Reliability Engineering and System Safety 151 (2016) 93–112

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Methods for building Conditional Probability Tables of Bayesian Belief Networks from limited judgment: An evaluation for Human Reliability Application

L. Mkrtchyan, L. Podofillini*, V.N. Dang

Paul Scherrer Institute, Switzerland

ARTICLE INFO

Available online 15 January 2016

Keywords:
Bayesian Belief Networks
Human Reliability Analysis
Expert judgment
Conditional Probability Tables

ABSTRACT

The present paper evaluates five methods for building Conditional Probability Tables (CPTs) of Bayesian Belief Networks (BBNs) from partial expert information: functional interpolation, the Elicitation BBN, the Cain calculator, Fenton et al. and Røed et al. methods. The evaluation considers application to a specific field of risk analysis, Human Reliability Analysis (HRA). The five methods are particularly suited for HRA models calculating the human error probability as a function of influencing factor assessments. The performance of the methods is evaluated on two simple examples, designed to test aspects relevant for HRA (but not exclusively): the representation of strong factor influences and interactions, the representation of uncertainty on the BBN relationships, and the method requirements as the BBN size increases. The evaluation underscores modelling limitations related to the treatment of multi-factor interdependencies and of different degrees of uncertainty in the factor relationships. The functional interpolation method is the least susceptible to these limitations; however, its elicitation requirements grow exponentially with the model size. Besides expert judgment, HRA applications of BBNs include the use of empirical data, combination of data and judgment, information from existing HRA methods: the building of the CPTs in these applications is outside the scope of the evaluation.

1. Introduction

Bayesian Belief Networks (BBNs) are increasingly being used in risk analysis applications to model the effect of multiple, diverse, inter-related influences on risk. Their ability to incorporate diverse types of factors has allowed the construction of comprehensive models for risk assessment including hardware, human, and organizational failures, as well as diverse risk-conditioning events, as in [1–3]. Risk analysis applications, dealing with rare events, often have to cope with the scarcity of data available to understand the complex interactions leading to failure events: in these applications, BBNs have proven useful to formalize, represent, and quantify subjective knowledge on uncertain events. At the other extreme, in data-rich applications (e.g. some medical diagnosis and financial applications), BBNs are typically used for data mining to make sense of causal or influencing relationships and build predictive models learnt from data [4]. Other applications fall between these categories and BBNs are generally developed by combining available data and expert judgment.

The focus of the present paper is on the development of BBNs from expert judgment, for cases in which data is not available or not adequate to determine the BBN relationships. While expert judgment may be used in many phases of BBN development (including the node and structure definition), the present paper focuses on the quantification of the BBN relationships, i.e. the Conditional Probability Distributions (CPDs). There is general agreement that this is the most delicate part of the BBN development [5]; although, concerns that some applications may lack transparency in the process of node and structure definition were raised recently [6,7].

In general, the elicitation of judgments to assess CPDs has followed three approaches, often combined: direct assessment of the probabilities by one or multiple experts; elicitation of probability ranking on a qualitative scale (to avoid the shortcomings, e.g. biases, of directly eliciting probabilities from experts); elicitation of selected model relationships (or, more generally, of partial model information) and derivation of the remaining relationships, to complete or fill up the Conditional Probability Tables (CPTs), via different methods. For the first two approaches, the issues to be addressed are typical of applications in which a large number of probabilities are elicited, e.g. avoiding different types of biases and ensuring consistency in the assessments (these issues are

* Corresponding author. Tel.: +41 56 310 53 56.
E-mail address: luca.podofillini@psi.ch (L. Podofillini).

http://dx.doi.org/10.1016/j.ress.2016.01.004
0951-8320/

presented in detail in [8,9]). Of course, when applying filling-up methods, these issues may also need to be addressed. These methods to populate CPTs on the basis of selected, elicited model relationships are referred to in this paper as CPT building methods.

The most popular method to populate CPTs from partial information is based on Noisy-OR gates [4,10,11]. The Noisy-OR model entails assessing the effect on the outcome of the presence of one factor at a time, with all other factors being absent. In their typical implementation, Noisy-OR gates require binary BBN nodes and model factor influences as independent of the presence of the other factors. A number of extensions of the Noisy-OR model have been developed, generally addressing either dependent influences or multi-state nodes (see e.g., [12] for a brief summary). Alternative methods adopt interpolation algorithms [13–15] that typically extract information on the factor influences from selected, specifically elicited CPDs. Fenton et al. [16] base their method on a further concept, according to which CPTs are derived from weighted functions of the influencing factors.

The available CPT building methods differ in their theoretical basis, the base information on which they derive CPDs, the elicitation requirements, and the interpretation and extrapolation of the factor influences. The present paper analyses a selection of these methods for use in Human Reliability Analysis (HRA) applications. HRA is the area of risk analysis dealing with identifying risk-significant human failure events, understanding and modelling their causes and influencing factors, and quantifying their probability. Besides sharing many aspects with BBNs for more general risk analysis applications, BBNs for HRA often attempt to formally combine cognitive models, empirical data, and expert judgment with the aim to enhance the empirical basis of HRA methods, [7,17]. BBNs have a number of attractive features for HRA and in general for fields with shortage of data and consequent reliance on subjective judgments: their intuitive graphical representation, the possibility of combining diverse sources of information, the use the probabilistic framework to characterize uncertainties. In terms of modelling capabilities, BBNs allow modelling strong factor effects and interactions: this potentially allows (provided that these effects can be quantified) to overcome the assumption of some methods of independent factor effects (e.g. SPAR-H [18], HEART [19]). Attractive features and research gaps of the BBN modelling framework are discussed in a recent review by the authors of the present paper [7]. Given the relevance of the issue of CPT building and the variety of available methods, it becomes important to clarify their strengths and limitations, to evaluate their suitability for HRA, and to identify gaps to be addressed by research.

The present analysis focuses on methods applicable to BBNs with multi-state nodes; these are better suited for HRA applications than those with binary nodes because the characterization of influencing factors in HRA methods generally involves multiple levels, e.g. SPAR-H [18], HEART [19] and CREAM [20]. The performance of five methods is examined: the functional interpolation method [14], the Elicitation BBN (EBBN) method [13], the Cain calculator [15], Fenton et al. [16] and Røed et al. [2]. Two small BBNs representing very simplified HRA models were used to benchmark the performance. Small BBNs were selected to allow a comprehensive comparison of the produced CPDs. The two examples are designed to test some aspects relevant for HRA modelling: the representation of strong factor influences and of factor interactions, the representation of uncertainty on the BBN relationships, and the method requirements as the size of the BBN increases.

Although on-going data collection efforts [21–23] aim at reducing the need for expert judgment in HRA, the need to combine empirical data and expert judgment is likely to persist for at least the middle term. This will be the case especially for application scopes for which data will continue to be difficult to obtain (e.g., in nuclear PSA applications, HRA for accident mitigation conditions and external initiating events) and also for industrial sectors in which the collection of HRA data may be less advanced than in the nuclear industry. Recent studies providing empirical human error probability estimates focus on data usable for quantification of first generation HRA methods [24–26]. Current efforts on data collection for newer generation methods (emphasizing the role of the context and decision-making errors) are not yet providing statistically solid figures [21,23,27]. This continued need for expert inputs to HRA methods motivates this work's focus on the use of expert data in the construction of BBNs for HRA. The choice of BBNs with multi-state nodes is also related to this focus. A motivation for the adoption of binary nodes in the recent efforts to use empirical data to develop BBNs for HRA [17,28] is that it makes data collection easier and statistically stronger. On the other hand, binary BBN nodes are generally associated with models based primarily on the presence or absence of an influencing factor. Reducing the number and complexity of the model relationships is an advantage but may make the expert elicitation more difficult because factors may be defined less specifically, their states enveloping broader sets of conditions.

Note that in the literature, the application of BBNs to HRA problems goes beyond models solely built on expert judgment, addressed by the present paper. For example, the already mentioned [17,28] quantify the CPDs from databases of human failure events (although in both [17,28] the available data is not sufficient to a statistically solid determination of the CPDs and expert judgment is used in combination). Other studies develop BBNs as extensions of existing HRA methods, such as SPAR-H [29]: in these cases the quantification of BBNs is based on the underling method relationships. CPT populating methods have also been applied in the HRA literature. A Noisy-Or filling up algorithm is used in [30], therefore assuming independent parent nodes. Reference [31] determines the CPDs by linear interpolation according to the number of states of the parent nodes representing positive conditions. As discussed in [7], the interpolation does not differentiate among configurations with different parent nodes in the same number of positive states. For HRA applications, however factors can be characterized by strong interactions so that their effect can be strongly dependent on which factors are in their positive and negative states. In [32], pre-defined triangular functions are associated to different factor strength category. The functions for each strength category differ by their variance: the stronger the influence, the smaller the variance. The functions for each parent are then weighted to derive the child CPT. The method in [32] shares similar characteristics with the Fenton et al. and Røed et al. methods (use of weights and influence functions) and therefore is not included in the present evaluation.

The rest of this paper is organized as follows. Section 2 briefly presents BBNs and introduces CPT building methods in general. Section 3 presents the design of the evaluation study. This is based on testing the methods on how they reproduce relevant modelling features for HRA, implemented on specific evaluation criteria. Two simple HRA models are introduced (BBN 1, BBN 2), so that the differences in the CPTs produced by the methods can be easily traced. Section 4 presents the application of the five methods to the two HRA models, on two analysis cases for each model (cases NOUNC and UNC). Section 5 presents the evaluation of the methods. Section 6 provides concluding remarks.

2. Bayesian Belief Networks: introduction

This section briefly introduces BBNs. Part of the section is taken from [7]. For extensive treatment of the BBN modelling framework the reader should refer to [4] and [10], for example.