Learning the Structure of Probabilistic SDDs

Jessa Bekker **Arthur Choi** **Guy Van den Broeck**

Abstract

Many domains, such as health care, gain benefit from machine learning if a certain degree of accuracy is guaranteed about the predictions. For techniques that model uncertainty, such as Bayesian networks and other graphical models, it is in general infeasible to do predictions with such guarantees. While we can measure the accuracy of the model, inference of simple queries such as marginals is #P-hard and therefore the predictions are often approximations with unknown accuracy. The domain of tractable learning provides a solution by restricting the learned models to those that do allow exact inference and therefore the predictions are as accurate as the learned model itself.

The key of tractable learning is the usage of a tractable representation for the model. A tractable representation essentially represents the calculations needed to do inference; the size of this representation is therefore the complexity of inference. By keeping this tractable representation small enough, exact inference will always be feasible. There exist different types of tractable representation that differ in the types of tractable queries, tractable operations (needed during learning) and compactness. Sentential Decision Diagrams (SDDs) support the widest range of tractable queries and operations but pay for this by being less compact [1]. Probabilistic SDDs (PSDDs) can be more compact than SDDs (at least as compact) while supporting the same tractable queries.

SDDs represent boolean formulas. An SDD that represents a probabilistic model has weights for its variables and supports efficient weighted model counting to do predictions. Combining two SDDs is also efficient, therefore any query can be answered by compiling it to an SDD and combining it with the model. PSDDs are a variation on SDDs with parameters on the edges instead of the variables. PSDDs are at least as compact as SDDs because any SDD can be represented by a PSDD but not the other way round. They can answer the same queries as SDDs because they too can be combined with an SDD that represent the query. Therefore PSDDs are an attractive tractable representation, but until now they were not used for tractable learning; a parameter learner does exists [2].

In this work, we present the first structure learner for PSDDs. It starts with an initial model and incrementally changes it to improve the accuracy. For the incremental changes, we introduce some new operators that change the distribution without changing the possible worlds. This method can naturally incorporate constraints by using them as the initial model and is therefore ideal to learn in structured spaces or on top of expert knowledge but it can also handle unconstrained cases.

References:

[1] J. Bekker, J. Davis, A. Choi, A. Darwiche, and G. Van den Broeck. Tractable Learning for Complex Probability Queries. In NIPS, 2015.
[2] A. Choi, G. Van den Broeck, and A. Darwich. Tractable Learning for Structured Probability Spaces: A Case Study in Learning Preference Distributions. In IJCAI, 2015.

Keywords:

Tractable Learning
Probabilistic Graphical Models
Structure Learning
Probabilistic Sentential Decision Diagrams