specific road scenarios and traffic interactions, and deploy it in the Autonomous Driving (AD) industry, as a component of a virtual environment able to interact in real time with AD software, taking the place of real humans in the simulations.

To achieve this, we implemented an advanced variant of Inverse Reinforcement Learning, based on adversarial training and powered by a deep Generative Adversarial Network. The algorithm is trained using traffic data supplied by Viscando (a Swedish company), in a collaboration aiming to create a model that can be used by industry in the near future.

Most of the experiments performed with IRL algorithms do not use real data with all the complications that it may pose. In this project we implement an algorithm tailored to the characteristics of real world data, and the goals of industry. The algorithm achieves a level of similarity to naturalistic data that is considered appropriate to start using in automated vehicle testing, at least at a stage of exploration.

A good portion of the variability and randomness in human behaviour is incorporated in the mechanisms of the model, which generates trajectories that for the most part resemble real expert demonstrations. Reasons for the situations where unwanted behaviour is present are studied.

Keywords: Inverse reinforcement learning · Adversarial training · Variational networks · Reward function.

Supervisor Yury Tarakanov, Viscando
yury@viscando.com

Examiner Adam Andersson, Department of Mathematical Sciences, Chalmers
adam.andersson@chalmers.se

References

1. Cs285, lectures 14 ; 15, http://rail.eecs.berkeley.edu/deeprlcourse-fa19/
2. Abbeel, P., Ng, A.Y.: Apprenticeship learning via inverse reinforcement learning. In: Proceedings of the Twenty-First International Conference on Machine Learning (2004)
3. Arora, S., Doshi, P.: A Survey of Inverse Reinforcement Learning: Challenges, Methods and Progress. arXiv e-prints p. arXiv:1806.06877 (2018)
4. Camerota, F.J.: Inverse Reinforcement Learning for Autonomous Vehicle Validation and Traffc Safety. Master's thesis (2020)
5. Finn, C., Christiano, P., Abbeel, P., Levine, S.: A Connection between Generative Adversarial Networks, Inverse Reinforcement Learning, and Energy-Based Models. arXiv e-prints arXiv:1611.03852 (2016)
6. Finn, C., Levine, S., Abbeel, P.: Guided Cost Learning: Deep Inverse Optimal Control via Policy Optimization. arXiv e-prints arXiv:1603.00448 (2016)
7. Finn, C., Levine, S., Abbeel, P.: Guided Cost Learning: Deep Inverse Optimal Control via Policy Optimization. arXiv e-prints arXiv:1603.00448 (2016)
8. Fu, J., Luo, K., Levine, S.: Learning Robust Rewards with Adversarial Inverse Reinforcement Learning. arXiv e-prints arXiv:1710.11248 (2017)
9. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2. p. 2672–2680 (2014)
10. Gruver, N., Song, J., Kochenderfer, M.J., Ermon, S.: Multi-agent adversarial inverse reinforcement learning with latent variables. p. 1855–1857 (2020)
11. Ho, J., Ermon, S.: Generative Adversarial Imitation Learning. arXiv e-prints arXiv:1606.03476 (2016)
12. Jeon, W., Barde, P., Nowrouzezahrai, D., Pineau, J.: Scalable Multi-Agent Inverse Reinforcement Learning via Actor-Attention-Critic. arXiv e-prints arXiv:2002.10525 (2020)
13. McCulloch, W.S., Pitts, W.: A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics **5**(4), 115–133 (1943)
14. Mombaur, K., Truong, A., Laumond, J.P.: From human to humanoid locomotion—an inverse optimal control approach. Autonomous Robots **28**(3), 369–383 (2010)
15. Ng, A.Y., Russell, S.J.: Algorithms for inverse reinforcement learning. In: Proceedings of the Seventeenth International Conference on Machine Learning. p. 663–670 (2000)
16. Peesapati, L.N., Hunter, M.P., Rodgers, M.O.: Can post encroachment time substitute intersection characteristics in crash prediction models? Journal of Safety Research **66**, 205–211 (2018)
17. Ramachandran, D., Amir, E.: Bayesian inverse reinforcement learning. p. 2586–2591 (2007)
18. Tak, S., Kim, S., Lee, D., Yeo, H.: A comparison analysis of surrogate safety measures with car-following perspectives for advanced driver assistance system. Journal of Advanced Transportation **2018**, 8040815 (2018)
19. Wang, C., Xie, Y., Huang, H., Liu, P.: A review of surrogate safety measures and their applications in connected and automated vehicles safety modeling. Accident Analysis Prevention **157**, 106157 (2021)
20. Yu, L., Song, J., Ermon, S.: Multi-Agent Adversarial Inverse Reinforcement Learning. arXiv e-prints arXiv:1907.13220 (Jul 2019)

21. Yu, L., Song, J., Ermon, S.: Multi-Agent Adversarial Inverse Reinforcement Learning. arXiv e-prints arXiv:1907.13220 (2019)
22. Yu, L., Yu, T., Finn, C., Ermon, S.: Meta-inverse reinforcement learning with probabilistic context variables. In: Advances in Neural Information Processing Systems. vol. 32, pp. 11772–11783 (2019)
23. Ziebart, B.D., Maas, A., Bagnell, J.A., Dey, A.K.: Maximum entropy inverse reinforcement learning. In: Proceedings of the 23rd National Conference on Artificial Intelligence - Volume 3. p. 1433–1438 (2008)