MAXIMUM ENTROPY APPROACHES FOR INVERSE REINFORCEMENT LEARNING

A. J. Snoswell, The University of Queensland, Australia, a.snoswell@uq.edu.au
S. P. N. Singh, The University of Queensland, Australia, spns@uq.edu.au
N. Ye, The University of Queensland, Australia, nan.ye@uq.edu.au

We make decisions to maximize our perceived reward, but handcrafting a reward function for an autonomous agent is challenging. Inverse Reinforcement Learning (IRL), which is concerned with learning a reward function from expert demonstrations, has recently attracted significant interest, with the Maximum Entropy (MaxEnt) approach being a popular method.

In this talk, we will explore and contrast a variety of MaxEnt IRL approaches. We show that in the presence of stochastic dynamics, a minimum KL-divergence condition provides a rigorous derivation of the MaxEnt model, improving over a prior heuristic derivation. Furthermore, we explore extensions of the MaxEnt IRL method to the case of unknown stochastic transition dynamics, including a generative model for trajectories, a discriminative model for action sequences, and a simple logistic regression model.

We will present evaluation results for simulated and real-world problems, including the UCI Taxi Service Trajectory dataset, which considers the problem of long-distance forecasting of driver behaviour.