A Bayesian Metareasoner for Algorithm Selection for Real-time Bayesian Network Inference Problems

Haipeng Guo

Laboratory for Knowledge Discovery in Databases
Department of Computing and Information Sciences, Kansas State University
hpguo@cis.ksu.edu http://www.kddresearch.org

1. Motivation

Bayesian network (BN) inference has long been seen as a very important and hard problem in AI. Both exact and approximate BN inference are NP-hard [Co90, Sh94]. To date researchers have developed many different kinds of exact and approximate BN inference algorithms. Each of these has different properties and works better for different classes of inference problems. Given a BN inference problem instance, it is usually hard but important to decide in advance which algorithm among a set of choices is the most appropriate. This problem is known as the algorithm selection problem [Ri76]. The goal of this research is to design and implement a meta-level reasoning system that acts as a "BN inference expert" and is able to quickly select the most appropriate algorithm for any given Bayesian network inference problem, and then predict the run time performance.

2. A Bayesian Metareasoner for Algorithm Selection for Bayesian Network Inference

I address this problem with a scheme based mainly on Bayesian methods [HRGK01]. Knowledge of dependencies among the characteristic of BN inference problem instances and the performance of the inference algorithms can be considered as dome kind of uncertain knowledge. This knowledge can be represented by a Bayesian network that we call the "inference expert network", or the metareasoner. The metareasoner is automatically learned from some training data with the guidance of some domain knowledge. To create a representative training data that contains the knowledge we are seeking, we are developing a controllable random BN inference problem generator. It could randomly generate BNs and inference problem instances from a description of the BN's characteristic. By controlling the characteristic parameters, we can generate many random representative BNs for both training and testing. Once we have synthetic real world BNs and synthetic instances of BN inference problems, we can then run the selected inference algorithms on these instances to generate the training data. These training data records the characteristic parameters of the problems, the algorithms being used, and the run time performance of the algorithms. They contain the knowledge of how well each algorithm matches each class of problems. The inference expert network, or the metareasoner, can be learned from the training data using some BNs learning algorithm such as K2 [CH92]. Having this inference expert network, we can then use it to select the right algorithm for a given Bayesian network inference problem instance and predict the run time performance of the algorithm on this problem instance. It can serve as a meta-level reasoner to selecting the "right" algorithm for real-time Bayesian network inference problems. The system should be easily extended in the future to include new characteristic parameters and new inference algorithms. The problem instances' characteristic parameters I consider include the size of the network, topology of the network, the extreme probabilities and the average skewness of the CPTs, the ratio of number of query nodes and evidence nodes, and so on. My library of candidate BN inference algorithms includes polytree algorithm, clique tree propagation algorithm, various stochastic sampling algorithms, genetic algorithms for inference, and a newly-designed search-based anytime inference algorithm.

References

[Co90] G. F. Cooper. The computational complexity of probabilistic infernece using bayesian belief networks. Artificial Intelligence, 42(2-3):393-405. Elsevier, 1990.

[HRGK01] E. Horvitz, Y. Ruan, C. Gomes, H. Kautz, B. Selman, D. M. Chickering. A Bayesian Approach to Tackling Hard Computational Problems. In UAI01, 2001.

[Ri76] Rice, J. R. The algorithm selection problem. Advances in computers 15:65-118, 1976.

[Sh94] Shimony, Solomon, E., "Finding MAPs for Belief Networks is NP-hard," Artificial Intelligence, 68, pp 399-410, 1994.