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

Contents lists available at ScienceDirect

Journal of Applied Logic

www.elsevier.com/locate/jal

Algebraic model counting

Angelika Kimmig [a,*], Guy Van den Broeck [b], Luc De Raedt [a]

[a] *Department of Computer Science, KU Leuven, Celestijnenlaan 200a, Box 2402, 3001 Heverlee, Belgium*
[b] *UCLA, Computer Science Department, 4531E Boelter Hall, Los Angeles, CA 90095-1596A, USA*

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Knowledge compilation
Model counting
Logic

ABSTRACT

Weighted model counting (WMC) is a well-known inference task on knowledge bases, and the basis for some of the most efficient techniques for probabilistic inference in graphical models. We introduce algebraic model counting (AMC), a generalization of WMC to a semiring structure that provides a unified view on a range of tasks and existing results. We show that AMC generalizes many well-known tasks in a variety of domains such as probabilistic inference, soft constraints and network and database analysis. Furthermore, we investigate AMC from a knowledge compilation perspective and show that all AMC tasks can be evaluated using `sd-DNNF` circuits, which are strictly more succinct, and thus more efficient to evaluate, than direct representations of sets of models. We identify further characteristics of AMC instances that allow for evaluation on even more succinct circuits.

1. Introduction

Today, some of the most efficient techniques for probabilistic inference employ reductions to weighted model counting (WMC) both for propositional and for relational probabilistic models [21,22,6,12,24]. The resulting weighted model counting task is often solved by a single pass over a propositional circuit, which is a compact graphical representation of the models of interest. This approach makes it possible to perform the possibly expensive knowledge compilation step, that is, the construction of the circuit, only once, and to then evaluate this circuit repeatedly, for instance, under different evidence or with different parameters.

On the other hand, it is well-known that probabilistic inference as well as many other tasks can be generalized to a sum of products computation over models with suitable operators from a semiring structure. This has led to common inference algorithms for a variety of different inference problems in many fields, including parsing [13], dynamic programming [10], constraint programming [20], databases [14], Bayesian inference [1], propositional logic [18], networks [2] and logic programming [17]. The work presented here

* Corresponding author.
E-mail addresses: angelika.kimmig@cs.kuleuven.be (A. Kimmig), guyvdb@cs.ucla.edu (G. Van den Broeck), luc.deraedt@cs.kuleuven.be (L. De Raedt).

http://dx.doi.org/10.1016/j.jal.2016.11.031
1570-8683/

Please cite this article in press as: A. Kimmig et al., Algebraic model counting, J. Appl. Log. (2016), http://dx.doi.org/10.1016/j.jal.2016.11.031

provides a unified view on these two lines of work by introducing both a general definition of model counting in a semiring setting and a solution approach for this task based on knowledge compilation.

As our first contribution, we introduce the task of *algebraic model counting (AMC)*. AMC generalizes weighted model counting to the semiring setting and supports various types of labels (or weights), including numerical ones as used in WMC, but also sets (e.g., to collect relevant variables), Boolean formulae (e.g., to obtain explicit representations of models), polynomials (e.g., for sensitivity analysis in probabilistic models), and many more. It thus provides a framework that covers many different tasks from a variety of different fields. As our second contribution, we investigate how to solve AMC problems using knowledge compilation. As AMC is defined in terms of the set of models of a propositional logic theory, we can exploit the succinctness results of the knowledge compilation map of Darwiche and Marquis [7]. We show that AMC can in general be evaluated using `sd-DNNF` circuits, which are more succinct, and thus more efficient to evaluate, than a direct representation of the set of models. Furthermore, we identify a number of characteristics of AMC tasks that allow for evaluation on even more succinct types of circuits. Our results provide a unified view on existing results, which also allows us to generalize well-known insights for satisfiability and model counting in circuits to broad classes of AMC tasks and to extend the task classification in algebraic Prolog [17] to more succinct types of circuits. As our third contribution, we further broaden the applicability of the AMC framework by linking it to semiring sums of products defined over derivations, that is, sequences of possibly repeated variables, instead of over models.

This paper is organized as follows. We introduce algebraic model counting in Section 2. Section 3 provides task characteristics that allow for correct evaluation on specific classes of circuits and shows how these generalize previous results. We discuss future work and conclude in Section 4.

2. Algebraic model counting

Our definition of algebraic model counting builds upon the well-known task of weighted model counting for propositional logic theories. Given a propositional logic theory T over a set of variables $\mathcal{V}$, an *interpretation* of $\mathcal{V}$ assigns a truth value from the set $\{\mathit{true}, \mathit{false}\}$ to every variable in $\mathcal{V}$. The set $\mathcal{M}(T)$ of *models* of theory T contains exactly those interpretations of $\mathcal{V}$ for which T evaluates to true. We here view interpretations (and models) as sets of literals, that is, for each variable $v \in \mathcal{V}$, an interpretation contains either the positive literal v or the negative literal $\neg v$. We use $\mathcal{L}$ to denote the set of literals for the variables in $\mathcal{V}$. In weighted model counting, non-negative real-valued weights are associated with all literals, and the weighted model count of a propositional theory is obtained by multiplying these weights for each model of the theory, and summing the results for all models.[1]

Definition 1 *(Weighted Model Counting (WMC)).* Given

- a *propositional logic theory* T over a set of variables $\mathcal{V}$ and
- a *weight function* $w : \mathcal{L} \to \mathbb{R}_{\geq 0}$, mapping literals of the variables in $\mathcal{V}$ to non-negative real-valued weights,

the task of *weighted model counting (WMC)* is to compute

$$\mathbf{WMC}(T) = \sum_{I \in \mathcal{M}(T)} \prod_{l \in I} w(l). \tag{1}$$

[1] The case where weights are associated with joint assignments to groups of variables, as in factor graphs, can be mapped to the case considered here; cf. Section 2.2.

Please cite this article in press as: A. Kimmig et al., Algebraic model counting, J. Appl. Log. (2016), http://dx.doi.org/10.1016/j.jal.2016.11.031