Combining Statistics and Semantics to Turn Data into Knowledge

Lise Getoor

University of California Santa Cruz
getoor@soe.ucsc.edu

Abstract. Addressing inherent uncertainty and exploiting structure are fundamental to turning data into knowledge. Statistical relational learning (SRL) builds on principles from probability theory and statistics to address uncertainty while incorporating tools from logic to represent structure. In this talk I will overview our recent work on probabilistic soft logic (PSL), an SRL framework for collective, probabilistic reasoning in relational domains. PSL is able to reason holistically about both entity attributes and relationships among the entities, along with ontological constraints. The underlying mathematical framework supports extremely efficient inference. Our recent results show that by building on state-of-the-art optimization methods in a distributed implementation, we can solve large-scale knowledge graph extraction problems with millions of random variables orders of magnitude faster than existing approaches.