Consolidation of Probabilistic Knowledge Bases by Inconsistency Minimization

Nico Potyka[1] and **Matthias Thimm**[2]

Abstract.   Consolidation describes the operation of restoring consistency in an inconsistent knowledge base. Here we consider this problem in the context of probabilistic conditional logic, a language that focuses on *probabilistic conditionals* (if-then rules). If a knowledge base, i. e., a set of probabilistic conditionals, is inconsistent traditional model-based inference techniques are not applicable. In this paper, we develop an approach to repair such knowledge bases that relies on a generalized notion of a model of a knowledge base that extends to classically inconsistent knowledge bases. We define a generalized approach to reasoning under maximum entropy on these generalized models and use it to repair the knowledge base. This approach is founded on previous work on inconsistency measures and we show that it is well-defined, provides a unique solution, and satisfies other desirable properties.

1   Introduction

Setting up a knowledge base for e. g. an expert system is usually a distributed task that involves merging information from different sources. In this process, inconsistencies easily arise as different experts may have different opinions or beliefs about their field of expertise. Although these inconsistencies often affect only a small portion of the joint knowledge base or emerge from only small differences in the experts' beliefs, they cause severe damage. Therefore, reasoning under inconsistency is an important field in knowledge representation and reasoning and there are many approaches that deal with this issue such as *paraconsistent* and *default logics* [14], *belief revision* and *information fusion* [4]. Here, we employ probabilistic conditional logic [7] for knowledge representation. The basic notion of probabilistic conditional logic is that of a *probabilistic conditional* which has the form $(\psi \mid \phi)[d]$ with the commonsense meaning "if ϕ is true then ψ is true with probability d". A popular choice for reasoning with sets of probabilistic conditionals is model-based inductive reasoning based on the principle of maximum entropy (ME-reasoning) [11, 7].

In this paper, we consider the problem of *consolidation* [4] in probabilistic conditional logic. Consolidation is the operation of minimally changing an inconsistent knowledge base (or belief set in a wider context) in order to restore consistency. Consolidation operators can (among others) be used to realize operators for merging by applying the consolidation operator to the join of a set of knowledge bases. For classical logics, this is usually handled by *removing* some minimal set of formulas from the knowledge base such that the remaining set is consistent, cf. [5]. In a probabilistic environment, there is another possibility for achieving consistency besides

removal of probabilistic conditionals, namely, *modification of probabilities*. More specifically, given an inconsistent knowledge base $\mathcal{K} = \{(\psi_1 \mid \phi_1)[d_1], \dots, (\psi_m \mid \phi_m)[d_m]\}$ we aim at finding a consistent knowledge base $\mathcal{K}' = \{(\psi_1 \mid \phi_1)[d'_1], \dots, (\psi_m \mid \phi_m)[d'_m]\}$ that is *qualitatively* the same as $\mathcal{K}$ and is *closest* to $\mathcal{K}$ given some notion of distance on the probabilities $d_1, \dots, d_n$. Here, we build on work on inconsistency measurement [17, 13] to define a consolidation operator. For this purpose, we generalize the notion of a model of a knowledge base by considering those probability functions that are as close as possible to satisfying a knowledge base in the classical sense. We use these generalized models to define a generalized version of ME-reasoning that is equivalent to classical ME-reasoning for consistent knowledge bases but uses the generalized models in the case of inconsistent knowledge bases. We then define a consolidation operator that modifies the original probabilities of the conditionals by taking the probabilities suggested by the generalized ME-reasoning approach. In summary, the contributions of this paper are as follows:

1. We formally introduce the problem of consolidation for probabilistic conditional logic and adapt desirable properties for consolidation operators from the literature on belief merging (Section 3).
2. We solve the problem of consolidation by exploiting previous work on inconsistency measurement and introducing an well-defined consolidation operator (Section 4). In particular:

   (a) we generalize the notion of a model and extend ME-reasoning to include inconsistent knowledge bases by considering probability functions that minimize inconsistency (Section 4.1).

   (b) we use generalized ME-reasoning for consolidating a knowledge base and show that it complies with most desirable properties for consolidation operators (Section 4.2).

   (c) we show that our approach has the same asymptotic worst-case complexity like classical ME-reasoning and provide a problem transformation that can be used to solve the problem more efficiently for certain distance measures (Section 4.3).

Proofs of technical results and links to the implementation can be found in an online appendix[3].

2   Probabilistic Conditional Logic

Let At be a propositional signature, i. e. a finite set of propositional atoms. Let $\mathcal{L}(\mathsf{At})$ be the corresponding propositional language generated by the atoms in At and the connectives $\wedge$ (*and*), $\vee$ (*or*), and $\neg$ (*negation*). For $\phi, \psi \in \mathcal{L}(\mathsf{At})$ we abbreviate $\phi \wedge \psi$ by $\phi\psi$ and $\neg\phi$ by $\overline{\phi}$. The symbols $\top$ and $\bot$ denote *tautology* and *contradiction*, respectively. We use *possible worlds* for interpreting sentences

---

[1] FernUniversität in Hagen, Germany, nico.potyka@FernUni-Hagen.de
[2] University of Koblenz-Landau, Germany, thimm@uni-koblenz.de

in $\mathcal{L}(\mathsf{At})$. A possible world ω is a complete conjunction, i.e. a conjunction that contains for each $a \in \mathsf{At}$ either a or $\neg a$. Let $\Omega(\mathsf{At})$ denote the set of all possible worlds. A possible world $\omega \in \Omega(\mathsf{At})$ *satisfies* an atom $a \in \mathsf{At}$, denoted by $\omega \models a$ if and only if a appears in ω. The entailment relation $\models$ is extended to arbitrary formulas in $\mathcal{L}(\mathsf{At})$ in the usual way. Formulas $\psi, \phi \in \mathcal{L}(\mathsf{At})$ are *equivalent*, denoted by $\phi \equiv \psi$, if and only if $\omega \models \phi$ whenever $\omega \models \psi$ for every $\omega \in \Omega(\mathsf{At})$.

The central notion of probabilistic conditional logic [7] is that of a *probabilistic conditional*.

Definition 1. If $\phi, \psi \in \mathcal{L}(\mathsf{At})$ and $d \in [0, 1]$ then $(\psi \,|\, \phi)[d]$ is called a *probabilistic conditional*.

A probabilistic conditional $c = (\psi \,|\, \phi)[d]$ is meant to describe a probabilistic *if-then* rule, i.e. the informal interpretation of c is that "If ϕ is true then ψ is true with probability d" (see below). If $\phi \equiv \top$ we abbreviate $(\psi \,|\, \phi)[d]$ by $(\psi)[d]$. Further, for $c = (\psi \,|\, \phi)[d]$ we denote with $\mathsf{pr}(c) = d$ the *probability* of c. Let $(\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}$ denote the set of all probabilistic conditionals wrt. $\mathcal{L}(\mathsf{At})$.

Definition 2. A *knowledge base* $\mathcal{K}$ is an ordered finite subset of $(\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}$, i.e., it is $\mathcal{K} = \langle c_1, \ldots, c_m \rangle$ for some $c_1, \ldots, c_n \in (\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}$.

We impose an ordering on the conditionals in a knowledge base $\mathcal{K}$ only for technical convenience. The order can be arbitrary and has no further meaning other than to enumerate the conditionals of a knowledge base in an unambiguous way. For knowledge bases $\mathcal{K} = \langle c_1, \ldots, c_{m_1} \rangle$, $\mathcal{K}' = \langle c'_1, \ldots, c'_{m_2} \rangle$ and a probabilistic conditional c we define $c \in \mathcal{K}$ via $c \in \{c_1, \ldots, c_{m_1}\}$, $\mathcal{K} \subseteq \mathcal{K}'$ via $\{c_1, \ldots, c_{m_1}\} \subseteq \{c'_1, \ldots, c'_{m_2}\}$, and $\mathcal{K} = \mathcal{K}'$ via $\{c_1, \ldots, c_{m_1}\} = \{c'_1, \ldots, c'_{m_2}\}$. The union of knowledge bases is defined via concatenation and removal of duplicates. Let $\mathsf{At}(\mathcal{K})$ denote the set of propositional atoms appearing in $\mathcal{K}$.

Semantics are given to probabilistic conditionals by *probability functions* on $\Omega(\mathsf{At})$. Let $\mathcal{P}(\mathsf{At})$ denote the set of all probability functions $P : \Omega(\mathsf{At}) \to [0, 1]$ with $\sum_{\omega \in \Omega(\mathsf{At})} P(\omega) = 1$. For $\phi \in \mathcal{L}(\mathsf{At})$, we define its probability as the probability of the satisfying worlds, i.e., $P(\phi) = \sum_{\omega \models \phi} P(\omega)$. If $P \in \mathcal{P}(\mathsf{At})$ then P satisfies a probabilistic conditional $(\psi \,|\, \phi)[d]$, denoted by $P \models^{pr} (\psi \,|\, \phi)[d]$, if and only if $P(\psi\phi) = dP(\phi)$. Note that we do not define probabilistic satisfaction via $P(\psi \,|\, \phi) = P(\psi\phi)/P(\phi) = d$ in order to avoid a case differentiation for $P(\phi) = 0$, see [11] for further justification. A probability function P satisfies a knowledge base $\mathcal{K}$ (or is a model of $\mathcal{K}$), denoted by $P \models^{pr} \mathcal{K}$, if and only if $P \models^{pr} c$ for every $c \in \mathcal{K}$. Let $\mathsf{Mod}(\mathcal{K}) \subseteq \mathcal{P}(\mathsf{At})$ be the set of models of $\mathcal{K}$. If $\mathsf{Mod}(\mathcal{K}) = \emptyset$ then $\mathcal{K}$ is *inconsistent*.

A probabilistic conditional $(\psi \,|\, \phi)[d]$ is *normal* [17] if and only if there are $\omega, \omega' \in \Omega(\mathsf{At})$ with $\omega \models \psi\phi$ and $\omega' \models \overline{\psi}\phi$. In other words, a probabilistic conditional c is normal if it is satisfiable but not tautological.

Example 1. The probabilistic conditionals $c_1 = (\top \,|\, a)[1]$ and $c_2 = (\overline{a} \,|\, a)[0.1]$ are *not* normal as c_1 is tautological (there is no $\omega \in \Omega(\mathsf{At})$ with $\omega \models \overline{\top}a$ as $\overline{\top}a \equiv \bot$) and c_2 is not satisfiable (there is no $\omega \in \Omega(\mathsf{At})$ with $\omega \models \overline{a}a$ as $\overline{a}a \equiv \bot$).

As a technical convenience, we consider only normal probabilistic conditionals here, so let $\mathbb{K}$ be the set of all non-empty knowledge bases of $(\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}$ that contain only normal probabilistic conditionals.

Knowledge bases $\mathcal{K}_1, \mathcal{K}_2$ are *extensionally equivalent*, denoted by $\mathcal{K}_1 \equiv^e \mathcal{K}_2$, if and only if $\mathsf{Mod}(\mathcal{K}_1) = \mathsf{Mod}(\mathcal{K}_2)$. Note that the

notion of extensional equivalence does not distinguish between inconsistent knowledge bases, i.e., for inconsistent $\mathcal{K}_1$ and $\mathcal{K}_2$ it always holds that $\mathcal{K}_1 \equiv^e \mathcal{K}_2$. Consequently, we also consider another equivalence relation for knowledge bases. Knowledge bases $\mathcal{K}_1, \mathcal{K}_2$ are *semi-extensionally equivalent*, denoted by $\mathcal{K}_1 \equiv^s \mathcal{K}_2$, if and only if there is a bijection $\sigma_{\mathcal{K}_1, \mathcal{K}_2} : \mathcal{K}_1 \to \mathcal{K}_2$ such that $\{c\} \equiv^e \{\sigma_{\mathcal{K}_1, \mathcal{K}_2}(c)\}$ for every $c \in \mathcal{K}_1$. Note that $\mathcal{K}_1 \equiv^s \mathcal{K}_2$ implies $\mathcal{K}_1 \equiv^e \mathcal{K}_2$ but the other direction is not true in general.

One way of reasoning with knowledge bases is by using model-based inductive reasoning techniques [11]. For example, reasoning based on the *principle of maximum entropy* selects among the models of a knowledge base $\mathcal{K}$ the unique probability function with maximum entropy. More formally, let the *entropy* $H(P)$ of a probability function $P \in \mathcal{P}(\mathsf{At})$ be defined as

$$H(P) = - \sum_{\omega \in \Omega} P(\omega) \log P(\omega)$$

Then the **ME**-model $\mathsf{ME}(\mathcal{K})$ of a consistent knowledge base $\mathcal{K}$ is defined as

$$\mathsf{ME}(\mathcal{K}) = \arg \max_{P \in \mathsf{Mod}(\mathcal{K})} H(P) \qquad (1)$$

Note that the ME-model $\mathsf{ME}(\mathcal{K})$ of a consistent knowledge base $\mathcal{K}$ always exists, is uniquely defined, and satisfies many commonsense reasoning properties [11, 7]. However, a necessary requirement for the application of model-based inductive reasoning techniques is the existence of at least one model of a knowledge base. In order to reason with inconsistent knowledge bases the inconsistency has to be resolved first.

3 Principles for Knowledge Base Consolidation

A *consolidation operator* [4] Γ is a function that maps a possibly inconsistent knowledge base $\mathcal{K}$ to a consistent knowledge base $\mathcal{K}' = \Gamma(\mathcal{K})$. In general, there are three basic approaches (and combinations thereof) for restoring consistency in probabilistic conditional logic, cf. [3]. First, one can remove conditionals such that $\Gamma(\mathcal{K}) \subseteq \mathcal{K}$. Second, one can modify the qualitative structure of conditionals, i.e., a conditional $(\psi \,|\, \phi)[d]$ is modified to $(\psi' \,|\, \phi')[d]$. Thirdly, one can modify the quantitative part of conditionals, i.e., a conditional $(\psi \,|\, \phi)[d]$ is modified to $(\psi \,|\, \phi)[d']$. Here, we follow the third paradigm and assume that for a consolidation function Γ and a knowledge base $\mathcal{K} = \langle (\psi_1 \,|\, \phi_1)[d_1], \ldots, (\psi_n \,|\, \phi_n)[d_n] \rangle$ it holds that $\Gamma(\mathcal{K}) = \langle (\psi_1 \,|\, \phi_1)[d'_1], \ldots, (\psi_n \,|\, \phi_n)[d'_n] \rangle$. Pursuing this approach is valid as there are always $d'_1, \ldots, d'_n$ such that $\langle (\psi_1 \,|\, \phi_1)[d'_1], \ldots, (\psi_n \,|\, \phi_n)[d'_n] \rangle$ is consistent [17]. These consolidation functions have the advantage of allowing a graded consolidation of a knowledge base, as opposed to the other two variants which can only change a knowledge base in a qualitative way. We will call this type of consolidation functions *quantitative consolidation functions*.

In the following, we present some principles that should be satisfied by a meaningful consolidation operator Γ. For that, we need some further notation. Let $\mathcal{K}$ be a knowledge base. If $\vec{x} \in [0, 1]^{|\mathcal{K}|}$ we denote by $\mathcal{K}[\vec{x}]$ the knowledge base that is obtained from $\mathcal{K}$ by replacing the probabilities of the conditionals in $\mathcal{K}$ by the values in $\vec{x}$, respectively. More precisely, if $\mathcal{K} = \langle (\psi_1 \,|\, \phi_1)[d_1], \ldots, (\psi_m \,|\, \phi_m)[d_m] \rangle$ then $\mathcal{K}[\vec{x}] = \langle (\psi_1 \,|\, \phi_1)[x_1], \ldots, (\psi_m \,|\, \phi_m)[x_m] \rangle$ for $\vec{x} = \langle x_1, \ldots, x_m \rangle \in [0, 1]^n$. Similarly, for a single probabilistic conditional $c = (\psi \,|\, \phi)[d]$ and $x \in [0, 1]$ we abbreviate $c[x] = (\psi \,|\, \phi)[x]$.

Knowledge bases $\mathcal{K}_1, \mathcal{K}_2$ are *qualitatively equivalent*, denoted by $\mathcal{K}_1 \cong^q \mathcal{K}_2$, if and only if $|\mathcal{K}_1| = |\mathcal{K}_2|$ and there is a $\vec{x} \in [0,1]^{|\mathcal{K}_1|}$ such that $\mathcal{K}_1 = \mathcal{K}_2[\vec{x}]$. Note that $\cong^q$ is an equivalence relation.

Definition 3. Let $\mathcal{K} = \mathcal{K}[\vec{x}]$ be a knowledge base. Let $\vec{y}, \vec{z} \in [0,1]^{|\mathcal{K}|}$ and let $\mathcal{K}_1 = \mathcal{K}[\vec{y}], \mathcal{K}_2 = \mathcal{K}[\vec{z}]$. Then $\mathcal{K}_1 \prec_{\mathcal{K}} \mathcal{K}_2$ if and only if for $1 \leq i \leq |\mathcal{K}|$ it holds that $|\vec{x}_i - \vec{y}_i| \leq |\vec{x}_i - \vec{z}_i|$ and for at least one i it holds that $|\vec{x}_i - \vec{y}_i| < |\vec{x}_i - \vec{z}_i|$.

The relation $\prec_{\mathcal{K}}$ is a partial order among qualitatively equivalent knowledge bases wrt. their overall *distance* to the knowledge base $\mathcal{K}$. In other words, it holds that $\mathcal{K}_1 \prec_{\mathcal{K}} \mathcal{K}_2$ if and only if the probability of each conditional in $\mathcal{K}_1$ is as least as close to the probability of the corresponding conditional in $\mathcal{K}$ as the corresponding conditional in $\mathcal{K}_2$ and there is at least one conditional in $\mathcal{K}_1$ with a probability strictly closer to the probability of the corresponding conditional in $\mathcal{K}$ as the corresponding conditional in $\mathcal{K}_2$.

We will consider some rationality postulates for consolidation operators from the field of belief merging. The following postulates are partially rephrased postulates from [6] and [9]. Let $\mathcal{K}, \mathcal{K}_1, \mathcal{K}_2 \in \mathbb{K}$.

Success. $\Gamma(\mathcal{K})$ is consistent.
Consistency. If $\mathcal{K}$ is consistent then $\Gamma(\mathcal{K}) = \mathcal{K}$.
Irrelevance of syntax. If $\mathcal{K}_1 \equiv^s \mathcal{K}_2$ then $\Gamma(\mathcal{K}_1) \equiv^s \Gamma(\mathcal{K}_2)$.
Non-dictatorship. If c is non-tautological then there is a $\mathcal{K}$ with $c \in \mathcal{K}$ such that $c \notin \Gamma(\mathcal{K})$.
Pareto-optimality. There is no consistent $\mathcal{K}'$ with $\mathcal{K}' \prec_{\mathcal{K}} \Gamma(\mathcal{K})$.
Weak IIA. If $\mathsf{At}(\mathcal{K}_1) \cap \mathsf{At}(\mathcal{K}_2) = \emptyset$ then $\Gamma(\mathcal{K}_1 \cup \mathcal{K}_2) \equiv^e \Gamma(\mathcal{K}_1) \cup \Gamma(\mathcal{K}_2)$.
IIA. If $\Gamma(\mathcal{K}_1) \cup \Gamma(\mathcal{K}_2)$ is consistent then $\Gamma(\mathcal{K}_1) \cup \Gamma(\mathcal{K}_2) \equiv^e \Gamma(\mathcal{K}_1 \cup \mathcal{K}_2)$.
Continuity. For each sequence $(\mathcal{K}[\vec{x}_n])_{n \in \mathbb{N}}$ of knowledge bases with $\lim_{n \to \infty} \mathcal{K}[\vec{x}_n] = \mathcal{K}[\vec{x}]$, it holds that $\lim_{n \to \infty} \Gamma(\mathcal{K}[\vec{x}_n]) = \Gamma(\mathcal{K}[\vec{x}])$.

The property *success* describes our basic demand for a consolidation function, i.e., that the result of the consolidation is consistent. The property *consistency* says that an already consistent knowledge base needs no modification. The property *irrelevance of syntax* demands that restoring consistency of semi-extensionally equivalent knowledge bases yield again semi-extensionally equivalent knowledge bases. Demanding *non-dictatorship* implies that there is no non-tautological probabilistic conditional that is never modified in any knowledge base. The property *Pareto-optimality* implements the *minimal change paradigm*: from all solutions to the consolidation problem the result should be as close to the original knowledge base as possible. The properties *Weak IIA* and *IIA* realize different views on the property *indifference of irrelevant alternatives* [6]. The property *Weak IIA* demands that for knowledge bases $\mathcal{K}_1$ and $\mathcal{K}_2$ that represent information about different topics, i.e., that do not share some proposition, the consolidation $\Gamma(\mathcal{K}_1 \cup \mathcal{K}_2)$ should be the same as $\Gamma(\mathcal{K}_1) \cup \Gamma(\mathcal{K}_2)$. The property *IIA* demands the same conclusion given that $\Gamma(\mathcal{K}_1) \cup \Gamma(\mathcal{K}_2)$ is consistent. The final property *continuity* demands that the consolidation function behaves continuously on changes of probabilities. Intuitively, this means that for $\mathcal{K}_1$ and $\mathcal{K}_2$ with $\mathcal{K}_1 \cong^q \mathcal{K}_2$ and $\mathcal{K}_1$ is close to $\mathcal{K}_2$ wrt. $\prec_{\mathcal{K}}$ then $\Gamma(\mathcal{K}_1)$ is also close to $\Gamma(\mathcal{K}_2)$ wrt. $\prec_{\mathcal{K}}$. Consider also the following relationship between the different notions of *indifference of irrelevant alternatives*.

Proposition 1. *Let* Γ *satisfy* success. *If* Γ *satisfies* IIA *then* Γ *satisfies* Weak IIA.

4 Probabilistic Knowledge Base Consolidation

We now present our solution to the problem of probabilistic knowledge base consolidation. The core idea of our approach relies on the use of inconsistency measures for probabilistic conditional logic [13, 17, 12] and the generalization of the ME-model to inconsistent knowledge bases. An inconsistency measure $\mathcal{I}$ is a function $\mathcal{I} : \mathbb{K} \to [0, \infty)$ that assigns to a knowledge base $\mathcal{K}$ a value $\mathcal{I}(\mathcal{K})$ with the intended meaning that the larger the value $\mathcal{I}(\mathcal{K})$ the larger the inconsistency in $\mathcal{K}$, with $\mathcal{I}(\mathcal{K}) = 0$ meaning that $\mathcal{K}$ is consistent. Inconsistency measures for probabilistic logics, such as the ones defined in [13, 17, 12], usually rely on the idea to measure the minimal changes needed to make an inconsistent knowledge base consistent. We exploit this idea here by considering those probability functions that are used for obtaining this minimal change as generalized models of the inconsistent knowledge base and use them to define the consolidated knowledge base by a generalized version of the ME-model.

4.1 Generalized ME-Reasoning

For defining generalized ME-Reasoning and thus our consolidation operator we use the minimal violation inconsistency measure proposed in [13] for two reasons. First, this measure is computationally attractive as it relies on solving convex optimization problems instead of non-convex ones as the measures proposed in [17, 12]. Second, this measure allows to uniquely define a consolidated knowledge base in an information-theoretic appealing way due to the properties of its solution space.

The minimal violation inconsistency measure is defined as follows. Recall that a probability function P satisfies a probabilistic conditional $(\psi \mid \phi)[d]$ if and only if $P(\psi\phi) = dP(\phi)$ which is equivalent to $P(\psi\phi) - dP(\phi) = 0$. Observe that if a knowledge base $\mathcal{K} = \langle (\psi_1 \mid \phi_1)[d_1], \ldots, (\psi_m \mid \phi_m)[d_m] \rangle$ is inconsistent there is no probability function P with $P(\psi_i\phi_i) - d_iP(\phi_i) = 0$ for all $i = 1, \ldots, m$. For each conditional $(\psi_i \mid \phi_i)[d_i]$ $(i = 1, \ldots, m)$ we introduce a variable x_i and set $P(\psi_i\phi_i) - d_iP(\phi_i) = x_i$. Given some real vector norm $\| \cdot \|$ we define the *distance* of a probability function P to a knowledge base $\mathcal{K}$ wrt. $\| \cdot \|$ as $\|\langle x_1, \ldots, x_n \rangle\|$. For the minimal violation inconsistency measure $\mathcal{I}_{\Pi}^p$ (for $p \geq 1$) we use the p-norm $\| \cdot \|_p$ for this purpose which is defined as

$$\|(x_1, \ldots, x_m)\|_p = \sqrt[p]{\sum_{j=1}^{m} |x_j|^p}$$

For $p \to \infty$ (we also write $p = \infty$) we obtain the maximum-norm:

$$\lim_{p \to \infty} \|(x_1, \ldots, x_m)\|_p = \|(x_1, \ldots, x_m)\|_\infty$$
$$= \max\{|x_1|, \ldots, |x_m|\}$$

Now, $\mathcal{I}_{\Pi}^p(\mathcal{K})$ for some knowledge base $\mathcal{K}$ is defined as the minimal distance among all probability functions. More formally, let $\mathcal{K} = \langle (\psi_1 \mid \phi_1)[d_1], \ldots, (\psi_m \mid \phi_m)[d_m] \rangle$ be a knowledge base and assume some canonical enumeration of the possible worlds of our language, i.e., $\Omega(\mathsf{At}) = \{\omega_1, \ldots, \omega_n\}$ with $|\Omega(\mathsf{At})| = n$. Let $\vec{x} = \langle x_1, \ldots, x_m \rangle$ be a vector of variables and consider

$$P(\psi_i\phi_i) - d_iP(\phi_i) = x_i$$

which is equivalent to

$$\sum_{j=1}^{n} P(\omega_j)(1_{\phi_i\psi_i}(\omega_j)(1 - d_i) - 1_{\overline{\psi_i}\phi_i}(\omega_j)d_i) = x_i \qquad (2)$$

where for a formula F the indicator function $1_F(\omega)$ maps to 1 iff $\omega \models F$ and to 0 otherwise. Note that (2) is a linear equation and the coefficients of $P(\omega_j)$ are fixed by the knowledge base $\mathcal{K}$. We can therefore write the set of equations (2) for $i = 1, \ldots, m$ as $A_\mathcal{K} P = \vec{x}$ where $A_\mathcal{K} = (a)_{ij} \in \mathbb{R}^{m \times n}$ is the *characteristic matrix* of $\mathcal{K}$ with

$$a_{ij} = (1_{\phi_i \psi_i}(\omega_j)(1 - d_i) - 1_{\overline{\psi_i} \phi_i}(\omega_j) d_i)$$

Then we define $\mathcal{I}_\Pi^p(\mathcal{K})$ for $p \geq 1$ through

$$\mathcal{I}_\Pi^p(\mathcal{K}) = \min\{\|\vec{x}\|_p \mid A_\mathcal{K} P = \vec{x} \text{ for some } P \in \mathcal{P}(\mathsf{At})\}$$

The measure $\mathcal{I}_\Pi^p$ satisfies a series of commonsense properties desirable for inconsistency measures and has been thoroughly investigated in [13]. The choice of the actual $p \geq 1$ influences how the *violation* of a probability function wrt. particular probabilistic conditionals is distributed. Note that the i-th component in the vector $A_\mathcal{K} P$ corresponds to the deviation of the i-th probabilistic conditional in $\mathcal{K}$ from 0 (its violation by P). For $p = 1$, a lower total violation might be obtained, but the violation of some conditionals can be rather extreme. As p grows, higher violations are penalized more heavily and the violation can be expected to be more distributed among the conditionals.

We use $\mathcal{I}_\Pi^p$ to define a generalized notion of a model of a knowledge base by considering those probability functions as generalized models that minimize the overall violation.

Definition 4. Let $\mathcal{K}$ be a knowledge base and $P \in \mathcal{P}(\mathsf{At})$ and $p \geq 1$. Then P is a *generalized model* of $\mathcal{K}$ wrt. to the p-norm, denoted by $P \approx_p \mathcal{K}$, if and only if $\|A_\mathcal{K} P\|_p = \mathcal{I}_\Pi^p(\mathcal{K})$. Let $\mathsf{GMod}_p(\mathcal{K}) = \{P \in \mathcal{P}(\mathsf{At}) \mid P \approx_p \mathcal{K}\}$ be the set of generalized models of $\mathcal{K}$.

The following proposition states that the generalized models $\mathsf{GMod}_p(\mathcal{K})$ indeed generalize the conventional definition of models.

Proposition 2. *For every $p \geq 1$, if $\mathcal{K}$ is consistent then $\mathsf{GMod}_p(\mathcal{K}) = \mathsf{Mod}(\mathcal{K})$, that is, for consistent $\mathcal{K}$ the generalized models of $\mathcal{K}$ are exactly the models of $\mathcal{K}$.*

The following lemma states that, for $1 < p < \infty$, the *violation vector* $\vec{x}_\mathcal{K}^p = A_\mathcal{K} P$ is identical for all $P \in \mathsf{GMod}_p(\mathcal{K})$. Hence, in this case, we can simplify the condition $\|A_\mathcal{K} P\|_p = \mathcal{I}_\Pi^p(\mathcal{K})$ in Definition 4 by the linear equation $A_\mathcal{K} P = \vec{x}_\mathcal{K}^p$.

Lemma 1. *Let $\mathcal{K}$ be a knowledge base and let $1 < p < \infty$. Let $P \in \mathsf{GMod}_p(\mathcal{K})$ be a generalized model and let $\vec{x} = A_\mathcal{K} P$. Then it holds $A_\mathcal{K} P' = \vec{x}$ for all $P' \in \mathsf{GMod}_p(\mathcal{K})$ and we call $\vec{x} = \vec{x}_\mathcal{K}^p$ the violation vector of $\mathcal{K}$.*

The set of models of a knowledge base $\mathsf{Mod}(\mathcal{K})$ has some features that makes it attractive for model-based reasoning techniques such as ME-reasoning. For instance, the existence and uniqueness of the ME-model is due to the *compactness* and *convexity* of $\mathsf{Mod}(\mathcal{K})$, cf. [11, 7]. Be reminded that a set $X \subseteq \mathbb{R}^n$ is *convex* if for $x_1, x_2 \in X$ it also holds that $\delta x_1 + (1 - \delta)x_2 \in X$ for every $\delta \in [0, 1]$ and X is *compact* if it is both *closed* and *bounded*. A set X is *closed* if for every converging sequence $x_1, x_2, \ldots$ with $x_i \in X$ ($i \in \mathbb{N}$) we have that $\lim_{i \to \infty} x_i \in X$ and X is *bounded* if it is contained in a ball $B_r \subseteq \mathbb{R}^n$ of finite radius r.

The next lemma states that $\mathsf{GMod}_p(\mathcal{K})$ has exactly the same desirable properties as the set of conventional models.

Lemma 2. *For every knowledge base $\mathcal{K}$ and $p \geq 1$ the set $\mathsf{GMod}_p(\mathcal{K})$ is compact and convex.*

As a consequence of Lemma 2, we can draw conclusions from inconsistent knowledge bases in a similar way like from consistent ones. For instance, we could compute probability intervals like in Nilsson's Probabilistic Logic [10] or select a best probability function among all generalized models with respect to a strictly convex (concave) evaluation function. We will generalize Maximum Entropy reasoning here.

Definition 5. Let $\mathcal{K}$ be a knowledge base and $p \geq 1$. The *generalized maximum entropy (ME) model* $\mathsf{GME}_p(\mathcal{K})$ of $\mathcal{K}$ wrt. the p-norm is defined as

$$\mathsf{GME}_p(\mathcal{K}) = \arg \max_{P \in \mathsf{GMod}_p(\mathcal{K})} H(P) \qquad (3)$$

Proposition 3. *For every knowledge base $\mathcal{K}$ and $p \geq 1$ the generalized ME-model $\mathsf{GME}_p(\mathcal{K})$ exists and is uniquely defined.*

Note that the definition of the generalized ME-model is exactly the same as the definition of the classical ME-model, cf. Equation (1), but instead of selecting it from the set of models of $\mathcal{K}$ we select it from the set $\mathsf{GMod}_p(\mathcal{K})$ of generalized models of $\mathcal{K}$. Therefore, Proposition 3 also applies to inconsistent knowledge bases while the classical ME-model exists only for consistent ones. However, the following corollary states that the generalized ME-model coincides with the conventional ME-model in the case of consistent knowledge bases.

Corollary 1. *For every $p \geq 1$, if $\mathcal{K}$ is consistent then $\mathsf{GME}_p(\mathcal{K}) = \mathsf{ME}(\mathcal{K})$.*

It follows that our framework allows to generalize classical ME-reasoning to inconsistent knowledge bases. Before we apply this approach to the problem of knowledge base consolidation we first give a small example that illustrates the behavior of $\mathsf{GME}_p(\mathcal{K})$ itself.

Example 2. Suppose, we want to define a simple spam filter. We consider the variables sp (spam), ss (suspicious subject), sc (suspicious content) and the following knowledge base

$$\mathcal{K} = \langle (sp)[0.25], (sp \mid ss)[0.8], (sp \mid sc)[0.6], (sc \mid sp)[0.7],$$
$$(ss \mid sp)[0.5], (sc \mid \overline{sp})[0.05], (ss \mid \overline{sp})[0.01] \rangle.$$

Note that $\mathcal{K}$ is inconsistent, i.e., there is no probability function that satisfies all conditionals in $\mathcal{K}$. However, the generalized ME-model $P_1^* = \mathsf{GME}_1(\mathcal{K})$ satisfies all conditionals except $(sp \mid ss)[0.8]$ and $(sp \mid sc)[0.6]$. For the latter it yields the conditional probabilities $P_1^*(sp \mid ss) = 0.94$ and $P_1^*(sp \mid sc) = 0.82$. Whereas most conditionals are satisfied, the probability of $(sp \mid sc)[0.6]$ is increased significantly by 22%. As we increase p, we get more "balanced" deviations. For instance, $P_2^* = \mathsf{GME}_2(\mathcal{K})$ violates all conditionals but the conditional probabilities deviate at most 14% from the original probabilities stated in $\mathcal{K}$. For $p = \infty$, the maximal deviation is only about 10%.

4.2 Generalized ME-Consolidation

The probability function $\mathsf{GME}_p(\mathcal{K})$ provides a means for consistent reasoning with inconsistent knowledge bases and will now serve as the basis for defining our *generalized ME-consolidation operator*.

Definition 6. Let $\mathcal{K} = \langle (\psi_1 \mid \phi_1)[d_1], \ldots, (\psi_m \mid \phi_m)[d_m] \rangle$ be a knowledge base, let $p \geq 1$ and let $P^* = \mathsf{GME}_p(\mathcal{K})$. Then the *generalized ME-consolidation operator* Γ_{ME}^p is defined as

$$\Gamma_{\mathsf{ME}}^p(\mathcal{K}) = \langle (\psi_1 \mid \phi_1)[d_1'], \ldots, (\psi_m \mid \phi_m)[d_m'] \rangle$$

with

$$d_i' = \begin{cases} P^*(\psi_i \mid \phi_i) & \text{if } P^*(\phi_i) > 0 \\ d_i & \text{otherwise} \end{cases}$$

for $i = 1, \ldots, m$.

If $P^*(\psi_i \mid \phi_i)$ is not defined (this is the case if $P^*(\phi_i) = 0$) we leave the probability of the conditional as it is, thus following the minimal change paradigm. Note that due to $P^*(\phi_i) = 0$ the probability function P^* is a model of $(\psi_i \mid \phi_i)[d_i']$ for every value of d_i'.

Example 3. Table 1 shows consolidated knowledge bases for the spam filter from Example 2 for different values of p. Note that no consolidation is Pareto-dominated by another one, that is, there are no two consolidated knowledge bases such that all probabilities in one are closer to the original probabilities than the probabilities in the other.

$\mathcal{K}$	$\Gamma_{\mathsf{ME}}^1(\mathcal{K})$	$\Gamma_{\mathsf{ME}}^2(\mathcal{K})$	$\Gamma_{\mathsf{ME}}^\infty(\mathcal{K})$
$(sp)[0.25]$	0.25	0.24	0.23
$(sp \mid ss)[0.8]$	0.94	0.88	0.8
$(sp \mid sc)[0.6]$	0.82	0.74	0.7
$(sc \mid sp)[0.7]$	0.7	0.65	0.61
$(ss \mid sp)[0.5]$	0.5	0.49	0.5
$(sc \mid \overline{sp})[0.05]$	0.05	0.07	0.08
$(ss \mid \overline{sp})[0.01]$	0.01	0.02	0.04

Table 1. Probabilities of consolidated knowledge bases for $p = 1, 2, \infty$.

The previous example showed that our approach provides meaningful results for probabilistic knowledge base consolidation. In fact, our approach also satisfies most of the rationality postulates discussed before.

Theorem 1. *Let $p \geq 1$. The consolidation operator Γ_{ME}^p satisfies* Success, Consistency, Irrelevance of syntax, Weak IIA, IIA *and* Continuity. *For $1 < p < \infty$, Γ_{ME}^p also satisfies* Non-dictatorship.

Arrow's impossibility result [6] suggests that Γ_{ME}^p cannot satisfy *Pareto-optimality* as the classical versions of *Non-dictatorship*, *IIA*, and *Pareto-optimality* are incompatible. However, note that we re-phrased the postulates to match our probabilistic framework and Arrow's impossibility result does not necessarily carry over to our framework. Indeed, empirical experiments suggest that Γ_{ME}^p also satisfies *Pareto-optimality* (for all $p \geq 1$). However, no formal proof has been found yet, so we can only give a conjecture on this.

Conjecture 1. *The consolidation operator Γ_{ME}^p satisfies* Pareto-optimality.

Before continuing with a discussion on the computational complexity and implementation of generalized ME-Reasoning and the generalized ME-consolidation operator, we conclude this subsection with a simple corollary that nicely illustrates the compatibility of classical ME-reasoning with generalized ME-reasoning and the generalized ME-consolidation operator.

Corollary 2. *Let $p \geq 1$ and let $\mathcal{K}$ be a knowledge base. Then*

$$GME_p(\mathcal{K}) = ME(\Gamma_{\mathsf{ME}}^p(\mathcal{K})).$$

The above corollary states that the generalized ME-model of any knowledge base (consistent or inconsistent) is the same as the classical ME-model of the consistent knowledge base obtained by consolidating the original knowledge base.

4.3 Computational Issues and Implementation

Consolidating knowledge bases in our framework consists of two phases. First, we compute the minimal violation measure with respect to some p-norm. Then, we compute the generalized maximum entropy model to consolidate the knowledge base.

The first problem, computing minimal violation measures, can be solved by means of convex optimization. For $p = 1$ and $p = \infty$ we obtain particular efficient linear programs, see [13] for details. For $p = 2$, computing the minimal violation measure is equivalent to a least-squares problem which can be solved by quadratic programming methods. The second problem, computing the generalized maximum entropy model, can also be solved by convex programming methods. The computational difference to classical ME-reasoning is that we need to compute $\mathcal{I}_\Pi^p(\mathcal{K})$ first. As computing minimal violation measures can be solved by convex programming techniques, and maximizing entropy corresponds to a convex programming problem, generalized ME-reasoning has the same asymptotic worst-case complexity as classical ME-reasoning. If we employ interior-point methods naively, we can expect 10 to 100 iterations with cost $\max\{|\Omega(\mathsf{At})|^3, |\Omega(\mathsf{At})|^2 \, |\mathcal{K}|\}$ per iteration [1].

For $1 < p < \infty$, Lemma 1 allows us to replace the convex constraint $\|A_\mathcal{K} P\|_p = \mathcal{I}_\Pi^p(\mathcal{K})$ by the affine constraint $A_\mathcal{K} P = \vec{x}_\mathcal{K}^p$. As a consequence, we can transform the corresponding optimization problem to an equivalent unconstrained problem that is easier to solve.

Proposition 4. *Let $\mathcal{K}$ be a knowledge base, let $1 < p < \infty$ and let $\vec{x}_\mathcal{K}^p$ be the violation vector of $\mathcal{K}$ with respect to p. If there is a positive generalized model $P \in \mathsf{GMod}_p(\mathcal{K})$ then $GME_p(\mathcal{K})$ is positive and can be computed by solving the following unconstrained optimization problem:*
Let $n = |\Omega(\mathsf{At})|$, $m = |\mathcal{K}|$, $A = A_\mathcal{K}$ and let $\vec{\pi} : \mathbb{R}^{m+1} \to \mathbb{R}^n$ be a function defined as

$$\vec{\pi}(\vec{\lambda}, \mu) = (\exp(\sum_{k=1}^m \vec{\lambda}_k A_{ki} + \mu - 1))_{1 \leq i \leq n} \tag{4}$$

and consider the optimization problem

$$\arg \min_{(\vec{\lambda}, \mu) \in \mathbb{R}^{m+1}} \sum_{1 \leq i \leq n} \vec{\pi}(\vec{\lambda}, \mu)_i - \vec{\lambda}^T \vec{x}_\mathcal{K}^p - \mu. \tag{5}$$

Then (5) has a solution $(\vec{\lambda}^*, \mu^*)$ and for each such solution it holds that

$$GME_p(\mathcal{K}) = \vec{\pi}(\vec{\lambda}^*, \mu^*).$$

Note that, in general, there might be no positive generalized model as required by Proposition 4 to be applicable. Still, the optimization problem yields a reasonable approximation in these cases. To get the intuition, note that the objective in (5) is convex in $\vec{\lambda}$ and μ. Therefore, the minimum is obtained in a stationary point with gradient zero. The partial derivative with respect to $\vec{\lambda}_k$ is $\sum_{1 \leq i \leq n} A_{ki} \vec{\pi}(\vec{\lambda}, \mu)_i - (\vec{x}_\mathcal{K}^p)_k$ and the partial derivative with respect to μ is $\sum_{1 \leq i \leq n} \vec{\pi}(\vec{\lambda}, \mu)_i - 1$. Hence, in a stationary point, all constraints are satisfied (the solution is in particular positive, because the exponential function yields only positive values). If there is no positive generalized model, some Lagrange multipliers $\vec{\lambda}_k$ will tend to infinity. However, a line-search method will follow a descent direction and as the objective decreases, the generalized ME-model is approximated better and better.

The minimal violation inconsistency measure $\mathcal{I}_\Pi^p$, the computation of the generalized ME-model $GME_p(\mathcal{K})$, and the generalized

ME-consolidation operator Γ^p_{ME} have been prototypically implemented using Java in *Tweety*[4] and are available under the GNU General Public License v3.0. We implemented general versions of these three components for computing minimal violations by employing the general optimization library OpenOpt[5]. As OpenOpt is a general optimization library that may exhibit some numerical oddities, we also implemented specialized versions of some of these components for the Manhattan norm and the Euclidean norm (i. e. for the p-norm with $p = 1, 2$) using the optimization solvers lpsolve[6] and the mathematical library ojAlgo[7] which (in general) provide better numerical performance. To avoid numerical inaccuracy, the examples in this work were computed with Matlab and CVX[8]. Links to the concrete packages and classes can be found in the online appendix[9].

5 Related Work

In [15] three approaches are proposed for restoring consistency in a probabilistic knowledge base $\mathcal{K}$. The first two approaches are very similar to each other but follow the paradigm of *qualitative modifications* of conditionals, cf. also [8]. In those approaches each probabilistic conditional $(\psi_i \,|\, \phi_i)[d_i] \in \mathcal{K}$ is extended to $(\psi_i \,|\, \phi_i \wedge w_i)[d_i]$ with a new proposition w_i for $i = 1, \ldots, m$. By doing so, inconsistencies in the former knowledge base are resolved and a model of the new knowledge base can be used to repair the probabilities in the old one. The third approach in [15] is a quantitative one and uses *generalized divergence* as a distance measure to determine new probabilities. The idea is similar in spirit to our approach, but instead of minimizing $|P(\psi\phi) - dP(\phi)|$, roughly speaking, the log-ratio $\log\big(\frac{P(\overline{\psi}\phi)}{P(\psi\phi)}\frac{d}{1-d}\big)$ is minimized. Note that this term is zero if $P(\overline{\psi}|\phi) = 1 - d$ and $P(\psi|\phi) = d$. Unfortunately, no justification and no evaluation of any of these approaches is given in [15].

The work [2] also considers the issue of extending reasoning based on the principle of maximum entropy to inconsistent knowledge bases. There, a fuzzy interpretation is used to define a *degree of satisfaction* for probability functions and knowledge bases.

The work [3] also proposes an approach to restore consistency in probabilistic knowledge bases by modification of probabilities. However, the approach follows a heuristic paradigm and has to be guided by the knowledge engineer by specifying importance of conditionals. The approach itself then changes the probabilities in a step-wise fashion, depending on the specified importance, until consistency is restored. In contrast, our approach is principled as it is based on a generalized approach to ME-reasoning. Furthermore, we have shown that our consolidation approach satisfies several quality criteria.

It has also been proposed to directly change the probabilities in an inconsistent knowledge base to a consistent one [16] or to relax them to consistent probability intervals [12] by minimizing the change of probabilities in the knowledge base. But while these approaches yield the best consolidation with respect to some distance measure by definition, they are not necessarily uniquely determined and are hard to compute in practice as they correspond to non-convex optimization problems that suffer from the existence of non-global local minima.

In [18] the authors discuss the problem of determining conditional probability tables for Bayesian Networks given possibly inconsistent information. While their motivation is similar to ours, the techniques used are quite different and neither relationships to other fields such as inconsistency minimization and belief merging nor an evaluation in terms of quality wrt. desirable properties is conducted.

6 Summary and Conclusion

We solved the problem of consolidation of probabilistic knowledge bases by proposing a new approach to inconsistency-tolerant reasoning based on the principle of maximal entropy. We introduced the notion of generalized models of a probabilistic knowledge base that is based upon inconsistency minimization and that enabled us to generalize the ME-model to inconsistent knowledge bases. We showed that our approach satisfies several desirable properties and discussed its computational complexity.

The generalization of ME-reasoning to inconsistent knowledge bases deserves a deeper investigation than that was possible within the scope of this paper. As part of future work we will investigate this reasoning approach in more depth.

References

[1] S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
[2] L. Daniel, *Paraconsistent Probabilistic Reasoning*, Ph.D. dissertation, L'École Nationale Supérieure des Mines de Paris, 2009.
[3] M. Finthammer, G. Kern-Isberner, and M. Ritterskamp, 'Resolving Inconsistencies in Probabilistic Knowledge Bases', in *KI 2007: Advances in Artificial Intelligence*, pp. 114–128, (2007).
[4] S. O. Hansson, 'A Survey of Non-Prioritized Belief Revision', *Erkenntnis*, **50**(2–3), 413–427, (1999).
[5] Sven Ove Hansson, *A Textbook of Belief Dynamics*, Kluwer Academic Publishers, Norwell, MA, USA, 2001.
[6] Jerry Kelly, *Social Choice Theory: An Introduction*, Springer-Verlag, 1988.
[7] G. Kern-Isberner, *Conditionals in Nonmonotonic Reasoning and Belief Revision*, number 2087 in LNCS, Springer-Verlag, 2001.
[8] G. Kern-Isberner and W. Rödder, 'Belief Revision and Information Fusion in a Probabilistic Environment', in *Proceedings of the 16th International FLAIRS Conference, (FLAIRS'03)*, pp. 506–510. AAAI Press, (2003).
[9] S. Konieczny and R. Pino Pérez, 'On the logic of merging', in *Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning (KR'98)*. Morgan Kaufmann, (1998).
[10] Nils J. Nilsson, 'Probabilistic logic', *Artificial Intelligence*, **28**, 71–88, (February 1986).
[11] J. B. Paris, *The Uncertain Reasoner's Companion – A Mathematical Perspective*, Cambridge University Press, 1994.
[12] D. Picado-Muiño, 'Measuring and Repairing Inconsistency in Probabilistic Knowledge Bases', *International Journal of Approximate Reasoning*, (2011).
[13] N. Potyka, 'Linear programs for measuring inconsistency in probabilistic logics', in *Proceedings of the 14th International Conf. on Principles of Knowledge Representation and Reasoning (KR'14)*, (2014).
[14] R. Reiter, 'A Logic for Default Reasoning', *Artificial Intelligence*, **13**(1–2), 81–132, (1980).
[15] W. Rödder and L. Xu, 'Elimination of Inconsistent Knowledge in the Probabilistic Expertsystem-Shell SPIRIT (in German)', in *Operations Research Proceedings: Selected Papers of the Symposium on Operations Research 2000*, pp. 260–265. Springer-Verlag, (2001).
[16] M. Thimm, *Probabilistic Reasoning with Incomplete and Inconsistent Beliefs*, volume 331 of *Dissertations in Artificial Intelligence*, IOS Press, December 2011.
[17] M. Thimm, 'Inconsistency measures for probabilistic logics', *Artificial Intelligence*, **197**, 1–24, (April 2013).
[18] Wim Wiegerinck and Tom Heskes, 'Probability assessment with maximum entropy in bayesian networks', in *Proceedings of the 33rd Symposium on the Interface*, volume 33 of *Computing Science and Statistics*, (2001).

---

[4] http://mthimm.de/projects/tweety/
[5] http://openopt.org/
[6] http://lpsolve.sourceforge.net
[7] http://ojalgo.org
[8] http://cvxr.com/cvx/
[9] http://www.mthimm.de/misc/pcons_ecai2014_proofs.pdf