Bayesian Networks in Reliability: Some Recent Developments

Helge Langseth
Dept. of Mathematical Sciences
Norwegian University of Science and Technology
NO-7034 Trondheim, Norway
helgel@math.ntnu.no

Abstract

In this talk I will introduce the basic properties of Bayesian network models, and discuss why BN models are well suited for applications in reliability.

1 Basic properties

A Bayesian Network (BN), (Pearl 1988; Cowell et al. 1999; Jensen 2001), is a compact representation of a multivariate statistical distribution function. A BN encodes the probability density function governing a set of random variables $\{X_1, \ldots, X_n\}$ by specifying a set of conditional independence statements together with a set of conditional probability functions. More specifically, a BN consists of a qualitative part, a *directed acyclic graph* where the nodes mirror the random variables X_i, and a quantitative part, the set of conditional probability functions. An example of a BN over the variables $\{X_1, \ldots, X_5\}$ is shown in Figure 1, only the qualitative part is given. We call the nodes with outgoing edges pointing into a specific node the *parents* of that node, and say that X_j is a *descendant* of X_i if and only if there exists a directed path from X_i to X_j in the graph. In Figure 1 X_1 and X_2 are the parents of X_3, written $\text{pa}(X_3) = \{X_1, X_2\}$ for short. Furthermore, $\text{pa}(X_4) = \{X_3\}$ and since there are no directed path from X_4 to any of the other nodes, the descendants of X_4 are given by the empty set and, accordingly, its non-descendants are $\{X_1, X_2, X_3, X_5\}$.

The edges of the graph represents the assertion that a variable is conditionally independent of its non-descendants in the graph given its parents in the same graph; other conditional independence statements can be read off the graph by using the rules of *d-separation* (Pearl 1988). The graph in Figure 1 does for instance assert that for all distributions compatible with it, we have that X_4 is conditionally independent of $\{X_1, X_2, X_5\}$ when conditioned on $\{X_3\}$.

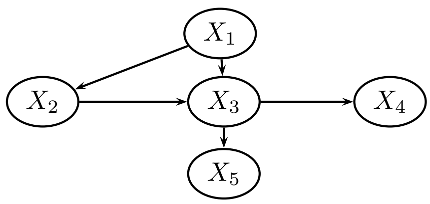

Figure 1: An example BN over the nodes $\{X_1, \ldots, X_5\}$. Only the qualitative part of the BN is shown.

When it comes to the quantitative part, each variable is described by the conditional probability function of that variable *given the parents* in the graph, i.e., the collection of conditional probability functions $\{f(x_i \,|\, \text{pa}(x_i))\}_{i=1}^n$ is required. The underlying assumptions of conditional independence encoded in the graph allow us to calculate the joint probability function as

$$f(x_1, , \ldots, x_n) = \prod_{i=1}^{n} f(x_i \,|\, \text{pa}(x_i)). \qquad (1)$$

Some of the main reasons why Bayesian networks have become widely used are (in the view of the author):

Efficient calculation scheme: Efficient algorithms for calculating arbitrary marginal distributions, say, $f(x_i, x_j, x_k)$ as well as conditional distributions, say, $f(x_i, x_j \mid x_k, x_\ell)$, make BNs well suited for modeling complex systems. Models over thousands of variables are not uncommon.

Intuitive representation: The qualitative part (the graph) has an intuitive interpretation as a model of causal influence. Although this interpretation is not necessarily entirely correct, it is helpful when the BN structure is to be elicited from experts. Furthermore, it can also be defended if some additional assumptions are made (Pearl 2000). To elicit the quantitative part from experts, one must acquire all conditional independencies ($f(x_i \mid \mathrm{pa}(x_i))$ for $i = 1, \ldots, n$ in Eq. 1), and once again the causal interpretation can come in as a handy tool. Alternatively, the expert can supply a mix of both marginal and conditional distributions, which can then be glued together by the IPFP algorithm (Whittaker 1990; Vomlel 1999).

Model fitting: Methods for estimating the quantitative part of the BN from data date back to the work of Spiegelhalter and Lauritzen (1990), see also Lauritzen (1995). A method for estimating the qualitative part from data was pioneered by Cooper and Herskovits (1992), and is still an active research area.

Compact representation: Through its factorized representation (Eq. 1), the BN attempts to represent the multi-dimensional distribution in a cost-efficient manner. The parametrization is however not optimized; it is merely defined to be sufficient to encode *any* distribution compatible with the conditional independence statements encoded in the graph. Many researchers, including Heckerman and Breese (1994) and Boutilier et al. (1996), have explored even more cost-efficient representations.

Expressive representation: A BN can represent or approximate *any* probability density function. The BN representation has traditionally been limited to only handle *discrete* distributions in addition to a particular class of hybrid distributions, the so-called *conditional Gaussian* distributions (Lauritzen and Wermuth 1989). Recently, Moral et al. (2001) have developed a framework for approximating any hybrid distribution arbitrarily well by employing *mixtures of truncated exponential* (MTE) distributions. They also show how the BN framework's efficient calculation scheme can be extended to handle the MTE distributions.

2 Bayesian Networks in reliability

The history of BNs in reliability can (at least) be traced back to Barlow (1988). More recently, BNs have found applications in, e.g., software reliability (Fenton et al. 1998; Gran 2002), fault finding systems (Jensen et al. 2001), maintenance modeling (Langseth and Lindqvist 2003), and general reliability modeling (Bobbio et al. 2001; Langseth 2002; Ingleby and West 2003).

References

Barlow, R. E. (1988). Using influence diagrams. In C. A. Clarotti and D. V. Lindley (Eds.), *Accelerated life testing and experts' opinions in reliability*, pp. 145–157.

Bobbio, A., L. Portinale, M. Minichino, and E. Ciancamerla (2001). Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. *Reliability Engineering and System Safety 71*(3), 249–260.

Boutilier, C., N. Friedman, M. Goldszmidt, and D. Koller (1996). Context-specific independence in Bayesian networks. In *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence*, Portland, Oregon, pp. 115–123.

Cooper, G. F. and E. Herskovits (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine Learning 9*, 309–347.

Cowell, R. G., A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter (1999). *Probabilistic Networks and Expert Systems*. Statistics for Engineering and Information Sciences. New York: Springer-Verlag.

Fenton, N., B. Littlewood, M. Neil, L. Strigini, A. Sutcliffe, and D. Wright (1998). Assessing dependability of safety critical systems using diverse evidence. *IEE Proceedings Software Engineering 145*(1), 35–39.

Gran, B. A. (2002). *The use of Bayesian Belief Networks for combining disparate sources of information in the safety assessment of software based systems*. Ph. D. thesis, Department of Mathematical Sciences, Norwegian University of Science and Technology. Doktor Ingeniør avhandling 2002:35.

Heckerman, D. and J. S. Breese (1994). A new look at causal independence. In *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*, San Francisco, CA., pp. 286–292. Morgan Kaufmann Publishers.

Ingleby, M. and M. West (2003). Causal influence coefficients: A localised maximum entropy approach to Bayesian inference. In K. Doksum and B. H. Lindqvist (Eds.), *Mathematical and Statistical Methods in Reliability*, Quality, Reliability and Engineering Statistics, Chapter 4, pp. 45–56. Singapore: World Scientific Publishing Co.

Jensen, F. V. (2001). *Bayesian Networks and Decision Graphs*. New York: Springer-Verlag.

Jensen, F. V., U. Kjærulff, B. Kristiansen, H. Langseth, C. Skaanning, J. Vomlel, and M. Vomlelová (2001). The SACSO methodology for troubleshooting complex systems. *Artificial Intelligence for Engineering, Design, Analysis and Manufacturing 15*(5), 321–333.

Langseth, H. (2002). *Bayesian Networks with Applications in Reliability Analysis*. Ph. D. thesis, Department of Mathematical Sciences, Norwegian University of Science and Technology. Doktor Ingeniør avhandling 2002:121.

Langseth, H. and B. H. Lindqvist (2003). A maintenance model for components exposed to several failure modes and imperfect repair. In K. Doksum and B. H. Lindqvist (Eds.), *Mathematical and Statistical Methods in Reliability*, Quality, Reliability and Engineering Statistics, Chapter 27, pp. 415–430. Singapore: World Scientific Publishing Co.

Lauritzen, S. L. (1995). The EM-algorithm for graphical association models with missing data. *Computational Statistics and Data Analysis 19*, 191–201.

Lauritzen, S. L. and N. Wermuth (1989). Graphical models for associations between variables, some of which are quantitative and some qualitative. *Annals of Statistics 17*, 31–57.

Moral, S., R. Rumí, and A. Salmerón (2001). Mixtures of truncated exponentials in hybrid Bayesian networks. In *Sixth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, Volume 2143 of *Lecture Notes in Artificial Intelligence*, pp. 145–167. Springer-Verlag.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA.: Morgan Kaufmann Publishers.

Pearl, J. (2000). *Causality – Models, Reasoning, and Inference*. Cambridge, UK: Cambridge University Press.

Spiegelhalter, D. J. and S. L. Lauritzen (1990). Sequential updating of conditional probabilities on directed graphical structures. *Networks 20*, 579–605.

Vomlel, J. (1999). *Methods of Probabilistic Knowledge Integration*. Ph. D. thesis, Faculty of Electrical Engineering, Czech Technical University.

Whittaker, J. (1990). *Graphical models in applied multivariate statistics*. Chichester: Wiley & Sons.