Proceedings of the Thirtieth International Florida Artificial Intelligence Research Society Conference

Typed Model Counting and Its Application to Probabilistic Conditional Reasoning at Maximum Entropy

Marco Wilhelm and **Gabriele Kern-Isberner**
Department of Computer Science
TU Dortmund, Germany

Abstract

Typed model counting expands model counting of propositional formulas by the ability to distinguish between certain types of models. Formally, we incorporate elements of a commutative monoid that represent these model types directly into the propositional formulas. An advantage of this approach is the ability of preserving information about which parts of a formula are satisfied by a certain type of model. We exploit this benefit when applying typed model counting to probabilistic conditional reasoning at maximum entropy. In particular, we address the task of determining the conditional structure induced by a reasoner's probabilistic conditional knowledge base in order to draw nonmonotonic inferences based on the maximum entropy distribution.

Introduction

The principle of maximum entropy (ME-principle) constitutes a most appropriate form of commensense probabilistic reasoning (Paris 1994) as it endows a reasoner with a nonmonotonic inference relation that allows for drawing inferences from incomplete knowledge. While fulfilling the paradigm of informational economy (Gärdenfors 1988), reasoning at maximum entropy is as cautious as possible. However, due to its non-transparency, the ME-principle is often perceived as a black box methodology.

In order to gain structural insights into ME-reasoning, we examine the conditional structure (Kern-Isberner 2001) that is induced by a probabilistic conditional knowledge base $\mathcal{KB}$, and we formulate ME-inferences based on this abstraction. A probabilistic conditional $(B|A)[p]$ is a statement of the form "if A holds, then B follows with probability p". Its logical part can be verified ($\omega \models A \wedge B$), falsified ($\omega \models A \wedge \neg B$), or found to be not applicable ($\omega \models \neg A$) with respect to a possible world ω. The conditional structure of ω encodes for every conditional in $\mathcal{KB}$ how its logical part is evaluated. As possible worlds with the same conditional structure also have the same ME-probability, the conditional structure allows for partitioning the set of possible worlds according to their influence on the ME-distribution and hence allows for a more generic view on drawing ME-inferences. We introduce the concept of typed model counting (TMC) that prove to be a convenient framework for examining the conditional structure. The formalization of TMC is the main contribution of this paper. It is based on a language $\mathcal{L}^{\mathcal{S}}$ that extends a propositional language by incorporating a commutative monoid, as for instance conditional structures. Hence, $\mathcal{L}^{\mathcal{S}}$ allows to represent both the knowledge stated in and the conditional structure imposed by a knowledge base within a shared framework. TMC then is the task of finding all models of a formula in $\mathcal{L}^{\mathcal{S}}$ and dividing them into certain types that are represented by the elements of the commutative monoid. With a view to the application of TMC to ME-reasoning, determining the equivalence classes of possible worlds induced by the conditional structure and calculating their cardinalities can be performed by TMC simultaneously. In this regard, TMC is an advancement of the multi-steps approach presented in (Wilhelm, Kern-Isberner, and Ecke 2016), where the determination of the equivalence classes and the actual model counting process were performed consecutively and the models were counted independently for every single equivalence class.

There is a strong connection between TMC and algebraic model counting (Kimmig, v. d. Broeck, and de Raedt 2012) as both map literals to elements of an algebraic structure. However, algebraic model counting does not allow formulas to contain elements of the algebraic structure themselves, which is necessary for computing conditional structures.

The rest of the paper is organized as follows: After recalling some basics of reasoning with probabilistic conditionals at maximum entropy, we present the idea of TMC. For this purpose, we introduce the structured language $\mathcal{L}^{\mathcal{S}}$ and show how TMC works. To this end, we introduce a normal form called sd-DNNF$^{\mathcal{S}}$ for formulas in $\mathcal{L}^{\mathcal{S}}$. Afterwards, we apply TMC to reasoning at maximum entropy and conclude. Due to space restrictions, we omit purely technical proofs.

Preliminaries

We consider a propositional language $\mathcal{L} = \mathcal{L}(\Sigma)$ over a finite set of atoms Σ. For formulas $A, B \in \mathcal{L}$, we write AB instead of $A \wedge B$ and $\overline{A}$ instead of $\neg A$. The set of atoms that appear in A is denoted with $\Sigma(A)$. A *literal* $\dot{\mathsf{a}}$ is either the atom a or its negation $\overline{\mathsf{a}}$. Thus, $\mathsf{Lit} = \{\mathsf{a} \mid \mathsf{a} \in \Sigma\} \cup \{\overline{\mathsf{a}} \mid \mathsf{a} \in \Sigma\}$ denotes the set of all literals. A subset of Lit is called *consistent* iff it does not contain both a and $\overline{\mathsf{a}}$ for any $\mathsf{a} \in \Sigma$.

Probabilistic Conditionals and Knowledge Bases

A *(probabilistic) conditional* $(B|A)[p]$ with $A, B \in \mathcal{L}$ and $p \in [0,1]$ is a formal representation of the statement "if A holds, then B follows with probability p". A finite set of conditionals is called a *knowledge base*.

Example 1. *Assume an agent believes that animals do fly or do not have wings with probability* 0.95*. Further, she is pretty sure that flying animals do have wings, namely with probability* 0.99*, and she thinks that animals that do not fly are likely to not be birds with probability* 0.8*. We consolidate these statements into the knowledge base $\mathcal{KB}_{fly}$ consisting of the three conditionals* $r_1 = (\mathsf{f} \vee \overline{\mathsf{w}}|\top)[0.95]$, $r_2 = (\mathsf{w}|\mathsf{f})[0.99]$, *and* $r_3 = (\overline{\mathsf{b}}|\overline{\mathsf{f}})[0.8]$.

The semantics of conditionals is based on possible worlds. Here, a *(possible) world* ω is represented by a *complete conjunction of literals*, i.e., every atom in Σ appears in ω exactly once, either positive or negated. The set of all possible worlds is denoted with Ω. There is a one-to-one correspondence between worlds and propositional interpretations, thus we may write $\mathcal{I}_\omega$ for the interpretation that is associated with the world ω. Further, we write $\omega \models A$ iff the formula $A \in \mathcal{L}$ is true in the interpretation $\mathcal{I}_\omega$.

A probability distribution $\mathcal{P}$ over Ω is a *model of the conditional* $(B|A)[p]$, written $\mathcal{P} \models (B|A)[p]$, iff $\mathcal{P}(A) > 0$ and $p = \frac{\mathcal{P}(AB)}{\mathcal{P}(A)}$, where the probability $\mathcal{P}(A)$ of a formula $A \in \mathcal{L}$ is defined by $\mathcal{P}(A) = \sum_{\omega \models A} \mathcal{P}(\omega)$. The probability distribution $\mathcal{P}$ is a *model of the knowledge base* $\mathcal{KB}$, also written $\mathcal{P} \models \mathcal{KB}$, iff $\mathcal{P}$ is a model of every conditional in $\mathcal{KB}$. If $\mathcal{KB}$ has at least one model, $\mathcal{KB}$ is called *consistent*.

Principle of Maximum Entropy (ME-Principle)

Usually, a consistent knowledge base $\mathcal{KB}$ has several models since the conditional probabilities in $\mathcal{KB}$ do not necessarily determine the probability of every single world. For reasoning tasks such as drawing inferences from $\mathcal{KB}$, it is useful to choose a certain model among them. From a logical point of view, the ME-*distribution* is preferable (Paris 1994), which is the unique distribution that maximizes the entropy $\mathcal{H}(\mathcal{KB}) = -\sum_{\omega \in \Omega} \mathcal{P}(\omega) \cdot \log(\mathcal{P}(\omega))$ among all models of $\mathcal{KB}$. Therefore, it adds the least amount of assumed information (Shore and Johnson 1980). Given the knowledge base $\mathcal{KB} = \{(B_1|A_1)[p_1], \dots, (B_n|A_n)[p_n]\}$, the ME-distribution can be obtained by calculating

$$\mathcal{P}_{\mathsf{ME}}(\mathcal{KB})(\omega) = \alpha_0 \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \alpha_i^{1-p_i} \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \alpha_i^{-p_i} \qquad (1)$$

for every $\omega \in \Omega$. Here, α_0 is a normalizing constant which ensures that $\mathcal{P}_{\mathsf{ME}}(\mathcal{KB})(\omega)$ is a probability, and the so-called *effects* α_i, $i = 1, \dots, n$, are impact values quantifying the influence of the i-th conditional on $\mathcal{P}_{\mathsf{ME}}(\mathcal{KB})(\omega)$. Formally, the effects are solutions of a nonlinear equation system. We refer to (Kern-Isberner 2001) for the technical details. Equation (1) implies that worlds which verify ($\omega \models A_i B_i$) and falsify ($\omega \models A_i \overline{B_i}$) the same conditionals also have the same ME-probability. Hence, it seems natural to build equivalence classes of worlds with respect to their evaluation of conditionals. Formally, this is done by the notion of *conditional structure* (Kern-Isberner 2001). The conditional structure of a world $\omega \in \Omega$ with respect to the knowledge base $\mathcal{KB} = \{(B_1|A_1)[p_1], \dots, (B_n|A_n)[p_n]\}$ is

$$\sigma_{\mathcal{KB}}(\omega) = \prod_{i=1}^{n} \sigma_i(\omega) = \prod_{i=1}^{n} \begin{cases} \mathbf{a_i^+} & \text{iff } \omega \models A_i B_i \\ \mathbf{a_i^-} & \text{iff } \omega \models A_i \overline{B_i} \\ \mathbf{1}_{\mathcal{G}} & \text{iff } \omega \models \overline{A_i} \end{cases}.$$

The symbols $\mathbf{a_i^+}$, $\mathbf{a_i^-}$, and $\mathbf{1}_{\mathcal{G}}$ reflect the three different ways in which (the logical part of) the i-th conditional in $\mathcal{KB}$ can be evaluated with respect to ω. More precisely, $\mathbf{a_i^+}$ stands for the verification of the i-th conditional, $\mathbf{a_i^-}$ for its falsification, and $\mathbf{1}_{\mathcal{G}}$ expresses that the conditional is not applicable in ω. Formally, the symbols are elements of a free abelian group $\mathcal{G}(\mathcal{KB})$ with identity element $\mathbf{1}_{\mathcal{G}}$ and basis $\{\mathbf{a_i^\pm} \mid i = 1, \dots, n,\ \pm \in \{+, -\}\}$. We omit the operation symbol and concatenate the group elements by juxtaposition. The conditional structure partitions Ω in the sense that two worlds ω and ω' are in the same equivalence class iff $\sigma_{\mathcal{KB}}(\omega) = \sigma_{\mathcal{KB}}(\omega')$. We denote the respective equivalence class with $[\omega]_{\sigma_{\mathcal{KB}}}$ and the set of all equivalence classes with $\Omega/\sigma_{\mathcal{KB}}$. Consequently, worlds within the same equivalence class, i.e., with the same conditional structure, also have the same ME-probability (cf. Eq. (1)). Therefore, it is not necessary to derive the ME-probability of every world but only of one representative per equivalence class, as well as the cardinalities of the equivalence classes, in order to fully determine $\mathcal{P}_{\mathsf{ME}}(\mathcal{KB})$. Besides a potential improvement of calculation performance (the number of worlds $|\Omega| = 2^{|\Sigma|}$ is exponentially large and thus a bottleneck for probabilistic reasoning), the transition to equivalence classes of worlds allows for a more structural view on the interdependencies between the conditionals in $\mathcal{KB}$ and on the ME-distribution itself. Therefore, it plays an important role in better understanding the ME-principle.

The ME-distribution yields a nonmonotonic inference relation which enables us to answer queries as per common sense. Given a knowledge base $\mathcal{KB}$ and a further conditional $(B|A)[p]$, we define the ME-*inference relation* $\mid\!\sim_{\mathsf{ME}}$ by $\mathcal{KB} \mid\!\sim_{\mathsf{ME}} (B|A)[p]$ iff $\mathcal{P}_{\mathsf{ME}}(\mathcal{KB}) \models (B|A)[p]$. It states that, given $\mathcal{KB}$, it is rational to assume that B follows from A with the conditional ME-probability $p = \frac{\mathcal{P}_{\mathsf{ME}}(\mathcal{KB})(AB)}{\mathcal{P}_{\mathsf{ME}}(\mathcal{KB})(A)}$. The fraction on the right-hand side can also be expressed in terms of equivalence classes of worlds:

$$p = \frac{\sum_{[\omega]_{\sigma_{\mathcal{KB}}} \in \Omega/\sigma_{\mathcal{KB}}} \mathcal{P}_{\mathsf{ME}}(\mathcal{KB})(\omega) \cdot c([\omega]_{\sigma_{\mathcal{KB}}}, AB)}{\sum_{[\omega]_{\sigma_{\mathcal{KB}}} \in \Omega/\sigma_{\mathcal{KB}}} \mathcal{P}_{\mathsf{ME}}(\mathcal{KB})(\omega) \cdot c([\omega]_{\sigma_{\mathcal{KB}}}, A)} \qquad (2)$$

where $c([\omega]_{\sigma_{\mathcal{KB}}}, A) = |\{\omega' \in [\omega]_{\sigma_{\mathcal{KB}}} \mid \omega' \models A\}|$ for $A \in \mathcal{L}$ (cf. (Wilhelm, Kern-Isberner, and Ecke 2016)). Besides calculating the ME-probability for a single representative of each equivalence class, there are two tasks to perform in order to draw inferences from $\mathcal{KB}$: (a) Find the equivalence classes of worlds by elaborating their conditional structure, and (b) count the worlds within these equivalence classes, i.e., worlds with the same conditional structure (that satisfy additional constraints, depending on the query). In the following, we address both tasks and solve them with typed model counting.

Typed Model Counting (TMC)

Typed model counting (TMC) extends model counting of propositional formulas by distinguishing between different *types* of models. These types are represented by elements of an algebraic structure the formulas are equipped with. Therewith, TMC allows for a more fine-grained evaluation of the formula. For instance, TMC can be used to preserve information about which parts of a formula are satisfied by a certain type of model. As a formal basis for TMC we define the *structured language* $\mathcal{L}^{\mathcal{S}}$ that is built upon the propositional language $\mathcal{L}$ and a commutative monoid $\mathcal{S}$. When applying TMC to ME-reasoning we will instantiate $\mathcal{S}$ with $\mathcal{G}(\mathcal{KB})$, i.e., with conditional structures.

Structured Language $\mathcal{L}^{\mathcal{S}}$

The structured language $\mathcal{L}^{\mathcal{S}}$ is obtained by incorporating an algebraic structure $(\mathcal{S}, \otimes)$ into the propositional language $\mathcal{L} = \mathcal{L}(\Sigma)$. More precisely, we require $(\mathcal{S}, \otimes)$ to be a commutative monoid and thus $\otimes$ to be associative and commutative. We denote the identity element with $\mathbf{1}_{\mathcal{S}}$ and usually write $\mathcal{S}$ instead of $(\mathcal{S}, \otimes)$. In order to combine elements from $\mathcal{S}$ with (propositional) formulas, we further introduce the outer operation $\circ : \mathcal{S} \times \mathcal{L}^{\mathcal{S}} \to \mathcal{L}^{\mathcal{S}}$.

Definition 1. *The* structured language $\mathcal{L}^{\mathcal{S}} = \mathcal{L}^{\mathcal{S}}(\mathcal{L}, \mathcal{S}, \circ)$ *is the smallest set such that*

1. *if* $A \in \mathcal{L}$, *then* $A \in \mathcal{L}^{\mathcal{S}}$,
2. *if* $\mathbf{s} \in \mathcal{S}$ *and* $B \in \mathcal{L}^{\mathcal{S}}$, *then* $\mathbf{s} \circ B \in \mathcal{L}^{\mathcal{S}}$,
3. *if* $B, C \in \mathcal{L}^{\mathcal{S}}$, *then* $B \wedge C \in \mathcal{L}^{\mathcal{S}}$ *and* $B \vee C \in \mathcal{L}^{\mathcal{S}}$.

Hence, $\mathcal{L}^{\mathcal{S}}$ consists of all formulas in $\mathcal{L}$, and additionally, elements from $\mathcal{S}$ may be concatenated to the left of any part of a formula as long as they are not in the scope of negations.

Example 2. *Let* $\mathsf{a}, \mathsf{b} \in \Sigma$ *and* $\mathbf{s_1}, \mathbf{s_2} \in \mathcal{S}$. *Then,* $\neg \mathbf{s_1} \circ \mathsf{a} \notin \mathcal{L}^{\mathcal{S}}$ *but* $A = \mathbf{s_1} \circ (\overline{\mathsf{a}} \vee \mathbf{s_2} \circ \mathsf{b}) \in \mathcal{L}^{\mathcal{S}}$.

In order to interpret *(structured) formulas* in $\mathcal{L}^{\mathcal{S}}$, we extend $\mathcal{S}$ by a distinguished element $\mathbf{0}_{\mathcal{S}} \notin \mathcal{S}$ which satisfies $\mathbf{0}_{\mathcal{S}} \otimes \mathbf{s} = \mathbf{s} \otimes \mathbf{0}_{\mathcal{S}} = \mathbf{0}_{\mathcal{S}}$ for all $\mathbf{s} \in \mathcal{S}$. The element $\mathbf{0}_{\mathcal{S}}$ is used to indicate that a formula in $\mathcal{L}^{\mathcal{S}}$ is not satisfied by a *structured interpretation* while elements in $\mathcal{S}$ are the model types of the considered formula.

Definition 2. *A (structured) interpretation* $\mathcal{I}^{\mathcal{S}}$ *on* Σ *is a mapping* $\Sigma \to \{\mathbf{1}_{\mathcal{S}}, \mathbf{0}_{\mathcal{S}}\}$. *We extend structured interpretations to deal with arbitrary formulas in* $\mathcal{L}^{\mathcal{S}}$ *and therefore to mappings* $\mathcal{I}^{\mathcal{S}} : \mathcal{L}^{\mathcal{S}} \to \mathcal{S} \cup \{\mathbf{0}_{\mathcal{S}}\}$ *by inductively defining*

1. *if* $A \in \mathcal{L}$, *then* $\begin{cases} \mathcal{I}^{\mathcal{S}}(\neg A) = \mathbf{1}_{\mathcal{S}} & \text{iff } \mathcal{I}^{\mathcal{S}}(A) = \mathbf{0}_{\mathcal{S}} \\ \mathcal{I}^{\mathcal{S}}(\neg A) = \mathbf{0}_{\mathcal{S}} & \text{iff } \mathcal{I}^{\mathcal{S}}(A) = \mathbf{1}_{\mathcal{S}} \end{cases}$
   (other cases do not occur as A *is purely propositional),*
2. *if* $A, B \in \mathcal{L}^{\mathcal{S}}$, *then* $\mathcal{I}^{\mathcal{S}}(A \wedge B) = \mathcal{I}^{\mathcal{S}}(A) \otimes \mathcal{I}^{\mathcal{S}}(B)$ *and*
   $$\mathcal{I}^{\mathcal{S}}(A \vee B) = \begin{cases} \mathcal{I}^{\mathcal{S}}(A) & \text{iff } \mathcal{I}^{\mathcal{S}}(B) = \mathbf{0}_{\mathcal{S}} \\ \mathcal{I}^{\mathcal{S}}(B) & \text{iff } \mathcal{I}^{\mathcal{S}}(A) = \mathbf{0}_{\mathcal{S}} \\ \mathcal{I}^{\mathcal{S}}(A) \otimes \mathcal{I}^{\mathcal{S}}(B) & \text{otherwise} \end{cases},$$
3. *if* $\mathbf{s} \in \mathcal{S}$ *and* $A \in \mathcal{L}^{\mathcal{S}}$, *then* $\mathcal{I}^{\mathcal{S}}(\mathbf{s} \circ A) = \mathbf{s} \otimes \mathcal{I}^{\mathcal{S}}(A)$.

Example 3. *We consider* $A = \mathbf{s_1} \circ (\overline{\mathsf{a}} \vee \mathbf{s_2} \circ \mathsf{b})$ *from Ex. 2 and the structured interpretation* $\mathcal{I}^{\mathcal{S}}$ *with* $\mathcal{I}^{\mathcal{S}}(\mathsf{a}) = \mathcal{I}^{\mathcal{S}}(\mathsf{b}) = \mathbf{1}_{\mathcal{S}}$. *Then,* $\mathcal{I}^{\mathcal{S}}(A) = \mathbf{s_1} \otimes \mathcal{I}^{\mathcal{S}}(\overline{\mathsf{a}} \vee \mathbf{s_2} \circ \mathsf{b}) = \mathbf{s_1} \otimes \mathcal{I}^{\mathcal{S}}(\mathbf{s_2} \circ \mathsf{b})$ *since* $\mathcal{I}^{\mathcal{S}}(\overline{\mathsf{a}}) = \mathbf{0}_{\mathcal{S}}$, *and thus* $\mathcal{I}^{\mathcal{S}}(A) = \mathbf{s_1} \otimes \mathbf{s_2} \otimes \mathcal{I}^{\mathcal{S}}(\mathsf{b}) = \mathbf{s_1} \otimes \mathbf{s_2}$.

There is a one-to-one correspondence between propositional and structured interpretations that can be observed by associating $\mathbf{1}_{\mathcal{S}}$ resp. $\mathbf{0}_{\mathcal{S}}$ with the truth values 0 resp. 1 assigned by propositional interpretations. Except for these replacements, propositional and structured interpretations interpret formulas in $\mathcal{L}$ in the same way. In particular, we can talk about the *structured interpretation induced by the world* ω, written $\mathcal{I}^{\mathcal{S}}_{\omega}$. For arbitrary formulas in $\mathcal{L}^{\mathcal{S}}$ we have the following definition.

Definition 3. *An* interpretation $\mathcal{I}^{\mathcal{S}}$ *is called a* typed model of $A \in \mathcal{L}^{\mathcal{S}}$ *iff* $\mathcal{I}^{\mathcal{S}}(A) \neq \mathbf{0}_{\mathcal{S}}$, *i.e., iff* $\mathcal{I}^{\mathcal{S}}(A) \in \mathcal{S}$. *It is called a* model of type $\mathbf{s} \in \mathcal{S}$ *iff* $\mathcal{I}^{\mathcal{S}}(A) = \mathbf{s}$. *The formula* $A \in \mathcal{L}^{\mathcal{S}}$ *is called* satisfiable *iff it has at least one typed model.*

Consequently, TMC is the task of finding all types of models for a given formula $A \in \mathcal{L}^{\mathcal{S}}$, as well as the frequencies every type can be observed with. To simplify TMC on $A \in \mathcal{L}^{\mathcal{S}}$, we $\mathcal{S}$*-equivalently* rewrite A until it is in a form that is especially suited for the TMC task. Two formulas $A, B \in \mathcal{L}^{\mathcal{S}}$ are called $\mathcal{S}$-equivalent, written $A \equiv_{\mathcal{S}} B$, iff $\mathcal{I}^{\mathcal{S}}(A) = \mathcal{I}^{\mathcal{S}}(B)$ for every interpretation $\mathcal{I}^{\mathcal{S}}$. In other words, every model of type $\mathbf{s}$ of A is a model of type $\mathbf{s}$ of B, and vice versa. We call A and B *mutually exclusive* iff $\mathcal{I}^{\mathcal{S}}(A) \otimes \mathcal{I}^{\mathcal{S}}(B) = \mathbf{0}_{\mathcal{S}}$ for every interpretation $\mathcal{I}^{\mathcal{S}}$, i.e., iff A and B cannot be satisfied at the same time.

Example 4. *The formula* $A = \mathbf{s_1} \circ (\overline{\mathsf{a}} \vee \mathbf{s_2} \circ \mathsf{b})$ *from Ex. 2 is* $\mathcal{S}$*-equivalent to* $(\mathbf{s_1} \circ \overline{\mathsf{a}}\,\overline{\mathsf{b}} \vee (\mathbf{s_1} \otimes \mathbf{s_2}) \circ \mathsf{b})$, *and it has three typed models, two of type* $\mathbf{s_1} \otimes \mathbf{s_2}$ *(when* $\mathcal{I}^{\mathcal{S}}(b) = \mathbf{1}_{\mathcal{S}}$*) and one of type* $\mathbf{s_1}$ *(when* $\mathcal{I}^{\mathcal{S}}(a) = \mathcal{I}^{\mathcal{S}}(b) = \mathbf{0}_{\mathcal{S}}$*).*

Formulas in $\mathcal{L}$ can be $\mathcal{S}$-equivalently transformed following the common equivalent transformations in $\mathcal{L}$. Some of them, namely commutativity, associativity, and neutrality ($A \wedge \top \equiv_{\mathcal{S}} A$ and $A \vee \bot \equiv_{\mathcal{S}} A$), do also hold for arbitrary formulas in $\mathcal{L}^{\mathcal{S}}$. However, idempotence, distributivity, and absorption only hold in certain cases due to the fact that elements in $\mathcal{S}$ are not necessarily idempotent (usually, $\mathcal{I}^{\mathcal{S}}(A \wedge A) = \mathcal{I}^{\mathcal{S}}(A) \otimes \mathcal{I}^{\mathcal{S}}(A) \neq \mathcal{I}^{\mathcal{S}}(A)$). De Morgan's laws and involution ($\neg(\neg A) \equiv_{\mathcal{S}} A$) are not defined for arbitrary formulas in $\mathcal{L}^{\mathcal{S}}$ due to the negations. We define a weakened version of distributivity that we make use of later on.

Proposition 1. *Let* $A, B, C \in \mathcal{L}^{\mathcal{S}}$. *If* A *and* B *are mutually exclusive, or if* $\mathcal{I}^{\mathcal{S}}(C)$ *is idempotent (e.g., if* $C \in \mathcal{L}$*), then*
$$(A \vee B) \wedge C \equiv_{\mathcal{S}} (A \wedge C) \vee (B \wedge C).$$

We refer to the property described in Prop. 1 as the rule of *weak distributivity*.

Now we introduce the *smooth deterministic decomposable structured negation normal form* (sd-DNNF$^{\mathcal{S}}$) which is the abovementioned form that is especially suited for TMC. The definition of sd-DNNF$^{\mathcal{S}}$ is very much the same as sd-DNNF for ordinary propositional formulas introduced in (Darwiche 2001), except that we consider *structured* formulas here, and hence we need to consider our adapted form of mutual exclusiveness. Formulas in sd-DNNF$^{\mathcal{S}}$ directly inherit their good properties for TMC from the fact that formulas in sd-DNNF are convenient for model counting.

Definition 4. *A formula* $A \in \mathcal{L}^{\mathcal{S}}$ *is in* structured negation normal form, *written* NNF$^{\mathcal{S}}$, *iff negations in* A *appear only directly in front of atoms. A formula* A *in* NNF$^{\mathcal{S}}$ *is said to be*

1. decomposable *iff for every conjunction* $\bigwedge_{i=1}^{m} A_i$ *in* A*, the sets* $\Sigma(A_1), \ldots, \Sigma(A_m)$ *are pairwise disjoint.*
2. deterministic *iff for every disjunction* $\bigvee_{i=1}^{m} A_i$ *in* A*, every two disjuncts* $A_i, A_j,\ i \neq j$*, are mutually exclusive, i.e., iff* $\mathcal{I}^{\mathcal{S}}(A_i) \otimes \mathcal{I}^{\mathcal{S}}(A_j) = \mathbf{0}_{\mathcal{S}}$ *for every interpretation* $\mathcal{I}^{\mathcal{S}}$*.*
3. smooth *iff for every disjunction* $\bigvee_{i=1}^{m} A_i$ *in* A*, it holds that* $\Sigma(A_i) = \Sigma(A_j)$ *for every* $i, j \in \{1, \ldots, m\}$*.*

The formula A *is in* $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$ *iff* A *is in* $\mathsf{NNF}^{\mathcal{S}}$ *and is smooth, deterministic, and decomposable.*

Example 5. *The formula* $A = \mathbf{s_1} \circ (\overline{\mathsf{a}} \vee \mathbf{s_2} \cdot \mathsf{b})$ *from Ex. 2 is not in* $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$ *(the disjunction is neither deterministic nor smooth), whereas* $B = \mathbf{s_1} \circ (\overline{\mathsf{a}}\,\overline{\mathsf{b}} \vee \mathbf{s_2} \circ ((\mathsf{a} \vee \overline{\mathsf{a}})\,\mathsf{b}))$ *is in* $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$*. The formulas* A *and* B *are* $\mathcal{S}$*-equivalent.*

The following proposition proves that every satisfiable formula in $\mathcal{L}^{\mathcal{S}}$ is $\mathcal{S}$-equivalent to a formula in $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$. The latter is not unique, as sd-DNNF's are not unique either.

Proposition 2. *Let* $A \in \mathcal{L}^{\mathcal{S}}$*. Then,* A *is not satisfiable or* A *is* $\mathcal{S}$*-equivalent to* $B(A) = \bigvee_{\omega \in \Omega \colon \mathcal{I}^{\mathcal{S}}_{\omega}(A) \neq \mathbf{0}_{\mathcal{S}}} \mathcal{I}^{\mathcal{S}}_{\omega}(A) \circ \omega$ *which is a structured formula in* $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$*.*

Proof. Let A be satisfiable, i.e., at least one typed model $\mathcal{I}^{\mathcal{S}}_{\omega'}$ of A exists and the disjunction in $B(A)$ is not vacuous. Further, let $\omega, \omega' \in \Omega$. Then, $\mathcal{I}^{\mathcal{S}}_{\omega'}(\omega) = \mathbf{0}_{\mathcal{S}}$ iff $\omega' \neq \omega$ and

$$\mathcal{I}^{\mathcal{S}}_{\omega'}\left(\bigvee\nolimits_{\omega \in \Omega \colon \mathcal{I}^{\mathcal{S}}_{\omega}(A) \neq \mathbf{0}_{\mathcal{S}}} \mathcal{I}^{\mathcal{S}}_{\omega}(A) \circ \omega\right)$$
$$= \mathcal{I}^{\mathcal{S}}_{\omega'}(\mathcal{I}^{\mathcal{S}}_{\omega'}(A) \circ \omega') = \mathcal{I}^{\mathcal{S}}_{\omega'}(A) \otimes \mathcal{I}^{\mathcal{S}}_{\omega'}(\omega') = \mathcal{I}^{\mathcal{S}}_{\omega'}(A)$$

which proves $A \equiv_{\mathcal{S}} B(A)$. In addition, worlds are decomposable and contain all atoms from Σ. Thus, disjunctions over worlds are smooth, deterministic, and decomposable. Therefore, $B(A)$ is in $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$. □

Given a formula $A \in \mathcal{L}^{\mathcal{S}}$, Prop. 2 implies a constructive method for deriving a formula in $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$ which is $\mathcal{S}$-equivalent to A. The drawback is that one has to know all typed models of A, and hence, this method undermines our intention behind considering formulas in $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$ which is unveiling the model types. When we apply TMC to ME-reasoning, we state more efficient methods for deriving formulas in $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$ in this concrete context. But first, we discuss how TMC based on formulas in $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$ works.

Typed Model Counting on Formulas in sd-DNNF$^{\mathcal{S}}$

In order to count the typed models of formulas in $\mathcal{L}^{\mathcal{S}}$, we have to enrich $(\mathcal{S} \cup \{\mathbf{0}_{\mathcal{S}}\}, \otimes)$ with a second associative and commutative operation $\oplus$ (addition). In detail, we build the closure $\mathcal{S}^{\oplus}$ of $\mathcal{S} \cup \{\mathbf{0}_{\mathcal{S}}\}$ under application of $\oplus$ while requiring that $\otimes$ and $\oplus$ satisfy the law of distributivity and $\mathbf{0}_{\mathcal{S}}$ serves as the identity element with respect to $\oplus$. Eventually, we obtain the commutative semiring $(\mathcal{S}^{\oplus}, \oplus, \otimes)$. We abbreviate $\bigoplus_{i=1}^{n} \mathbf{s}$ with $n\,\mathbf{s}$ where $\mathbf{s} \in \mathcal{S}^{\oplus}$ and $n \in \mathbb{N}_0$.

Once a formula is in $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$, TMC is an easy task and proceeds similar to model counting based on propositional formulas in sd-DNNF. For ordinary model counting, a formula $A \in \mathcal{L}$ which is in sd-DNNF is transformed into an arithmetic statement by substituting every literal in A by 1, every $\wedge$ by $\cdot$, and every $\vee$ by $+$. For TMC, formulas are transformed into algebraic statements instead.

Definition 5. *Let* $A \in \mathcal{L}^{\mathcal{S}}$ *be a formula in* $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$*. The element* $\mathbf{s}(A) \in \mathcal{S}^{\oplus}$ *obtained by substituting every literal in* A *by* $\mathbf{1}_{\mathcal{S}}$*, every occurence of* $\wedge$ *and of* $\circ$ *by* $\otimes$*, and every occurence of* $\vee$ *by* $\oplus$*, is called the* structure element of A*.*

Structure elements are *the* essentials of TMC.

Typed Model Counting. Let $A \in \mathcal{L}^{\mathcal{S}}$ be in $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$, and let $\mathsf{M}_{\mathbf{s}}(A)$ be the number of models of type $\mathbf{s}$ of A. Counting the typed models of A means deriving $\mathbf{s}(A)$, as

$$\mathbf{s}(A) = \bigoplus_{\omega \in \Omega} \mathcal{I}^{\mathcal{S}}_{\omega}(A) = \bigoplus_{\mathbf{s} \in \mathcal{S}} \mathsf{M}_{\mathbf{s}}(A)\ \mathbf{s}.$$

TMC holds for the same reasons for which model counting holds. Basically, determinism guarantees that no model is counted twice, decomposability of deterministic formulas ensures that only interpretations are considered (as multiplication of counts arising from inconsistent sets of literals cannot happen), and the smoothness guarantees that every model is taken into account. The only difference to ordinary model counting is the splitting of the models into different types which does not affect the way of counting itself but preserves the information encoded in the algebraic type.

Example 6. *Recall the formula* $B = \mathbf{s_1} \circ (\overline{\mathsf{a}}\,\overline{\mathsf{b}} \vee \mathbf{s_2} \circ ((\mathsf{a} \vee \overline{\mathsf{a}})\mathsf{b}))$ *from Ex. 5 which is in* $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$*. The counts of the typed models of* B *are encoded in*

$$\mathbf{s}(B) = \mathbf{s_1} \otimes ((\mathbf{1}_{\mathcal{S}} \otimes \mathbf{1}_{\mathcal{S}}) \oplus (\mathbf{s_2} \otimes (\mathbf{1}_{\mathcal{S}} \oplus \mathbf{1}_{\mathcal{S}}) \otimes \mathbf{1}_{\mathcal{S}}))$$
$$= \mathbf{s_1} \otimes (\mathbf{1}_{\mathcal{S}} \oplus 2\,\mathbf{s_2}) = \mathbf{s_1} \oplus 2\,(\mathbf{s_1} \otimes \mathbf{s_2}).$$

From $\mathbf{s}(B)$*, we can not only read the number of models of* B *(i.e., 3) but also how many models of a specific type* B *has.*

Conditioned Typed Model Counting. It is possible to focus on certain models when counting the typed models of a formula $A \in \mathcal{L}^{\mathcal{S}}$ which is in $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$. This is of interest when only those typed models of A are sought that are also models of an additional formula. We face this task when drawing ME-inferences later on. In order to focus on models, we extend the concept of structure elements. For this, let $\mathcal{C}$ be a consistent set of literals. We define $\mathbf{s}_{\mathcal{C}}(A) \in \mathcal{S}^{\oplus}$ in the same way as $\mathbf{s}(A)$ except that the literals in $\{\mathsf{c} \in \mathsf{Lit} \mid \overline{\mathsf{c}} \in \mathcal{C}\}$ are replaced by $\mathbf{0}_{\mathcal{S}}$ instead of $\mathbf{1}_{\mathcal{S}}$. Then, *counting the typed models of* A *conditioned on* $\mathcal{C}$, i.e., calculating $\mathbf{s}_{\mathcal{C}}(A)$, yields

$$\mathbf{s}_{\mathcal{C}}(A) = \mathbf{s}\left(A \wedge \bigwedge_{\mathsf{c} \in \mathcal{C}} \mathsf{c}\right).$$

In plain words, counting the typed models of the formula $A \wedge \bigwedge_{\mathsf{c} \in \mathcal{C}} \mathsf{c}$ can be performed by counting the typed models of A with the modification that now not all literals in A are substituted by $\mathbf{1}_{\mathcal{S}}$, but those literals in A that contradict $\bigwedge_{\mathsf{c} \in \mathcal{C}} \mathsf{c}$ are substituted by $\mathbf{0}_{\mathcal{S}}$. The benefit of this method is that one does not have to transform $A \wedge \bigwedge_{\mathsf{c} \in \mathcal{C}} \mathsf{c}$ into $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$ when A already is in $\mathsf{sd\text{-}DNNF}^{\mathcal{S}}$.

In particular, conditioned typed model counting is suitable for interpreting a formula. Let $\omega \in \Omega$ be an arbitrary world and $\mathcal{C} = \{\mathsf{c} \in \mathsf{Lit} \mid \omega \models \mathsf{c}\}$. Then, $\mathbf{s}_{\mathcal{C}}(A) = \mathcal{I}^{\mathcal{S}}_{\omega}(A)$.

Next, we apply TMC to simplify ME-reasoning.

Application to ME-Reasoning

We show that typed model counting (TMC) is convenient for solving both open tasks for drawing ME-inferences, namely finding the equivalence classes $[\omega]_{\sigma_{\mathcal{KB}}}$ as well as calculating their cardinalities. Therewith, we are able to draw ME-inferences by exploiting Eq. (2). In order to apply TMC to ME-reasoning, we have to convert knowledge bases into structured formulas. More precisely, we consider the knowledge base $\mathcal{KB}$ and extend $\mathcal{L}$ by conditional structures. Thus, we consider the structured language $\mathcal{L}^{\mathcal{G}} = \mathcal{L}^{\mathcal{G}}(\mathcal{L}, \mathcal{G}, \circ)$ with $\mathcal{G} = \mathcal{G}(\mathcal{KB})$ and define the formula $\phi(\mathcal{KB}) \in \mathcal{L}^{\mathcal{G}}$ by

$$\phi(\mathcal{KB}) = \bigwedge_{i=1}^{n} \left[(\mathbf{a_i^+} \circ A_i B_i) \vee (\mathbf{a_i^-} \circ A_i \overline{B_i}) \vee \overline{A_i} \right]. \quad (3)$$

Here, $\mathbf{a_i^+}$ and $\mathbf{a_i^-}$ for $i = 1, \ldots, n$ are the generators of $\mathcal{G}$. The next proposition states that conditional structures serve as the semantics of these formulas.

Proposition 3. *Let $\mathcal{KB}$ be a knowledge base and let $\omega \in \Omega$. Then, $\mathcal{I}_\omega^{\mathcal{G}}(\phi(\mathcal{KB})) = \sigma_{\mathcal{KB}}(\omega)$.*

Proof. Since the outer disjunctions in $(A_i B_i \vee A_i \overline{B_i} \vee \overline{A_i})$ are deterministic for $i = 1, \ldots, n$, it follows that

$$\mathcal{I}_\omega^{\mathcal{G}}(\phi(\mathcal{KB})) = \mathcal{I}_\omega^{\mathcal{G}}\left(\bigwedge_{i=1}^{n} ((\mathbf{a_i^+} \circ A_i B_i) \vee (\mathbf{a_i^-} \circ A_i \overline{B_i}) \vee \overline{A_i}) \right)$$

$$= \prod_{i=1}^{n} \begin{cases} \mathbf{a_i^+} & \text{iff } \omega \models A_i B_i \\ \mathbf{a_i^-} & \text{iff } \omega \models A_i \overline{B_i} \\ \mathbf{1}_{\mathcal{G}} & \text{iff } \omega \models \overline{A_i} \end{cases} = \prod_{i=1}^{n} \sigma_i(\omega) = \sigma_{\mathcal{KB}}(\omega). \qquad \square$$

As a consequence of Prop. 3, calculating all equivalence classes $[\omega]_{\sigma_{\mathcal{KB}}}$ and their cardinalities reduces to counting the typed models of $\phi(\mathcal{KB})$, as

$$s(\phi(\mathcal{KB})) = \bigoplus_{\omega \in \Omega} \sigma_{\mathcal{KB}}(\omega) = \bigoplus_{[\omega]_{\sigma_{\mathcal{KB}}} \in \Omega / \sigma_{\mathcal{KB}}} |[\omega]_{\sigma_{\mathcal{KB}}}| \, \sigma_{\mathcal{KB}}(\omega).$$

Hence, it is essential to find a $\mathcal{G}$-equivalent formula in sd-DNNF$^{\mathcal{G}}$ for $\phi(\mathcal{KB})$. We show how this can be achieved by applying the rule of weak distributivity (cf. Prop. 1) while avoiding the intractable procedure from Prop. 2.

As the outer disjunctions in $\phi(\mathcal{KB})$ are deterministic, we may apply weak distributivity to $\phi(\mathcal{KB})$ in Eq. (3) and get

$$\phi(\mathcal{KB}) \equiv_{\mathcal{G}} \bigvee_{\substack{\lambda = (\lambda_1, \ldots, \lambda_n) \\ \lambda_i \in \{\mathbf{a_i^+}, \mathbf{a_i^-}, \mathbf{1}_{\mathcal{G}}\}}} \left(\bigotimes_{i=1}^{n} \lambda_i \right) \circ \underbrace{\bigwedge_{i=1}^{n} \begin{cases} A_i B_i & \text{iff } \lambda_i = \mathbf{a_i^+} \\ A_i \overline{B_i} & \text{iff } \lambda_i = \mathbf{a_i^-} \\ \overline{A_i} & \text{iff } \lambda_i = \mathbf{1}_{\mathcal{G}} \end{cases}}_{=: A(\lambda)},$$

The outer disjunction on the right-hand side is deterministic and $A(\lambda)$ is in $\mathcal{L}$ for $\lambda \in \Lambda$. Now, we expand $A(\lambda)$ by those atoms $\mathsf{a}_{i(\lambda)} \in \Sigma$ that do not appear in $A(\lambda)$ but in Σ by adding the conjunction $\bigwedge_{i(\lambda)} (\mathsf{a}_{i(\lambda)} \vee \overline{\mathsf{a}}_{i(\lambda)})$ to $A(\lambda)$. Afterwards, the outer disjunction is also smooth, and we just have to transform $A(\lambda) \wedge \bigwedge_{i(\lambda)} (\mathsf{a}_{i(\lambda)} \vee \overline{\mathsf{a}}_{i(\lambda)}) \in \mathcal{L}$ into (standard) sd-DNNF for each $\lambda \in \Lambda$ separately. Weak distributivity can be applied iteratively to pairs of disjuncts, and, on every step of the iteration, the obtained formulas can be tested for satisfiability. Unsatisfiable disjuncts may be removed in order to avoid transforming them into sd-DNNF unnecessarily. This strategy coincides with the algorithm CONDSTRUCTOR presented in (Wilhelm, Kern-Isberner, and Ecke 2016).

Example 7. *We continue Ex. 1 from the introduction and consider $\mathcal{KB}_{fly} = \{r_1, r_2, r_3\}$ with $r_1 = (\mathsf{f} \vee \overline{\mathsf{w}} | \top)[0.95]$, $r_2 = (\mathsf{w}|\mathsf{f})[0.99]$, and $r_3 = (\overline{\mathsf{b}}|\overline{\mathsf{f}})[0.8]$. We have*

$$\begin{aligned} \phi(\mathcal{KB}_{fly}) = \quad & \left[(\mathbf{a_1^+} \circ (\mathsf{f} \vee \overline{\mathsf{w}})) \vee (\mathbf{a_1^-} \circ \overline{\mathsf{f}}\,\mathsf{w}) \vee \bot \right] \\ \wedge & \left[(\mathbf{a_2^+} \circ \mathsf{f}\,\mathsf{w}) \quad \vee (\mathbf{a_2^-} \circ \mathsf{f}\,\overline{\mathsf{w}}) \vee \overline{\mathsf{f}} \right] \\ \wedge & \left[(\mathbf{a_3^+} \circ \overline{\mathsf{f}}\,\overline{\mathsf{b}}) \quad \vee (\mathbf{a_3^-} \circ \overline{\mathsf{f}}\,\mathsf{b}) \ \vee \mathsf{f} \right] \end{aligned}$$

and iteratively apply weak distributivity as well as smoothing. We get

$$\begin{aligned} \phi(\mathcal{KB}_{fly}) \equiv_{\mathcal{G}} & \left[(\mathbf{a_1^+}\,\mathbf{a_2^+} \circ \mathsf{f}\,\mathsf{w}) \vee (\mathbf{a_1^+}\,\mathbf{a_2^-} \circ \mathsf{f}\,\overline{\mathsf{w}}) \vee (\mathbf{a_1^+} \circ \overline{\mathsf{f}}\,\overline{\mathsf{w}}) \right. \\ & \left. \vee (\mathbf{a_1^-} \circ \overline{\mathsf{f}}\,\mathsf{w}) \right] \wedge \left[(\mathbf{a_3^+} \circ \overline{\mathsf{f}}\,\overline{\mathsf{b}}) \vee (\mathbf{a_3^-} \circ \overline{\mathsf{f}}\,\mathsf{b}) \vee \mathsf{f} \right] \\ \equiv_{\mathcal{G}} & \left[\mathbf{a_1^+}\,\mathbf{a_2^+} \circ \mathsf{f}\,\mathsf{w}\,(\mathsf{b} \vee \overline{\mathsf{b}}) \right] \vee \left[\mathbf{a_1^+}\,\mathbf{a_2^-} \circ \mathsf{f}\,\overline{\mathsf{w}}\,(\mathsf{b} \vee \overline{\mathsf{b}}) \right] \\ & \vee \left[\mathbf{a_1^+}\,\mathbf{a_3^+} \circ \overline{\mathsf{f}}\,\overline{\mathsf{w}}\,\overline{\mathsf{b}} \right] \vee \left[\mathbf{a_1^+}\,\mathbf{a_3^-} \circ \overline{\mathsf{f}}\,\overline{\mathsf{w}}\,\mathsf{b} \right] \\ & \vee \left[\mathbf{a_1^-}\,\mathbf{a_3^+} \circ \overline{\mathsf{f}}\,\mathsf{w}\,\overline{\mathsf{b}} \right] \vee \left[\mathbf{a_1^-}\,\mathbf{a_3^-} \circ \overline{\mathsf{f}}\,\mathsf{w}\,\mathsf{b} \right]. \end{aligned} \quad (4)$$

Counting typed models eventually yields

$$\begin{aligned} \mathbf{s}(\phi(\mathcal{KB}_{fly})) = \quad & \left[2\,\mathbf{a_1^+}\,\mathbf{a_2^+} \right] \oplus \left[2\,\mathbf{a_1^+}\,\mathbf{a_2^-} \right] \oplus \left[\mathbf{a_1^+}\,\mathbf{a_3^+} \right] \\ & \oplus \left[\mathbf{a_1^+}\,\mathbf{a_3^-} \right] \ \oplus \left[\mathbf{a_1^-}\,\mathbf{a_3^+} \right] \ \oplus \left[\mathbf{a_1^-}\,\mathbf{a_3^-} \right]. \end{aligned}$$

Thus, e.g., there are two models of type $\mathbf{a_1^+}\mathbf{a_2^+}$ of $\phi(\mathcal{KB}_{fly})$ that correspond to two worlds in which the first two conditionals of $\mathcal{KB}$ are verified (and the third is not applicable).

When $\phi(\mathcal{KB})$ consists of two decomposable conjuncts, one can apply weak distributivity to both conjuncts separately. This is the case when $\mathcal{KB}$ can be divided into two sets of conditionals that do not share any atoms. However, knowledge bases typically do not show this property. With the aid of conditioning (cf. (Darwiche 1999) for a more general definition) this property can be produced artificially.

Definition 6. *Let $A \in \mathcal{L}^{\mathcal{S}}$ and $\mathsf{a} \in \Sigma$. Then $(A|\dot{\mathsf{a}}) \in \mathcal{L}^{\mathcal{S}}$ is the formula that is obtained by replacing every occurrence of a in A by $\top$ iff $\dot{\mathsf{a}} = \mathsf{a}$ and by $\bot$ iff $\dot{\mathsf{a}} = \overline{\mathsf{a}}$. We call $(A|\dot{\mathsf{a}})$ the* literal conditioning of A by $\dot{\mathsf{a}}$.

The next proposition makes use of literal conditioning and is closely related to the Shannon expansion (Shannon 1949).

Proposition 4. *Let $A \in \mathcal{L}^{\mathcal{S}}$ and $\mathsf{a} \in \Sigma$. Then,*

$$A \equiv_{\mathcal{S}} (\mathsf{a} \wedge (A|\mathsf{a})) \vee (\overline{\mathsf{a}} \wedge (A|\overline{\mathsf{a}})).$$

Note that the disjunction in $(\mathsf{a} \wedge (A|\mathsf{a})) \vee (\overline{\mathsf{a}} \wedge (A|\overline{\mathsf{a}}))$ is deterministic and smooth, and the conjunctions are decomposable. Thus, transforming A into sd-DNNF$^{\mathcal{S}}$ is reduced to the two smaller problems of transforming $(A|\mathsf{a})$ and $(A|\overline{\mathsf{a}})$ into sd-DNNF$^{\mathcal{S}}$. We demonstrate the concept of literal conditioning and show how TMC can help drawing ME-inferences by means of the following example.

$[\omega]_{\sigma_{\mathcal{KB}_{fly}}}$	$\lvert[\omega]_{\sigma_{\mathcal{KB}_{fly}}}\rvert$	$\sigma_{\mathcal{KB}_{fly}}(\omega)$	$\mathcal{P}_{\mathsf{ME}}(\mathcal{KB}_{fly})(\omega)$	$c([\omega]_{\sigma_{\mathcal{KB}_{fly}}}, \mathsf{b})$	$c([\omega]_{\sigma_{\mathcal{KB}_{fly}}}, \mathsf{bf})$
$\{\mathsf{b\,f\,w}, \overline{\mathsf{b}}\,\mathsf{f\,w}\}$	2	$\mathbf{a_1^+}\,\mathbf{a_2^+}$	0.25	1	1
$\{\mathsf{b\,f}\,\overline{\mathsf{w}}, \overline{\mathsf{b}}\,\mathsf{f}\,\overline{\mathsf{w}}\}$	2	$\mathbf{a_1^+}\,\mathbf{a_2^-}$	0.02	1	1
$\{\overline{\mathsf{b}}\,\overline{\mathsf{f}}\,\overline{\mathsf{w}}\}$	1	$\mathbf{a_1^+}\,\mathbf{a_3^+}$	0.33	0	0
$\{\mathsf{b}\,\overline{\mathsf{f}}\,\overline{\mathsf{w}}\}$	1	$\mathbf{a_1^+}\,\mathbf{a_3^-}$	0.08	1	0
$\{\overline{\mathsf{b}}\,\overline{\mathsf{f}}\,\mathsf{w}\}$	1	$\mathbf{a_1^-}\,\mathbf{a_3^+}$	0.04	0	0
$\{\mathsf{b}\,\overline{\mathsf{f}}\,\mathsf{w}\}$	1	$\mathbf{a_1^-}\,\mathbf{a_3^-}$	0.01	1	0

Table 1: Equivalence classes $[\omega]_{\sigma_{\mathcal{KB}_{fly}}} \in \Omega/[\omega]_{\sigma_{\mathcal{KB}_{fly}}}$ and their relevant properties for drawing ME-inferences.

Example 8. *We consider $\phi(\mathcal{KB}_{fly})$ from Ex. 7 and apply literal conditioning. More precisely, we condition by $\dot{\mathsf{f}}$ and perform obvious simplifications. We get*

$$\begin{aligned}\phi(\mathcal{KB}_{fly}) &\equiv_{\mathcal{G}} [\mathsf{f} \wedge (\phi(\mathcal{KB}_{fly})|\mathsf{f})] \vee \left[\overline{\mathsf{f}} \wedge (\phi(\mathcal{KB}_{fly})|\overline{\mathsf{f}})\right] \\ &\equiv_{\mathcal{G}} \left[\mathsf{f} \wedge (\mathbf{a_1^+} \circ \top) \wedge ((\mathbf{a_2^+} \circ \mathsf{w}) \vee (\mathbf{a_2^-} \circ \overline{\mathsf{w}}))\right] \vee \left[\overline{\mathsf{f}} \wedge \right. \\ &\qquad \left. ((\mathbf{a_1^+} \circ \overline{\mathsf{w}}) \vee (\mathbf{a_1^-} \circ \mathsf{w})) \wedge ((\mathbf{a_3^+} \circ \overline{\mathsf{b}}) \vee (\mathbf{a_3^-} \circ \mathsf{b}))\right].\end{aligned}$$

Afterwards, we establish the smoothness property:

$$\begin{aligned}&\phi(\mathcal{KB}_{fly}) \equiv_{\mathcal{G}} \\ &\quad \left[\mathsf{f} \wedge (\mathbf{a_1^+} \circ \top) \wedge ((\mathbf{a_2^+} \circ \mathsf{w}) \vee (\mathbf{a_2^-} \circ \overline{\mathsf{w}})) \wedge (\mathsf{b} \vee \overline{\mathsf{b}})\right] \\ &\quad \vee \left[\overline{\mathsf{f}} \wedge ((\mathbf{a_1^+} \circ \overline{\mathsf{w}}) \vee (\mathbf{a_1^-} \circ \mathsf{w})) \wedge ((\mathbf{a_3^+} \circ \overline{\mathsf{b}}) \vee (\mathbf{a_3^-} \circ \mathsf{b}))\right].\end{aligned}$$

The resulting formula is in $\mathsf{sd\text{-}DNNF}^{\mathcal{G}}$ *and can be used in the same way as Eq. (4) to determine the equivalence classes of worlds regarding $\mathcal{KB}_{fly}$ as well as their cardinalities. The equivalence classes, their cardinalities, and the appropriate* ME*-probabilities are shown in Tab. 1.*

Once the ME*-probabilities are given, we can draw inferences from $\mathcal{KB}$ via Eq. (2). For example, a possible query could be: With which probability do birds fly? In order to exploit Eq. (2) we have to count the number of worlds within each equivalence class that additionally satisfy* b *resp.* bf*, i.e., we have to calculate $c([\omega]_{\sigma_{\mathcal{KB}_{fly}}}, \mathsf{b})$ and $c([\omega]_{\sigma_{\mathcal{KB}_{fly}}}, \mathsf{bf})$ for every $[\omega]_{\sigma_{\mathcal{KB}_{fly}}}$. These counts can be derived by counting the typed models of $\phi(\mathcal{KB}_{fly})$ conditioned on $\{\mathsf{b}\}$ resp. $\{\mathsf{b}, \mathsf{f}\}$, which leads to*

$$\begin{aligned}\mathbf{s}_{\{\mathsf{b}\}}(\phi(\mathcal{KB}_{fly})) &= \mathbf{a_1^+}\mathbf{a_2^+} \oplus \mathbf{a_1^+}\mathbf{a_2^-} \oplus \mathbf{a_1^+}\mathbf{a_3^-} \oplus \mathbf{a_1^-}\mathbf{a_3^-}, \\ \mathbf{s}_{\{\mathsf{b},\mathsf{f}\}}(\phi(\mathcal{KB}_{fly})) &= \mathbf{a_1^+}\mathbf{a_2^+} \oplus \mathbf{a_1^+}\mathbf{a_2^-}.\end{aligned}$$

For example, $\mathbf{s}_{\{\mathsf{b}\}}(\phi(\mathcal{KB}_{fly}))$ states, among other things, that there is one model of type $\mathbf{a_1^+}\mathbf{a_2^+}$ of $\phi(\mathcal{KB}_{fly})$ that is also a model of b *(the same holds with respect to* bf*). See the last two columns of Tab. 1 for all counts. We finally derive $\mathcal{P}_{\mathsf{ME}}(\mathcal{KB}_{fly})(\mathsf{f}|\mathsf{b}) = \frac{0.27}{0.36} = 0.75$ from Eq. (2), i.e., birds do fly with probability $p = 0.75$ with respect to $\mathcal{KB}_{fly}$.*

Conclusion

We introduced the concept of typed model counting and pointed out its benefits by applying it to probabilistic conditional reasoning. Typed model counting extends model counting for propositional formulas as it allows for distinguishing between certain types of models. In order to do this, formulas are equipped with elements from a commutative monoid. As a consequence, typed models can reflect the structure of the formula to some extent. We used this capability in the context of reasoning under the principle of maximum entropy (ME-principle). Because ME-probabilities follow conditional structures, we incorporated the conditional structure induced by a knowledge base directly into the propositional language in which the knowledge is represented. Typed model counting then yielded the equivalence classes of possible worlds regarding the conditional structure, as well as their cardinalities. We exploited this in order to draw nonmonotonic inferences at maximum entropy.

In future work, we aim to adopt the notion of typed models to first-order sentences in order to draw lifted inferences at maximum entropy based on the ideas of first order model counting stated in (v. d. Broeck 2013). Besides structural insights, we legitimately expect significant performance improvements as in (Finthammer and Beierle 2012).

Acknowledgments. This research was supported by the German National Science Foundation (DFG) research unit FOR 1513 on Hybrid Reasoning for Intelligent Systems.

References

Darwiche, A. 1999. Compiling knowledge into decomposable negation normal form. In *Proceedings of the 16th IJCAI Conference*.

Darwiche, A. 2001. On the tractable counting of theory models and its application to belief revision and truth maintenance. *Journal of Applied Non-Classical Logics*.

Finthammer, M., and Beierle, C. 2012. Using equivalences of worlds for aggregation semantics of relational conditionals. In *Proceedings of the 35th KI Conference*. Springer.

Gärdenfors, P. 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press.

Kern-Isberner, G. 2001. *Conditionals in nonmonotonic reasoning and belief revision*. Springer.

Kimmig, A.; v. d. Broeck, G.; and de Raedt, L. 2012. Algebraic model counting. *Computing Research Repository*.

Paris, J. B. 1994. *The uncertain reasoner's companion – A mathematical perspective*. Cambridge University Press.

Shannon, C. E. 1949. The synthesis of two-terminal switching circuits. *Bell System Technical Journal*.

Shore, J. E., and Johnson, R. W. 1980. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory*.

v. d. Broeck, G. 2013. *Lifted Inference and Learning in Statistical Relational Models*. Ph.D. Dissertation, KU Leuven.

Wilhelm, M.; Kern-Isberner, G.; and Ecke, A. 2016. Propositional probabilistic reasoning at maximum entropy modulo theories. In *Proceedings of the 29th FLAIRS Conference*.