Abductive inference in bayesian networks

David Quesada López
Computational Intelligence Group, Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, Spain

Abstract

Abductive inference in bayesian networks solves the problem of obtaining the most probable explanation (MPE) of a network given some evidence of its nodes. This inference can be total, if you aim to obtain the MPE of the whole network, or partial, if you are only interested in some of the nodes. In this state of the art we will cover both approaches and the methods used to solve them.

KEY WORDS: Bayesian networks; Abductive inference; Approximate inference

1 Introduction

A Bayesian network is a directed acyclic graph (DAG) where you represent each node as a random variable and the edges represent the dependence or independence among them. Each node also has associated a conditional probability distribution conditioned to its parents that represents the probability distribution of the node's variable.

One of the problems associated with BNs is abductive inference, which consists in finding the maximum *a posteriori* probability (MAP) state of the network given some evidence on the state of variables of the network. In general, we are interested in the K most probable explanations (k-MPE). This problem is known to be NP-hard (Shimony (1994)), and most of the times trying to obtain an exact MPE turns out to be intractable. This is the reason why most researchers focus on getting K approximate MPEs instead of the exact one.

Abductive inference can also be called total, when we want to know the whole network's k-MPEs, or partial, when we only want the K most probable state assignments for a subset of the variables known as the explanation set (Fortier *et al.* (2013)).

In this state of the art, we will review some methods to perform total and partial abductive inference on a Bayesian network.

2 State-of-the-art

We can formally describe the problem of the k-MPE as: given a set of unobserved variables X_u, a set of observed variables X_o, we want to find the configuration x_u^* such that:

$$x_u^* = arg \max_{x_u} P(x_u|x_o) \tag{1}$$

Which means we are aiming to maximize the MAP state of the network by finding an appropiate x_u^* (De Campos *et al.* (2002)).

Traditionally, there has been three solutions to this problem (De Campos *et al.* (2001)): Through the computation of the *chain rule*, which is intractable in even the simplest of networks, through belief propagation and through abductive inference.

2.1 Total abductive inference

It has been proven that belief propagation and abductive inference are both equal in complexity when finding the MPE of a network (Dawid (1992)), however the problem of finding the k-MPE is a more complex one.

There are two ways of tackling this problem:

- With exact computation. This is the case of Dawid (1992), were a belief propagation algorithm was adapted to find the MPE of a network by replacing the summation by maximum in the message passing. This method was able to only find the first 3 MPEs of the network (Nilsson (1998)), so it had to be modified to be able to find the k-MPE.
- With approximate results. Here, we treat the problem as an optimization one. One approach was performed by Gelsema (1995) using a genetic algorithm (GA). In this method, each individual of the population represents a possible state of the network by representing the presence or absence of the nodes with a binary vector. Genetic programming is also used by Rojas and Kramer (1993) to solve this problem. Here, they represent the network's graph in each individual. Another method was proposed by Welch (1996), where he used a Montecarlo simulation to initialize the population and then used a GA similar to that of Gelsema. The most recent approach to this problem is with the use of overlapping swarm intelligence (Fortier *et al.* (2013)), in wich each swarm tries to find a MPE of the network.

2.2 Partial abductive inference

In partial abductive inference, instead of having a set of unobserved variables X_u, we have an explanation set X_E of the variables we are interested in and a set X_R such as $X_R = X_u \backslash X_E$. With this, the the configuration x_E^* we want to find is:

$$x_E^* = arg\max_{x_E} P(x_E|x_o) = arg\max_{x_E} \sum_{x_R} P(x_E, x_R|x_o) \tag{2}$$

This process is more complex than total abductive inference (De Campos *et al.* (2002)). They both are similar, but not all the clique trees obtained from the BN are valid, so the tree generation has to be modified from what we have in total abductive inference.

To solve this problem, there are the same approaches as in the total one:

- Trying to get the exact MPE. Here, there are methods that rely on variable elimination, like Dechter (1998) which uses his bucket elimination method to find the MPE and then we can use Li and D'Ambrosio (1993) linear algorithm k-1 times to find the k-MPE. A variation of Nilsson (1998) regarding the formation of the trees can be used to solve partial abduction.
- Aproximating the best k-MPE. The main researcher in this area was performed De Campos *et al.*, who started his research adapting a genetic algorithm to solve this problem as in the method followed by Gelsema (De Campos *et al.* (1999)). In this case, the chain rule cannot be used to evaluate the fitness of each individual, so it uses probability propagation. He also focused on improving the efficiency of this GA's operators (De Campos *et al.* (2002)). Another alternative proposed was the use of simulated annealing (De Campos *et al.* (2001)) to move through the solution space

3 Conclusions and future research

Abductive inference in Bayesian networks is a useful in many real case scenarios, such as diagnosis. Total abductive inference has been very studied by the literature and has a wide range of both exact

and approximate methods to get the k-MPE of a network. Partial abductive inference is not as widely studied, but it seems like its main idea could be more useful than the total one. In both cases, the most recent advances are in the approximate branch, which is more feasible for bigger Bayesian networks.

The main open lines of research in this problem seem to be the development of more refined genetic algorithms for both total and partial, or the application of different methods of moving through the solution space of the maximization problem, as in the case of simulated annealing and overlapping swarm intelligence.

References

Dawid A. P. (1992). Applications of a general propagation algorithm for probabilistic expert systems. Statistics and computing, 2(1), 25-36.

De Campos L. M., Gámez J. A., Moral S. (1999). Partial abductive inference in Bayesian belief networks using a genetic algorithm. Pattern Recognition Letters, 20(11), 1211-1217.

De Campos L. M., Gámez J. A., Moral S. (2001). Partial abductive inference in Bayesian belief networks by simulated annealing. International Journal of Approximate Reasoning, 27(3), 263-283.

De Campos L. M., Gamez J. A., Moral S. (2002). Partial abductive inference in Bayesian belief networks-an evolutionary computation approach by using problem-specific genetic operators. IEEE Transactions on Evolutionary Computation, 6(2), 105-131.

Dechter R. (1998). Bucket elimination: A unifying framework for probabilistic inference. In Learning in graphical models (pp. 75-104). Springer, Dordrecht.

Fortier N., Sheppard J., Pillai K. G. (2013, April). Bayesian abductive inference using overlapping swarm intelligence. In Swarm Intelligence (SIS), 2013 IEEE Symposium on (pp. 263-270). IEEE.

Gelsema E. S. (1995). Abductive reasoning in Bayesian belief networks using a genetic algorithm. Pattern Recognition Letters, 16(8), 865-871.

Li Z., D'Ambrosio B. (1993, July). An efficient approach for finding the MPE in belief networks. In Proceedings of the Ninth international conference on Uncertainty in artificial intelligence (pp. 342-349). Morgan Kaufmann Publishers Inc..

Nilsson D. (1998). An efficient algorithm for finding the M most probable configurationsin probabilistic expert systems. Statistics and computing, 8(2), 159-173.

Rojas-Guzman C., Kramer M. A. (1993, July). Galgo: A genetic algorithm decision support tool for complex uncertain systems modeled with Bayesian belief networks. In Proceedings of the Ninth international conference on Uncertainty in artificial intelligence (pp. 368-375). Morgan Kaufmann Publishers Inc..

Seroussi B., Golmard J. L. (1994). An algorithm directly finding the K most probable configurations in Bayesian networks. International Journal of Approximate Reasoning, 11(3), 205-233.

Shimony S. (1994). Finding maps for belief networks is NP-hard. Artificial Intelligence 68, 399-410.

Welch R. L. (1996, August). Real time estimation of Bayesian networks. In Proceedings of the Twelfth international conference on Uncertainty in artificial intelligence (pp. 533-544). Morgan Kaufmann Publishers Inc..