Proceedings of the 20th International Conference on Principles of Knowledge Representation and Reasoning
Main Track

Integrating Linear Arithmetic Constraints Into Conditional Maximum Entropy Reasoning

Marco Wilhelm
Department of Computer Science, TU Dortmund University, Dortmund, Germany
marco.wilhelm@tu-dortmund.de

Abstract

The principle of maximum entropy (MaxEnt principle) constitutes a valuable methodology for probabilistic commonsense reasoning by adding missing information to probabilistic conditional belief bases in an information theoretically optimal way. In this paper, we integrate linear arithmetic constraints over the integers and reals into propositional probabilistic conditionals in order to be able to formalize uncertain beliefs about arithmetic expressions. The satisfiability of (sets of) constraints is decided modulo theory such that probabilistic reasoning stays finite although the constraints range over infinite domains. Therewith, we provide a novel extension of the MaxEnt principle to beliefs about infinite domains.

1 Introduction

Probability theory constitutes one of the most powerful and widely used frameworks for quantitative uncertain reasoning. Probabilistic conditionals of the form $(B|A)[\xi]$ with the meaning "if A holds, then B follows with probability ξ" are a concise means to make beliefs explicit, and inductive reasoning methodologies like probabilistic networks or the principle of maximum entropy (MaxEnt principle, (Paris 1994)) provide convenient methods for answering queries. This paper focuses on the MaxEnt principle as a most appropriate form of probabilistic commonsense reasoning (Paris 1998) which endows the reasoner with a nonmonotonic inference relation that allows for drawing inferences from probabilistic conditional belief bases. By fulfilling the paradigm of informational economy (Gärdenfors 1988), reasoning at maximum entropy adds as little information as possible and, thus, is as cautious as possible.

While MaxEnt reasoning wrt. finite domains is understood quite well, it is an open question how to extend the MaxEnt principle to beliefs about infinite domains. The naïve approach of considering probabilistic conditionals which are defined over a first-order language with infinitely many constants runs into the fundamental problem that there is no uniform distribution on countably infinite sample spaces. Therefore, common approaches to MaxEnt reasoning wrt. infinite domains consider limit processes (Barnett and Paris 2008; Paris and Rad 2010; Thimm and Kern-Isberner 2012) or come up with a sophisticated comparative notion of maximality (Landes and Williamson 2015; Williamson 2008) without a generally accepted policy so far.

In this paper, we integrate (linear arithmetic) constraints over the integers and reals into propositional probabilistic conditionals in order to be able to formalize uncertain beliefs about arithmetic expressions. Therewith, we subtly open the door for beliefs about infinite domains, both countably infinite ($\mathbb{Z}$) and uncountably infinite ($\mathbb{R}$). We give a running example for this paper which is about uncertain beliefs that involve such constraints.

Example 1. *A shareholder wonders if the company in which she holds shares will pay a dividend at the end of the year. She believes with a probability of* 0.8 *that this is the case if the company makes high profit, i.e., if the income exceeds the expenses by more than* $\$\,1,000,000$*. So far, the company has produced an income of* $\$\,10,000,000$*. There are further payments pending, but it is unclear whether they will be executed this year. The expenses are unlikely to exceed* $\$\,12,000,000$ *as the main investments have already been finalized. Thus, with a probability of* 0.9 *the expenses will stay below the threshold of* $\$\,12,000,000$*.*

For instance, in Ex. 1 the condition that "the income exceeds the expenses by more than $\$\,1,000,000$" is formalizable by the constraint $x_{inc} - y_{exp} > 1,000,000$.

The satisfiability of constraints within conditionals is tested modulo theory (known as SMT, (Barrett et al. 2021)) such that the conditional language remains finite and no conceptual adaption of the MaxEnt principle has to be made. Our approach is in line with (de Salvo Braz et al. 2016) where probabilistic inferences modulo theory were introduced before. Novel here is the combination with the MaxEnt principle. A major part of the paper deals with the influence of the syntactic representation of constraints on reasoning which is also not a subject in (de Salvo Braz et al. 2016). By aggregating possible worlds with the same conditional structure (Kern-Isberner 2004), we come up with a notion of aggregated MaxEnt models which is independent of the syntactic representation of constraints.

The rest of the paper is organized as follows. First, we recall some preliminaries on probabilistic conditional reasoning in general and on MaxEnt reasoning in particular before we briefly discuss the theory of linear arithmetic. After that, we introduce a propositional logic of linear arithmetic constraints and use this logic to define probabilistic conditionals with constraint representatives. Thereafter, we show how MaxEnt reasoning based on such conditionals works and how it can be made independent of the syntactic representation of constraints. Eventually, we discuss related work and conclude with an outlook.

2 Preliminaries

We consider a *propositional language* $\mathcal{L}(\Sigma)$ which is defined over a finite *signature* Σ. Elements in Σ are called *atoms*. *Formulas* in $\mathcal{L}(\Sigma)$ are either atoms or compounded formulas of the form $A \wedge B$ (*conjunction*), $A \vee B$ (*disjunction*), or $\neg A$ (*negation*) where A and B are formulas. A *literal* is either an atom or its negation. The semantics of formulas is given by *interpretations* $I \in \mathsf{Int}(\Sigma)$ which assign a truth value from $\{0, 1\}$ to formulas as usual. To shorten expressions, we abbreviate conjunctions $A \wedge B$ with AB, negations $\neg A$ with $\overline{A}$, and tautologies $A \vee \overline{A}$ with $\top$.

Probabilistic conditionals $(B|A)[\xi]$ consisting of formulas $A, B \in \mathcal{L}(\Sigma)$ and a probability $\xi \in [0, 1]$ formalize uncertain statements of the form: "If A holds, then B follows with probability ξ." With $\mathcal{PCL}(\Sigma)$ we denote the set of all conditionals over Σ. A finite set of conditionals is called *belief base*. The semantics of conditionals is given by probability distributions over *possible worlds*. Here, possible worlds are in a one-to-one correspondence with interpretations and written as conjunctions of those literals which hold in the respective interpretation. That is, a possible world is a complete conjunction of literals in which every atom from Σ occurs once, either positive or negated. The set of all possible worlds is denoted with $\Omega(\Sigma)$. For $\omega \in \Omega(\Sigma)$ and $\Sigma' \subseteq \Sigma$,

$$\omega_{|\Sigma'} = \bigwedge\nolimits_{a \in \Sigma' :\, \omega \models a} a \wedge \bigwedge\nolimits_{a \in \Sigma' :\, \omega \models \bar{a}} \bar{a}$$

is the *marginalization* of ω on Σ'.

A probability distribution $\mathcal{P} \colon \Omega(\Sigma) \to [0, 1]$ which assigns a probability to each possible world in $\Omega(\Sigma)$ is a *probabilistic model* of a conditional $(B|A)[\xi]$, written

$$\mathcal{P} \models (B|A)[\xi], \text{ if } \mathcal{P}(A) > 0 \text{ and } \mathcal{P}(B|A) = \xi,$$

whereby the probability of a formula A is the sum of the probabilities of its models, $\mathcal{P}(A) = \sum_{\omega \in \Omega(\Sigma) :\, \omega \models A} \mathcal{P}(\omega)$, and $\mathcal{P}(B|A) = \mathcal{P}(AB) \cdot \mathcal{P}(A)^{-1}$ is the conditional probability of B given A. Further, $\mathcal{P}$ is a *probabilistic model* of a belief base $\mathcal{R}$ if $\mathcal{P}$ models every conditional in $\mathcal{R}$. If a belief base $\mathcal{R}$ has a model, then $\mathcal{R}$ is called *consistent*.

Consistent belief bases usually have infinitely many probabilistic models. For reasoning tasks such as drawing inferences, it is useful to select a unique model among them as calling for inferences that hold in all models often leads to only little and uninformative new information (Wilhelm et al. 2022). From a commonsense perspective, the *maximum entropy (MaxEnt) model* $\mathsf{ME}_{\mathcal{R}}$ of a consistent belief base $\mathcal{R}$ is the preferable model (Paris 1998). The basic idea of the underlying *MaxEnt principle* is to determine the undefined probabilities while adding as little information as possible. Technically, $\mathsf{ME}_{\mathcal{R}}$ is the unique model of $\mathcal{R}$ which maximizes the *entropy* $\mathcal{H}(\mathcal{P}) = -\sum_{\omega \in \Omega(\Sigma)} \mathcal{P}(\omega) \cdot \log \mathcal{P}(\omega)$, i.e., which is given by

$$\mathsf{ME}_{\mathcal{R}} = \arg\max_{\mathcal{P} \models \mathcal{R}} \mathcal{H}(\mathcal{P}),$$

where the convention $0 \cdot \log 0 = 0$ applies.

For consistent belief bases $\mathcal{R}$ and conditionals $(B|A)[\xi]$, the MaxEnt principle yields the *nonmonotonic relation*

$$\mathcal{R} \mathrel{|\!\sim_{\mathsf{ME}}} (B|A)[\xi] \text{ if } \mathsf{ME}_{\mathcal{R}} \models (B|A)[\xi] \tag{1}$$

which satisfies the properties *Direct Inference* (DI) and *Trivial Vacuity* (TV),

- if $(B|A)[\xi] \in \mathcal{R}$, then $\mathcal{R} \mathrel{|\!\sim_{\mathsf{ME}}} (B|A)[\xi]$, (DI)
- if $\mathcal{R} = \emptyset$, then $\mathcal{R} \mathrel{|\!\sim_{\mathsf{ME}}} (B|A)[1]$ only if $A \models B$, (TV)

and, hence, is an *inductive inference relation* according to (Kern-Isberner, Beierle, and Brewka 2020).

The goal of this paper is to integrate *linear arithmetic constraints* into conditional MaxEnt reasoning. With *linear arithmetic constraints*, or *constraints* for short, we refer to mathematical expressions of the form

$$a_1 \cdot x_1 + \ldots + a_m \cdot x_m \bowtie a_0,$$

where $m \in \mathbb{N}_{>0}$, the x_i's are numeric variables, the a_i's are constants from $\mathbb{Z}$ or $\mathbb{R}$, and where $\bowtie \in \{<, \leq, =, \neq, \geq, >\}$ (cf. (Barrett et al. 2021)). Depending on whether the variables and constants range over the integers or reals, we refer to the arithmetic of such constraints with *linear integer arithmetic ($\mathcal{LIA}$)* or *linear real arithmetic ($\mathcal{LRA}$)*. Note that albeit both arithmetics are defined over infinite domains they are decidable. More precisely, the satisfiability of any finite set of constraints from $\mathcal{LRA}$ can be decided in polynomial time (Karmarkar 1984) whereas the satisfiability of $\mathcal{LIA}$-constraints is NP-complete (Papadimitriou 1981). The arithmetic $\mathcal{LIA}$ is also known as *Presburger arithmetic* (Presburger 1929).

3 Logic of Linear Arithmetic Constraints

As preparatory work for integrating (linear arithmetic) constraints into conditionals, we define a *propositional language of constraints*. For this, let $\mathcal{LA} = \mathcal{LIA} \cup \mathcal{LRA}$ be the set of constraints from either $\mathcal{LIA}$ or $\mathcal{LRA}$. Obviously, $\mathcal{LA}$ is decidable just as $\mathcal{LIA}$ and $\mathcal{LRA}$. With $\mathsf{Sol}(\mathcal{C})$ we denote the set of the *common solutions* of any set of constraints $\mathcal{C} \subseteq \mathcal{LA}$. Further, for each constraint $c \in \mathcal{LA}$ we introduce a fresh symbol a_c and subsume all these symbols within the (uncountably infinite) set $\Sigma_{\mathcal{LA}}$. We call a_c the *constraint representative* of c.

Definition 1 (Propositional Language of Constraints)**.** *Let $\Sigma^f_{\mathcal{LA}} \subset \Sigma_{\mathcal{LA}}$ be a finite set of constraint representatives. We call the propositional language $\mathcal{L}(\Sigma^f_{\mathcal{LA}})$ in which the constraint representatives from $\Sigma^f_{\mathcal{LA}}$ play the role of atoms the* propositional language of constraints *over* $\Sigma^f_{\mathcal{LA}}$.

Formulas in $\mathcal{L}(\Sigma^f_{\mathcal{LA}})$ and their interpretations are defined as usual in propositional logic. However, if $\mathcal{L}(\Sigma^f_{\mathcal{LA}})$ shall be coherent with $\mathcal{LA}$, we have to have a closer look at the semantics of formulas in $\mathcal{L}(\Sigma^f_{\mathcal{LA}})$ because the constraint representatives in $\Sigma^f_{\mathcal{LA}}$ are not independent when understood as constraints in $\mathcal{LA}$ in general. Instead, the satisfiability of these constraints can depend on each other. In particular, for each constraint

$$a_1 \cdot x_1 + \ldots + a_m \cdot x_m \bowtie a_0$$

in $\mathcal{LA}$, there is a *complementary constraint*

$$a_1 \cdot x_1 + \ldots + a_m \cdot x_m \; \phi(\bowtie) \; a_0$$

where ϕ maps $<$ to $\geq$, $\leq$ to $>$, $=$ to $\neq$, and vice versa. We denote the *complement* of a constraint $c \in \mathcal{LA}$ with $\widehat{c}$. Obviously, the solution sets of a constraint c and its complement $\widehat{c}$ are disjoint and, hence, $\mathsf{Sol}(\{c, \widehat{c}\}) = \emptyset$. Contrary to that, if $a_c, a_{\widehat{c}} \in \Sigma^f_{\mathcal{LA}}$, then there is an interpretation $I \in \mathsf{Int}(\Sigma^f_{\mathcal{LA}})$ with $I(a_c) = 1$ and $I(a_{\widehat{c}}) = 1$ which states that a_c and $a_{\widehat{c}}$ hold at the same time. In order to exclude such "improper" interpretations, we introduce a notion of coherency between $\mathcal{L}(\Sigma^f_{\mathcal{LA}})$ and $\mathcal{LA}$ on the level of interpretations.

Definition 2 (Coherent Interpretation)**.** *Let $\Sigma^f_{\mathcal{LA}}$ be a finite set of constraint representatives, and for $I \in \mathsf{Int}(\Sigma^f_{\mathcal{LA}})$ let*

$$\mathcal{CS}(I) = \{c \in \mathcal{LA} \mid \exists a_c \in \Sigma^f_{\mathcal{LA}} \colon I(a_c) = 1\} \cup \{\widehat{c} \in \mathcal{LA} \mid \exists a_c \in \Sigma^f_{\mathcal{LA}} \colon I(a_c) = 0\} \quad (2)$$

be the translation of the interpretation I into the corresponding set of constraints ("constraint set"*) from $\mathcal{LA}$. We call I* coherent *if $\mathcal{CS}(I)$ is satisfiable and denote with $\mathsf{Int}_c(\Sigma^f_{\mathcal{LA}})$ the set of all coherent interpretations over $\Sigma^f_{\mathcal{LA}}$.*

Coherent interpretations respect $\mathcal{LA}$ in that they are not self-contradictory when understood as constraint sets. That is, $\mathsf{Sol}(\mathcal{CS}(I)) \neq \emptyset$ holds iff $I \in \mathsf{Int}_c(\Sigma^f_{\mathcal{LA}})$. In particular, there is no coherent interpretation $I \in \mathsf{Int}_c(\Sigma^f_{\mathcal{LA}})$ with $I(a_c) = I(a_{\widehat{c}}) = 1$ (or with $I(a_c) = I(a_{\widehat{c}}) = 0$) because $\mathsf{Sol}(\{c, \widehat{c}\}) = \emptyset$. As a consequence, $\neg a_c \equiv a_{\widehat{c}}$ holds if logical equivalence is defined wrt. coherent interpretations only.

We call

$$\mathsf{vconf}(I) = \mathsf{Sol}(\mathcal{CS}(I))$$

the set of the *variable configurations* which are *admissible* in $I \in \mathsf{Int}_c(\Sigma^f_{\mathcal{LA}})$. Coherent interpretations have at least one admissible variable configuration. With

$$\mathsf{vconf}(A) = \bigcup_{I \in \mathsf{Int}_c(\Sigma^f_{\mathcal{LA}}) \colon I \models A} \mathsf{vconf}(I)$$

we generalize the notion of admissible variable configurations to formulas $A \in \mathcal{L}(\Sigma^f_{\mathcal{LA}})$.

Example 2. *We refer to Ex. 1 and consider the constraints*

$$\begin{aligned} p\colon & \quad x - y > 1{,}000{,}000 \\ r\colon & \quad x = 10{,}000{,}000 \\ o\colon & \quad x > 10{,}000{,}000 \\ e\colon & \quad y \leq 12{,}000{,}000 \end{aligned}$$

where p stands for high profit *("The income x exceeds the expenses y by more than \$ 1,000,000."), r for* received income *("The received income is \$ 10,000,000."), o for* income with outstanding payments *("The income with outstanding payments is higher than \$ 10,000,000."), and e for* maximal expenses *("The expenses are not higher than \$ 12,000,000."). For simplicity, the numeric variables x (income) and y (expenses) are assumed to be non-negative integers. The corresponding constraint representatives are*

$$\Sigma^{f,\mathsf{ex}}_{\mathcal{LA}} = \{a_p, a_r, a_o, a_e\}.$$

Not all interpretations in $\mathsf{Int}(\Sigma^{f,\mathsf{ex}}_{\mathcal{LA}})$ are coherent. For example, the constraints r and o cannot be satisfied jointly and interpretations $I \in \mathsf{Int}(\Sigma^{f,\mathsf{ex}}_{\mathcal{LA}})$ with $I(a_r) = I(a_o) = 1$ are not coherent. Likewise, interpretations with $I(a_p) = 1$ but $I(a_o) = I(a_e) = 0$ are not coherent because

$$\mathsf{Sol}(\{p, \widehat{o}, \widehat{e}\}) = \mathsf{Sol}(\{x - y > 1{,}000{,}000, \; x \leq 10{,}000{,}000, \; y > 12{,}000{,}000\}) = \emptyset.$$

The coherent interpretations (denoted as complete conjunctions of the literals which are true in the interpretation) are

$$\mathsf{Int}_c(\Sigma^{f,\mathsf{ex}}_{\mathcal{LA}}) = \{a_p a_r \overline{a_o} a_e, \; a_p \overline{a_r} a_o a_e, \; a_p \overline{a_r} a_o \overline{a_e}, \; a_p \overline{a_r}\,\overline{a_o} a_e, \; \overline{a_p} a_r \overline{a_o} a_e, \; \overline{a_p} a_r \overline{a_o}\,\overline{a_e}, \; \overline{a_p}\,\overline{a_r} a_o a_e, \; \overline{a_p}\,\overline{a_r} a_o \overline{a_e}, \; \overline{a_p}\,\overline{a_r}\,\overline{a_o} a_e, \; \overline{a_p}\,\overline{a_r}\,\overline{a_o}\,\overline{a_e}\}.$$

For instance, we have

$$\mathsf{vconf}(a_p a_r \overline{a_o} a_e) = \{(x, y) \mid x = 10 \cdot 10^6 \wedge y < 9 \cdot 10^6\}.$$

Focusing on coherent interpretations yields a notion of *satisfiability modulo theory* (SMT, cf. (Barrett et al. 2021)).

Definition 3 (Satisfiability Modulo $\mathcal{LA}$)**.** *Let $\Sigma^f_{\mathcal{LA}}$ be a finite set of constraint representatives. A formula A in $\mathcal{L}(\Sigma^f_{\mathcal{LA}})$ is* satisfiable modulo $\mathcal{LA}$ *if there is a coherent interpretation $I \in \mathsf{Int}_c(\Sigma^f_{\mathcal{LA}})$ with $I(A) = 1$.*

The coherency of interpretations has a much bigger influence than just excluding self-contradictory interpretations as the next proposition shows.

Proposition 1. *Let $\Sigma^f_{\mathcal{LA}}$ be a finite set of constraint representatives, and let $I, J \in \mathsf{Int}_c(\Sigma^f_{\mathcal{LA}})$ be coherent interpretations with $I \neq J$. Then, $\mathsf{vconf}(I) \cap \mathsf{vconf}(J) = \emptyset$.*

Proof. Because $I \neq J$, there is a constraint representative $a_c \in \Sigma^f_{\mathcal{LA}}$ with $I(a_c) \neq J(a_c)$. Without loss of generality, we assume $I(a_c) = 1$ and $J(a_c) = 0$. Then, $I(a_c) = 1$ implies $c \in \mathcal{CS}(I)$ and $\mathsf{vconf}(I) \subseteq \mathsf{Sol}(\{c\})$ follows. Note that the variables which are not mentioned in c but in $\mathcal{CS}(I)$ are not restricted by the solutions in $\mathsf{Sol}(\{c\})$ but can be chosen freely so that $\mathsf{vconf}(I)$ and $\mathsf{Sol}(\{c\})$ refer to the same set of variables. On the other hand, $J(a_c) = 0$ implies $\widehat{c} \in \mathcal{CS}(J)$ and $\mathsf{vconf}(J) \subseteq \mathsf{Sol}(\{\widehat{c}\})$ holds. Because $\mathsf{Sol}(\{c\}) \cap \mathsf{Sol}(\{\widehat{c}\}) = \emptyset$, we can conclude $\mathsf{vconf}(I) \cap \mathsf{vconf}(J) \subseteq \mathsf{Sol}(\{c\}) \cap \mathsf{Sol}(\{\widehat{c}\}) = \emptyset$. ☐

According to Prop. 1, different coherent interpretations have disjoint sets of admissible variable configurations. Because these sets are non-empty by the definition of coherency, we can directly conclude that they are distinct. Therefore, we can say that coherent interpretations are unique semantic entities.

Corollary 1. *Let $\Sigma^f_{\mathcal{LA}}$ be a finite set of constraint representatives, and let $I, J \in \mathsf{Int}_c(\Sigma^f_{\mathcal{LA}})$ be coherent interpretations with $I \neq J$. Then, $\mathsf{vconf}(I) \neq \mathsf{vconf}(J)$.*

To summarize, the benefit of treating constraint representatives as propositional atoms and focusing on coherent interpretations is that we achieve the ability to build formulas, i.e., Boolean combinations of constraint representatives and test their satisfiability modulo the theory of linear arithmetic. Considering all interpretations would undermine the fact that the underlying constraints are not necessarily semantically independent. In contrast to that, coherent interpretations can be understood as possible worlds due to their semantic uniqueness.

4 Integrating Linear Arithmetic Constraints Into Conditionals

Now we extend conditionals from $\mathcal{PCL}(\Sigma)$ by (linear arithmetic) constraints. For this, let $\Sigma^f_{\mathcal{LA}}$ be a finite set of constraint representatives. We assume that $\Sigma \cap \Sigma^f_{\mathcal{LA}} = \emptyset$ holds. Otherwise, this can be established by renaming the atoms in Σ. We set $\Sigma_\cup = \Sigma \,\dot\cup\, \Sigma^f_{\mathcal{LA}}$ and consider the propositional language $\mathcal{L}(\Sigma_\cup)$ which draws on both the atoms from Σ and the constraint representatives from $\Sigma^f_{\mathcal{LA}}$. An interpretation $I \in \mathsf{Int}(\Sigma_\cup)$ is *coherent* if the restriction of I on $\Sigma^f_{\mathcal{LA}}$, in symbols $I_{|\Sigma^f_{\mathcal{LA}}}$, is coherent. The set of all coherent interpretations over $\Sigma_\cup$ is denoted with $\mathsf{Int}_c(\Sigma_\cup)$.

Definition 4 (Conditionals with Constraint Representatives)**.** *Let $\Sigma_\cup = \Sigma \,\dot\cup\, \Sigma^f_{\mathcal{LA}}$ be a finite set of atoms and constraint representatives. A* conditional with constraint representatives *is an expression of the form $(B|A)[\xi]$ where $A, B \in \mathcal{L}(\Sigma_\cup)$ and $\xi \in [0,1]$ is a probability. The set of all conditionals with constraint representatives is $\mathcal{PCL}(\Sigma_\cup)$.*

Conditionals in $\mathcal{PCL}(\Sigma_\cup)$ express probabilistic beliefs about qualitative logical statements combined with statements about admissible variable configurations. We use the term *belief base* for finite subsets of $\mathcal{PCL}(\Sigma_\cup)$, too.

Example 3. *The beliefs of the shareholder from Ex. 1 can be formalized in the belief base $\mathcal{R}^{\mathsf{ex}} = \{r_1, r_2, r_3\}$ with*

$$r_1 = (d|a_p)[0.8],\ r_2 = (a_r \vee a_o|\top)[1],\ r_3 = (a_e|\top)[0.9].$$

The meaning of the constraint representatives a_p, a_r, a_o, and a_e is discussed in Ex. 2. The additional atom d stands for dividend is paid*. Altogether, the signature is*

$$\Sigma^{\mathsf{ex}}_\cup = \Sigma^{\mathsf{ex}} \,\dot\cup\, \Sigma^{f,\mathsf{ex}}_{\mathcal{LA}} = \{d\} \,\dot\cup\, \{a_p, a_r, a_o, a_e\}.$$

Conditional r_1 states that if the profit of the company is high (a_p), then a dividend is paid (d) with probability 0.8, where the meaning of "high profit" is specified in the constraint p. Conditional r_2 states that the income of the company is $\$\,10,000,000$ (a_r) or higher (a_o), and conditional r_3 states that the company's expenses do not exceed $\$\,12,000,000$ with a probability of 0.9.

The definition of probabilistic models of belief bases $\mathcal{R} \subseteq \mathcal{PCL}(\Sigma_\cup)$ is analogous to the respective definition in the purely propositional setting albeit we have to clarify the notion of possible worlds. As argued in Sec. 3, especially in Cor. 1, it is not appropriate to consider all interpretations over $\Sigma_\cup$ as possible worlds but one needs to focus on coherent interpretations.

Definition 5 (Possible Worlds (With Constraint Representatives))**.** *Let $\Sigma_\cup = \Sigma \,\dot\cup\, \Sigma^f_{\mathcal{LA}}$ be a finite set of atoms and constraint representatives, and let $I \in \mathsf{Int}_c(\Sigma_\cup)$ be a coherent interpretation. Then, the* possible world *wrt. I is*

$$\omega_I = \bigwedge\nolimits_{a \in \Sigma_\cup:\, I(a)=1} a \wedge \bigwedge\nolimits_{a \in \Sigma_\cup:\, I(a)=0} \bar{a}.$$

We denote the set of all possible worlds over $\Sigma_\cup$ (which are defined wrt. coherent interpretations) with $\Omega_c(\Sigma_\cup)$ and with $\mathcal{CS}(\omega_I) = \mathcal{CS}(I_{|\Sigma^f_{\mathcal{LA}}})$ we denote the constraint set which refers to $\omega_I \in \Omega_c(\Sigma_\cup)$.

Possible worlds with constraint representatives are in a one-to-one correspondence with coherent interpretations. If $\Sigma^f_{\mathcal{LA}} = \emptyset$, then $\Sigma_\cup = \Sigma$ holds and the interpretations in $\mathsf{Int}(\Sigma_\cup)$ are trivially coherent such that Def. 5 subsumes the standard definition of possible worlds in $\Omega(\Sigma)$.

All notions from probabilistic conditional reasoning carry over from $\mathcal{PCL}(\Sigma)$ to $\mathcal{PCL}(\Sigma_\cup)$ when replacing the set of possible worlds $\Omega(\Sigma)$ by $\Omega_c(\Sigma_\cup)$. Also the definition of the MaxEnt model $\mathsf{ME}_\mathcal{R}$ is the same. Note that the probability $\mathsf{ME}_\mathcal{R}(a_c)$ of a constraint representative a_c reflects the subjective probability with which the MaxEnt reasoner with belief base $\mathcal{R}$ believes in the satisfaction of the constraint c. It does not depend on a measure on admissible variable configurations. In particular, it does not depend on the number of solutions in $\mathsf{Sol}(\{c\})$.

Example 4. *The MaxEnt model $\mathsf{ME}_{\mathcal{R}^{\mathsf{ex}}}$ of the belief base $\mathcal{R}_{\mathsf{ex}}$ from Ex. 3 is shown in Tab. 1. It can be computed with the software system* `Spirit` *(Rödder and Meyer 1996), for instance. With regard to the MaxEnt inference relation (1), we can infer the conditional $(d|\top)[0.63]$ from $\mathcal{R}_{\mathsf{ex}}$. That is, following the MaxEnt principle and her beliefs as formalized in $\mathcal{R}_{\mathsf{ex}}$, the shareholder should believe in the payout of a dividend with probability 0.63.*

In the reminder of the paper, we address the issue that reasoning with probabilistic conditionals in $\mathcal{PCL}(\Sigma_\cup)$ depends on the selection of the constraints which are represented in $\Sigma^f_{\mathcal{LA}}$. First, we show how reasoning can be made independent of this selection and, therewith, independent of the syntactic representation of admissible variable configurations. Afterwards, we discuss an extension of the MaxEnt inference relation (1) which allows us to infer probability bounds for constraints that are not represented in $\Sigma^f_{\mathcal{LA}}$.

5 Becoming Independent of the Syntactic Representation of Variable Configurations

Once a set $\Sigma_\cup = \Sigma \,\dot\cup\, \Sigma^f_{\mathcal{LA}}$ of atoms and constraint representatives is fixed, MaxEnt reasoning in $\mathcal{PCL}(\Sigma_\cup)$ is independent of the syntactic representation of admissible variable configurations because the coherent interpretations in $\mathsf{Int}_c(\Sigma_\cup)$ and, thus, the possible worlds in $\Omega_c(\Sigma_\cup)$ are semantically unique entities which differ in their admissible variable configurations (cf. Cor. 1). However, in $\mathcal{LA}$ it is possible to determine one and the same solution set with different sets of constraints and, hence, the selection of the constraint representatives in $\Sigma^f_{\mathcal{LA}}$ is not unique. For instance, all the constraints $n \cdot x = n$ with $n \in \mathbb{N}_{>0}$ have the same solution $x = 1$. As a consequence, the choice of $\Sigma^f_{\mathcal{LA}}$ can affect the MaxEnt model of a belief base in $\mathcal{PCL}(\Sigma_\cup)$.

ω	$\mathsf{ver}_{\mathcal{R}^{\mathsf{ex}}}(\omega)$	$\mathsf{fal}_{\mathcal{R}^{\mathsf{ex}}}(\omega)$	$\mathsf{ME}_{\mathcal{R}^{\mathsf{ex}}}(\omega)$	ω	$\mathsf{ver}_{\mathcal{R}^{\mathsf{ex}}}(\omega)$	$\mathsf{fal}_{\mathcal{R}^{\mathsf{ex}}}(\omega)$	$\mathsf{ME}_{\mathcal{R}^{\mathsf{ex}}}(\omega)$
$da_pa_r\overline{a_o}a_e$	$\{r_1, r_2, r_3\}$	$\emptyset$	0.166	$\bar{d}a_pa_r\overline{a_o}a_e$	$\{r_2, r_3\}$	$\{r_1\}$	0.042
$da_p\overline{a_r}a_oa_e$	$\{r_1, r_2, r_3\}$	$\emptyset$	0.166	$\bar{d}a_p\overline{a_r}a_oa_e$	$\{r_2, r_3\}$	$\{r_1\}$	0.042
$da_p\overline{a_r}a_o\overline{a_e}$	$\{r_1, r_2\}$	$\{r_3\}$	0.018	$\bar{d}a_p\overline{a_r}a_o\overline{a_e}$	$\{r_2\}$	$\{r_1, r_3\}$	0.005
$da_p\overline{a_r}\,\overline{a_o}a_e$	$\{r_1, r_3\}$	$\{r_2\}$	0	$\bar{d}a_p\overline{a_r}\,\overline{a_o}a_e$	$\{r_3\}$	$\{r_1, r_2\}$	0
$d\overline{a_p}a_r\overline{a_o}a_e$	$\{r_2, r_3\}$	$\emptyset$	0.126	$\bar{d}\overline{a_p}a_r\overline{a_o}a_e$	$\{r_2, r_3\}$	$\emptyset$	0.126
$d\overline{a_p}a_r\overline{a_o}\,\overline{a_e}$	$\{r_2\}$	$\{r_3\}$	0.014	$\bar{d}\overline{a_p}a_r\overline{a_o}\,\overline{a_e}$	$\{r_2\}$	$\{r_3\}$	0.014
$d\overline{a_p}\,\overline{a_r}a_oa_e$	$\{r_2, r_3\}$	$\emptyset$	0.126	$\bar{d}\overline{a_p}\,\overline{a_r}a_oa_e$	$\{r_2, r_3\}$	$\emptyset$	0.126
$d\overline{a_p}\,\overline{a_r}a_o\overline{a_e}$	$\{r_2\}$	$\{r_3\}$	0.014	$\bar{d}\overline{a_p}\,\overline{a_r}a_o\overline{a_e}$	$\{r_2\}$	$\{r_3\}$	0.014
$d\overline{a_p}\,\overline{a_r}\,\overline{a_o}a_e$	$\{r_3\}$	$\{r_2\}$	0	$\bar{d}\overline{a_p}\,\overline{a_r}\,\overline{a_o}a_e$	$\{r_3\}$	$\{r_2\}$	0
$d\overline{a_p}\,\overline{a_r}\,\overline{a_o}\,\overline{a_e}$	$\emptyset$	$\{r_2, r_3\}$	0	$\bar{d}\overline{a_p}\,\overline{a_r}\,\overline{a_o}\,\overline{a_e}$	$\emptyset$	$\{r_2, r_3\}$	0

Table 1: MaxEnt model $\mathsf{ME}_{\mathcal{R}^{\mathsf{ex}}}$ of the belief base $\mathcal{R}^{\mathsf{ex}}$ from Ex. 3 as well as the conditional structures of the possible worlds in $\Omega_c(\Sigma_{\cup}^{\mathsf{ex}})$.

Example 5. *In $\mathcal{R}^{\mathsf{ex}}$ from Ex. 3 the formula $a_r \vee a_o$ occurs which states that the company's income is equal to or higher than \$ $10,000,000$. By replacing $a_r \vee a_o$ with the constraint representative a_t which refers to the constraint*

$$t: \qquad x \geq 10,000,000,$$

the belief base $\mathcal{R}^{\mathsf{ex}}$ can be rewritten to $\mathcal{R}^{\mathsf{ex2}} = \{r_1, r_2', r_3\}$ which mentions the conditional $r_2' = (a_t|\top)[1]$ instead of r_2 and is based on the signature $\Sigma_{\cup}^{\mathsf{ex2}} = \{d\} \,\dot{\cup}\, \{a_p, a_t, a_e\}$. The MaxEnt model $\mathsf{ME}_{\mathcal{R}^{\mathsf{ex2}}}$ of $\mathcal{R}^{\mathsf{ex2}}$ differs from $\mathsf{ME}_{\mathcal{R}^{\mathsf{ex}}}$ not only in the number of possible worlds a probability is assigned to (because of the modification of the signature) but also the inferences that can be drawn slightly differ in their probabilities. For example, the conditional $r = (d|\top)[\xi]$ can be MaxEnt inferred from $\mathcal{R}^{\mathsf{ex2}}$ with a probability $\xi = 0.635$. With respect to $\mathcal{R}^{\mathsf{ex}}$, the probability would be different, namely $\xi = 0.63$ (cf. Ex. 4), even though r does not mention any constraint representative.

A conceivable way of determining the set of constraint representatives $\Sigma_{\mathcal{LA}}^{f}$ would be to rely on the signature induced by the belief base $\mathcal{R}$. However, this delegates the responsibility of a proper selection of constraint representatives to the knowledge engineer. If this selection is purely pragmatic and without a profound justification, one probably would still like to become independent of it. Therefore, we show how it is possible to completely abstract from the syntactic representation of admissible variable configurations in conditionals. The basic idea is to aggregate possible worlds to equivalence classes and define probabilistic models wrt. these equivalence classes. Because we define these equivalence classes semantically regarding the evaluation of the conditionals in $\mathcal{R}$, reasoning based on these equivalence classes is independent of the syntax, particularly of the selection of the constraint representatives in $\Sigma_{\mathcal{LA}}^{f}$. Note that our elaborations are related to and generalize the *principle of atomicity* from (Paris 1994) which states that a fragmentation of atoms should not alter the inferences.

Definition 6 (Conditional Structure (Kern-Isberner 2004))**.** *Let $\Sigma_{\cup} = \Sigma \,\dot{\cup}\, \Sigma_{\mathcal{LA}}^{f}$ be a finite set of atoms and constraint representatives, and let $\mathcal{R} \subseteq \mathcal{PCL}(\Sigma_{\cup})$ be a consistent belief base. For a possible world $\omega \in \Omega_c(\Sigma_{\cup})$, we define the sets of conditionals from $\mathcal{R}$ which are* verified *($\mathsf{ver}_{\mathcal{R}}(\omega)$) resp.* falsified *($\mathsf{fal}_{\mathcal{R}}(\omega)$) in ω by*

$$\mathsf{ver}_{\mathcal{R}}(\omega) = \{(B|A)[\xi] \in \mathcal{R} \mid \omega \models AB\},$$
$$\mathsf{fal}_{\mathcal{R}}(\omega) = \{(B|A)[\xi] \in \mathcal{R} \mid \omega \models A\overline{B}\}.$$

The tuple $\sigma_{\mathcal{R}}(\omega) = (\mathsf{ver}_{\mathcal{R}}(\omega), \mathsf{fal}_{\mathcal{R}}(\omega))$ is called the conditional structure *of ω wrt. $\mathcal{R}$.*

Possible worlds with the same conditional structure evaluate the conditionals in $\mathcal{R}$ in the same way and, hence, there is good reason for probabilistic models of $\mathcal{R}$ to assign them the same probability. This principle is called *conditional indifference* and is fulfilled by the MaxEnt model $\mathsf{ME}_{\mathcal{R}}$ (Kern-Isberner 2004).

Example 6. *We recall $\mathcal{R}^{\mathsf{ex}}$ from Ex. 3. The conditional structures of the possible worlds in $\Omega_c(\Sigma_{\cup}^{\mathsf{ex}})$ wrt. $\mathcal{R}^{\mathsf{ex}}$ are shown in Tab. 1. For instance, for $\omega' = \bar{d}a_pa_r\overline{a_o}a_e$ and $\omega'' = \bar{d}a_p\overline{a_r}a_oa_e$ we have $\sigma_{\mathcal{R}^{\mathsf{ex}}}(\omega') = \sigma_{\mathcal{R}^{\mathsf{ex}}}(\omega'') = (\{r_2, r_3\}, \{r_1\})$ because both ω' and ω'' verify r_2 as well as r_3 and falsify r_1. As ω' and ω'' have the same conditional structure, their MaxEnt probabilities are equal.*

Note that conditional structures are defined semantically and induce an equivalence relation on the set of possible worlds as introduced in the following definition.

Definition 7 (Conditional Equivalence (based on (Kern-Isberner 2004)))**.** *Let $\Sigma_{\cup} = \Sigma \,\dot{\cup}\, \Sigma_{\mathcal{LA}}^{f}$ be a finite set of atoms and constraint representatives, and let $\mathcal{R} \subseteq \mathcal{PCL}(\Sigma_{\cup})$ be a consistent belief base. We say that two possible worlds $\omega, \omega' \in \Omega_c(\Sigma_{\cup})$ are* conditionally equivalent *wrt. $\mathcal{R}$, in symbols $\omega \sim_{\mathcal{R}} \omega'$, if they have the same conditional structure, $\sigma_{\mathcal{R}}(\omega) = \sigma_{\mathcal{R}}(\omega')$, and they agree on Σ, i.e., $\omega_{|\Sigma} = \omega'_{|\Sigma}$.*

Conditionally equivalence constitutes an equivalence relation between possible worlds, and we denote the respective equivalence classes with

$$[\omega]_{\mathcal{R}} = \{\omega' \in \Omega_c(\Sigma_{\cup}) \mid \omega' \sim_{\mathcal{R}} \omega\}.$$

The set of all these equivalence classes is

$$\Omega_{\mathcal{R}}^{\sim}(\Sigma_{\cup}) = \{[\omega]_{\mathcal{R}} \mid \omega \in \Omega_c(\Sigma_{\cup})\}.$$

Eventually, because $\sigma_{\mathcal{R}}(\omega) = \sigma_{\mathcal{R}}(\omega')$ for all $\omega' \in [\omega]_{\mathcal{R}}$, we may define $\sigma_{\mathcal{R}}([\omega]_{\mathcal{R}}) = \sigma_{\mathcal{R}}(\omega)$ for $[\omega]_{\mathcal{R}} \in \Omega_{\mathcal{R}}^{\sim}(\Sigma_{\cup})$.

In Def. 7 we differentiate between atoms and constraint representatives and aggregate possible worlds with the same conditional structure to equivalence classes only if they agree on the evaluation of the atoms. Therewith, we maintain atoms as semantic entities but merge constraint representatives.

An equivalence class $[\omega']_{\mathcal{R}}$ can be understood as the interpretation $I \in \mathsf{Int}(\Sigma)$ of the atoms in Σ with $\omega_I = \omega'_{|\Sigma}$, together with the set of admissible variable configurations

$$\mathsf{vconf}([\omega']_{\mathcal{R}}) = \bigcup_{\omega \in [\omega']_{\mathcal{R}}} \mathsf{vconf}(\omega),$$

where $\mathsf{vconf}(\omega) = \mathsf{vconf}(\omega_{|\Sigma^f_{\mathcal{LA}}})$. That is, equivalence classes $[\omega]_{\mathcal{R}}, [\omega']_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)$ with $[\omega]_{\mathcal{R}} \neq [\omega']_{\mathcal{R}}$ differ in their evaluation of the propositional atoms in Σ, $\omega_{|\Sigma} \neq \omega'_{|\Sigma}$, or in their admissible variable configurations, $\mathsf{vconf}([\omega]_{\mathcal{R}}) \neq \mathsf{vconf}([\omega']_{\mathcal{R}})$.

Example 7. *We recall $\mathcal{R}^{\mathsf{ex}}$ from Ex. 3. The equivalence classes in $\Omega^{\sim}_{\mathcal{R}^{\mathsf{ex}}}(\Sigma^{\mathsf{ex}}_\cup)$ are shown in Tab. 2. They aggregate possible worlds in such a way that the two constraints $r\colon x = 10,000,000$ and $o\colon x > 10,000,000$ are combined. For example, we have*

$$\begin{aligned}\mathsf{vconf}([\omega_1]_{\mathcal{R}^{\mathsf{ex}}}) &= \mathsf{Sol}(\{p,\ x = 10,000,000,\ e\}) \\ &\quad \cup \mathsf{Sol}(\{p,\ x > 10,000,000,\ e\}) \\ &= \mathsf{Sol}(\{p,\ x \geq 10,000,000,\ e\}).\end{aligned}$$

The aggregation is possible because the differentiation between the constraints r and o is irrelevant for the evaluation of the conditionals in $\mathcal{R}^{\mathsf{ex}}$. It matches the idea of replacing $a_r \vee a_o$ by a_t as proposed in Ex. 5.

Now we discuss how probabilistic reasoning in general and MaxEnt reasoning in particular works when considering equivalence classes of possible worlds instead of the possible worlds themselves.

Definition 8 (Aggregated (MaxEnt) Model)**.** *Let $\Sigma_\cup$ be a finite set of atoms and constraint representatives, and let $\mathcal{R} \subseteq \mathcal{PCL}(\Sigma_\cup)$ be a consistent belief base. A probability distribution $\mathcal{P}\colon \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup) \to [0,1]$ which assigns a probability to each equivalence class in $\Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)$ is an* aggregated model *of $\mathcal{R}$ if, for every conditional $r = (B|A)[\xi]$ in $\mathcal{R}$,*

$$\sum_{[\omega]_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)\colon\, r \in \mathsf{ver}_{\mathcal{R}}(\omega) \cup \mathsf{fal}_{\mathcal{R}}(\omega)} \mathcal{P}([\omega]_{\mathcal{R}}) > 0$$

and

$$\frac{\sum_{[\omega]_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)\colon\, r \in \mathsf{ver}_{\mathcal{R}}(\omega)} \mathcal{P}([\omega]_{\mathcal{R}})}{\sum_{[\omega]_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)\colon\, r \in \mathsf{ver}_{\mathcal{R}}(\omega) \cup \mathsf{fal}_{\mathcal{R}}(\omega)} \mathcal{P}([\omega]_{\mathcal{R}})} = \xi.$$

If $\mathcal{P}$ is an aggregated model of $\mathcal{R}$, then we write $\mathcal{P} \models^{\sim} \mathcal{R}$.

The aggregated MaxEnt model *of a consistent belief base $\mathcal{R}$ is defined by*

$$\mathsf{ME}^{\sim}_{\mathcal{R}} = \arg\max_{\mathcal{P} \models^{\sim} \mathcal{R}} - \sum_{[\omega]_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)} \mathcal{P}([\omega]_{\mathcal{R}}) \cdot \log \mathcal{P}([\omega]_{\mathcal{R}}).$$

Note that in Def. 8 we do not take the cardinalities of the equivalence classes into account. With the aggregation of possible worlds to equivalence classes and the disregarding of the cardinalities, we abstract from which and how many constraint representatives are used to describe admissible variable configurations. Consequently, for the aggregated MaxEnt model, only the conditional structures of the equivalence classes are relevant and not their compositions.

Proposition 2. *Let $\Sigma_\cup = \Sigma \,\dot\cup\, \Sigma^f_{\mathcal{LA}}$ and $\Sigma'_\cup = \Sigma \,\dot\cup\, \Sigma^{f'}_{\mathcal{LA}}$ be finite sets of atoms and constraint representatives, and let $\mathcal{R} = \{(B_i|A_i)[\xi_i] \mid i = 1,\dots,n\} \subseteq \mathcal{PCL}(\Sigma_\cup)$ and $\mathcal{R}' = \{(B'_i|A'_i)[\xi'_i] \mid i = 1,\dots,m\} \subseteq \mathcal{PCL}(\Sigma'_\cup)$ be consistent belief bases. If $n = m$, $\xi_i = \xi'_i$ for $i = 1,\dots,n$, and there is a bijection $\beta\colon \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup) \to \Omega^{\sim}_{\mathcal{R}'}(\Sigma'_\cup)$ such that $\sigma_{\mathcal{R}'}(\beta([\omega]_{\mathcal{R}})) = \sigma_{\mathcal{R}}([\omega]_{\mathcal{R}})$ for all $[\omega_{\mathcal{R}}] \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)$, then*

$$\mathsf{ME}^{\sim}_{\mathcal{R}}([\omega]_{\mathcal{R}}) = \mathsf{ME}^{\sim}_{\mathcal{R}'}(\beta([\omega]_{\mathcal{R}})).$$

Proof Sketch. Let $\mathcal{R}'' = \{(B_1|A_1)[\xi_1], \dots, (B_n|A_n)[\xi_n]\}$ be a belief base which consists of conditionals without constraint representatives, i.e., $\mathcal{R}'' \subseteq \mathcal{PCL}(\Sigma)$. Then, the MaxEnt model $\mathsf{ME}_{\mathcal{R}''}$ yields a product representation which is of the form (Kern-Isberner 2004)

$$\mathsf{ME}_{\mathcal{R}''}(\omega) = \alpha_0 \prod_{\substack{i=1,\dots,n\colon\\ \omega \models A_i B_i}} \alpha_i^{1-\xi_i} \prod_{\substack{i=1,\dots,n\colon\\ \omega \models A_i \overline{B_i}}} \alpha_i^{-\xi_i}$$

with *effects* $\alpha_0, \alpha_1, \dots, \alpha_n \in \mathbb{R}^{\infty}_{\geq 0}$ for $\omega \in \Omega(\Sigma)$. This product representation carries over to belief bases with constraint representatives and the aggregated MaxEnt model. Moreover, because $|\Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)| = |\Omega^{\sim}_{\mathcal{R}'}(\Sigma'_\cup)|$ and the effects depend on the conditional structures and the probabilities $\xi_1, \dots, \xi_n$ only, we have

$$\mathsf{ME}^{\sim}_{\mathcal{R}}([\omega]_{\mathcal{R}}) = \alpha_0 \prod_{\substack{i=1,\dots,n\colon\\ r_i \in \mathsf{ver}_{\mathcal{R}}(\omega)}} \alpha_i^{1-\xi_i} \prod_{\substack{i=1,\dots,n\colon\\ r_i \in \mathsf{fal}_{\mathcal{R}}(\omega)}} \alpha_i^{-\xi_i}$$

for $[\omega]_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)$ as well as

$$\mathsf{ME}^{\sim}_{\mathcal{R}'}([\omega']_{\mathcal{R}'}) = \alpha_0 \prod_{\substack{i=1,\dots,n\colon\\ r_i \in \mathsf{ver}_{\mathcal{R}'}(\omega')}} \alpha_i^{1-\xi_i} \prod_{\substack{i=1,\dots,n\colon\\ r_i \in \mathsf{fal}_{\mathcal{R}'}(\omega')}} \alpha_i^{-\xi_i}$$

for $[\omega']_{\mathcal{R}'} \in \Omega^{\sim}_{\mathcal{R}'}(\Sigma'_\cup)$ with the same effects. Eventually, because $\sigma_{\mathcal{R}'}([\omega']_{\mathcal{R}'}) = \sigma_{\mathcal{R}}([\omega]_{\mathcal{R}})$ for $[\omega']_{\mathcal{R}'} = \beta([\omega]_{\mathcal{R}})$ and this implies $\mathsf{ver}_{\mathcal{R}'}(\omega') = \mathsf{ver}_{\mathcal{R}}(\omega)$ and $\mathsf{fal}_{\mathcal{R}'}(\omega') = \mathsf{fal}_{\mathcal{R}}(\omega)$, it follows that $\mathsf{ME}^{\sim}_{\mathcal{R}}([\omega]_{\mathcal{R}}) = \mathsf{ME}^{\sim}_{\mathcal{R}'}(\beta([\omega]_{\mathcal{R}}))$ holds. □

Prop. 2 proves that aggregated MaxEnt reasoning is independent of the selection of the constraint representatives in $\Sigma^f_{\mathcal{LA}}$ as long as the selection does not influence the conditional structures and, therewith, the meaning of the belief base.

Example 8. *The aggregated MaxEnt model of the belief base $\mathcal{R}^{\mathsf{ex}}$ from Ex. 3 is shown in Tab. 2. It coincides with the (aggregated) MaxEnt model of $\mathcal{R}^{\mathsf{ex2}}$ (cf. Ex. 5).*

A drawback of aggregated models is that they do not provide a probability assignment to single possible worlds. Next, we tackle this problem and extend the MaxEnt inference relation (1) such that it assigns lower and upper probability bounds to possible worlds. Our extension of (1) is also capable of assigning probabilities to constraints which are not represented in $\Sigma^f_{\mathcal{LA}}$ so that one is not limited to the syntax of the belief base when formulating queries.

	$[\omega]_{\mathcal{R}^{\mathsf{ex}}}$	$\omega_{\mid\Sigma^{\mathsf{ex}}}$	$\mathsf{ver}_{\mathcal{R}^{\mathsf{ex}}}([\omega]_{\mathcal{R}^{\mathsf{ex}}})$	$\mathsf{fal}_{\mathcal{R}^{\mathsf{ex}}}([\omega]_{\mathcal{R}^{\mathsf{ex}}})$	$\mathsf{ME}^{\sim}_{\mathcal{R}^{\mathsf{ex}}}([\omega]_{\mathcal{R}^{\mathsf{ex}}})$
$[\omega_1]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{da_pa_r\overline{a_o}a_e,\ da_p\overline{a_r}a_oa_e\}$	d	$\{r_1, r_2, r_3\}$	$\emptyset$	0.325
$[\omega_2]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{da_p\overline{a_r}a_o\overline{a_e}\}$	d	$\{r_1, r_2\}$	$\{r_3\}$	0.036
$[\omega_3]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{da_p\overline{a_r}\overline{a_o}a_e\}$	d	$\{r_1, r_3\}$	$\{r_2\}$	0
$[\omega_4]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{d\overline{a_p}a_r\overline{a_o}a_e,\ d\overline{a_p}\overline{a_r}a_oa_e\}$	d	$\{r_2, r_3\}$	$\emptyset$	0.247
$[\omega_5]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{d\overline{a_p}a_r\overline{a_o}\overline{a_e},\ d\overline{a_p}\overline{a_r}a_o\overline{a_e}\}$	d	$\{r_2\}$	$\{r_3\}$	0.027
$[\omega_6]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{d\overline{a_p}\overline{a_r}\overline{a_o}a_e\}$	d	$\{r_3\}$	$\{r_2\}$	0
$[\omega_7]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{d\overline{a_p}\overline{a_r}\overline{a_o}\overline{a_e}\}$	d	$\emptyset$	$\{r_2, r_3\}$	0
$[\omega_8]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{\bar{d}a_pa_r\overline{a_o}a_e,\ \bar{d}a_p\overline{a_r}a_oa_e\}$	$\bar{d}$	$\{r_2, r_3\}$	$\{r_1\}$	0.081
$[\omega_9]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{\bar{d}a_p\overline{a_r}a_o\overline{a_e}\}$	$\bar{d}$	$\{r_2\}$	$\{r_1, r_3\}$	0.009
$[\omega_{10}]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{\bar{d}a_p\overline{a_r}\overline{a_o}a_e\}$	$\bar{d}$	$\{r_3\}$	$\{r_1, r_2\}$	0
$[\omega_{11}]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{\bar{d}\overline{a_p}a_r\overline{a_o}a_e,\ \bar{d}\overline{a_p}\overline{a_r}a_oa_e\}$	$\bar{d}$	$\{r_2, r_3\}$	$\emptyset$	0.247
$[\omega_{12}]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{\bar{d}\overline{a_p}a_r\overline{a_o}\overline{a_e},\ \bar{d}\overline{a_p}\overline{a_r}a_o\overline{a_e}\}$	$\bar{d}$	$\{r_2\}$	$\{r_3\}$	0.027
$[\omega_{13}]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{\bar{d}\overline{a_p}\overline{a_r}\overline{a_o}a_e\}$	$\bar{d}$	$\{r_3\}$	$\{r_2\}$	0
$[\omega_{14}]_{\mathcal{R}^{\mathsf{ex}}} =$	$\{\bar{d}\overline{a_p}\overline{a_r}\overline{a_o}\overline{a_e}\}$	$\bar{d}$	$\emptyset$	$\{r_2, r_3\}$	0

Table 2: Equivalence classes of possible worlds in $\Omega^{\sim}_{\mathcal{R}^{\mathsf{ex}}}(\Sigma^{\mathsf{ex}}_{\cup})$ wrt. $\mathcal{R}^{\mathsf{ex}}$ from Ex. 3 and their aggregated MaxEnt probabilities.

6 Drawing Inferences From the Aggregated Maximum Entropy Model

When reason based on the aggregated MaxEnt model $\mathsf{ME}^{\sim}_{\mathcal{R}}$ of a consistent belief base $\mathcal{R} \subseteq \mathcal{PCL}(\Sigma_{\cup})$, one is restricted to ask queries their probability can be calculated from the probability assignment to the equivalence classes of possible worlds in $\Omega^{\sim}_{\mathcal{R}}(\Sigma_{\cup})$. We overcome this restriction by deriving lower and upper bounds for the MaxEnt probabilities of arbitrary possible worlds which are defined wrt. the atoms in Σ and any finite set of constraints representatives $\Sigma^{f'}_{\mathcal{LA}}$. The lower and upper probability bounds have a similar meaning as the beliefs and plausibilities in evidence theory (Dempster 1967). Based on the probability bounds, we extend the MaxEnt inference relation (1) to conditionals which mention arbitrary atoms and constraint representatives. To do so, we have to relate atoms, constraint representatives, and possible worlds to the equivalence classes in $\Omega^{\sim}_{\mathcal{R}}(\Sigma_{\cup})$ which brings us to the notions of *conflict-free* and *refining (equivalence-classes of) possible worlds*.

Definition 9 (Conflict-free and Refining Possible World)**.** *Let $\Sigma_{\cup} = \Sigma \,\dot{\cup}\, \Sigma^{f}_{\mathcal{LA}}$ and $\Sigma'_{\cup} = \Sigma \,\dot{\cup}\, \Sigma^{f'}_{\mathcal{LA}}$ be finite sets of atoms and constraint representatives. For possible worlds $\omega \in \Omega_c(\Sigma_{\cup})$ and $\omega' \in \Omega_c(\Sigma'_{\cup})$ we say that*

- *ω is* conflict-free *with ω' if $\omega_{\mid\Sigma} = \omega'_{\mid\Sigma}$ and*
$$\mathsf{vconf}(\omega) \cap \mathsf{vconf}(\omega') \neq \emptyset, \tag{3}$$
- *ω* refines *ω' if $\omega_{\mid\Sigma} = \omega'_{\mid\Sigma}$ and*
$$\mathsf{vconf}(\omega) \subseteq \mathsf{vconf}(\omega'). \tag{4}$$

Note that both conditions (3) and (4) can be expressed in terms of satisfiability. In particular, condition (4) holds iff $\mathsf{vconf}(\omega) \cap \bigcup_{\omega'' \in \Omega(\Sigma'_{\cup}) \setminus \{\omega'\}} \mathsf{vconf}(\omega'') = \emptyset$.

If $\Sigma_{\cup} = \Sigma'_{\cup}$, the notions of conflict-free and refining possible worlds become trivial.

Proposition 3. *Let $\Sigma_{\cup}$ be a finite set of atoms and constraint representatives, and let $\omega, \omega' \in \Omega_c(\Sigma_{\cup})$. If ω is conflict-free with ω' or ω refines ω', then $\omega = \omega'$.*

Proof. Without loss of generality, let $\Sigma_{\cup} = \Sigma \,\dot{\cup}\, \Sigma^{f}_{\mathcal{LA}}$. In both cases, if ω is conflict-free with ω' or if ω refines ω', $\omega_{\mid\Sigma} = \omega'_{\mid\Sigma}$ holds. It remains to show that ω and ω' agree on $\Sigma^{f}_{\mathcal{LA}}$. If ω is conflict-free with ω', then $\mathsf{vconf}(\omega) \cap \mathsf{vconf}(\omega') \neq \emptyset$. If ω refines ω', then $\mathsf{vconf}(\omega) \subseteq \mathsf{vconf}(\omega')$. In both cases, $\mathsf{vconf}(\omega) \cap \mathsf{vconf}(\omega') = \emptyset$ does not hold. With Prop. 1 it directly follows that $\omega_{\mid\Sigma^{f}_{\mathcal{LA}}} = \omega'_{\mid\Sigma^{f}_{\mathcal{LA}}}$ must be true. Altogether, $\omega = \omega'$ follows. □

The notions of conflict-free and refining possible worlds can be extended to equivalence classes of possible worlds relative to formulas $A \in \mathcal{L}(\Sigma'_{\cup})$ where $\Sigma'_{\cup} = \Sigma \,\dot{\cup}\, \Sigma^{f'}_{\mathcal{LA}}$ involves arbitrary constraint representatives. Because possible worlds, atoms, and also constraint representatives can all be understood as formulas $A \in \mathcal{L}(\Sigma'_{\cup})$, the following definition subsumes all these cases.

Definition 10 (Conflict-free and Refining Equivalence Class of Possible Worlds)**.** *Let $\Sigma_{\cup}$ and $\Sigma'_{\cup}$ as in Def. 9. Further, let $\mathcal{R} \subseteq \mathcal{PCL}(\Sigma_{\cup})$ be a consistent belief base. For $[\omega]_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_{\cup})$ and $A \in \mathcal{L}(\Sigma'_{\cup})$ we say that*

- *$[\omega]_{\mathcal{R}}$ is* conflict-free *with A if there are $\omega' \in [\omega]_{\mathcal{R}}$ and $\omega'' \in \Omega_c(\Sigma'_{\cup})$ with $\omega'' \models A$ such that ω' is conflict-free with ω''.*
- *$[\omega]_{\mathcal{R}}$* refines *A if for all $\omega' \in [\omega]_{\mathcal{R}}$ there is $\omega'' \in \Omega_c(\Sigma'_{\cup})$ with $\omega'' \models A$ such that ω' refines ω''.*

We denote with $\mathsf{cf}_{\mathcal{R}}(A)$ the set of equivalence classes from $\Omega^{\sim}_{\mathcal{R}}(\Sigma_{\cup})$ which are conflict-free with A and with $\mathsf{rf}_{\mathcal{R}}(A)$ the set of equivalence classes which refine A.

In plain words, if $[\omega]_{\mathcal{R}}$ refines A, then A holds in every possible world in $[\omega]_{\mathcal{R}}$, and if $[\omega]_{\mathcal{R}}$ is conflict-free with A, then A cannot be proven wrong based on the possible worlds in $[\omega]_\mathcal{R}$. All remaining equivalence classes from $\Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)$ are in conflict with A. Hence, A should hold at least with the probability of its refining equivalence classes and at most with the probability of its conflict-free equivalence classes. We will formalize this observation in the following. Before, we prove that $\mathsf{rf}_\mathcal{R}(A) \subseteq \mathsf{cf}_\mathcal{R}(A)$ holds such that the equivalence classes in $\mathsf{rf}_\mathcal{R}(A)$ indeed yield a lower probability for A than the equivalence classes in $\mathsf{cf}_\mathcal{R}(A)$.

Proposition 4. *Let $\Sigma_\cup$, $\Sigma'_\cup$, $\mathcal{R}$, and A as in Def. 10.*

1. *One has $\mathsf{rf}_\mathcal{R}(A) \subseteq \mathsf{cf}_\mathcal{R}(A)$.*
2. *If $\Sigma_\cup = \Sigma'_\cup = \Sigma$, i.e., $\Sigma_\cup$ and $\Sigma'_\cup$ are free of constraint representatives, then*

$$\mathsf{rf}_\mathcal{R}(A) = \mathsf{cf}_\mathcal{R}(A) = \{\{\omega\} \mid \omega \in \Omega(\Sigma)\colon \omega \models A\}.$$

Proof. 1. Let $[\omega]_\mathcal{R} \in \mathsf{rf}_\mathcal{R}(A)$. Then, there are $\omega' \in [\omega]_\mathcal{R}$ and $\omega'' \in \Omega_c(\Sigma'_\cup)$ with $\omega'' \models A$ such that ω' refines ω''. That is, $\omega'_{|\Sigma} = \omega''_{|\Sigma}$ and $\mathsf{vconf}(\omega') \subseteq \mathsf{vconf}(\omega'')$ hold from which $\mathsf{conf}(\omega') = \mathsf{vconf}(\omega') \cap \mathsf{vconf}(\omega'')$ follows. Because of $\mathsf{vconf}(\omega') \neq \emptyset$ for possible worlds $\omega' \in \Omega_c(\Sigma_\cup)$, we deduce $\mathsf{vconf}(\omega') \cap \mathsf{vconf}(\omega'') \neq \emptyset$ and with $\omega'_{|\Sigma} = \omega''_{|\Sigma}$ it follows that ω' is conflict-free with ω'' which proves that $[\omega]_\mathcal{R}$ is conflict-free with A.

2. In the absence of constraint representatives, the equivalence classes in $\Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)$ are unit sets because for every two possible worlds $\omega, \omega' \in \Omega_c(\Sigma)$ it is $\omega_{|\Sigma} = \omega'_{|\Sigma}$ if and only if $\omega = \omega'$. It directly follows that also the sets of conflict-free and refining equivalence classes of possible worlds are unit sets and coincide. □

Now we are able to extend the MaxEnt inference relation (1) to arbitrary formulas $A \in \mathcal{L}(\Sigma'_\cup)$ by

$$\mathcal{R} \mathrel{|\!\!\sim^{\sim}_{\mathsf{ME}}} A[l;u] \text{ if } l = \sum_{[\omega]_\mathcal{R} \in \mathsf{rf}_\mathcal{R}(A)} \mathsf{ME}^{\sim}_{\mathcal{R}}([\omega]_\mathcal{R}) \text{ and } u = \sum_{[\omega]_\mathcal{R} \in \mathsf{cf}_\mathcal{R}(A)} \mathsf{ME}^{\sim}_{\mathcal{R}}([\omega]_\mathcal{R}).$$

The relation $|\!\!\sim^{\sim}_{\mathsf{ME}}$ states that based on the belief base $\mathcal{R}$ and according to the MaxEnt principle, A holds at least with l and at most with u, where l is the MaxEnt probability of the equivalence classes of possible worlds which refine A and u is the MaxEnt probability of the equivalence classes which are conflict-free with A. If $l = u$, which holds in the absence of constraint representatives, for instance (cf. Prop. 4(2.)), then we write $A[l]$ instead of $A[l;u]$.

Example 9. *We consider the aggregated MaxEnt model of the belief base $\mathcal{R}^{\mathsf{ex}}$ from Ex. 3 as shown in Tab. 2. The shareholder wants to investigate how likely a dividend payout is, if the expenses do not exceed \$ 13,000,000. Therefore, she formulates the constraint*

$$m\colon \qquad y \leq 13,000,000$$

stating that "the expenses are at most \$ 13,000,000." With respect to the new signature $\Sigma'^{\mathsf{ex}}_\cup = \Sigma^{\mathsf{ex}}_\cup \cup \{a_m\}$, we obtain

$$\mathsf{rf}(da_m) = \{[\omega_1]_{\mathcal{R}^{\mathsf{ex}}},\ [\omega_3]_{\mathcal{R}^{\mathsf{ex}}},\ [\omega_4]_{\mathcal{R}^{\mathsf{ex}}},\ [\omega_6]_{\mathcal{R}^{\mathsf{ex}}}\},$$

basically because $e\colon y \leq 12,000,000$ implies m. Further,

$$\mathsf{cf}(da_m) = \mathsf{rf}(da_m) \cup \{[\omega_2]_{\mathcal{R}^{\mathsf{ex}}},\ [\omega_5]_{\mathcal{R}^{\mathsf{ex}}},\ [\omega_7]_{\mathcal{R}^{\mathsf{ex}}}\},$$

because $\widehat{e}\colon y > 12,000,000$ and m are not in conflict. As a consequence, we can infer

$$\mathcal{R}_{\mathsf{ex}} \mathrel{|\!\!\sim^{\sim}_{\mathsf{ME}}} (da_m)[0.572, 0.635]$$

which means that based on $\mathcal{R}_{\mathsf{ex}}$, it holds with a MaxEnt probability between 0.572 and 0.635 that both the expenses do not exceed \$ 13,000,000 and a dividend is paid. Analogously, we have

$$\mathsf{rf}(\bar{d}a_m) = \{[\omega_8]_{\mathcal{R}^{\mathsf{ex}}},\ [\omega_{10}]_{\mathcal{R}^{\mathsf{ex}}},\ [\omega_{11}]_{\mathcal{R}^{\mathsf{ex}}},\ [\omega_{13}]_{\mathcal{R}^{\mathsf{ex}}}\},$$
$$\mathsf{cf}(\bar{d}a_m) = \mathsf{rf}(da_m) \cup \{[\omega_9]_{\mathcal{R}^{\mathsf{ex}}},\ [\omega_{12}]_{\mathcal{R}^{\mathsf{ex}}},\ [\omega_{14}]_{\mathcal{R}^{\mathsf{ex}}}\},$$

and, thus, $\mathcal{R}_{\mathsf{ex}} \mathrel{|\!\!\sim^{\sim}_{\mathsf{ME}}} (\bar{d}a_m)[0.328, 0.364]$.

In order to generalize the relation $|\!\!\sim^{\sim}_{\mathsf{ME}}$ to conditionals $(B|A) \in \mathcal{PCL}(\Sigma'_\cup)$, we exploit the lower and upper probability bounds of the verification and the falsification of $(B|A)$. Let $\mathcal{R} \mathrel{|\!\!\sim^{\sim}_{\mathsf{ME}}} AB[l_{AB}; u_{AB}]$ and $\mathcal{R} \mathrel{|\!\!\sim^{\sim}_{\mathsf{ME}}} A\overline{B}[l_{A\overline{B}}; u_{A\overline{B}}]$. Then, we define

$$\mathcal{R} \mathrel{|\!\!\sim^{\sim}_{\mathsf{ME}}} (B|A)[l;u] \text{ if } l = \frac{l_{AB}}{l_{AB} + u_{A\overline{B}}} \text{ and } u = \frac{u_{AB}}{u_{AB} + l_{A\overline{B}}}.$$

Example 10. *We refer to Ex. 9 and obtain*

$$\mathcal{R}^{\mathsf{ex}} \mathrel{|\!\!\sim^{\sim}_{\mathsf{ME}}} (d|a_m)[0.611; 0.659].$$

That is, under the shareholder's assumption that the company's expenses stay less than \$ 13,000,000 she should believe in the payout of the dividend with a probability of at least 0.611 and at most 0.659 according to her prior beliefs and the (aggregated) MaxEnt principle.

We finally prove that $|\!\!\sim^{\sim}_{\mathsf{ME}}$ behaves well in the sense of inductive inference relations.

Proposition 5. *1. The relation $|\!\!\sim^{\sim}_{\mathsf{ME}}$ satisfies Direct Inference and Trivial Vacuity.*

2. If $\mathcal{R}$ is a consistent belief base without constraint representatives, i.e., $\mathcal{R} \subseteq \mathcal{PCL}(\Sigma)$, and if $A, B \in \mathcal{L}(\Sigma)$, then

$$\mathcal{R} \mathrel{|\!\!\sim^{\sim}_{\mathsf{ME}}} (B|A)[\xi] \text{ iff } \mathcal{R} \mathrel{|\!\!\sim_{\mathsf{ME}}} (B|A)[\xi],$$

Proof. 1. Let $\Sigma_\cup = \Sigma \,\dot\cup\, \Sigma^f_{\mathcal{LA}}$ be a finite set of atoms and constraint representatives, and let $\mathcal{R} \subseteq \mathcal{PCL}(\Sigma_\cup)$ be a consistent belief base.

Direct Inference: Let $r = (B|A)[\xi]$ be an arbitrary conditional in $\mathcal{R}$. Then, $AB, A\overline{B} \in \mathcal{L}(\Sigma_\cup)$ and, by the definition of aggregated models,

$$\frac{\sum_{[\omega]_\mathcal{R} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)\colon r \in \mathsf{ver}_\mathcal{R}(\omega)} \mathsf{ME}^{\sim}_{\mathcal{R}}([\omega_\mathcal{R}])}{\sum_{[\omega]_\mathcal{R} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)\colon r \in \mathsf{ver}_\mathcal{R}(\omega) \cup \mathsf{fal}_\mathcal{R}(\omega)} \mathsf{ME}^{\sim}_{\mathcal{R}}([\omega_\mathcal{R}])} = \xi.$$

If $[\omega]_\mathcal{R} \in \mathsf{rf}_\mathcal{R}(AB)$, then for all $\omega' \in [\omega]_\mathcal{R}$ there is $\omega'' \in \Omega_c(\Sigma_\cup)$ with $\omega'' \models AB$ such that ω' refines ω''. According to Prop. 3, $\omega' = \omega''$ holds. Hence, for all $\omega' \in [\omega]_\mathcal{R}$ it holds that $\omega' \models AB$, i.e., $r \in \mathsf{ver}_\mathcal{R}(\omega')$. Otherwise, if

$[\omega]_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)$ such that there is $\omega' \in [\omega]_{\mathcal{R}}$ with $r \in \mathsf{ver}_{\mathcal{R}}(\omega')$, then $r \in \mathsf{ver}_{\mathcal{R}}(\omega')$ for all $\omega' \in [\omega]_{\mathcal{R}}$ because of the definition of $[\omega]_{\mathcal{R}}$. Consequently, for all $\omega' \in [\omega]_{\mathcal{R}}$ there is $\omega'' \in \Omega_c(\Sigma_\cup)$ with ω' refines ω'', namely $\omega'' = \omega'$, and

$$\mathsf{rf}_{\mathcal{R}}(AB) = \{[\omega]_{\mathcal{R}} \mid \omega \in \Omega_c(\Sigma_\cup) \colon r \in \mathsf{ver}_{\mathcal{R}}(\omega)\}.$$

If $[\omega]_{\mathcal{R}} \in \mathsf{cf}_{\mathcal{R}}(AB)$ then, there are $\omega' \in [\omega]_{\mathcal{R}}$ and $\omega'' \in \Omega_c(\Sigma_\cup)$ with $\omega'' \models AB$ such that ω' is conflict-free with ω''. With Prop. 3, $\omega' = \omega''$ and, consequently, $\omega' \models AB$ follow, i.e., $r \in \mathsf{ver}_{\mathcal{R}}(\omega')$. Otherwise, if $[\omega]_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)$ such that there is $\omega' \in [\omega]_{\mathcal{R}}$ with $r \in \mathsf{ver}_{\mathcal{R}}(\omega')$, then there is $\omega'' \in \Omega_c(\Sigma_\cup)$ with $\omega'' \models AB$ such that ω' is conflict-free with ω'', namely $\omega'' = \omega'$. Hence,

$$\mathsf{cf}_{\mathcal{R}}(AB) = \{[\omega]_{\mathcal{R}} \mid \omega \in \Omega_c(\Sigma_\cup) \colon r \in \mathsf{ver}_{\mathcal{R}}(\omega)\}$$

holds, too. Analogously, we can show that

$$\mathsf{rf}_{\mathcal{R}}(A\overline{B}) = \mathsf{cf}_{\mathcal{R}}(A\overline{B}) = \{[\omega]_{\mathcal{R}} \mid \omega \in \Omega_c(\Sigma_\cup) \colon r \in \mathsf{fal}_{\mathcal{R}}(\omega)\}.$$

Altogether, we can MaxEnt infer from $\mathcal{R}$:

$$\mathcal{R} \mid\!\sim^{\sim}_{\mathsf{ME}} AB[l_{AB}; u_{AB}] \text{ with } l_{AB} = u_{AB},$$

$$\mathcal{R} \mid\!\sim^{\sim}_{\mathsf{ME}} A\overline{B}[l_{A\overline{B}}; u_{A\overline{B}}] \text{ with } l_{A\overline{B}} = u_{A\overline{B}}.$$

That is, the inferences hold with precise probabilities and we obtain

$$\frac{u_{AB}}{u_{AB} + l_{A\overline{B}}} = \frac{l_{AB}}{l_{AB} + u_{A\overline{B}}} = \xi.$$

Eventually, $\mathcal{R} \mid\!\sim^{\sim}_{\mathsf{ME}} (B|A)[\xi]$ follows.

Trivial Vacuity: Because $\mathcal{R} = \emptyset$, all possible worlds $\omega \in \Omega_c(\Sigma_\cup)$ have the same conditional structure $\sigma_{\mathcal{R}}(\omega) = (\emptyset, \emptyset)$ and the equivalence classes in $\Omega_{\mathcal{R}}(\Sigma_\cup)$ are of the form

$$[\omega]_{\mathcal{R}} = \{\omega' \mid \omega' \in \Omega_c(\Sigma_\cup) \colon \omega'_{|\Sigma} = \omega_{|\Sigma}\}.$$

Further, we have $\mathsf{ME}^{\sim}_{\mathcal{R}}([\omega]_{\mathcal{R}}) = |\Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)|^{-1} > 0$ for all $[\omega]_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)$ because the MaxEnt model is the uniform distribution in the absence of any beliefs. Now, let $\mathcal{R} \mid\!\sim^{\sim}_{\mathsf{ME}} AB[l_{AB}; u_{AB}]$ and $\mathcal{R} \mid\!\sim^{\sim}_{\mathsf{ME}} A\overline{B}[l_{A\overline{B}}; u_{A\overline{B}}]$. Then, we have $\mathcal{R} \mid\!\sim^{\sim}_{\mathsf{ME}} (B|A)[1]$ only if

$$\frac{l_{AB}}{l_{AB} + u_{A\overline{B}}} = \frac{u_{AB}}{u_{AB} + l_{A\overline{B}}} = 1$$

which implies $l_{A\overline{B}} = u_{A\overline{B}} = 0$. In particular, no equivalence class $[\omega]_{\mathcal{R}} \in \Omega^{\sim}_{\mathcal{R}}(\Sigma_\cup)$ is conflict-free with $A\overline{B}$ and, thus, there is no possible world $\omega \in \Omega_c(\Sigma_\cup)$ which is conflict-free with any model of $A\overline{B}$. This, however, can only happen if $A\overline{B}$ has no model which directly implies $A \models B$.

2. The equivalence

$$\mathcal{R} \mid\!\sim^{\sim}_{\mathsf{ME}} (B|A)[\xi] \text{ iff } \mathcal{R} \mid\!\sim_{\mathsf{ME}} (B|A)[\xi]$$

in $\mathcal{PCL}(\Sigma)$ directly follows from the fact that in case of $\Sigma_\cup = \Sigma$ one has $\Omega^{\sim}_{\mathcal{R}}(\Sigma) = \{\{\omega\} \mid \omega \in \Omega(\Sigma)\}$ such that each equivalence class $[\omega]_{\mathcal{R}} = \{\omega\}$ is in one-to-one correspondence with the possible world ω and Def. 8 coincides with the standard definition of probabilistic models. Also the notions of conflict-free and refining possible worlds become trivial (cf. Prop. 4(2.)) and, thus, $l = u$ holds for all inferences $A[l; u]$ drawn from $\mathcal{R}$ with $\mid\!\sim^{\sim}_{\mathsf{ME}}$. □

Prop. 5 proves that $\mid\!\sim^{\sim}_{\mathsf{ME}}$ is an inductive inference relation (1.) and also a proper generalization of $\mid\!\sim_{\mathsf{ME}}$ as it coincides with $\mid\!\sim_{\mathsf{ME}}$ in $\mathcal{PCL}(\Sigma)$ (2.).

7 Discussion and Related Work

Our integration of linear arithmetic constraints into probabilistic conditionals combines qualitative statements in terms of propositional formulas, mathematical calculations in terms of arithmetic constraint sets, and uncertain beliefs in terms of probabilistic conditionals within a single inference formalism. In particular, the approach allows for MaxEnt inferences involving statements over infinite domains ($\mathbb{Z}$ and $\mathbb{R}$) without the need to extend the MaxEnt principle to the infinite case which turned out to be difficult (Barnett and Paris 2008; Paris and Rad 2010; Landes and Williamson 2015; Williamson 2008) and still lacks a generally accepted policy.

The work in (de Salvo Braz et al. 2016) already introduces probabilistic inferences modulo linear arithmetic. However, in (de Salvo Braz et al. 2016) there is a fixed probability distribution given which ignores the influence of the syntactic representation of admissible variable configurations on the inferences. The combination of linear arithmetic and MaxEnt reasoning as well as our subsequent investigation on becoming independent of the syntax is a novel contribution.

Note that our approach is not limited to the MaxEnt principle, although we recommend to use the MaxEnt model for reasoning. The only assumption which we have made and which is necessary to show the irrelevance of the syntactic representation of variable configurations is the principle of conditional indifference which holds for the MaxEnt model but which is not exclusive for it. Actually, it is a desirable property of probabilistic models in general.

8 Conclusions and Future Work

In this paper, we integrated linear arithmetic constraints over the integers and reals into probabilistic conditional belief bases from which we inferred uncertain beliefs based on the principle of maximum entropy (MaxEnt principle). Therewith, it was possible to express uncertain beliefs about admissible configurations of numeric variables where the admissibility was decided based on satisfiability tests in the background theory of linear arithmetic. This implementation of the satisfiability modulo theory (SMT) principle into MaxEnt reasoning allowed us to make qualitative statements about infinite domains ($\mathbb{Z}$ and $\mathbb{R}$) without a need to adapt the MaxEnt model to the infinite case.

A major part of the paper was concerned about the influence of the syntactic representation of variable configurations on the MaxEnt inferences. Based on the notion of conditional structures, we were able to define aggregated models of belief bases as well as a generalized MaxEnt inference relation which both are completely independent of the syntactic representation of the variable configurations.

In future work, we want to further investigate the properties of the generalized MaxEnt inference relation, compare our approach with alternative approaches on MaxEnt reasoning wrt. infinite domains, and we want to apply our approach to broader classes of conditionals like relational probabilistic conditionals and conditionals based on Description Logics. We also want to implement our approach.

Acknowledgements

This work was supported by Grant KE 1413/14-1 of the German Research Foundation (DFG).

References

Barnett, O., and Paris, J. B. 2008. Maximum entropy inference with quantified knowledge. *Log. J. IGPL* 16(1):85–98.

Barrett, C. W.; Sebastiani, R.; Seshia, S. A.; and Tinelli, C. 2021. Satisfiability modulo theories. In Biere, A.; Heule, M.; van Maaren, H.; and Walsh, T., eds., *Handbook of Satisfiability - Second Edition*, volume 336 of *Frontiers in Artificial Intelligence and Applications*. IOS Press. 1267–1329.

de Salvo Braz, R.; O'Reilly, C.; Gogate, V.; and Dechter, R. 2016. Probabilistic inference modulo theories. In Kambhampati, S., ed., *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*, 3591–3599. IJCAI/AAAI Press.

Dempster, A. P. 1967. Upper and Lower Probabilities Induced by a Multivalued Mapping. *The Annals of Mathematical Statistics* 38(2):325 – 339.

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

Karmarkar, N. 1984. A new polynomial-time algorithm for linear programming. *Comb.* 4(4):373–396.

Kern-Isberner, G.; Beierle, C.; and Brewka, G. 2020. Syntax splitting = relevance + independence: New postulates for nonmonotonic reasoning from conditional belief bases. In Calvanese, D.; Erdem, E.; and Thielscher, M., eds., *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020, Rhodes, Greece, September 12-18, 2020*, 560–571.

Kern-Isberner, G. 2004. A thorough axiomatization of a principle of conditional preservation in belief revision. *Ann. Math. Artif. Intell.* 40(1-2):127–164.

Landes, J., and Williamson, J. 2015. Justifying objective bayesianism on predicate languages. *Entropy* 17(4):2459–2543.

Papadimitriou, C. H. 1981. On the complexity of integer programming. *J. ACM* 28(4):765–768.

Paris, J. B., and Rad, S. R. 2010. A note on the least informative model of a theory. In Ferreira, F.; Löwe, B.; Mayordomo, E.; and Gomes, L. M., eds., *Programs, Proofs, Processes, 6th Conference on Computability in Europe, CiE 2010, Ponta Delgada, Azores, Portugal, June 30 - July 4, 2010. Proceedings*, volume 6158 of *Lecture Notes in Computer Science*, 342–351. Springer.

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

Paris, J. B. 1998. Common sense and maximum entropy. *Synth.* 117(1):75–93.

Presburger, M. 1929. über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. In *Comptes Rendus du I congrès de Mathématiciens des Pays Slaves*, 92–101.

Rödder, W., and Meyer, C. 1996. Coherent knowledge processing at maximum entropy by SPIRIT. In Horvitz, E., and Jensen, F. V., eds., *UAI '96: Proceedings of the Twelfth Annual Conference on Uncertainty in Artificial Intelligence, Reed College, Portland, Oregon, USA, August 1-4, 1996*, 470–476. Morgan Kaufmann.

Thimm, M., and Kern-Isberner, G. 2012. On probabilistic inference in relational conditional logics. *Log. J. IGPL* 20(5):872–908.

Wilhelm, M.; Howey, D.; Kern-Isberner, G.; Sauerwald, K.; and Beierle, C. 2022. Integrating cognitive principles from ACT-R into probabilistic conditional reasoning by taking the example of maximum entropy reasoning. In Barták, R.; Keshtkar, F.; and Franklin, M., eds., *Proceedings of the Thirty-Fifth International Florida Artificial Intelligence Research Society Conference, FLAIRS 2022, Hutchinson Island, Jensen Beach, Florida, USA, May 15-18, 2022*.

Williamson, J. 2008. Objective bayesian probabilistic logic. *J. Algorithms* 63(4):167–183.