

Andreas Krause

Safe and Efficient Exploration in Model-Based Reinforcement Learning

How can we enable systems to efficiently and safely learn online, from interaction with the real world? I will first discuss safe Bayesian optimization, where we quantify uncertainty in the unknown reward function and constraints, and, under some regularity conditions, can guarantee both safety and convergence to a natural notion of reachable optimum. I will then consider Bayesian model-based deep reinforcement learning, where we use the epistemic uncertainty in the dynamics model to guide exploration while ensuring safety. Lastly I will discuss how we can meta-learn flexible data-driven priors from related tasks and simulations, and discuss several applications in robotics and science.