Journal of Applied Logic ••• (••••) •••–•••

Contents lists available at ScienceDirect

Journal of Applied Logic

www.elsevier.com/locate/jal

An overview of algorithmic approaches to compute optimum entropy distributions in the expert system shell MECore (extended version)

Nico Potyka*, Engelbert Mittermeier, David Marenke

Dept. of Computer Science, FernUniversität in Hagen, 58084 Hagen, Germany

A R T I C L E I N F O

Article history:
Available online xxxx

Keywords:
Probabilistic Reasoning
Probabilistic Conditional Logic
Maximum Entropy
Uncertain Reasoning

A B S T R A C T

The expert system shell MECore provides a series of knowledge management operations to define probabilistic knowledge bases and to reason under uncertainty. To provide a reference work for MECore algorithmics, we bring together results from different sources that have been applied in MECore and explain their intuitive ideas. Additionally, we report on our ongoing work regarding further development of MECore's algorithms to compute optimum entropy distributions and provide some empirical results. Altogether this paper explains the intuition of important theoretical results and their practical implications, compares old and new algorithmic approaches and points out their benefits as well as possible limitations and pitfalls.

1. Introduction

When designing expert systems, classical logics fail to meet the demands of reality, as they cannot deal with uncertain information. Probabilistic logics [33,2,36] provide a natural extension to deal with uncertainty. Numerous approaches have been developed to deal with the accompanying computational complexity [25,17,40] or to deal with the combination of statistical data and expert knowledge [35]. We focus on probabilistic conditional logics [42,27,22] here. Experts can define intuitive rules of the form 'if A then B with probability x', which are formalized by *conditionals* $(B|A)[x]$. A and B can be arbitrary logical formulas and x is a probability that can express an expert's degree of belief or can be obtained by statistical means.

The expert system shell MECore [14] allows the knowledge engineer to enter conditional knowledge bases and to accomplish probabilistic reasoning under the principle of optimum entropy [36,22]. A comprehensive case study can be found in [3], where a knowledge base for analyzing brain tumor data is designed by

* Corresponding author.
E-mail address: nico.potyka@uni-osnabrueck.de (N. Potyka).

http://dx.doi.org/10.1016/j.jal.2016.05.003

combining both expert knowledge and statistical data. One of the most important operations for MECore is to adapt the current epistemic state, which is basically represented by a probability distribution, with respect to new knowledge. This is accomplished by computing I-projections, i.e., by computing a new epistemic state that satisfies the new knowledge and is closest to the old state. The 'distance' between two such epistemic states is measured by the relative entropy [7].

We report on our ongoing work on improving the efficiency of computing I-projections in the context of conditional knowledge bases. In Section 2, we recall the probabilistic logical framework and the basics of the principle of optimum entropy as they are needed in the following. In Section 3, we provide some intuition for MECore's original I-projection algorithm (*IIP*) and a novel implementation based on *L-BFGS* [26]. In Section 4, we summarize some approaches to speed up the computation by taking the structure of conditional knowledge bases into account [22,13]. We also provide some experimental results to illustrate the benefit and shortcomings of different approaches. This article is an extended version of [39].

2. Basics

2.1. Probabilistic conditional logic and optimum entropy

We consider a propositional conditional language $\mathcal{L}$ built up over a finite set of propositional variables Σ using logical connectives like conjunction and disjunction in the usual way. For formulas $\psi, \phi \in \mathcal{L}$, we abbreviate negation $\neg\psi$ by an overbar $\bar{\psi}$ and conjunction $\psi \wedge \phi$ by juxtaposition $\psi\phi$. A *possible world* is a classical logical interpretation $\omega : \Sigma \to \{0, 1\}$ assigning a truth value to each propositional variable. We also allow multi-valued variables that are associated with a domain and an interpretation assigns to each such variable a value from the corresponding domain. Let Ω denote the set of all possible worlds. An atom $\mathfrak{a}$ is satisfied by ω iff $\omega(\mathfrak{a}) = 1$. More generally, for a multi-valued variable X, $x \in \text{domain}(X)$, the multi-valued atom $X = x$ is satisfied by ω iff $\omega(X) = x$. The definition is extended to complex formulas in the usual way. The *classical models* $\text{Mod}(\psi)$ of a formula ψ are the possible worlds $\omega \in \Omega$ that satisfy ψ.

A probabilistic conditional language over $\mathcal{L}$ thus arises as follows: $(\mathcal{L}|\mathcal{L}) := \{(\psi|\phi)[x] \mid \psi, \phi \in \mathcal{L}, x \in [0, 1]\}$. Intuitively a conditional $(\psi|\phi)[x]$ expresses that our belief in ψ given that ϕ holds is x. A conditional knowledge base $\mathcal{R} \subset (\mathcal{L}|\mathcal{L})$ is a set of conditionals.

Example 1. In [3], the brain tumor domain is modeled by 9 propositional variables. Two boolean variables *warningSymptoms* and *icpSymptoms* indicate the presence of important symptoms. The remaining variables are multi-valued. For instance, the 3-valued variable *age* can take the values *le20*, *20to80*, *ge80* representing three age groups ($age \leq 20$, $20 < age < 80$, $age \geq 80$). The variable *diagnosis* can take 11 values corresponding to different brain tumor types. The remaining variables model the malignancy of the tumor, the physical fitness of the patients, possible therapies, possible complications and possible outcomes for the health of the patient after inpatient stay.

The knowledge base contains both statistical knowledge and subjective expert beliefs. The statistical knowledge includes empirical frequencies of certain brain tumor types. For instance, the conditionals

$$(\textit{diagnosis} = \textit{meningeoma} \,|\, !(\textit{age}{=}\textit{le20}))[0.2]$$

$$(\textit{diagnosis} = \textit{medulloblastoma} \,|\, !(\textit{age}{=}\textit{le20}))[0.07]$$

express relative frequencies for two brain tumor types among people older than 20 (the exclamation mark ! expresses negation in MECore). The subjective knowledge includes conditionals like

$$(\textit{diagnosis} = \textit{glioblastoma} \,|\, !(\textit{age}{=}\textit{le20}) \textit{ and warningSymptoms})[0.2]$$

expressing possible diagnoses for typical states of patients.

Please cite this article in press as: N. Potyka et al., An overview of algorithmic approaches to compute optimum entropy distributions in the expert system shell MECore (extended version), J. Appl. Log. (2016), http://dx.doi.org/10.1016/j.jal.2016.05.003