Probabilistic Logic with Conditional Independence Formulae

Magdalena Ivanovska and **Martin Giese**[1]

Abstract. We investigate a probabilistic propositional logic that is an extension of the logic proposed by Fagin et al. [4] with conditional independence formulae. We give an axiomatization which we show to be complete for the kind of inferences allowed by Bayesian networks.

1 Introduction

We report[2] work carried out within the CODIO project on COllaborative Decision support for Integrated Operations. The goal of this project is to develop a decision support system for petroleum drilling operations. As one part of the CODIO project, we have designed a system based on a Bayesian network [9] model [6]. When testing this approach on case data from a major oil company, Bayesian networks as a modeling tool showed some shortcomings. It seems as if a logic-based approach to decision support would make it possible to resolve some of the issues like inability to cope with imprecise information, time, and the fact that different people with different information are involved in the decision making.

We are currently in the process of investigating logics to express decision problems. The topic of this paper is a small but important part of this endeavor, namely the combination of quantitative reasoning about uncertainty with the qualitative reasoning about conditional independence which is central to approaches based on Bayesian networks, but which has received comparatively little attention from the logical side. There have previously been attempts at a combination of the expressiveness and the inference mechanisms of probabilistic logics and probabilistic networks, see e.g. [7, 3, 1, 2]. Our work differs from these in that we represent all information within the logical formalism, while most others represent independence using a network. Moreover, we describe inference with an axiomatic system instead of referring to linear or non-linear programming. While this surely doesn't give the most efficient reasoning, we consider the sufficiency of a relatively small axiomatic system to be interesting in its own right. It will, for instance make it easier to combine the formalism with temporal, epistemic, or other kinds of reasoning.

2 Probabilistic Propositional Logic with Independence Formulae

Following Fagin, Halpern, and Megiddo [4], we consider a probabilistic propositional logic obtained by augmenting propositional logic with *linear likelihood formulae*

$$a_1\ell(\varphi_1)+\cdots+a_k\ell(\varphi_k)\geq a \quad ,$$

where $a_1,\ldots,a_k,a$ are real numbers and $\varphi_1,\ldots,\varphi_k$ are *pure propositional formulae*. The intention is that $\ell(\varphi)$ expresses the probability of φ being true, and the language allows expressing arbitrary linear relationships between such probabilities.

This logic is interpreted over is a set of possible worlds, μ is a probability measure that assigns a value in $[0,1]$ to any subset of W, and π is an interpretation function. To each element $w\in W$, the function π assigns a truth-value function π_w that fixes the interpretation of propositional letters. We extended π_w to arbitrary formulae in the usual way, where the interpretation of linear likelihood formulae is defined as follows:

$\pi_w(a_1\ell(\varphi_1)+\cdots+a_k\ell(\varphi_k)\geq a)=1$ iff $a_1\mu(\varphi_1^M)+\cdots+a_k\mu(\varphi_k^M)\geq a$, where $\varphi^M:=\{w|\pi_w(\varphi)=1\}$ for any φ.

Conditional likelihood formulae can be introduced as abbreviations: $\ell(\varphi/\psi)\geq c$ is defined as $\ell(\varphi\wedge\psi)-c\ell(\psi)\geq 0$. We also introduce $\ell(\varphi/\psi)\leq c$ and $\ell(\varphi/\psi)=c$ in the obvious way. Linear combinations of conditional likelihood terms are not allowed.

Fagin et al. [4] give a sound and complete axiomatization consisting of the following axioms and inference rules:

Prop All substitution instances of tautologies in propositional logic;
QU1 $\ell(\varphi)\geq 0$;
QU2 $\ell(\top)=1$;
QU3 $\ell(\varphi)=\ell(\varphi\wedge\psi)+\ell(\varphi\wedge\neg\psi)$, where φ and ψ are pure propositional formulae;
Ineq All substitution instances of valid linear inequality formulae;
MP From f and $f\Rightarrow g$ infer g;
QUGen From $\varphi\Leftrightarrow\psi$ infer $\ell(\varphi)=\ell(\psi)$.

What is *not* expressible in this logic is stochastic independence of formulae. In fact, it is not hard to see that any statement about independence leads to non-linear statements about probabilities.

We choose to represent conditional independence by defining *conditional independence formulae* (CI-formulae)

$$I(\mathbf{X}_1,\mathbf{X}_2/\mathbf{X}_3) \quad ,$$

where $\mathbf{X}_1$, $\mathbf{X}_2$, and $\mathbf{X}_3$ are sets of propositional letters.

We interpret the CI-formulae in a structure $M=(W,\mu,\pi)$ in the following way: $\pi_w(I(\mathbf{X}_1,\mathbf{X}_2/\mathbf{X}_3))=1$ iff $I_\mu(\mathbf{X}_1^M,\mathbf{X}_2^M/\mathbf{X}_3^M)$, where $\mathbf{X}_i^M=\{X^M\mid X\in\mathbf{X}_i\}$.

We denote the logic given by this syntax and semantics by $\mathbf{L}$.

3 Expressing Bayesian Networks

A Bayesian network (BN) consists of a DAG (directed acyclic graph) with a set of nodes V representing random variables and cpts (conditional probability tables) for every node (variable) in the graph.

---

[1] University of Oslo, Norway, email: magdalei/martingi@ifi.uio.no

[2] See [8] for a longer and formally more accurate version of this paper including proofs.

A cpt gives the probability of every value of a node's variable, conditional on every possible combination of values of the parent node's variables. The graph structure captures qualitative information about conditional independencies among the random variables in V, whilst the cpt entries determine conditional probability distributions of these random variables. It can be shown that each BN determines a unique joint probability distribution of its variables; we will refer to it as to the *probability distribution described by the network*. We refer to the literature on BNs for details [9]. We restrict ourselves to binary BNs, where variables can have only two states 0 and 1.

Both qualitative and quantitative information embedded in a binary BN can be appropriately represented in our logical language **L**. For example, if a network consists of the nodes A, B, and C, and edges (B,A) and (C,A) and the cpt contains the following information about the conditional probability distribution: $p(A=1/B=1,C=1)=a_1$, $p(A=1/B=1,C=0)=a_2$, $p(A=1/B=0,C=1)=a_3$ and $p(A=1/B=0,C=0)=a_4$, then this information can be given in our language in the following way: $\ell(A/B\wedge C)=a_1$, $\ell(A/B\wedge\neg C)=a_2$, $\ell(A/\neg B\wedge C)=a_3$ and $\ell(A/\neg B\wedge\neg C)=a_4$. In addition to this we add the conditional independence formulae $I(A,\{B,C\}/\{B,C\})$, $I(B,C/\emptyset)$ and $I(C,B/\emptyset)$ to represented the qualitative information about p.

We generalize the idea in the following definition. Before we give it, we define an *atom* to be a conjunction of possibly negated propositional letters. An *S-atom* for some set $S\subseteq\mathbf{P}$ is an atom that contains all the letters from S. It is easy to see that every atom corresponds to one assignment of the variables represented by the letters in it. In what follows we will use the same notation for both the atoms and the corresponding assignments.

Definition 1 *Let* **BN** *be the class of all binary BNs and* **F** *be the set of all formulae in* **L**. *The* specific axioms function, $Ax:\mathbf{BN}\to 2^{\mathbf{F}}$ *is a function that to each BN $\mathscr{B}$ assigns the set of formulae containing*

- $\ell(X/\delta)=c$ *for every node $X\in V$ and every $Pa(X)$-atom δ, where $c=p(X=1/\delta)$ and $Pa(X)$ are the parents of X, and*
- $I(X,ND(X)/Pa(X))$ *for every node $X\in V$.*

Bayesian networks capture only the probabilities of combinations of certain events identified by the nodes, and not a probability measure on some underlying set of worlds or elementary outcomes. Also the formulae in our logic only describe certain properties of models. It is easy to show however that any conditional probability of conjunctions of literals that can be computed from a BN can also be inferred from the network's set of specific axioms.

Theorem 1 *Let $\mathscr{B}$ be a binary BN and let p be the probability distribution described by $\mathscr{B}$. If $p(\varphi/\psi)=b$, then $Ax(\mathscr{B})\models\ell(\varphi/\psi)=b$, where φ and ψ are atoms (resp. the corresponding assignments).*

While Theorem 1 guarantees the semantic connection between a BN $\mathscr{B}$ and its axiomatization $Ax(\mathscr{B})$ in our logic **L**, it says nothing about the derivability of conditional likelihood formulae. This will be covered in the following section.

4 Axiomatic System and Theorem for Syntactic Entailment

The axiomatic system given in Sect. 2 is complete for reasoning in a logic without CI-formulae. The axiomatization of conditional independence has been the subject of a certain amount of research, see e.g. [5] for a survey, and it has been shown that no complete axiomatization of independence statements exists.

In this work, we are less interested in deriving new CI formulae. We want to mimic the kind of reasoning possible with a BN, so we want a system that allows to derive arbitrary statements about conditional likelihood of propositional formulae from the specific axioms $Ax(\mathscr{B})$ of a BN.

Our axiomatic system consists of four parts, each dealing with a different type of reasoning: propositional reasoning, reasoning about probability, reasoning about linear inequalities, and reasoning about conditional independence. For the first three parts, we use the axioms from the system AX_{MEAS} in [4], as given in Sect. 2. For conditional independence reasoning, we add the following inference rules:

SYM From $I(\mathbf{X}_1,\mathbf{X}_2/\mathbf{X}_3)$ infer $I(\mathbf{X}_2,\mathbf{X}_1/\mathbf{X}_3)$
DEC From $I(\mathbf{X}_1,\mathbf{X}_2\cup\mathbf{X}_3/\mathbf{X}_4)$ infer $I(\mathbf{X}_1,\mathbf{X}_2/\mathbf{X}_4)$
IND From $I(\mathbf{X}_1,\mathbf{X}_2/\mathbf{X}_3)$ and $\ell(\varphi_1/\varphi_3)\leq(\geq)a$ infer $\ell(\varphi_1/\varphi_2\wedge\varphi_3)\leq(\geq)a$, where φ_i is an arbitrary $\mathbf{X}_i$-atom, for $i\in\{1,2,3\}$.

Theorem 2 *Let $\mathscr{B}$ be a binary BN and let p be the probability distribution described by $\mathscr{B}$. If $p(\varphi/\psi)=b$, then $Ax(\mathscr{B})\vdash\ell(\varphi/\psi)=b$, where φ and ψ are atoms (resp. the corresponding assignments).*

5 Future Work

Credal networks [2] allow to give ranges of probabilities in cpts instead of sharp values. This can be useful in applications when probability elicitation is difficult. While such ranges can be expressed in our logic, the given axiom system is not complete for the inference of tight bounds on arbitrary conditional probabilities. We would like to investigate whether our calculus can be extended so as to obtain completeness also in this case. We also want to explore more efficient reasoning techniques, possibly by recasting ideas from Bayesian networks into a logic setting.

We are currently in the process of extending our logic with a mechanism for expressing decision scenarios, including options and utilities, which will make it possible to infer optimal strategies from a set of observations.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous referees for their useful comments. The CODIO project is partially funded by grants from the Research Council of Norway, Petromaks Project 175899/S30.

REFERENCES

[1] K. A. Andersen and J. N. Hooker, 'Bayesian logic', *Decis. Support Syst.*, **11**(2), 191–210, (1994).
[2] F. G. Cozman, 'Credal networks', *Artif. Intell.*, **120**(2), 199–233, (2000).
[3] F. G. Cozman, C. P. de Campos, and J. C. Ferreira da Rocha, 'Probabilistic logic with independence', *Int. J. Approx. Reasoning*, **49**(1), 3–17, (2008).
[4] R. Fagin, J. Y. Halpern, and N. Megiddo, 'A logic for reasoning about probabilities', *Information and Computation*, **87**, 78–128, (1990).
[5] D. Geiger and J. Pearl, 'Logical and algorithmic properties of conditional independence and graphical models', *Ann. Statist.*, **21**(4), 2001–2021, (1993).
[6] M. Giese and R. B. Bratvold, 'Probabilistic modelling for decision support in drilling operations, SPE 127761', in *Proc. SPE Intelligent Energy Conference*, (2010).
[7] R. Haenni, J. W. Romeijn, G. Wheeler, and J. Williamson. Probabilistic logic and probabilistic networks, December 2008. Draft.
[8] Magdalena Ivanovska and Martin Giese, 'Probabilistic logic with conditional independence formulae', Technical Report 395, Univ. of Oslo, Dept. of Informatics, (2010). available from `http://heim.ifi.uio.no/martingi/pub/ecai2010long.pdf`.
[9] R. E. Neapolitan, *Learning Bayesian Networks*, Prentice Hall, 2003.