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

A Software System for the Computation, Visualization, and Comparison of Conditional Structures for Relational Probabilistic Knowledge Bases

Christoph Beierle
University of Hagen
Hagen, Germany

Steffen Kuche
University of Hagen
Hagen, Germany

Marc Finthammer
University of Hagen
Hagen, Germany

Gabriele Kern-Isberner
Univ. of Technology Dortmund
Dortmund, Germany

Abstract

Combining logic with probabilities is a core idea to uncertain reasoning. Recently, approaches to probabilistic conditional logics based on first-order languages have been proposed that employ the principle of maximum entropy (ME), e.g. the logic FO-PCL. In order to simplify the ME model computation, FO-PCL knowledge bases can be transformed so that they become parametrically uniform. On the other hand, conditional structures have been proposed as a structural tool for investigating properties of conditional knowledge bases. In this paper, we present a software system for the computation, visualization, and comparison of conditional structures for relational probabilistic knowledge bases as they evolve in the transformation process that achieves parametric uniformity.

1 Introduction

A popular approach to the representation of and the reasoning with uncertain knowledge is to use probabilistic conditionals of the form *If A then B with probability x*. Generally, combining logic with probabilities has a long tradition (e.g. (Nilsson 1986; Fagin, Halpern, and Megiddo 1990; Fagin and Halpern 1994)). For probabilistic conditionals based on an underlying propositional logic, the principle of Maximum Entropy (ME) is a well-established method for the selection of a "best", i.e. most unbiased model, providing an excellent basis for model-based reasoning (Paris 1994; Kern-Isberner and Lukasiewicz 2004; Kern-Isberner 1998; 2001). Recently, several approaches employing the ME principle to a first-order setting have been proposed (e.g. (Kern-Isberner and Thimm 2010; Fisseler 2012)).

In this paper, we address two key concepts that have been developed for studying the structural properties of ME reasoning, *conditional structures* (Kern-Isberner 2001) and *parametric uniformity* (Fisseler 2012). Both concepts induce certain equivalences on constants, worlds, conditionals, or on groundings of conditionals, and both concepts have been utilized successfully for simplifying the ME model computation, cf. (Finthammer and Beierle 2012b; 2012a). However, it is still an open question how exactly these two concepts relate to each other. For a better understanding of ME reasoning, it will be helpful to know how conditional structures evolve when transforming a knowledge base to an equivalent one that is parametrically uniform. An advanced insight in the structures of the corresponding equivalence classes is essential for developing more efficient approaches to characterize the classes and to determine their cardinalities, being key ingredients for the computation of the ME model. In order to support the investigation of these open problems, we developed and implemented a software system called *CSPU* (*Conditional Structures and Parametric Uniformity*). *CSPU* serves for the computation, visualization, and comparison of conditional structures for relational probabilistic FO-PCL knowledge bases (*First-Order Probabilistic Conditional Logic*) (Fisseler 2012) as they evolve in the transformation process that achieves parametric uniformity (Beierle and Krämer 2015).

In Sec. 2, we briefly recall the background of FO-PCL and parametric uniformity. In Sec. 3 we introduce conditional structures for FO-PCL. In Sec. 4 we present a detailed example illustrating the functionalities of our system *CSPU*, and in Sec. 5 we conclude and point out further work.

2 FO-PCL and Parametric Uniformity

FO-PCL Syntax FO-PCL uses function-free signatures of the form $\Sigma = (S, D, Pred)$ where S is a set of sorts, $D = \bigcup_{s \in S} D^{(s)}$ is a finite set of (disjoint) sets of sorted constant symbols, and $Pred$ is a set of predicate symbols, each coming with an arity of the form $s_1 \times \ldots \times s_n \in S^n$ indicating the required sorts for the arguments. Variables $\mathcal{V}$ also have a unique sort, and all formulas and variable substitutions must obey the obvious sort restrictions. In the following, we will adopt the unique names assumption, i. e. different constants denote different elements.

An FO-PCL *conditional* $R = \langle(\phi_R|\psi_R)[\xi_R], C_R\rangle$ is composed of a *premise* ψ_R and a *conclusion* ϕ_R, which are quantifier and function free first-order formulas (over Σ and $\mathcal{V}$) without equality, a probability value $\xi_R \in [0, 1]$, and a *constraint formula* C_R which is a quantifier-free first-order formula using only the equality predicate. For $\neg(V = X)$ we also write $(V \neq X)$, and $\top$ resp. $\bot$ denote a tautology resp. a contradiction. An *FO-PCL knowledge base* is a pair $(\Sigma, \mathcal{R})$ where $\mathcal{R}$ is a set of conditionals over $\Sigma, \mathcal{V}$. In the following, we will call just $\mathcal{R}$ a knowledge base and Σ will be given by the context. We will use the notation $C_R \models V \neq c$ to express that under the unique names assumption, the constraint formula C_R of a conditional R entails the constraint

$V \neq c$, and $C_R \not\models V \neq c$ means that C_R does not entail the constraint.

The constraint formula makes it possible to explicitly express that a generic conditional is not applicable with respect to a particular individual. Without constraint formulas, having a generic conditional and a corresponding conditional for a specific individual, that specific conditional might formally contradict the general one when considering all instances.

Example 1 (Misanthrope). *The knowledge base* $\mathcal{R} = \{R_1, R_2\}$*, adapted from (Fisseler 2012), models friendship relations within a group of people, with one exceptional member, a misanthrope. In general, if a person V likes another person U, then it is very likely that U likes V, too. But there is one person, the misanthrope, who generally does not like other people:*

$R_1 : \langle(\mathit{likes}(U,V)|\mathit{likes}(V,U))[0.9], U \neq V\rangle$
$R_2 : \langle(\mathit{likes}(a,V))[0.05], V \neq a\rangle$

Please note that given the set of contants $D = \{a, b, c\}$, instantiating the conditionals in $\mathcal{R}$ without considering their constraint formulas would yield the contradictory conditional $\langle(\mathit{likes}(a,a)|\mathit{likes}(a,a))[0.9]$.

When the constraint formula of a ground instance of R evaluates to $true$, that instance is called *admissible*, and $gnd(R)$ denotes the set of all admissible instances of R (over Σ), in the following also just called instances.

FO-PCL Models The Herbrand base $\mathcal{H}(\mathcal{R})$ is the set of all atoms in all $gnd(R_k)$ with $R_k \in \mathcal{R}$, and every subset $x \subseteq \mathcal{H}(\mathcal{R})$ is a Herbrand interpretation, also called *world*, defining a logical semantics for $\mathcal{R}$. The set $X(\mathcal{R}) = \{x \mid x \subseteq \mathcal{H}(\mathcal{R})\}$ denotes the set of all Herbrand interpretations. The probabilistic semantics of $\mathcal{R}$ is a possible world semantics (Halpern 2005) where the ground atoms in $\mathcal{H}(\mathcal{R})$ are binary random variables. An FO-PCL *interpretation* $p_{X(\mathcal{R})}$ of $\mathcal{R}$ is thus a probability distribution over $X(\mathcal{R})$.

For $R_k \in \mathcal{R}$ and every $g_{R_k} \in gnd(R_k)$, let θ_{R_k} be an admissible ground substitution for the variables in R_k so that $g_{R_k} = \langle(\theta_{g_{R_k}}(\phi_{R_k}) \mid \theta_{g_{R_k}}(\psi_{R_k}))[\xi_{R_k}], \top\rangle$. Then $p_{X(\mathcal{R})}$ satisfies R_k iff for every instance $g_{R_k} \in gnd(R_k)$ we have: $p_{X(\mathcal{R})}(\theta_{R_k}(\phi_{R_k}) \wedge \theta_{R_k}(\psi_{R_k})) = \xi_{R_k} \cdot p_{X(\mathcal{R})}(\theta_{R_k}(\psi_{R_k}))$. Note that for the case of $p_{X(\mathcal{R})}(\theta_{R_k}(\psi_{R_k})) > 0$, this equation is equivalent to $\frac{p_{X(\mathcal{R})}(\theta_{R_k}(\phi_{R_k}) \wedge \theta_{R_k}(\psi_{R_k}))}{p_{X(\mathcal{R})}(\theta_{R_k}(\psi_{R_k}))} = \xi_{R_k}$ and thus to $p_{X(\mathcal{R})}((\theta_{g_{R_k}}(\phi_{R_k}) \mid \theta_{g_{R_k}}(\psi_{R_k}))) = \xi_{R_k}$, expressing conditional probability. $p_{X(\mathcal{R})}$ is a *model* of $\mathcal{R}$, denoted by $p_{X(\mathcal{R})} \models \mathcal{R}$, if it satisfies every $R_k \in \mathcal{R}$.

Maximum Entropy Model and Parametric Uniformity A knowledge base $\mathcal{R} = \{R_1, \ldots, R_m\}$ may have many different models, and the principle of maximum entropy (Paris 1994; Kern-Isberner and Lukasiewicz 2004; Kern-Isberner 1998; 2001) provides a method to select a model that is optimal in the sense that it is the most unbiased one. The uniquely determined model

$$p^*_{X(\mathcal{R})} = \arg \max_{p_{X(\mathcal{R})} \models \mathcal{R}} H(p_{X(\mathcal{R})}) \quad (1)$$

of $\mathcal{R}$ having maximum entropy $H(p^*_{X(\mathcal{R})})$ can be represented by a Gibbs distribution (Geman and Geman 1984):

$$p^*_{X(\mathcal{R})}(x) = \frac{1}{Z} \exp\left(\sum_{k=1}^{m} \sum_{g_{R_k} \in gnd(R_k)} \lambda_{g_{R_k}} f_{g_{R_k}}(x)\right) \quad (2)$$

where $f_{g_{R_k}}$ is the feature function determined by g_{R_k}, $\lambda_{g_{R_k}}$ is a Lagrange multiplier (Boyd and Vandenberghe 2004) and Z is a normalization constant. We will not elaborate on the details of Equation (2) as they are not important for the rest of this work (see (Fisseler 2012) for a detailed explanation). What is important to note is that according to Equation (2), one optimization parameter $\lambda_{g_{R_k}}$ has to be determined for *each single ground instance* g_{R_k} of each conditional R_k. This readily yields a computationally infeasible optimization problem for larger knowledge bases because there might be just too many ground instances.

However, there are FO-PCL knowledge bases for which the ground instances of a conditional share the same entropy-optimal parameter. Parametric uniformity (Fisseler 2012) means that for each conditional all its ground instances share the same entropy-optimal parameter value. The advantage of the semantical notion of parametric uniformity is that just *one* optimization parameter λ_{R_k} per conditional R_k has to be computed instead of one parameter per ground instance :

$$p^*_{X(\mathcal{R})}(x) = \frac{1}{Z} \exp\left(\sum_{k=1}^{m} \lambda_{R_k} \sum_{g_{R_k} \in gnd(R_k)} f_{g_{R_k}}(x)\right) \quad (3)$$

Parametric uniformity indicates identical knowledge about all ground instances of the same conditional for an FO-PCL knowledge base $\mathcal{R}$. Due to this, one should be able to transpose two ground instances g_{R_k}, g'_{R_k} of a conditional in $\mathcal{R}$ without changing the joint probability function with maximum entropy. In this case the transposed ground instances must possess the same entropy optimal parameter, as the Gibbs distribution in Equation (2) is determined by a unique set of Lagrange multipliers (Fisseler 2012).

In the following, the ME model of an FO-PCL knowledge base $\mathcal{R}$ will be denoted by $P^*_{\mathcal{R}}$.

Interactions and Transformation Rules Luckily, each FO-PCL knowledge base $\mathcal{R}$ that is not parametrically uniform can be transformed into an equivalent $\mathcal{R}$' that is parametrically uniform. This is achieved by the set of transformations rules $\mathcal{PU}$ developed in (Beierle and Krämer 2015). In (Beierle and Krämer 2015), the reasons causing $\mathcal{R}$ to be not parametrically uniform are investigated in detail and the syntactic criterion of *inter-rule* and *intra-rule interactions* is introduced. For each of the different types of interactions, there is a corresponding interaction removing transformation rule in $\mathcal{PU}$ (Beierle and Krämer 2015), for instance:

$$(TE_1)\ \frac{\mathcal{R} \cup \{R_1, R_2\}}{\mathcal{R} \cup \{R_1\} \cup \nu\{\sigma(R_2), \overline{\sigma}(R_2)\}} \quad \begin{array}{l} R_1 \leftarrow \langle P\rangle_{V,c} \rightarrow R_2, \\ \sigma = \{V/c\} \end{array}$$

Example 2 (Application of $\mathcal{PU}$). *Among the conditionals* R_1 *and* R_2 *from Example 1 there is an inter-rule interaction of type 1, denoted by* $R_2 \leftarrow \langle \mathit{likes}\rangle_{V,a} \rightarrow R_1$*. The transfor-*

mation rule (TE_1) *removes it by replacing* R_1 *with*

$R_{1\_1} : \langle (likes(U, a) \,|\, likes(a, U))\, [0.9]\, , U \neq a\rangle$

$R_{1\_2} : \langle (likes(U, V) \,|\, likes(V, U))\, [0.9]\, , U \neq V \wedge V \neq a\rangle.$

Thus, the transformation rule (TE_1) removes an inter-rule interaction of type 1 by replacing a conditional R with two new conditionals $\nu(\sigma(R))$ and $\nu(\bar{\sigma}(R))$, where $\sigma(R)$ is the result of applying the variable substitution $\sigma = \{V/c\}$ to R, and $\bar{\sigma}(R)$ is the result of adding the constraint $V \neq c$ to the constraint formula of R. The operator ν transforms a conditional in *constraint normal form*. Similarly, there are two transformation rules (TE_2) and (TE_3) that remove inter-rule interactions of type 2 and 3. The three different types of intra-rule interactions occur within a single conditional and are removed by one of the three rules (TA_1), (TA_2), (TA_3) in $\mathcal{PU}$ (Beierle and Krämer 2015).

Proposition 1. (Beierle and Krämer 2015) *Applying* $\mathcal{PU}$ *to a knowledge base* $\mathcal{R}$ *terminates and yields a knowledge base* $\mathcal{PU}(\mathcal{R})$ *having the same maximum-entropy model and* $\mathcal{PU}(\mathcal{R})$ *is parametrically uniform.*

We refer to (Beierle and Krämer 2015) for further details of $\mathcal{PU}$, including many examples, formal definitions and proofs.

3 FO-PCL and Conditional Structures

Kern-Isberner (2001) investigates the behaviour of worlds with respect to conditionals and introduces the concept of *conditional structure* of a world with respect to a set of propositional conditionals $\mathcal{R}^{\mathrm{pr}}$. Formally, the conditional structure of ω with respect to $\mathcal{R}^{\mathrm{pr}}$ is given by a product in a free abelian group with generators indicating that ω verifies resp. falsifies the i-th conditional in $\mathcal{R}^{\mathrm{pr}}$. Kern-Isberner's idea of a conditional structure carries over to the relational case by employing functions counting the number of verifying and falsifying groundings (Kern-Isberner and Thimm 2012). Since each free abelian group is isomorphic to a cartesian product of copies of the set of integers $\mathbb{Z}$, the extension of a conditional structure of a world to the relational setting can also be defined by using ordered tuples instead of a free abelian group notation (Finthammer and Beierle 2012b). Both in (Kern-Isberner and Thimm 2012) and (Finthammer and Beierle 2012b), conditional structures are defined for relational conditionals under *aggregation semantics* (Kern-Isberner and Thimm 2010). In the following, we will adapt this notion to the FO-PCL case by taking the constraint formulas of FO-PCL conditionals into account, and as in (Finthammer and Beierle 2012b), we will use ordered tuples for representing conditonal structures.

Definition 1 (FO-PCL Counting Functions). *For an FO-PCL* conditional $R_i = \langle(\phi_R|\psi_R)[\xi_R], C_R\rangle$ *the* counting functions $ver_i, fal_i : \Omega \to \mathbb{N}_0$ *are given by:*

$$ver_i(\omega) = |\{\langle(\phi_g|\psi_g)[\xi], \top\rangle \in gnd(R_i) \mid \omega \models \psi_g \wedge \phi_g\}|$$

$$fal_i(\omega) = |\{\langle(\phi_g|\psi_g)[\xi], \top\rangle \in gnd(R_i) \mid \omega \models \psi_g \wedge \neg\phi_g\}|$$

$ver_i(\omega)$ *indicates the number of groundings of* R_i *which are* verified *by* ω*, whereas* $fal_i(\omega)$ *specifies the number of groundings of* R_i *which are* falsified *by* ω*.*

Considering the counting functions of all conditionals in a knowledge base $\mathcal{R}$ yields the conditional structure of $\mathcal{R}$.

Definition 2 (vf-Pair, FO-PCL Conditional Structure). *For a world* ω*, the pair* $(ver_i(\omega), fal_i(\omega)) \in \mathbb{N}_0 \times \mathbb{N}_0$ *is called the* vf-pair of ω with respect to conditional R_i*. The* conditional structure $\gamma_{\mathcal{R}}(\omega)$ *of* ω *with respect to an FO-PCL knowledge base* $\mathcal{R} = \{R_1 \ldots, R_m\}$ *is the m-tuple:*

$$\gamma_{\mathcal{R}}(\omega) := (\, (ver_1(\omega), fal_1(\omega))\, , \ldots, \\ (ver_m(\omega), fal_m(\omega))\,) \in (\mathbb{N}_0 \times \mathbb{N}_0)^m \quad (4)$$

Example 3. *For* $\mathcal{R}$ *from Example 1 with* $D = \{a, b, c\}$*, the world* $\omega = \{likes(a, b), likes(b, a)\}$ *e.g. falsifies the grounding* $\langle(likes(a, c))[0.05], \top\rangle$ *of* R_2*. Its conditional structure is* $\gamma_{\mathcal{R}}(\omega) = ((2, 0),\, (1, 1))$.

Note that the conditional structure $\gamma_{\mathcal{R}}(\omega)$ does not take any probabilities into account, i.e., it just considers the logical part of the conditionals in $\mathcal{R}$. This logical structure induces an equivalence relation on worlds.

Definition 3 (Structural Equivalence). *Two worlds* $\omega_1, \omega_2 \in \Omega$ *are* structurally equivalent *with respect to* $\mathcal{R}$*, denoted by* $\omega_1 \equiv_{\mathcal{R}} \omega_2$*, iff* $\gamma_{\mathcal{R}}(\omega_1) = \gamma_{\mathcal{R}}(\omega_2)$*. With* $[\omega_l]_{\equiv_{\mathcal{R}}} := \{\omega \in \Omega \mid \omega \equiv_{\mathcal{R}} \omega_l\}$ *we denote the equivalence class of* $\omega_l \in \Omega$*, with* $\Omega/\equiv_{\mathcal{R}} := \{[\omega]_{\equiv_{\mathcal{R}}} \mid \omega \in \Omega\}$ *the set of all equivalence classes, and with* $|[\omega_l]_{\equiv_{\mathcal{R}}}|$ *the cardinality of the equivalence class* $[\omega_l]_{\equiv_{\mathcal{R}}}$.

An important observation is that the ME distribution $P^*_{\mathcal{R}}$ of $\mathcal{R}$ respects the conditional structure for parametrically uniform $\mathcal{R}$: While in general, a probability distribution satisfying $\mathcal{R}$ may assign different probabilities to worlds having the same conditional structure, such worlds always have the same probability under $P^*_{\mathcal{R}}$.

Proposition 2. *If* $\mathcal{R}$ *is parametrically uniform, then for* $\omega_1, \omega_2 \in \Omega$*,* $\omega_1 \equiv_{\mathcal{R}} \omega_2$ *implies* $P^*_{\mathcal{R}}(\omega_1) = P^*_{\mathcal{R}}(\omega_2)$.

For proving Proposition 2, the feature functions for a conditional R used in Equations (2) and (3) are expressed in terms of the FO-PCL counting functions for R and the probability of R, similar as it is done in (Finthammer and Beierle 2012b) for aggregation semantics.

The observations outlined above strongly indicate that both conditional structures and parametric uniformity are crucial concepts for describing properties of FO-PCL knowledge bases with ME semantics. Moreover, both concepts can be exploited successfully in ME model computations (e.g. (Finthammer and Beierle 2012b; 2012a)). Still, up to now the exact relationship between the two concepts is not known, e.g., to which extent can both notions be expressed in terms of the other one? How do conditional structures evolve along the transformation steps when applying $\mathcal{PU}$? How do the induced equivalence classes of worlds change accordingly? For supporting the study of these questions we designed and implemented a system for the computation and comparison of conditional structures in connection with the $\mathcal{PU}$ transformation process. This system is part of the KREATOR environment[1] (Finthammer and Thimm 2012), an integrated development environment for relational probabilistic logic; its functionality is illustrated in the following section.

[1] KREATOR can be found at http://kreator-ide.sourceforge.net/

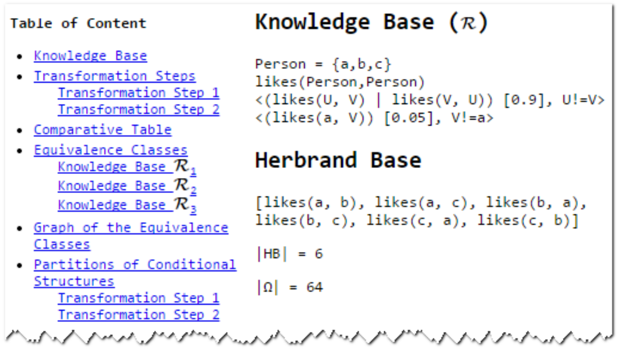


Figure 1: Navigation for $\mathcal{PU}$ transformation report

4 System Illustration and Examples

The KREATOR extension *CSPU* performs the $\mathcal{PU}$ transformation process and calculates the conditional structures of all knowledge bases arising from it. The complete information of the $\mathcal{PU}$ transformation process of a knowledge base is summarized in a $\mathcal{PU}$ *transformation process report* or simply $\mathcal{PU}$ *report*. One part of this $\mathcal{PU}$ report is a table wherein each knowledge base is represented as a column and each world is represented as a row, with the number of worlds typically being very large. Furthermore, a graph representing the relations among the equivalence classes of worlds may contain thousands of nodes and edges, which is also not representable in a static textual representation format. Therefore, HTML was chosen as the format of the report. With the help of HTML it is possible to manage large tables and graphs by providing appropriate navigation and visualization tools. Of course, the size of an HTML table is also limited. Therefore, if a knowledge base has more than 10000 worlds, the table is split into multiple files, where each file contains at most 10000 worlds. Moreover, the dynamic nature of HTML is utilized, e.g., to fade in additional information on demand, or to provide a paging mechanism on the rows of the table. In the following, we will illustrate the $\mathcal{PU}$ report for the knowledge base $\mathcal{R}$ from Example 1.

The first part of the report consists of the knowledge base and its Herbrand base, along with a navigation panel (Figure 1). Subsequently, the individual steps of the $\mathcal{PU}$ transformation process are shown, with the following information being displayed for each transformation step (cf. Figure 2): The substituted conditional, the two newly added conditionals, the resulting knowledge base, the interaction type being resolved by the transformation step, and a brief explanation of the interaction. Another part of the report is a table representing all knowledge bases arising from the transformation process and the corresponding conditional structures for all worlds (Figure 3). For each world, its probability under the ME model $P^*_{\mathcal{R}}$ is given. Detailed information is available on demand, e.g., a listing of all groundings of a conditional being verified or falsified by a particular world (Figure 4).

The induced equivalence classes of each knowledge base arising from the transformation process are represented in separate tables. Each row represents one equivalence class

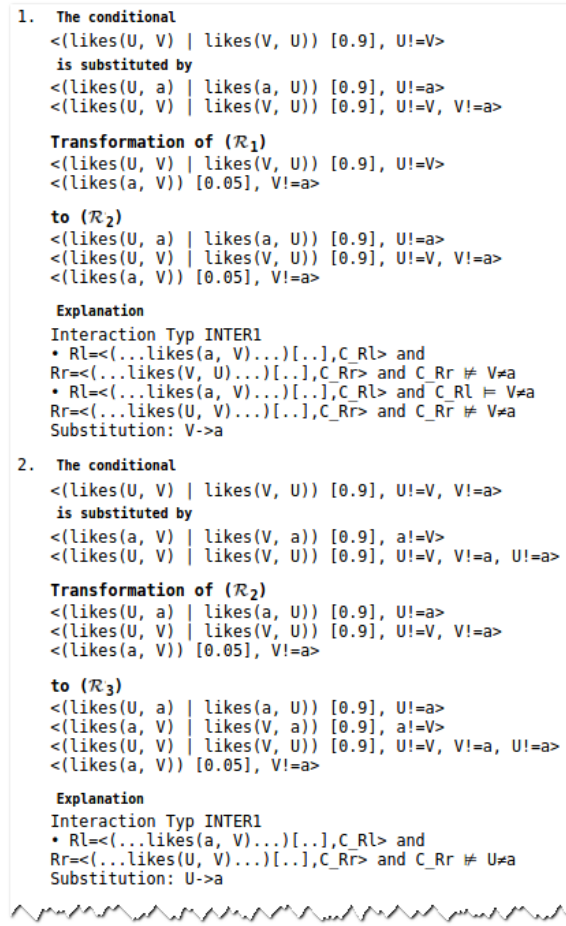


Figure 2: $\mathcal{PU}$ transformation steps from $\mathcal{R} = \mathcal{R}_1$ to $\mathcal{R}_2$ and from $\mathcal{R}_2$ to $\mathcal{R}_3 = \mathcal{PU}(\mathcal{R})$ for $\mathcal{R}$ from Example 1

and includes the conditional structure, its worlds, and the number of different probabilities of its worlds. It is possible to fade in the occurring probabilities and the ground atoms

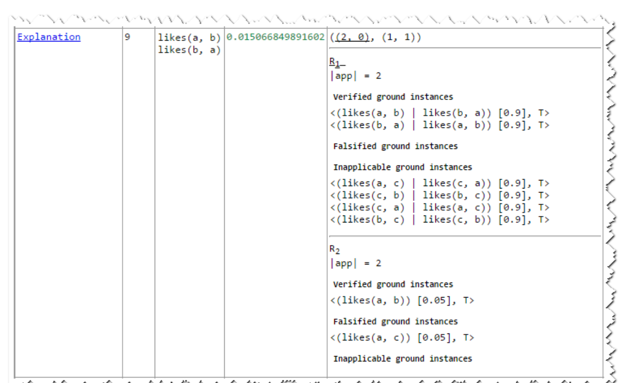


Figure 4: Explanation of the conditional of structure of the knowledge base $\mathcal{R} = \{R_1, R_2\}$ w.r.t. a specific world

An underlined conditional is substituted in the next transformation step. Bold conditionals were added to $\mathcal{R}_X$ because of the substitutions in the previous transformation step.
An underlined vf-pair corresponds to the conditional that is substituted in the next transformation step. Bold vf-pairs correspond to the conditionals, which were added to $\mathcal{R}_X$ because of the previous substitution.

Show 25 ▾ entries

	Index	ω	Probability	Conditional Structure $\gamma_{\mathcal{R}_1}(\omega)$	Conditional Structure $\gamma_{\mathcal{R}_2}(\omega)$	Conditional Structure $\gamma_{\mathcal{R}_3}(\omega)$

				$\mathcal{R}_1$ = {
				R1 = <(likes(U, V) | likes(V, U)) [0.9], U!=V>
				R2 = <(likes(a, V)) [0.05], V!=a>
				}
					$\mathcal{R}_2$ = {
					R1-1 = <(likes(U, a) | likes(a, U)) [0.9], U!=a>
					R1-2 = <(likes(U, V) | likes(V, U)) [0.9], U!=V, V!=a>
					R2 = <(likes(a, V)) [0.05], V!=a>
					}
						$\mathcal{R}_3$ = {
						R1-1 = <(likes(U, a) | likes(a, U)) [0.9], U!=a>
						R1-2-1 = <(likes(a, V) | likes(V, a)) [0.9], a!=V>
						R1-2-2 = <(likes(U, V) | likes(V, U)) [0.9], U!=V, V!=a, U!=a>
						R2 = <(likes(a, V)) [0.05], V!=a>
						}

Number of equivalence classes				22	26	30
Display gnd($\mathcal{R}$)						
Explanation	1		0.316403847513587	((0, 0), (0, 2))	((0, 0), (0, 0), (0, 2))	((0, 0), (0, 0), (0, 0), (0, 2))
Explanation	2	likes(a, b)	0.001674094422022	((0, 1), (1, 1))	((0, 1), (0, 0), (1, 1))	((0, 1), (0, 0), (0, 0), (1, 1))
Explanation	3	likes(a, c)	0.001674094422022	((0, 1), (1, 1))	((0, 1), (0, 0), (1, 1))	((0, 1), (0, 0), (0, 0), (1, 1))
Explanation	4	likes(b, a)	0.001674094443827	((0, 1), (0, 2))	((0, 0), (0, 1), (0, 2))	((0, 0), (0, 1), (0, 0), (0, 2))
Explanation	5	likes(b, c)	0.052420104764137	((0, 1), (0, 2))	((0, 0), (0, 1), (0, 2))	((0, 0), (0, 0), (0, 1), (0, 2))
Explanation	6	likes(c, a)	0.001674094443827	((0, 1), (0, 2))	((0, 0), (0, 1), (0, 2))	((0, 0), (0, 1), (0, 0), (0, 2))
Explanation	7	likes(c, b)	0.052420104764137	((0, 1), (0, 2))	((0, 0), (0, 1), (0, 2))	((0, 0), (0, 0), (0, 1), (0, 2))
Explanation	8	likes(a, b) likes(a, c)	0.000008857642396	((0, 2), (2, 0))	((0, 2), (0, 0), (2, 0))	((0, 2), (0, 0), (0, 0), (2, 0))
Explanation	9	likes(a, b) likes(b, a)	0.015066849891602	((2, 0), (1, 1))	((1, 0), (1, 0), (1, 1))	((1, 0), (1, 0), (0, 0), (1, 1))
Explanation	10	likes(a, b) likes(b, c)	0.000277355050127	((0, 2), (1, 1))	((0, 1), (0, 1), (1, 1))	((0, 1), (0, 0), (0, 1), (1, 1))

Figure 3: Comparison of conditional structures induced by $\mathcal{PU}$ transformation process

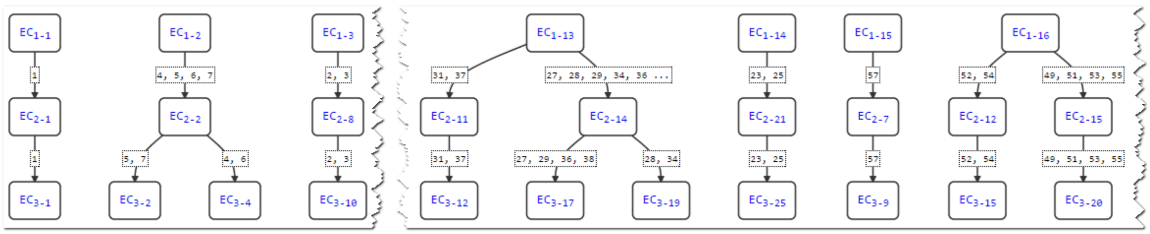


Figure 6: Visualization of relations among equivalence classes of worlds induced by a $\mathcal{PU}$ transformation process

	Equivalence Class	Conditional Structure	World Indices	Number of different Probabilities	Cardinality
Display Atoms Display Probabilities	EC1-1	((0, 0), (0, 2))	1	1	1
Display Atoms Display Probabilities	EC1-2	((0, 1), (0, 2))	4, 5, 6, 7	2	4
Hide Atoms Hide Probabilities	EC1-3	((0, 1), (1, 1))	2, 3 World 2 likes(a, b) World 3 likes(a, c)	1 0.001674094422022 0.001674094422022	2
Display Atoms Display Probabilities	EC1-4	((0, 2), (0, 2))	17, 18, 19, 20, 22	2	5

Figure 5: Visualization of equivalence classes of worlds

of each world in one equivalence class (Figure 5). E.g., the class EC_{1-2} (w.r.t. $\mathcal{R}_1$) contains the worlds 4,5,6,7 showing two different probabilities. Figure 3 reveals that these worlds have the same conditional structure w.r.t. both $\mathcal{R}_1$ and $\mathcal{R}_2$. However, as implied by Proposition 2, the conditional structure w.r.t. $\mathcal{R}_3$ accurately distinguishes the worlds 4, 6 and 5, 7, respectively, since $\mathcal{R}_3$ is parametrically uniform.

The relations of equivalence classes among each other are visualized by a graph (Figure 6), with the classes in the i-th row corresponding to the knowledge base before the i-th transformation step. For instance, both the classes EC_{1-2} induced by $\mathcal{R}_1$ and EC_{2-2} induced by $\mathcal{R}_2$ contain the worlds 4, 5, 6, 7. EC_{2-2} is then split up into EC_{3-2} with worlds 5, 7 and EC_{3-4} with worlds 4, 6. It is possible to navigate directly to an equivalence class by clicking on the label of a node. An edge is labeled with the indices of the first five common worlds of the corresponding equivalence classes.

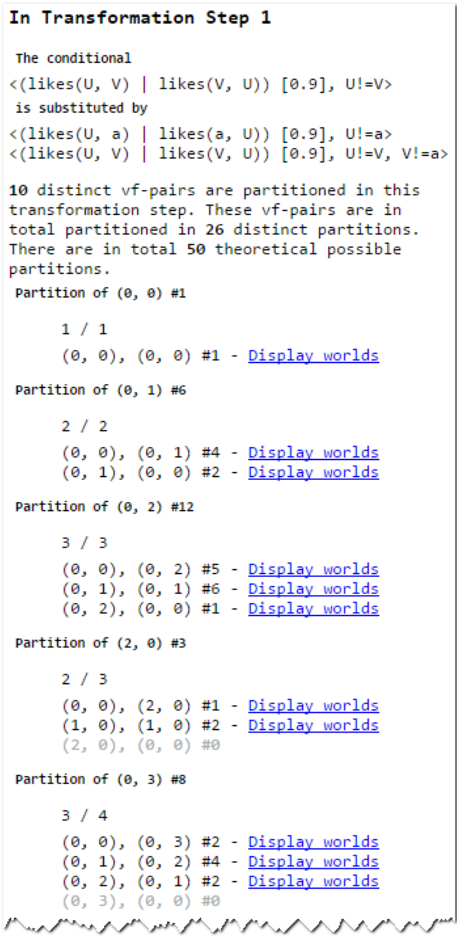
Since in each $\mathcal{PU}$ transformation step one conditional is replaced by two conditionals, one vf-pair of each conditional structure of a knowledge base is partitioned into two vf-pairs by each transformation step (Figure 7).

Example 4. *The first $\mathcal{PU}$ transformation step for $\mathcal{R}$ from Example 1 replaces R_1 by two conditionals (cf. Example 2). The conditional structure of $\omega = \{likes(a, b), likes(b, a)\}$ is transformed correspondingly from $((\mathbf{2}, \mathbf{0}),\ (1, 1))$ for $\mathcal{R}$ (cf. Example 3) to $((\mathbf{1}, \mathbf{0}),\ (\mathbf{1}, \mathbf{0}),\ (1, 1))$ for the knowledge base resulting from the first $\mathcal{PU}$ transformation step.*

The other two possible partitionings of the vf-pair $(2, 0)$ are $((0, 0),\ (2, 0))$ and $((2, 0),\ (0, 0))$, where the latter does not occur. This information is also given, together with the information how many (and which) worlds share the same vf-pairs, both before and after the transformation step.

5 Conclusions and Further Work

In order to support a better understanding of conditional structures and of parametric uniformity – two key concepts for ME model computation and reasoning for relational probabilistic knowledge bases – we presented the software system *CSPU* computing, visualizing, and analyzing them. *CSPU* has already been a valuable tool in experimenting with different knowledge bases and evaluating them with respect to several important parameters. These directly address the questions raised at the end of Section 3 and include the number of equivalence classes and of vf-pairs, the partitioning of vf-pairs, and the comparison of first-order conditionals and their instances under maximum entropy


In Transformation Step 1

The conditional
<(likes(U, V) | likes(V, U)) [0.9], U!=V>
is substituted by
<(likes(U, a) | likes(a, U)) [0.9], U!=a>
<(likes(U, V) | likes(V, U)) [0.9], U!=V, V!=a>

10 distinct vf-pairs are partitioned in this transformation step. These vf-pairs are in total partitioned in 26 distinct partitions. There are in total 50 theoretical possible partitions.

Partition of (0, 0) #1

1 / 1
(0, 0), (0, 0) #1 - Display worlds

Partition of (0, 1) #6

2 / 2
(0, 0), (0, 1) #4 - Display worlds
(0, 1), (0, 0) #2 - Display worlds

Partition of (0, 2) #12

3 / 3
(0, 0), (0, 2) #5 - Display worlds
(0, 1), (0, 1) #6 - Display worlds
(0, 2), (0, 0) #1 - Display worlds

Partition of (2, 0) #3

2 / 3
(0, 0), (2, 0) #1 - Display worlds
(1, 0), (1, 0) #2 - Display worlds
(2, 0), (0, 0) #0

Partition of (0, 3) #8

3 / 4
(0, 0), (0, 3) #2 - Display worlds
(0, 1), (0, 2) #4 - Display worlds
(0, 2), (0, 1) #2 - Display worlds
(0, 3), (0, 0) #0

Figure 7: Partitionings of vf-pairs induced by $\mathcal{PU}$

semantics (Beierle, Finthammer, and Kern-Isberner 2015). Our future work also includes the investigation of the exact relationship of techniques of lifted inference (Poole 2003; de Salvo Braz, Amir, and Roth 2005) to the concepts presented in this paper.

References

Beierle, C., and Krämer, A. 2015. Achieving parametric uniformity for knowledge bases in a relational probabilistic conditional logic with maximum entropy semantics. *Annals of Mathematics and Artificial Intelligence* 73(1-2):5–45.

Beierle, C.; Finthammer, M.; and Kern-Isberner, G. 2015. Relational probabilistic conditionals and their instantiations under maximum entropy semantics for first-order knowledge bases. *Entropy* 17(2):852–865.

Boyd, S., and Vandenberghe, L. 2004. *Convex Optimization*. New York, NY, USA: Cambridge University Press.

de Salvo Braz, R.; Amir, E.; and Roth, D. 2005. Lifted first-order probabilistic inference. In *IJCAI-05*, 1319–1325. Professional Book Center.

Fagin, R., and Halpern, J. Y. 1994. Reasoning about knowledge and probability. *J. ACM* 41(2):340–367.

Fagin, R.; Halpern, J. Y.; and Megiddo, N. 1990. A logic for reasoning about probabilities. *Information and Computation* 87(1-2):78 – 128.

Finthammer, M., and Beierle, C. 2012a. How to exploit parametric uniformity for maximum entropy reasoning in a relational probabilistic logic. In *JELIA 2012*, volume 7519 of *LNAI*, 189–201. Springer.

Finthammer, M., and Beierle, C. 2012b. Using equivalences of worlds for aggregation semantics of relational conditionals. In *KI 2012*, LNAI 7526, 49–60. Springer.

Finthammer, M., and Thimm, M. 2012. An integrated development environment for probabilistic relational reasoning. *Logic Journal of the IGPL* 20(5):831–871.

Fisseler, J. 2012. First-order probabilistic conditional logic and maximum entropy. *Logic Journal of the IGPL* 20(5):796–830.

Geman, S., and Geman, D. 1984. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6:721–741.

Halpern, J. 2005. *Reasoning About Uncertainty*. MIT Press.

Kern-Isberner, G., and Lukasiewicz, T. 2004. Combining probabilistic logic programming with the power of maximum entropy. *Artif. Intell.* 157(1-2):139–202.

Kern-Isberner, G., and Thimm, M. 2010. Novel semantical approaches to relational probabilistic conditionals. In *Proc. KR-2010*, 382–391. Menlo Park, CA: AAAI Press.

Kern-Isberner, G., and Thimm, M. 2012. A ranking semantics for first-order conditionals. In *ECAI-2012*, 456–461. IOS Press.

Kern-Isberner, G. 1998. Characterizing the principle of minimum cross-entropy within a conditional-logical framework. *Artificial Intelligence* 98:169–208.

Kern-Isberner, G. 2001. *Conditionals in nonmonotonic reasoning and belief revision*. Springer, Lecture Notes in Artificial Intelligence LNAI 2087.

Nilsson, N. 1986. Probabilistic logic. *Artificial Intelligence* 28:71–87.

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

Poole, D. 2003. First-order probabilistic inference. In Gottlob, G., and Walsh, T., eds., *Proc. IJCAI-03*, 985–991. Morgan Kaufmann.