Happy Mittal. PhD Thesis: Abstract

Title: EFFECTIVE INFERENCE & LEARNING IN STATISTICAL RELATIONAL MODELS

Abstract:

A vast majority of traditional Machine Learning algorithms make the i.i.d. assumption, i.e., instances are assumed to be sampled independently from an identical distribution. However, the real world data is relational i.e. there exist various types of relations among the instances which need to be captured. Moreover, these relations also have a certain degree of uncertainty that needs to be modeled. The field of Statistical Relational Learning (SRL) achieves the merger between the two by explicitly modeling relations as well as uncertainty among them. Due to the underlying template structure, SRL models often result in the construction of large ground networks posing serious computational challenges. Therefore, designing effective inference and learning algorithms becomes critical for the success of these models. This thesis focuses on proposing multiple such inference and learning techniques: (a) We first propose two different algorithms for lifted inference, i.e., techniques which exploit the symmetry of the underlying relational representation. In the first approach, we exploit the underlying symmetry using a rule called Single Occurrence (SO) for MAP inference. SO rule results in inference time which is independent of domain size under certain conditions. Second, we propose a language for specifying constraints for constructing the underlying ground theory and extend existing lifting techniques to work with these constraints. (b) In the second part of this thesis, we present two different techniques for effective learning in SRL models. First, we present an algorithm for fine-grained learning, where we learn a different weight for different subsets of the groundings of the same logical formula. We show that this can result in improved prediction accuracy by reducing bias. Second, we present a novel re-parameterization technique to generalize the weights learned for logical formulas across domains of varying sizes. We also present various theoretical properties of our formalism. In each of the above cases, we present experimental results validating the efficacy of proposed algorithms on real as well as artificial datasets. Much of our work is presented in the context of Markov Logic Networks (MLNs), which is a popular SRL formalism though our techniques are general enough and can be extended to other similar representations.