Colloquium

Decision Making and Inference under Limited Information and Large Dimensionality

Stefano Ermon
Cornell University

Abstract: Statistical inference in high-dimensional probabilistic models (i.e., with many variables) is one of the central problems of statistical machine learning and stochastic decision making. To date, only a handful of distinct methods have been developed, most notably (MCMC) sampling, decomposition, and variational methods. In this talk, I will introduce a fundamentally new approach based on random projections and combinatorial optimization. Our approach provides provable guarantees on accuracy, and outperforms traditional methods in a range of domains, in particular those involving combinations of probabilistic and causal dependencies (such as those coming from physical laws) among the variables. This allows for a tighter integration between inductive and deductive reasoning, and offers a range of new modeling opportunities. As an example, I will discuss an application in the emerging field of Computational Sustainability aimed at discovering new fuel-cell materials where we greatly improved the quality of the results by incorporating prior background knowledge of the physics of the system into the model.

Friday, February 21, 2014, 3:00 pm
Mathematics and Science Center: W201

Mathematics and Computer Science
Emory University