TECHNISCHE UNIVERSITÄT DARMSTADT

Thesis (B.Sc. / M.Sc.)
Inverse Reinforcement Learning for Recovering Learning Properties

Fachbereich 18
Elektrotechnik und Informationstechnik
Bioinspirierte Kommunikationssysteme

Department 18
Electrical Engineering and Information Technology
Bioinspired Communication Systems

Prof. Dr. Heinz Koeppl
Head of lab

For some autonomous learning systems, the learning process can be observed, yet the exact learning method is not directly accessible. An example might be an autonomous vehicle that learns to adapt its driving style to the passenger. The manufacturer invested a lot of financial resources to equip the vehicle with an elaborated learning method and of course wants to keep the exact learning method confidential. For admitting such a vehicle, the public authority has to investigate whether the vehicle conforms with specific regulations and might also have to check properties of the internal hidden learning algorithm. As the method is not revealed by the manufacturer, the agency just has the ability to get properties of the learning method indirectly by watching the learning process of the vehicle. How to recover properties of the learning method just by having observations is an open problem up to now.

For inverting control processes, so-called Inverse Reinforcement Learning (IRL) methods [1, 2, 3] have been proposed. Their goal lies usually on imitation learning where an autonomous system tries to imitate human behavior. Given demonstrations, IRL learns a generalizable objective function which can be used to reproduce similar behavior. Instead of learning the objective function of demonstrations in standard IRL, one could also aim to learn another latent quantity – the properties of a learner – and thus solve the problem described in the previous paragraph.

In this project, we are going to investigate how one can adapt current IRL methods to obtain properties of a learner given learning demonstrations. Ideally, the prospective student has already basic knowledge of Reinforcement Learning and Bayesian statistics and Python programming skills.

For further information, please contact Matthias Schultheis.

Matthias Schultheis
Project supervisor

Rundeturmstraße 12
64283 Darmstadt

Phone: +49 6151 16 - 57 253
matthias.schultheis@bcs.tu-darmstadt.de
https://www.bcs.tu-darmstadt.de

September 2020

References

[1] Andrew Y Ng, Stuart J Russell, et al. Algorithms for inverse reinforcement learning. In *Icml*, volume 1, page 2, 2000.

[2] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. In *Aaai*, volume 8, pages 1433–1438. Chicago, IL, USA, 2008.

[3] Constantin A Rothkopf and Christos Dimitrakakis. Preference elicitation and inverse reinforcement learning. In *Joint European conference on machine learning and knowledge discovery in databases*, pages 34–48. Springer, 2011.