Ordering-Based Search for Tractable Bayesian Networks

Jessa Bekker **Guy Van den Broeck** **Jesse Davis**

Tractable Bayesian network learning's goal is to learn Bayesian networks (BNs) where inference is guaranteed to be efficient. The state of the art techniques for tractable learning are based on knowledge compilation: they keep a tractable representation of the network next to the classic graphical model. The tractable representation allows inference with a complexity polynomial in the size of this representation and therefore provide a measure for tractability. The current best tractable BN learner is ACBN (also known as LearnAC) which uses Arithmetic Circuits (ACs) as the tractable representation. It greedily splits the conditional distributions. The splits are scored on the increase in likelihood with the AC size as penalty. We propose the usage Sentential Decision Diagrams (SDDs) which is a subclass of ACs.

SDDs have desirable properties that are absent in more general representations such as ACs. The additional properties translate among other to the ability of not only efficiently adding but also deleting edges in the corresponding Bayesian network. Deleting edges is out of the question with ACs. This makes it possible to use other than greedy learning algorithms for learning tractable BNs. Using SDDs we can thus transform existing BN learners to tractable BN learners.

In this work we investigate using the ordering-based search algorithm for learning tractable BNs. In classic ordering-based search, one is looking for a Bayesian network that optimizes a decomposable score function such as BIC, MDL or BDe. The key idea of ordering-based search is that even though the task of finding the structure that optimizes the score function is NP-hard, the same task becomes efficient when the variable ordering is fixed. Because of the decomposability of the score function one only needs to look for the optimal set of parents for each node, which is trivial. Ordering-based search searches over the space of variable orderings by swapping adjacent variables.

In tractable learning, a different scoring function is used. It is a trade-off between the likelihood of the model and the inference complexity, i.e. the size of the tractable representation. Unfortunately the SDD size is not decomposable. Different BN edges change the SDD size in different ways, dependent on the current SDD structure. We therefore adjust ordering-based search with an extra search dimension: the number of edges in the tree CPT of each node.

We propose SDDBN, an algorithm that searches for a BN structure that optimizes a trade-off between likelihood and tractability. It uses two search operators: "Swap" and "Split". "Swap" swaps two adjacent variables in the variable order. It adds parents to the swapped nodes until the size of the SDD is about the same as before the swap. This allows a fair comparison between the two orderings. "Split" adds a parent to a node by adding a split to the tree CPT. We use random restarts to cope with local maxima.

Another contribution is an empirical evaluation. We conduct experiments that compare the performance of the models learned with SDDBN with the models learned by LearnAC. The preliminary results look promising.