Artificial Intelligence in Medicine 66 (2016) 41–52

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

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

Value of information analysis for interventional and counterfactual Bayesian networks in forensic medical sciences

Anthony Costa Constantinou*, Barbaros Yet, Norman Fenton, Martin Neil, William Marsh

Risk and Information Management Research Group, School of Electronic Engineering and Computer Science, Queen Mary University of London, Mile End Road, Mile End Campus, Computer Science Building, E1 4NS London, UK

ARTICLE INFO

Article history:
Received 28 March 2015
Received in revised form 18 August 2015
Accepted 2 September 2015

Keywords:
Causal inference
Bayesian networks
Interventional analysis
Counterfactual analysis
Value of Information
Forensic medicine

ABSTRACT

Objectives: Inspired by real-world examples from the forensic medical sciences domain, we seek to determine whether a decision about an interventional action could be subject to amendments on the basis of some incomplete information within the model, and whether it would be worthwhile for the decision maker to seek further information prior to suggesting a decision.
Method: The method is based on the underlying principle of *Value of Information* to enhance decision analysis in *interventional* and *counterfactual* Bayesian networks.
Results: The method is applied to two real-world Bayesian network models (previously developed for decision support in forensic medical sciences) to examine the average gain in terms of both *Value of Information* (average relative gain ranging from 11.45% and 59.91%) and decision making (potential amendments in decision making ranging from 0% to 86.8%).
Conclusions: We have shown how the method becomes useful for decision makers, not only when decision making is subject to amendments on the basis of some unknown risk factors, but also when it is not. Knowing that a decision outcome is independent of one or more unknown risk factors saves us from the trouble of seeking information about the particular set of risk factors. Further, we have also extended the assessment of this implication to the counterfactual case and demonstrated how answers about interventional actions are expected to change when some unknown factors become known, and how useful this becomes in forensic medical science.

1. Introduction

Value of Information (*VoI*) is a technique initially proposed in economics [1] for the purposes of:

1. determining the amount a decision maker would be willing to pay for further information; and
2. prioritising unobserved model factors for acquiring information based on their impact against a desired utility value or probability distribution.

VoI analysis has subsequently been adopted in a number of domains including finance [2], supply chain management [3], pharmaceuticals [4], and health care [5].

An especially important application domain is medicine. For example, *VoI* has been used:

1. as a decision analytic approach to clinical trial design and research priority-setting, by taking into consideration the costs of sampling, the benefits of the sample information, and the decision rules of the cost-effectiveness analysis [6];
2. to determine optimal sample size for clinical trials as an alternative to the more traditional null hypothesis methods [7–10];
3. for the development and evaluation of clinical trials [11,12];
4. to investigate the expected value of partial perfect information, and the research decision it can address in medical decision making [13];
5. as a guide to evaluate decision support for differential diagnosis [14];
6. as a decision analysis technique to identify the most beneficial factors in health economic models [5,15–17].

* Corresponding author. Tel.: +44 7903292836.
E-mail address: anthony@constantinou.info (A.C. Constantinou).

http://dx.doi.org/10.1016/j.artmed.2015.09.002
0933-3657/

For a comprehensive review of *VoI* analyses related to health risk management see [18].

In this paper, we are interested in using *VoI* to determine whether missing information can lead to different interventional actions in decision analysis with Bayesian networks (BNs). The use of *VoI* for interventions has previously been explored in [19] where *VoI* is used to identify novel actions (a process which the authors call *search for opportunities*) in influence diagrams, in the sense that interventions are identified to improve a desirable utility function. More recently, in [20] *VoI* is also considered as an evaluation method for interventional strategies in epidemiology, under competing models, and to quantify the benefit of adaptive versus static intervention strategies. Our major contribution here is to extend *VoI* for interventional decision analysis to the counterfactual setting. This allows decision makers to compare the observed results of the actual world to those of a hypothetical world; i.e. what would have happened had we proposed treatment (or intervention) *B* instead of treatment *A*. To the best of our knowledge, there have been no previous attempts to incorporate the concept of *VoI* to counterfactual problems with BNs.

Our application of *VoI* is motivated by real-world problems in forensic medical sciences in which BNs were developed for decision making. BNs are based on sound foundations of causality and conditional probability theory. Our objective is to show how *VoI* can be applied to BNs to make them especially suitable for simulating interventions and inferring answers from counterfactual questions.

The paper is structured as follows: Section 2 describes the forensic medical science problem motivating this work; Section 3 provides the necessary background overview of the methods: *VoI*, BNs, interventional and counterfactual analysis; Section 4 demonstrates the modelling process of integrating *VoI* analysis into interventional and counterfactual BN decision analysis models; Section 5 demonstrates and discusses the results generated by applying the method to two real-world forensic medical case studies; we provide our concluding remarks in Section 6.

2. Motivation: the forensic mental health problem

Forensic medical practitioners and scientists based at the Violence Prevention Research Unit[1] (VPRU); Queen Mary University of London have, for several years, sought improved decision support for determining care and release of people with mental health problems. In particular, they are interested in managing the risk of violent reoffending by releasing such convicted prisoners from prison and discharging such patients from medium secure services [21]. In collaboration with the medical practitioners we have developed two BN models for this purpose – one for prisoners and one for patients [22,23]. These models delivered significantly improved predictive accuracy with respect to whether a prisoner/patient is determined suitable for release/discharge (hereafter referred to simply as 'release'). The models also provided the additional benefits that causal BN models provide over and above black-box decision models (see Chapters 2 and 3 of [24] for a detailed discussion). However, while, those models were developed for the purpose of simulating interventions (i.e. treatments/therapies) for violence risk management, prior to releasing an individual, they did not consider the possibility that decisions about release could be subject to amendments on the basis of some incomplete information within the model. The BN models were large and complex. Consequently, when assessing an individual for release, information was very often missing for variables that could have been observed.[2]

Specifically, a decision maker (such as a probation officer or a clinician) has to determine whether to release a prisoner/patient based on the probability distribution (or the expected value) of the hypothesis variable; i.e. the risk of violence assuming release. Prior to deciding on release, the decision maker has the option to simulate various interventions for the purpose of determining whether an individual's risk of violence can be managed to acceptable levels. Additionally, the decision maker may have the option to gather further information about the individual. While any set of unknown information can still be estimated on the basis of Bayesian inference (via observations provided to other relevant factors within the BN model) it is still possible that knowing (rather than estimating) one or more of these unobserved factors, may lead to amendments in the probation officer's original decision about release.

3. Methods

While a detailed description of the four constituent methods, BNs, *VoI* analysis, interventional and counterfactual analysis is beyond the scope of this paper, this section provides sufficient background to understand the modelling process demonstrated in Section 4.

3.1. Bayesian networks (BNs)

BNs, also sometimes known as *belief networks* or *causal probabilistic networks*, are directed acyclic graphical models [26]. They consist of nodes which represent uncertain variables, and arcs which represent causal or influential relationships between the variables. The 'Bayesian' in BNs is due to the use of Bayes' theorem for revising probabilities. Bayes' theorem is a simple equation that specifies how to calculate conditional probabilities:

$$p(A \mid B) = \frac{p(B \mid A) \times p(A)}{p(B)}$$

where $p(A)$ is the *prior* probability of A and $p(B|A)$ is the likelihood of B given A. The probability $p(A|B)$ is called the *posterior* probability of A. In its prior state all of the variables in a BN are uncertain and assumed to be provisional upon experience/data gained to date. This prior probability is then revised based on new experience/data, to provide the updated *posterior* probability.

Fig. 1 presents a very simple BN with just two variables and one dependency. The example is based on a well-known probability problem [25], where a test to detect a disease whose prevalence is 1 in a 1000 has a false-positive rate of 5%. Fig. 1.1 presents this problem with both variables being unknown (i.e. the prior marginal probabilities reflecting the average individual). Further, Fig. 1.2 presents the posterior probabilities for *Test* given the two possible knowns for *Disease*, whereas Fig. 1.3 presents the posterior probabilities for *Disease* given the two possible knowns for *Test*. While case (2) demonstrates how the cause node affects the probabilities of the effect node, case (3) demonstrates how inference propagates backwards to the cause node having observed the effect, and this is what makes BNs unique for decision analysis. For further reading in BNs see [24,27].

[1] Formerly known as Forensic Psychiatry Research Unit (FPRU).

[2] Some variables in BNs are supposed to be unobserved. For instance, specific type of latent or uncertain synthetic variables. These also include variables representing symptoms post-treatment, on the basis of some imperfect intervention (see Fig. 3). In this paper we are only interested in variables with missing information; i.e. those that are not observed, but could have been observed.