Proceedings of the Twenty-Ninth International
Florida Artificial Intelligence Research Society Conference

Propositional Probabilistic Reasoning at Maximum Entropy Modulo Theories

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

Andreas Ecke
Department of Computer Science
TU Dresden, Germany

Abstract

The principle of maximum entropy (MaxEnt principle) provides a valuable methodology for reasoning with probabilistic conditional knowledge bases realizing an idea of information economy in the sense of adding a minimal amount of assumed information. The conditional structure of such a knowledge base allows for classifying possible worlds regarding their influence on the MaxEnt distribution. In this paper, we present an algorithm that determines these equivalence classes and computes their cardinality by performing satisfiability tests of propositional formulas built upon the premises and conclusions of the conditionals. An example illustrates how the output of our algorithm can be used to simplify calculations when drawing nonmonotonic inferences under maximum entropy. For this, we use a characterization of the MaxEnt distribution in terms of conditional structure that completely abstracts from the propositional logic underlying the conditionals.

Introduction

Probability theory is one of the most powerful and popular frameworks for nonmonotonic reasoning, and the principle of maximum entropy (MaxEnt principle) constitutes a most appropriate form of common-sense probabilistic reasoning when the given knowledge is incomplete (Paris 1999; Shore and Johnson 1980; Jaynes 1983). However, the MaxEnt principle is often constituted as a black box methodology due to its non-transparency. In this paper, we systematically exploit the conditional structure of knowledge bases consisting of probabilistic conditionals of the form $(B|A)[x]$ with the meaning "if A holds, then B follows with probability x" in order to make MaxEnt calculations more understandable and simple. Formally, the MaxEnt distribution can be obtained by solving a nonlinear equation system depending on the evaluation of the stated conditionals within possible worlds (Kern-Isberner 2001). Here, we give a characterization of the MaxEnt distribution that abstracts from these possible worlds and does only depend on equivalence classes induced by the conditional structure. The main contribution of this paper is an algorithm that computes the equivalence classes (and their cardinalities) of possible worlds with respect to the conditional structure. Starting from a (qualitative) knowledge base, it iteratively splits the set of possible worlds into disjoint subsets depending on whether the possible worlds satisfy the premises and/or the conclusions of the conditionals in the knowledge base. In order to avoid unnecessary computations, consistency is tested at every step. The algorithm can be seen as a base algorithm that provides various means of modification and optimization. Although there are investigations on extracting the conditional structure from a given knowledge base (Beierle et al. 2015), to our knowledge, this is the first approach for systematically computing the aforementioned equivalence classes without having to look at every possible world.

The rest of the paper is organized as follows: First, we give some preliminary notes, especially on the concept of conditional structure and the MaxEnt principle. Then, we present our algorithm for calculating the equivalence classes of possible worlds. Afterwards, the output of the algorithm is used for answering an inference query by way of illustration. Finally, we conclude and point out future work.

Conditionals and the MaxEnt Principle

We consider a propositional language $\mathcal{L}_\mathcal{V}$ over a finite alphabet $\mathcal{V}$ that is equipped with the logical connectives $\wedge$ (*and*), $\vee$ (*or*), and $\neg$ (*negation*). Roman uppercase letters denote atoms or formulas in $\mathcal{L}_\mathcal{V}$. We write AB instead of $A \wedge B$ and $\overline{A}$ instead of $\neg A$ to shorten formulas. Let $\Omega_\mathcal{V}$ be the set of *possible worlds*, i.e., a complete set of interpretations of $\mathcal{L}_\mathcal{V}$. We identify each possible world $\omega \in \Omega_\mathcal{V}$ with the complete conjunction that has exactly ω as a model. If ω satisfies a formula A, we write $\omega \models A$ and call ω a *model* of A. A *conditional* is of the form $(B|A)$ where A, B are propositional formulas, and $(B|A)$ has the intention of a default option "if A, then typically B". Formally, conditionals lead to a three-valued logic as $(B|A)$ is *verified* in $\omega \in \Omega_\mathcal{V}$ iff $\omega \models AB$, it is *falsified* iff $\omega \models A\overline{B}$, and it is *not applicable* iff $\omega \models \overline{A}$. By assigning each conditional a probability, we obtain a *probabilistic conditional language*

$$(\mathcal{L}_\mathcal{V}|\mathcal{L}_\mathcal{V})^{prob} = \{(B|A)[x] \mid A, B \in \mathcal{L}_\mathcal{V}, x \in [0,1]\}.$$

Let $\mathcal{P}$ be a probability distribution over $\Omega_\mathcal{V}$ where possible worlds are understood as elementary events and atoms in $\mathcal{L}_\mathcal{V}$ as random variables. Then, every formula A can be assigned a probability via $\mathcal{P}(A) = \sum_{\omega \models A} \mathcal{P}(\omega)$, and satisfaction of a conditional is defined by

$$\mathcal{P} \models (B|A)[x] \text{ iff } \mathcal{P}(A) > 0 \text{ and } x = \frac{\mathcal{P}(AB)}{\mathcal{P}(A)}.$$

A finite set $\mathcal{KB} = \{(B_1|A_1)[x_1], \ldots, (B_n|A_n)[x_n]\}$ of probabilistic conditionals is called a *knowledge base*. $\mathcal{P}$ satisfies $\mathcal{KB}$ iff $\mathcal{P}$ satisfies every conditional in $\mathcal{KB}$. If a probability distribution exists that satisfies $\mathcal{KB}$, the former is called a *model* of $\mathcal{KB}$ and the latter *consistent*. For $\mathcal{KB}$, we define its *qualitative* counterpart $\mathcal{KB}^q = \{(B_1|A_1), \ldots, (B_n|A_n)\}$.

The *conditional structure* (Kern-Isberner 2001) of a possible world $\omega \in \Omega_{\mathcal{V}}$ with respect to $\mathcal{KB}^q$ is defined by $\sigma_{\mathcal{KB}^q}(\omega) = \prod_{i=1}^n \sigma_i^{\mathcal{KB}^q}(\omega)$ where

$$\sigma_i^{\mathcal{KB}^q}(\omega) = \begin{cases} \mathbf{a_i^+} & \text{iff } \omega \models A_i B_i \\ \mathbf{a_i^-} & \text{iff } \omega \models A_i \overline{B_i} \\ 1 & \text{iff } \omega \models \overline{A_i} \end{cases}, \quad i = 1, \ldots, n.$$

The symbols $\mathbf{a_i^+}$ (resp. $\mathbf{a_i^-}$) indicate wether the i-th conditional in $\mathcal{KB}^q$ is verified (resp. falsified) in ω. The conditional structure induces an equivalence relation $\omega \equiv_{\mathcal{KB}^q} \omega'$ on $\Omega_{\mathcal{V}}$ which holds iff $\sigma_{\mathcal{KB}^q}(\omega) = \sigma_{\mathcal{KB}^q}(\omega')$. We denote the equivalence class regarding ω with $[\omega] = [\omega]_{\equiv_{\mathcal{KB}^q}}$ and the set of all equivalence classes with $[\Omega_{\mathcal{V}}] = \Omega_{\mathcal{V}} / \equiv_{\mathcal{KB}^q}$. As there are three different ways a conditional can be evaluated within a possible world, there are at most 3^n different equivalence classes in $[\Omega_{\mathcal{V}}]$, where n is the number of conditionals in $\mathcal{KB}^q$. Note that this bound is not sharp, as Ex. 1 shows, and that $|[\Omega_{\mathcal{V}}]|$ is independent of the size of $\mathcal{V}$. With

$$v_i(\omega) = \begin{cases} 1 \text{ iff } \omega \models A_i B_i \\ 0 \text{ iff } \omega \not\models A_i B_i \end{cases}, \quad a_i(\omega) = \begin{cases} 1 \text{ iff } \omega \models A_i \\ 0 \text{ iff } \omega \not\models A_i \end{cases},$$

the conditional structure of a possible world ω is given by

$$\sigma_{\mathcal{KB}^q}(\omega) = \prod_{i=1}^n (\mathbf{a_i^+})^{v_i(\omega)} (\mathbf{a_i^-})^{a_i(\omega) - v_i(\omega)}. \quad (1)$$

Thus, it is completely determined by the functions v_i and a_i.

If $\mathcal{KB} = \{(B_1|A_1)[x_1], \ldots, (B_n|A_n)[x_n]\}$ is a consistent knowledge base, there may exist several models of $\mathcal{KB}$. The *principle of maximum entropy (MaxEnt principle)* selects a unique model among these which shows particularly good properties for drawing inferences from $\mathcal{KB}$ (Kern-Isberner 2001; Paris 1994). This *MaxEnt distribution* is defined by

$$\mathcal{P}_{\mathsf{ME}} = \mathcal{P}_{\mathsf{ME}}(\mathcal{KB}) = \arg \max_{\mathcal{P} \models \mathcal{KB}} - \sum_{\omega \in \Omega_{\mathcal{V}}} \mathcal{P}(\omega) \log \mathcal{P}(\omega).$$

Defining $\infty^0 = 1$, $\infty^{-1} = 0$, and $0^0 = 1$, $\mathcal{P}_{\mathsf{ME}}$ satisfies

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

where α_0 is a *normalizing constant* and the so-called *effects* α_i for $i = 1, \ldots, n$ fulfill the *adjustment condition*

$$(1 - x_i)\, \alpha_i^{1-x_i} \sum_{\omega \models A_i B_i} \prod_{\substack{j \ne i \\ \omega \models A_j B_j}} \alpha_j^{1-x_j} \prod_{\substack{j \ne i \\ \omega \models A_j \overline{B_j}}} \alpha_j^{-x_j}$$
$$= x_i\, \alpha_i^{-x_i} \sum_{\omega \models A_i \overline{B_i}} \prod_{\substack{j \ne i \\ \omega \models A_j B_j}} \alpha_j^{1-x_j} \prod_{\substack{j \ne i \\ \omega \models A_j \overline{B_j}}} \alpha_j^{-x_j}$$

as well as the *positivity condition*

$$\alpha_i \begin{cases} > 0 & \text{iff } x_i \in (0, 1) \\ = \infty & \text{iff } x_i = 1 \\ = 0 & \text{iff } x_i = 0 \end{cases}.$$

In terms of the conditional structure (1), the MaxEnt distribution $\mathcal{P}_{\mathsf{ME}}$ may be rewritten to

$$\mathcal{P}_{\mathsf{ME}}(\omega) = \alpha_0 \prod_{i=1}^n \alpha_i^{v_i(\omega) - x_i a_i(\omega)}, \quad (2)$$

and the adjustment condition leads to

$$\sum_{[\omega] \in [\Omega_{\mathcal{V}}]} (v_i(\omega) - x_i\, a_i(\omega))\, |[\omega]| \prod_{j=1}^n \alpha_j^{v_j(\omega) - x_j a_j(\omega)} = 0$$

for $i = 1, \ldots, n$. Since the MaxEnt distribution assigns the same probability to each possible world within the same equivalence class, $\mathcal{P}_{\mathsf{ME}}([\omega]) = |[\omega]|\, \mathcal{P}_{\mathsf{ME}}(\omega)$ holds. Thus, $\mathcal{P}_{\mathsf{ME}}$ and particularly the adjustment condition no longer depend on possible worlds but on equivalence classes of possible worlds induced by the conditional structure. Finally, the MaxEnt distribution yields a *nonmonotonic inference relation* $\mid\!\sim_{\mathsf{ME}}$ with $\mathcal{KB} \mid\!\sim_{\mathsf{ME}} (B|A)[x]$ iff $\mathcal{P}_{\mathsf{ME}} \models (B|A)[x]$. For any MaxEnt computation or MaxEnt inference task, being able to use the class-based representation of $\mathcal{P}_{\mathsf{ME}}$ yields significant advantages because of the reduction of the complexity. However, while the α_i's can be found by solving a nonlinear equation system, computing the equivalence classes of possible worlds and their respective sizes is non-trivial. In the next section, we present an algorithm that systematically reproduces the conditional structure.

An Algorithm for Clustering Possible Worlds

The algorithm $\textsc{CondStructor}(\mathcal{KB}^q, \mathcal{V})$ (Fig. 1) computes the equivalence classes in $[\Omega_{\mathcal{V}}]$ as well as their cardinalities. It follows ideas from (de Salvo Braz et al. 2015), where an algorithm is presented that allows for probabilistic inferences modulo theories in general. Our approach instead takes advantage of the specific algebraic structure induced by conditional knowledge.

The core idea behind the tree-based algorithm is as follows: Starting from the complete qualitative knowledge base $\mathcal{KB}^q$, it will iteratively select an unresolved conditional $(B|A)$ and split the current case into three more specific cases: One where $(B|A)$ is verified, one where it is falsified, and one where the conditional is not applicable. In order to do that, it checks if the premise A or its negation $\overline{A}$ is consistent with the previous selections of the algorithm and creates new nodes for those cases. If the premise can be satisfied in the current node, the algorithm will check if the conclusion B or its negation $\overline{B}$ can be satisfied as well. The selection of an unresolved conditional and the splitting into new nodes is done until all conditionals have been resolved. Finally, for each leaf node, the algorithm counts the number of worlds that satisfy the accumulated condition.

We now highlight some technical details of $\textsc{CondStructor}(\mathcal{KB}^q, \mathcal{V})$. Its main method initializes the set S with a tuple $(\mathcal{KB}^q, 1, \top, \mathcal{V})$ mainly containing

$\textsc{CondStructor}(\mathcal{KB}^q = \{(B_1|A_1), \ldots, (B_n|A_n)\}, \mathcal{V})$
Input: a qualitative knowledge base $\mathcal{KB}^q$ and an alphabet $\mathcal{V}$
Output: a set of tuples S where every tuple comprises an equivalence class and its cardinality

$S \leftarrow \{(\mathcal{KB}^q, 1, \top, \mathcal{V})\}$
WHILE exists $(E, L, C, P) \in S$ with $E \neq \emptyset$
  $S \leftarrow \textsc{Split}((E, L, C, P), S)$
FOR each $(\emptyset, L, C, P) \in S$
  $(\emptyset, L, C, P) \leftarrow (\emptyset, \textsc{SatCalc}_{\mathcal{V}}(C)\ L, C, \emptyset)$
RETURN S

$\textsc{Split}((E, L, C, P), S)$
Input: a tuple $(E, L, C, P) \in S$ and the set S itself
Output: an updated version of S where (E, L, C, P) is split according to a conditional in E

$\textsc{Select}((B_j|A_j), E)$
$S \leftarrow S \setminus \{(E, L, C, P)\}$
$E \leftarrow E \setminus \{(B_j|A_j)\}$
IF $\textsc{Sat}_{\mathcal{V}}(CA_j) = \mathsf{true}$ **THEN**
  IF $\textsc{Sat}_{\mathcal{V}}(CA_jB_j) = \mathsf{true}$ **THEN**
    $S \leftarrow S \cup \{(E, L\alpha_j^+, CA_jB_j, P)\}$
  IF $\textsc{Sat}_{\mathcal{V}}(CA_j\overline{B_j}) = \mathsf{true}$ **THEN**
    $S \leftarrow S \cup \{(E, L\alpha_j^-, CA_j\overline{B_j}, P)\}$
IF $\textsc{Sat}_{\mathcal{V}}(C\overline{A_j}) = \mathsf{true}$ **THEN**
  $S \leftarrow S \cup \{(E, L, C\overline{A_j}, P)\}$
RETURN S

Figure 1: Pseudo code of $\textsc{CondStructor}(\mathcal{KB}^q, \mathcal{V})$ and $\textsc{Split}((E, L, C, P), S)$

the knowledge base $\mathcal{KB}^q$ and the set of propositions $\mathcal{V}$ and iteratively selects an element (E, L, C, P) of S to invoke the method $\textsc{Split}((E, L, C, P), S)$. The latter non-deterministically selects an unresolved conditional $(B_i|A_i)$ from E by invoking the method $\textsc{Select}((B_j|A_j), E)$ and splits (E, L, C, P) into three new tuples as explained above, each corresponding to possible worlds in which $(B_i|A_i)$ is either verified, falsified, or not applicable. The latter is ascertained by a SAT solver $\textsc{Sat}_{\mathcal{V}}(C)$. For each case, the conjunction of the constraint C and A_iB_i, $A_i\overline{B_i}$, or $\overline{A_i}$ is built, and the appropriate factor $\mathbf{a_i^+}$, $\mathbf{a_i^-}$, or 1 is added to L. $\textsc{Split}((E, L, C, P), S)$ also checks if the updated constraint is satisfiable. If so, the relevant tuple is added to S. The algorithm terminates when there is no tuple in S with $E \neq 0$ which always happens since $\textsc{Split}((E, L, C, P), S)$ updates S with tuples whose entry E is reduced by an element every time. Finally, $\textsc{SatCalc}_{\mathcal{V}}(C)$ counts all models that satisfy the constraint C with respect to the set of propositions $\mathcal{V}$, i.e., it solves the #SAT problem (Biere et al. 2009). At this point, each remaining tuple $(\emptyset, L, C, \emptyset) \in S$ corresponds to an equivalence class $[\omega] \in [\Omega_{\mathcal{V}}]$, in particular to the one with $|[\omega]|\ \sigma_{\mathcal{KB}^q}([\omega]) = L$. Thus, the polynomial $\Phi(\mathcal{KB}^q, \mathcal{V}) = \sum_{(\emptyset, L, C, \emptyset) \in S} L$ represents all equivalence classes and their cardinalities in principle. However, it is useful to store the respective constraint C for every equivalence class, too, as this allows us to determine the corresponding equivalence class for every possible world. This capability will be necessary for drawing inferences.

We now demonstrate the functionality of our algorithm by means of an example.

Example 1. *Let $\mathcal{V}_{ex} = \{b, f, i, s\}$ and $\mathcal{KB}_{ex} = \{r_1, r_2\}$ be with $r_1 = (b \vee i|f)[0.9]$ and $r_2 = (\overline{f}s|b)[0.1]$. The conditional r_1 states that flying individuals are birds or insects with a probability 0.9. The conditional r_2 stands for the statement that birds do not fly but swim with probability 0.1. Fig. 2 shows one possible computation tree of the algorithm for the knowledge base $\mathcal{KB}_{ex}$; in this computation, the algorithm first resolves the conditional $(b \vee i|f)$ followed by $(\overline{f}s|b)$. For a more readable visualization of the algorithm, we adopted a sum-like notation for the tuples (E, L, C, P):*

$$\sum_{\substack{v_1, \ldots, v_k \in \{\mathsf{true}, \mathsf{false}\} \\ C = \mathsf{true}}} L \prod_{(B_i|A_i) \in E} (B_i|A_i)_i, \quad v_1, \ldots, v_k \in \mathcal{V}.$$

The indices of the conditionals in Fig. 2 illustrate the correspondence between r_i and $\mathbf{a_i^{\pm}}$. For example, the first equivalence class in the computation tree, which is $[\omega_1]$ with $\sigma_{\mathcal{KB}}([\omega_1]) = \mathbf{a_1^+ a_2^-}$, consists of the four possible worlds that satisfy the condition $f(b \vee i)b(f \vee \overline{s})$. In total, we get the following six equivalence classes:

$$\begin{aligned}
&[\omega_1] = \{fbis, fbi\overline{s}, fb\overline{i}s, fb\overline{i}\overline{s}\}, && [\omega_2] = \{f\overline{b}is, f\overline{b}i\overline{s}\}, \\
&[\omega_3] = \{f\overline{b}\overline{i}s, f\overline{b}\overline{i}\overline{s}\}, && [\omega_4] = \{\overline{f}bis, \overline{f}b\overline{i}s\}, \\
&[\omega_5] = \{\overline{f}bi\overline{s}, \overline{f}b\overline{i}\overline{s}\}, && [\omega_6] = \{\overline{f}\,\overline{b}is, \overline{f}\,\overline{b}i\overline{s}, \overline{f}\,\overline{b}\overline{i}s, \overline{f}\,\overline{b}\overline{i}\overline{s}\}.
\end{aligned}$$

Solving the nonlinear equation system given in the preliminaries, we get the following MaxEnt probabilities:

$$\begin{aligned}
&\mathcal{P}_{\mathit{ME}}(\mathcal{KB}_{ex})([\omega_1]) = 0.329, && \mathcal{P}_{\mathit{ME}}(\mathcal{KB}_{ex})([\omega_2]) = 0.148, \\
&\mathcal{P}_{\mathit{ME}}(\mathcal{KB}_{ex})([\omega_3]) = 0.053, && \mathcal{P}_{\mathit{ME}}(\mathcal{KB}_{ex})([\omega_4]) = 0.053, \\
&\mathcal{P}_{\mathit{ME}}(\mathcal{KB}_{ex})([\omega_5]) = 0.148, && \mathcal{P}_{\mathit{ME}}(\mathcal{KB}_{ex})([\omega_6]) = 0.268.
\end{aligned}$$

Drawing Inferences

As we have seen, the equivalence classes derived by the algorithm can be used to compute MaxEnt distributions. Another application is to draw inferences, i.e., deciding for which probability x the query $\mathcal{P}_{\mathsf{ME}} \models (B|A)[x]?$ holds, given a knowledge base $\mathcal{KB}$ and a conditional $(B|A)$. For this, we need to count how many possible worlds in each equivalence class satisfy the premise and the conclusion of the conditional $(B|A)$, i.e., for each tuple $(\emptyset, L, C, \emptyset)$ in the output S of $\textsc{CondStructor}(\mathcal{KB}^q, \mathcal{V})$ with the corresponding equivalence class $[\omega_i]$, we need to count the models of CA and CAB, denoted by $c([\omega_i], A)$ and $c([\omega_i], AB)$. Then $\mathcal{P}_{\mathsf{ME}} \models (B|A)[x]$ holds for

$$x = \frac{\sum_{[\omega_i] \in [\Omega_{\mathcal{V}}]} \mathcal{P}_{\mathsf{ME}}([\omega_i]) \frac{c(\omega_i, AB)}{|[\omega_i]|}}{\sum_{[\omega_i] \in [\Omega_{\mathcal{V}}]} \mathcal{P}_{\mathsf{ME}}([\omega_i]) \frac{c(\omega_i, A)}{|[\omega_i]|}}.$$

Example 2. *Consider $\mathcal{KB}_{ex}$ from Ex. 1 and the query $\mathcal{P}_{\mathit{ME}} \models (s|b)[x]?$, i.e., we are interested in the probability that a bird can swim, given the prior knowledge from $\mathcal{KB}_{ex}$. Since the premise of $(s|b)$ is the same as that of r_2, it is fulfilled in exactly those possible worlds in any $[\omega]$ with*

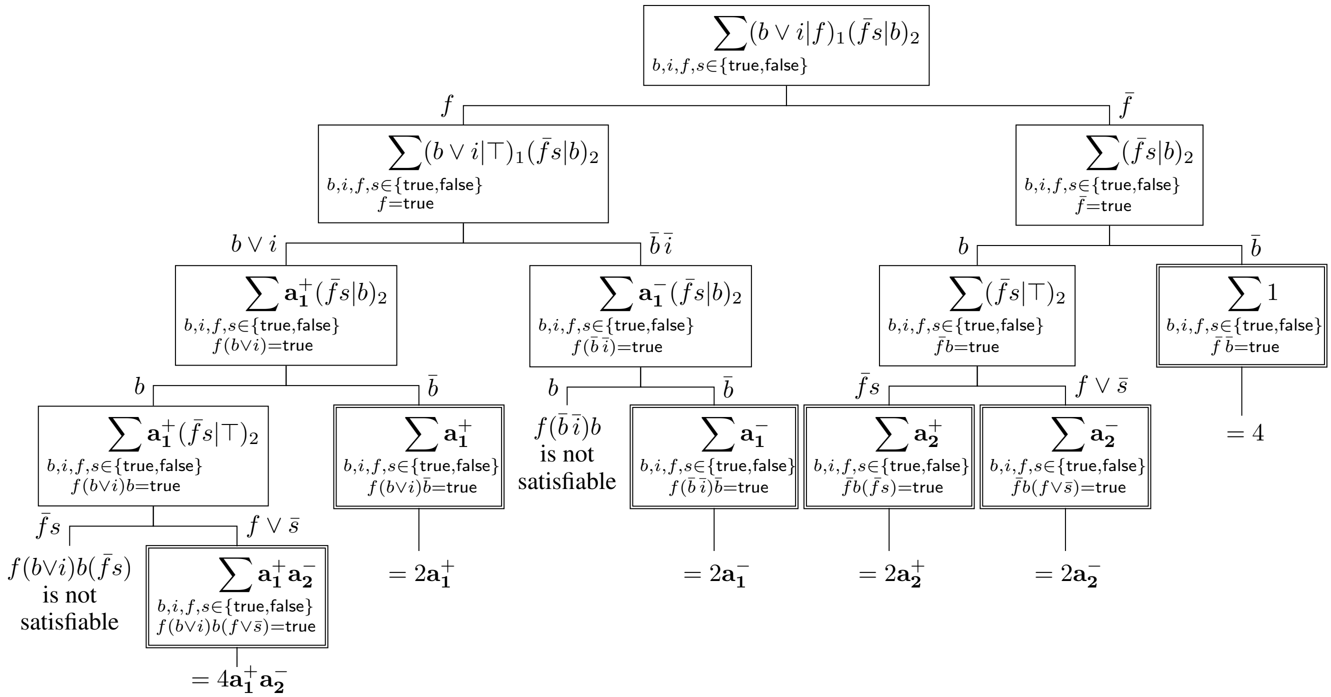

Figure 2: Stepwise computation of the equivalence classes of possible worlds by splitting

$a_2([\omega]) = 1$. *When counting how many possible worlds in each of these equivalence classes additionally satisfy the conclusion s, we get $c([\omega_1], sb) = 2$, $c([\omega_4], sb) = 2$, and $c([\omega_i], sb) = 0$ for all other equivalence classes. Thus, given $\mathcal{KB}_{ex}$, the most reasonable probability for birds to be able to swim is* $x = (0.5{\cdot}0.329{+}0.053)/(0.329{+}0.053{+}0.148) = 0.41$.

Conclusion and Future Work

We presented an algorithm for computing the equivalence classes of possible worlds with respect to the conditional structure induced by a knowledge base $\mathcal{KB}$. Notably, the output of the algorithm can be used both for establishing the adjustment condition of the MaxEnt distribution regarding $\mathcal{KB}$ (Kern-Isberner 2001), as well as for answering MaxEnt inference queries. The form of the algorithm allows for various modifications and optimizations: For example, the splitting procedure may be performed with respect to the truth values of atoms, conditionals with zero probability may be used to ignore irrelevant possible worlds, and factorization can help decomposing the conditional structure efficiently. We are currently working on implementations that incorporate these approaches, and then, we want to test our implementation on real life problems. Although this paper focuses on propositional knowledge bases, our algorithm might principally be lifted to the relational case since the concept of conditional structure carries over to it (Finthammer and Beierle 2012; Beierle et al. 2015). In future work, we will pursue investigations in this direction, and we are hopeful that an adaption of our algorithm is able to offset the massive increase of the number of possible worlds when increasing the domain size in first-order settings.

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

References

Beierle, C.; Kuche, S.; Finthammer, M.; and Kern-Isberner, G. 2015. A software system for the computation, visualization, and comparison of conditional structures for relational probabilistic knowledge bases. In *Proc. of FLAIRS'15*.

Biere, A.; Heule, M.; Van Maaren, H.; and Walsh, T. 2009. *Handbook of Satisfiability*. IOS Press.

de Salvo Braz, R.; O'Reilly, C.; Gogate, V.; and Dechter, R. 2015. Probabilistic inference modulo theories. In *Workshop on Hybrid Reasoning at IJCAI'15*.

Finthammer, M., and Beierle, C. 2012. Using equivalences of worlds for aggregation semantics of relational conditionals. In *Proc. of KI'12*, volume 7526 of *LNCS*, 49–60.

Jaynes, E. 1983. *Papers on Probability, Statistics and Statistical Physics*. D. Reidel Publishing Company.

Kern-Isberner, G. 2001. *Conditionals in Nonmonotonic Reasoning and Belief Revision*. Springer.

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

Paris, J. 1999. Common sense and maximum entropy. *Synthese* 117:75–93.

Shore, J., and Johnson, R. 1980. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory* IT-26:26–37.