On the Problem of Grounding a Relational Probabilistic Conditional Knowledge Base

Sebastian Loh and **Matthias Thimm** and **Gabriele Kern-Isberner**
Technische Universität Dortmund, Department of Computer Science, Germany

Abstract

For first-order probabilistic knowledge representation, grounding is an important means to define a semantics for knowledge bases which extends the propositional semantics. However, naive approaches to grounding may give rise to conflicts and inconsistencies, in particular, if the formalism involves point probabilities, in contrast to using approximative or interval-based probabilities. In this paper, we formulate properties that can guide the search for suitable grounding operators. Moreover, we present three operators the most sophisticated of which implements a stratified use of a specificity relation so that more specific information on objects is given priority over less specific information.

Introduction

Probabilistic reasoning in relational representations of knowledge is a very active and controversy research area. In the past few years the fields of *probabilistic inductive logic programming* and *statistical relational learning* put forth a lot of proposals that deal with combining traditional probabilistic models of knowledge like Bayes nets or Markov nets (Pearl, 1998) with first-order logic, see (Getoor and Taskar, 2007) for an introduction. In most of these proposals this is done by appropriately grounding the parts of the knowledge base that are needed for answering a particular query and treating this grounded parts as a propositional knowledge base. In this paper, we address the general problem of grounding relational knowledge bases in order to apply a propositional probabilistic reasoner. In particular, we are interested in finding strategies that yield consistent ground knowledge bases for problematic relational representations of knowledge. To this end we employ a relational extension of probabilistic conditional logic.

In (propositional) probabilistic conditional logic knowledge is captured using conditionals of the form $(\phi \,|\, \psi)[\alpha]$ with some formulas ϕ, ψ of a given propositional language and $\alpha \in [0,1]$. A probabilistic conditional of this form partially describes an (unknown) probability distribution P^* by stating that $P^*(\phi \,|\, \psi) = \alpha$ holds. In contrast to Bayes nets probabilistic conditional logic does not demand to fully describe a probability distribution but only to state constraints on it. Reasoning can be performed in probabilistic conditional logic by performing model-based inference based on the one probability distribution with maximum entropy (Kern-Isberner, 2001) which satisfies several desirable properties for commonsense reasoning.

The contribution of this paper is threefold. First, we introduce a relational extension of probabilistic conditional logic and define its semantics. Second, we employ the principle of maximum entropy to reason within this relational extension of probabilistic conditional logic. The final and most significant contribution of this paper lies in investigating the problem of grounding relational knowledge bases and thus in filling the gap between the first two contributions. Usually, applying universal instantiation on probabilistic first-order knowledge bases yields an inconsistent ground knowledge base that is useless for reasoning. Consider the probabilistic knowledge base $\{(\mathit{likes}(X,Y) \,|\, \mathit{elephant}(X) \wedge \mathit{keeper}(X))[0.9], (\mathit{likes}(X,\mathit{fred}) \,|\, \mathit{elephant}(X))[0.3]\}$ representing the commonsense knowledge that elephants usually like their keeper, but they mostly do not like keeper Fred. When instantiating variable X with any elephant, and variable Y with $fred$, then a contradiction arises, as the probabilities of the corresponding rules are not compatible. Our approach here lies in constraining the instantiations of probabilistic rule schemata in order to prevent the grounded knowledge base from becoming inconsistent. To this end, we develop a list of desirable properties of such "grounding operators" that capture the idea that the grounding of a relational knowledge base should be consistent and maximal. We propose three strategies that aim at removing instances of conditionals that are problematic for consistency in some sense. The final of these applies an implicit specificity ordering of conditionals to prioritize their instantiations, thus realizing the possibility to model default rules with exceptions.

The rest of this paper is organized as follows. In the next section we introduce a first-order conditional logic and present our semantical approach. After that, we discuss the problem of grounding knowledge bases by first developing a list of desirable properties of grounding operators and then propose three different variants of grounding operators. Finally, we briefly review some related work and conclude.

First-Order Probabilistic Conditional Logic

As a base of our probabilistic language, we use a fragment $\mathcal{L}$ of a first-order language over a signature Σ containing only predicates and constants, the formulas of which are wellformed according to the usual standards, but without any quantifiers. $\mathcal{L}$ may be sorted, i. e. the constants $U = U_{\mathcal{L}}$ and the variables $V = V_{\mathcal{L}}$ of $\mathcal{L}$ are partitioned into different sorts, and the arguments of the predicates may be sorted as well, so that only part of all possible instantiations via grounding substitutions are allowed. A grounding substitution $\theta : V_{\mathcal{L}} \to U_{\mathcal{L}}$ instantiates variables with constants. It is extended to formulas in the usual way, e. g. we define $\theta(p(X,Y) \wedge q(X)) = p(\theta(X), \theta(Y)) \wedge q(\theta(X))$. A grounding substitution θ is *legal* if any variable of sort S in r is mapped to a constant of sort S. We extend this relational language $\mathcal{L}$ to a probabilistic conditional language by introducing conditionals and probabilities.

Definition 1. A *relational probabilistic conditional* r is an expression of the form $r = (\phi \,|\, \psi)[\alpha]$ with formulas $\phi, \psi \in \mathcal{L}$ and $\alpha \in [0,1]$. The conditional and the probabilistic parts of r are denoted by $Cnd(r) = (\phi \,|\, \psi)$ and $Pr(r) = \alpha$, respectively.

Conditionals $r = (\phi \,|\, \psi)[\alpha]$ and $r' = (\phi' \,|\, \psi')[\alpha']$ are *qualitatively equivalent*, denoted by $Cnd(r) \equiv Cnd(r')$, if both $\psi \equiv \psi'$ and $\psi \wedge \phi \equiv \psi' \wedge \phi'$ hold, cf. (Kern-Isberner, 2001).

A conditional r is called *ground* iff r contains no variables. Non-ground conditionals can be grounded by legal grounding substitutions. The language of all relational probabilistic conditionals is denoted by $(\mathcal{L} \,|\, \mathcal{L})$, and the restricted language on all ground conditionals using constants from U is denoted as $(\mathcal{L} \,|\, \mathcal{L})_U$. A set $\mathcal{R}$ of relational probabilistic conditionals is called a *knowledge base*.

Example 1. As a running example in this paper, we consider the language $\mathcal{L}_{zoo}$ with a set of constants $U = U_{Elephant} \cup U_{Keeper}$ with $U_{Elephant} = \{dumbo, clyde, tuffi\}$, $U_{Keeper} = \{fred, hank\}$ and predicates $likes(Elephant, Keeper)$ and $givesPeanuts(Keeper, Elephant)$, the arguments of which have to be of the proper sort. Then both $likes(dumbo, hank)$ and $\neg likes(dumbo, fred) \vee likes(clyde, hank)$ are formulas in $\mathcal{L}_{zoo}$ while $likes(hank, hank)$ is not. Moreover, $likes(X,Y)$ with $X \in V_{Elephant}$ and $Y \in V_{Keeper}$ is a formula of the language.

Let us assume that usually elephants like their keepers. But if a keeper does not give his elephant peanuts on a regular basis he makes himself unpopular. Then our knowledge base can be given as

$$\mathcal{R} = \{(likes(X,Y) \,|\, \neg givesPeanuts(Y,X))[0.1], \\ (likes(X,Y))[0.9], (givesPeanuts(hank, X))[1.0]\}$$

Ground conditionals $r \in (\mathcal{L} \,|\, \mathcal{L})_U$ can be interpreted as in the propositional case, i. e. $P \models (\phi|\psi)[\alpha]$ iff $P(\phi|\psi) = \alpha$ and $P(\psi) > 0$, where P is a probability distribution over the Herbrand base of $\mathcal{L}$, and $P(\phi) = \sum_{\omega \in \Omega, \omega \models \phi} P(\omega)$ for any classical ground formula ϕ. Here Ω is the set of all Herbrand interpretations of $\mathcal{L}$ and provides a possible worlds semantics for the classical part of $\mathcal{L}$. Non-conditional formulas $(\phi)[\alpha]$ can be considered consistently as conditionals with tautological premise $(\phi \,|\, \top)[\alpha]$, so that no explicit distinction between conditionals and flat formulas is necessary in the following. In the general case, if r contains variables, different groundings may yield different conditional probability values. This could be handled by assigning probabilistic intervals to open conditionals, as is done e. g. in (Kern-Isberner and Lukasiewicz, 2004). In our framework, however, we stick to pointwise probabilities, so basically, all groundings of a relational conditional should hold with the same probability. But this may give rise to conflicts among the formulas of a knowledge base, as the following example shows.

Example 2. We consider the knowledge base $\mathcal{R}$ from Ex. 1 and extend $\mathcal{R}$ with the specific information about the keeper Fred. We observed that Fred is not that popular among the elephants. Therefore we add the conditional $(likes(X, fred))[0.3]$ to the knowledge base $\mathcal{R}$. But instantiations of this new conditional for elephants X conflict with instantiations of $(likes(X,Y)[0.9]$ for $Y = fred$. For instance, for $X = dumbo$, the constraints in $\mathcal{R}$ imply $P(likes(dumbo, fred)) = 0.3$ and $P(likes(dumbo, fred)) = 0.9$, which can not hold at the same time.

In the example above, obviously, Fred is an exceptional keeper, and a plain grounding is not able to handle such exceptions appropriately. Hence we need more elaborate grounding strategies to base semantics for the first-order probabilistic language $(\mathcal{L} \,|\, \mathcal{L})$ on propositional grounding. For any set S, let $\mathfrak{P}(S)$ denote the power set of a S.

Definition 2. A *grounding operator* (GOP) $\mathcal{G}$ is a function $\mathcal{G} : \mathfrak{P}((\mathcal{L} \,|\, \mathcal{L})) \to \mathfrak{P}((\mathcal{L} \,|\, \mathcal{L})_U)$.

A GOP $\mathcal{G}$ takes a general relational knowledge base $\mathcal{R}$ and maps it to a ground one $\mathcal{G}(\mathcal{R})$ by instantiating variables according to the language of the knowledge base and some strategy. By doing so we may use the propositional probabilistic semantics for the first-order case. If P is a probability distribution over the Herbrand base of $\mathcal{L}$ and r is a relational probabilistic conditional, then we define $P \models_{\mathcal{G}} r$ iff $P \models r^*$ for all $r^* \in \mathcal{G}(\{r\})$. This is a common method in relational probabilistic knowledge representation, cf. e. g. the Markov logic networks (Getoor and Taskar, 2007; Ch. 12), or (Fisseler, 2009). However, in our framework of using declarative probabilistic conditionals as constraints, we have to find reasonable grounding strategies. The actual definition of a GOP relies on grounding substitutions for variables. For a conditional r let $\Gamma(r)$ denote the set of a legal grounding substitutions for r. The most simple approach to ground a knowledge base is *universal instantiation* which naively instantiates every variable with any constant of the same sort.

Definition 3. The *naive grounding operator* $\mathcal{G}_U$ is defined as $\mathcal{G}_U(\mathcal{R}) := \{\theta(r) \mid r \in \mathcal{R}, \theta \in \Gamma(r)\}$.

Although the universal instantiation may give rise to conflicts, as Ex. 2 shows, it will serve as a foundation for other grounding strategies to be developed in the next section.

As a (consistent) knowledge base $\mathcal{R}$ usually specifies incomplete information, one often is interested in applying inductive representation techniques that help computing a single probability distribution which describes $\mathcal{R}$ in a most appropriate way, giving a complete description of the problem area at hand. This can be done using methods based on maximum entropy, which feature several nice properties (Kern-Isberner, 2001). The *entropy* H is an information-theoretic measure on probability distributions and is defined as $H(P) = -\sum_{\omega\in\Omega} P(\omega)\log P(\omega)$. By employing the *principle of maximum entropy* one can determine the single probability distribution that is the optimal model for a knowledge base R in an information-theoretic sense:

$$P_R^{ME} = \arg\max_{P\models R} \mathcal{H}(P)$$

This principle can also be applied to consistent ground relational knowledge bases $\mathcal{G}(\mathcal{R})$, therefore $P^{ME}_{\mathcal{G}(\mathcal{R})}$ can be calculated in this case. So, putting together the grounding strategy employed by an operator $\mathcal{G}$ and the maximum entropy principle, we obtain the expressive *relational maximum entropy semantics* (RME) for conditional probabilistic knowledge bases: A conditional $q \in (\mathcal{L}\,|\,\mathcal{L})$ is *RME-entailed by the knowledge base* $\mathcal{R}$ *under the GOP* $\mathcal{G}$ iff

$$\mathcal{R} \models^{ME}_{\mathcal{G}} q \quad \text{iff} \quad P^{ME}_{\mathcal{G}(\mathcal{R})} \models_{\mathcal{G}} q, \tag{1}$$

i. e. iff for all $q^* \in \mathcal{G}(\{q\})$, it holds that $P^{ME}_{\mathcal{G}(\mathcal{R})} \models q^*$.

Reasoning under RME semantics can be divided into three steps: 1.) ground the knowledge base with a GOP $\mathcal{G}$, 2.) calculate the probability distribution $P^{ME}_{\mathcal{G}(R)}$ with maximum entropy for the grounded instance $\mathcal{G}(R)$, and 3.) calculate whether the grounding of a given query is fulfilled by the distribution $P^{ME}_{\mathcal{G}(R)}$. An overview of this process is given in Fig. 1.

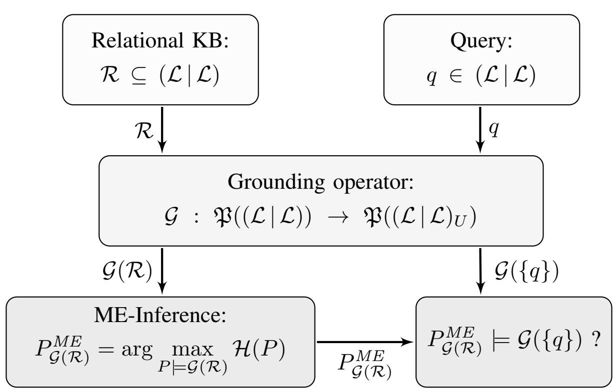

Figure 1: The RME inference process

Grounding Relational Knowledge Bases

In this section we investigate several possibilities for grounding relational knowledge bases. We already introduced the most simple grounding strategy—the universal instantiation—in the previous section. But when dealing with uncertain and incomplete information, universal instantiation might not always be an appropriate means to treat relational probabilistic conditionals. As in default logic, one often wants to be able to represent exceptions to rules and universal instantiation often yields an inconsistent and therefore useless ground knowledge.

Example 3. We continue Ex. 2 and consider the knowledge base $\mathcal{R} = \{r_1, r_2, r_3\}$ with $r_1 = (\mathit{likes}(X,Y))[0.9]$, $r_2 = (\mathit{likes}(X, \mathit{fred}))[0.3]$, and $r_3 = (\mathit{likes}(\mathit{clyde}, \mathit{fred}))[1]$. Obviously, the naive grounding of $\mathcal{R}$ is inconsistent. In the following, we will use this example to evaluate our GOPs.

Grounding a relational knowledge base with a specific grounding strategy is of crucial concern as the result, i. e. the ground knowledge base, directly influences the possible inferences of the original knowledge base via $\models^{ME}_{\mathcal{G}}$. In the following, we develop several common sense properties a meaningful GOP $\mathcal{G}$ should satisfy. To this end, let $\mathcal{R}$ denote a knowledge base and let $\mathcal{G} : \mathfrak{P}((\mathcal{L}\,|\,\mathcal{L})) \to \mathfrak{P}((\mathcal{L}\,|\,\mathcal{L})_U)$ be a GOP.

Our first property relates the relational case of probabilistic reasoning to the propositional case.

(Compatibility) If r is ground then $\mathcal{G}(\{r\}) = \{r\}$.

Property (Compatibility) ensures that ground knowledge should always be preserved when grounding a knowledge base.

The rationale behind the next property is that ground knowledge represents factual knowledge on specific individuals and should not be neglected.

(Stability) If $r \in \mathcal{R}$ is ground then $r \in \mathcal{G}(\mathcal{R})$.

Notice the difference between the properties (Compatibility) and (Stability). While (Compatibility) is only concerned with the behavior of a single ground rule, (Stability) ensures that ground probabilistic knowledge is preserved in the whole knowledge base.

The following property constrains our investigation to natural GOPs, in the sense that conditionals are just instantiated but no probabilities are modified.

(Structural Preservation) For each $r^* \in \mathcal{G}(\mathcal{R})$ there is an $r \in \mathcal{R}$ and $\theta \in \Gamma(r)$ such that $r^* = \theta(r)$.

Satisfaction of (Structural Preservation) ensures that the qualitative part of a conditional is only modified by instantiating variables and the quantitive part is not modified under the grounding process. If (Structural Preservation) is fulfilled for $\mathcal{G}$ each ground conditional in $\mathcal{G}(\mathcal{R})$ can be traced back to at least one relational conditional from which it was instantiated. In this case, we denote by $r \to^{\mathcal{R}}_{\mathcal{G}} r^*$ that conditional $r^* \in \mathcal{G}(\mathcal{R})$ has been instantiated from conditional $r \in \mathcal{R}$ (there may be multiple r satisfying this property).

Our first three properties are related as follows.

Proposition 1. *If* $\mathcal{G}$ *satisfies (Stability) and (Structural Preservation), then* $\mathcal{G}$ *satisfies (Compatibility).*

While the above properties are concerned with minimality conditions on the result of grounding the next property deals with a maximum condition.

(Upper Bound) It holds that $\mathcal{G}(\mathcal{R}) \subseteq \mathcal{G}_U(\mathcal{R})$.

The property (Upper Bound) ensures that the universal instantiation is an upper bound for any reasonable GOP. In particular, this means that no new conditional is "invented" by $\mathcal{G}$. One can see there is a trivial relationship between the property (Upper Bound) and (Structural Preservation).

Proposition 2. *A GOP $\mathcal{G}$ satisfies (Structural Preservation) if and only if $\mathcal{G}$ satisfies (Upper Bound).*

Grounding a knowledge base makes only sense when there is a chance that the resulting ground knowledge base is consistent. An obvious flaw of a knowledge base is the existence of *direct conflicts*, i. e. conditionals that are qualitatively equivalent but differ in their probabilities.

Definition 4. Two conditionals r, r' are in *direct conflict*, denoted by $r \perp r'$, if $Cnd(r) \equiv Cnd(r')$ and $Pr(r) \neq Pr(r')$.

For example, the two conditionals $(p(X) \mid q(X))[0.3]$ and $(p(Y) \mid q(Y))[0.4]$ are in direct conflict. A knowledge base R has a direct conflict if there are two conditionals $r, r' \in R$ that are in direct conflict. The existence of conditionals that are in direct conflict renders these two conditionals useless and we require any GOP to map such corrupt knowledge bases to the empty knowledge base.

(Rationality) If $\mathcal{R}$ has a direct conflict then $\mathcal{G}(\mathcal{R}) = \emptyset$.

The demand for (Rationality) makes clear that we do not address merging problems here, our strategies aim at avoiding syntactical conflicts.

A crucial property of a GOP $\mathcal{G}$ is that it achieves a consistent grounding, i. e. a ground knowledge that has at least one model.

(Consistency) $\mathcal{G}(\mathcal{R})$ is consistent for any $\mathcal{R}$.

However, consistency is hard to achieve as it does not only depend on the grounding strategy but on the probabilistic structure of the knowledge base itself. Consider the following example.

Example 4. Let $\mathcal{R} = \{(a \mid b)[1], (a)[0], (b)[1]\}$. $\mathcal{R}$ does not contain any predicates with arity greater zero. So any GOP $\mathcal{G}$ that satisfies at least (Stability) yields $\mathcal{R} \subseteq \mathcal{G}(\mathcal{R})$. Furthermore, $\mathcal{R}$ is inherently inconsistent and without neglecting one (ground) conditional or manipulating the probabilities, consistency cannot be achieved in $\mathcal{G}(\mathcal{R})$.

Example 4 suggests that the demand for general consistency is too hard for reasonable GOPs. In the following we propose two weakened forms of consistency that are more rational for a meaningful GOP. The first property ensures that there are no direct conflicts in the ground knowledge base, i. e. conditionals that have the same qualitative structure but differ in their quantitative structure. Obviously, any ground knowledge base that has a direct conflict is inherently inconsistent.

(Conflict Freeness) $\mathcal{G}(\mathcal{R})$ has no direct conflicts.

(Conflict Freeness) is a strictly weaker property as (Consistency) as the following proposition shows.

Proposition 3. *If $\mathcal{G}$ satisfies (Consistency) then $\mathcal{G}$ satisfies (Conflict Freeness).*

The other direction is not always true as Ex. 4 showed. There, $\mathcal{G}(\mathcal{R})$ has no direct conflict but still is inconsistent.

(Maximality) Let $\mathcal{G}$ satisfy (Conflict Freeness). For all $r \in \mathcal{G}_U(\mathcal{R}) \setminus \mathcal{G}(\mathcal{R})$ it holds that $\mathcal{G}(\mathcal{R}) \cup \{r\}$ has a direct conflict.

The property (Maximality) further refines the notion of conflict freeness as we further impose the grounding to be maximal.

The above properties are reasonable demands for a GOP. In the rest of this subsection, we will discuss further properties that appear reasonable at first sight. One of the simplest properties for functions in general is monotonicity.

(Monotonicity) For any $\mathcal{R} \subseteq \mathcal{R}'$ it holds $\mathcal{G}(\mathcal{R}) \subseteq \mathcal{G}(\mathcal{R}')$.

For the same reasons as in default logic, (Monotonicity) is not a desirable property when dealing with rules with exceptions. A GOP that satisfies (Monotonicity) is not able to appropriately reason, e. g., with the knowledge base in Ex. 3. There, a knowledge base just consisting of the conditional $(likes(X, Y))[0.9]$ should yield the universal instantiation as grounding for any rational GOP. As a consequence, any operator fulfilling (Monotonicity) and at least (Conflict Freeness) must completely neglect the exceptional knowledge of any extension of this knowledge base. This discussion also leads to the consideration of the following property.

(Minimality) Let $\mathcal{G}$ satisfy (Structural Preservation). For each conditional $r \in \mathcal{R}$ there exists an $r^* \in \mathcal{G}(\mathcal{R})$ with $r \rightarrow_{\mathcal{G}}^{\mathcal{R}} r^*$.

(Minimality) ensures that no conditional is ignored when grounding a knowledge base. In general, this demand seems appropriate but highly depends on the population under consideration. As a counterexample to this demand (taken from (Delgrande, 1998)) consider the conditional "Lemons are yellow". But due to a rare disease, every lemon in our domain is actually green and thus an exception to the conditional. Although the conditional represents default knowledge the actual population might not lead to any true instance.

Cautious Grounding

Recalling Ex. 3 one can see that the instantiations of different conditionals using the constants $clyde$ and $fred$ are in direct conflict. As Clyde and Fred are exceptional individuals the instantiation of conditionals r_1 and r_2 using these constants should be prohibited. Our first approach to achieve this is to completely ignore individuals that already appear within the knowledge base when instantiating the conditionals. By doing so, only conditionals that mention these individuals in the first place are carried over to the ground knowledge base. We formalize this intuition by defining the *cautious grounding operator* as follows. For this let $im(f)$ denote the image of a function f and $Con(\mathcal{R})$ the set of constants that appear in a knowledge base $\mathcal{R}$.

Definition 5. The *cautious grounding operator* $\mathcal{G}_{ca}$ is defined as

$$\mathcal{G}_{ca}(\mathcal{R}) = \{\theta(r) \mid r \in \mathcal{R},\ \theta \in \Gamma(r),\ im(\theta) \cap Con(\mathcal{R}) = \emptyset\}$$

if $\mathcal{R}$ has no direct conflicts, and $\mathcal{G}_{ca}(\mathcal{R}) = \emptyset$ otherwise.

The cautious GOP is very rigorous in the selection of suitable instantiations for the individual conditionals as no constant is instantiated in any other conditional, even if these conditionals might not cause conflicts in the final ground knowledge base.

Example 5. We continue Ex. 3. When grounding $\mathcal{R}$ with the cautious GOP $\mathcal{G}_{ca}$ this yields the ground knowledge base $\mathcal{G}_{ca}(\mathcal{R})$ depicted in the right column of Tab. 1. Note, that predicates and constants have been abbreviated by their first letters, respectively.

$\mathcal{R}$		$\mathcal{G}_{ca}(\mathcal{R})$
$r_1 : (l(X,Y))[0.9]$	$\rightarrow^{\mathcal{R}}_{\mathcal{G}_{ca}}$	$(l(d,h))[0.9]$
		$(l(g,h))[0.9]$
$r_2 : (l(X,f))[0.1]$	$\rightarrow^{\mathcal{R}}_{\mathcal{G}_{ca}}$	$(l(d,f))[0.3]$
		$(l(g,f))[0.3]$
$r_3 : (l(c,f))[1]$	$\rightarrow^{\mathcal{R}}_{\mathcal{G}_{ca}}$	$(l(c,f))[1]$

Table 1: Cautious grounding

Theorem 1. *The cautious GOP $\mathcal{G}_{ca}$ satisfies (Compatibility), (Stability), (Structural Preservation), (Upper Bound), (Rationality), and (Conflict Freeness).*

Obviously, $\mathcal{G}_{ca}$ does not satisfy (Consistency), cf. Ex. 4. Furthermore, $\mathcal{G}_{ca}$ does not satisfy (Maximality) as in Ex. 5 the instance $(likes(c,h))[0.9]$ of $(likes(X,Y))[0.9]$ can be added to $\mathcal{G}_{ca}(\mathcal{R})$ without violating (Conflict Freeness). The operator $\mathcal{G}_{ca}$ also does not satisfy (Monotonicity) and (Minimality).

Conservative Grounding

The cautious grounding is very rigorous in removing instances of conditionals as Ex. 5 illustrates. A major drawback of this strategy is that it fails to model exceptions adequately. As Ex. 3 shows, the major culprit for an inconsistent ground knowledge base are instances that are in direct conflict. In contrast to the cautious grounding strategy the *conservative grounding strategy* is not founded on the handling of exceptional constants of the knowledge base but on direct conflicts of instances of conditionals itself.

Definition 6. Let $\mathcal{R}$ be a knowledge base. The *conflict set* $\odot(\mathcal{R})$ of $\mathcal{R}$ consists of all conditionals of $\mathcal{R}$ that are in direct conflict: $\odot(\mathcal{R}) = \{r \in \mathcal{R} \mid \exists r' \in \mathcal{R} : r \perp r'\}$.

The conservative grounding strategy simply removes all direct conflicts from the universal instantiation.

Definition 7. The *conservative grounding operator $\mathcal{G}_{co}$* is defined as

$$\mathcal{G}_{co}(\mathcal{R}) = \mathcal{G}_U(\mathcal{R}) \setminus \odot(\mathcal{G}_U(\mathcal{R}))$$

if $\mathcal{R}$ has no direct conflicts, and $\mathcal{G}_{co}(\mathcal{R}) = \emptyset$ otherwise.

Example 6. We continue Ex. 3. When grounding $\mathcal{R}$ with the conservative GOP $\mathcal{G}_{co}$ this yields the ground knowledge base $\mathcal{G}_{co}(\mathcal{R})$ depicted in the right column of Tab. 2.

Theorem 2. *The conservative GOP $\mathcal{G}_{co}$ satisfies (Compatibility), (Structural Preservation), (Upper Bound), (Rationality), and (Conflict Freeness).*

In contrast to the cautious grounding, the conservative grounding does not satisfy (Stability) as can be seen in Ex. 6.

$\mathcal{R}$		$\mathcal{G}_{co}(\mathcal{R})$
$r_1 : (l(X,Y))[0.9]$	$\rightarrow^{\mathcal{R}}_{\mathcal{G}_{co}}$	$(l(c,h))[0.9]$
		$(l(d,h))[0.9]$
$r_2 : (l(X,f))[0.3]$	$\rightarrow^{\mathcal{R}}_{\mathcal{G}_{co}}$	–
$r_3 : (l(c,f))[1]$	$\rightarrow^{\mathcal{R}}_{\mathcal{G}_{co}}$	–

Table 2: Conservative grounding

Specificity Grounding

Our final approach for a rational GOP takes a more sophisticated direction when determining instances of conditionals that should be neglected in the final ground knowledge base. More precisely, we employ *specificity* (Delgrande and Schaub, 1997) as a means to sort out conditionals that carry outdated information when more specific information is available. Our understanding of specificity relies on the subset relation of the universal instantiations of different conditionals. A knowledge base $\mathcal{R}$ is a *qualitative subset* of a knowledge base $\mathcal{R}'$, denoted by $\mathcal{R} \sqsubseteq \mathcal{R}'$, if for every $r \in \mathcal{R}$ we can find an $r' \in \mathcal{R}'$ such that $Cnd(r) \equiv Cnd(r')$. $\mathcal{R}$ is a *strict qualitative subset* of $\mathcal{R}'$, denoted by $\mathcal{R} \sqsubset \mathcal{R}'$, if $\mathcal{R} \sqsubseteq \mathcal{R}'$ and $\mathcal{R}' \not\sqsubseteq \mathcal{R}$.

Definition 8. A conditional r is *less specific* than a conditional r', denoted by $r \prec r'$, if and only if $\mathcal{G}_U(\{r'\}) \sqsubset \mathcal{G}_U(\{r\})$.

Notice that the relation $\prec$ is both asymmetric and transitive. If r is no less specific than r' and r' is no less specific than r, then r and r' are *incomparable*, denoted by $r \nsim r'$.

Example 7. We continue Ex. 3. There, the conditional $(likes(X, fred))[0.3]$ is less specific than $(likes(clyde, fred))[1]$ and $(likes(X,Y))[0.9]$ is less specific than $(likes(X, fred))[0.3]$.

Our aim is to define the specificity GOP to prefer instances of more specific conditionals over instances of less specific conditionals. For instance, in Ex. 3 we want to remove all instances of the least specific conditional $(likes(X,Y))[0.9]$ that are in conflict with instances of the conditional $(likes(X, fred))[0.3]$. However, some conditionals cannot be compared by the specificity relation but their instances might still be conflicting.

Example 8. Consider the conditionals $r'_1 = (likes(clyde, Y))[\alpha]$ and $r'_2 = (likes(X, fred))[\beta]$ with $\alpha \neq \beta$. It holds $r'_1 \nsim r'_2$ but there are also conflicting instances $(likes(clyde, fred))[\alpha]$ and $(likes(clyde, fred))[\beta]$.

To deal with incomparable conditionals we choose a conservative approach. When instances of conditionals are in conflict we choose the instance of the more specific conditional, if possible, and otherwise use the conservative GOP and remove all conflicting instances. Thus, the problematic cases are instances which directly conflict and whose original conditionals are incomparable.

Definition 9. Let $r_1 \rightarrow^{\mathcal{R}}_{\mathcal{G}} r_1^*$ and $r_2 \rightarrow^{\mathcal{R}}_{\mathcal{G}} r_2^*$. If $r_1 \nsim r_2$ and $r_1^* \perp r_2^*$ we say the instance r_1^* is *incomparably conflicting* to the instance r_2^*, denoted by $r_1^* \times r_2^*$.

Notice that $\times$ is symmetric as both $\perp$ and $\not\sim$ are symmetric.

To be able to order instances of conditionals by specificity of these conditionals we start by considering the universal instantiation $\mathcal{G}_U(\mathcal{R})$ of a knowledge base $\mathcal{R}$ and remove any incomparably conflicting instances.

$$\mathcal{R}^{\times} = \mathcal{G}_U(\mathcal{R}) \setminus \{r \in \mathcal{G}_U(\mathcal{R}) \mid \exists r' \in \mathcal{G}_U(\mathcal{R}) : r \times r'\}$$

Now, we can partition $\mathcal{R}^{\times}$ by specificity as follows.

Definition 10. Let $\mathcal{R}$ be a knowledge base. A *legal partitioning* A of $\mathcal{R}^{\times}$ is a partitioning $A = \{A_1, \ldots, A_n\}$ of $\mathcal{R}^{\times}$ such that the following condition holds: for any $r_1^* \in A_i$ and $r_2^* \in A_j$ and any $r_1 \to^{\mathcal{R}}_{\mathcal{G}_U} r_1^*$ resp. $r_2 \to^{\mathcal{R}}_{\mathcal{G}_U} r_2^*$ and $r_1 \prec r_2$, it holds that $i < j$.

A legal partitioning of $\mathcal{R}^{\times}$ ensures that instances of exceptional rules are placed in a higher indexed partition than the instances of its less specific general rules.

Example 9. We extend the knowledge base $\mathcal{R}$ of Ex. 3 with the additional conditional $r_4 = (\mathit{likes}(\mathit{dumbo}, Y))[0.8]$ to demonstrate how the specificity GOP handles incomparably conflicting conditionals. The extended knowledge base $\mathcal{S}$ contains the four conditionals $r_1 = (\mathit{likes}(X, Y))[0.9]$, $r_2 = (\mathit{likes}(X, \mathit{fred}))[0.3]$, $r_3 = (\mathit{likes}(\mathit{clyde}, \mathit{fred}))[1]$, and $r_4 = (\mathit{likes}(\mathit{dumbo}, Y))[0.8]$. Since r_2 and r_4 are incomparable ($r_2 \not\sim r_4$), $\mathcal{S}^{\times}$ does not contain their two conflicting instances $(\mathit{likes}(\mathit{dumbo}, \mathit{fred}))[0.3]$ and $(\mathit{likes}(\mathit{dumbo}, \mathit{fred}))[0.8]$. However, $\mathcal{S}^{\times}$ does contain the instance of r_1 $(\mathit{likes}(\mathit{dumbo}, \mathit{fred}))[0.9]$, because there does not exist a conditional in $\mathcal{S}$ that is incomparably conflicting to r_1. Figure 2 shows the two possible legal partitions of $\mathcal{S}^{\times}$. A directed edge from a node r_i to r_j indicates the specificity relation between r_i and r_j, i. e. $r_j \prec r_i$. The node r_i represents each instance r_i^* of r_i ($r_i \to^{\mathcal{S}}_{\mathcal{G}} r_i^*$) that occurs in $\mathcal{S}^{\times}$.

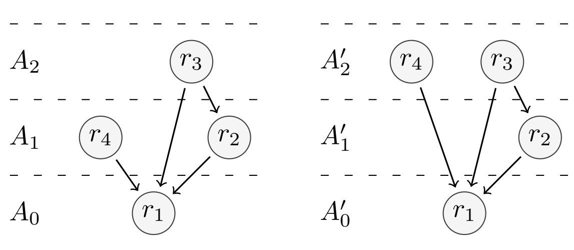

Figure 2: Two legal partitions of $\mathcal{S}^{\times}$

Now we are able to define the specificity GOP which favors instances of more specific conditionals to instances of less specific conditionals.

Definition 11. Let $\mathcal{R}$ be a knowledge base and let $A = \{A_1, \ldots, A_n\}$ be a legal partitioning of $\mathcal{R}^{\times}$. The *specificity grounding operator* $\mathcal{G}^A_{sp}$ is inductively defined as follows

$$\begin{aligned}
\mathcal{G}^{n,A}_{sp}(\mathcal{R}) &= A_n \\
\mathcal{G}^{i-1,A}_{sp}(\mathcal{R}) &= \mathcal{G}^{i,A}_{sp}(\mathcal{R}) \cup (A_i \setminus \odot(A_i \cup \mathcal{G}^{i,A}_{sp}(\mathcal{R}))) \\
& \qquad (1 < i < n) \\
\mathcal{G}^{A}_{sp}(\mathcal{R}) &= \mathcal{G}^{1,A}_{sp}(\mathcal{R}),
\end{aligned}$$

if $\mathcal{R}$ has no direct conflicts, and $\mathcal{G}^A_{sp}(\mathcal{R}) = \emptyset$ otherwise.

$\mathcal{S}$		$\mathcal{G}_{sp}(\mathcal{S})$
$r_1 : (l(X, Y))[0.9]$	$\to^{\mathcal{S}}_{\mathcal{G}_{sp}}$	$(l(c, h))[0.9]$
		$(l(d, f))[0.9]$
		$(l(t, h))[0.9]$
$r_2 : (l(X, f))[0.3]$	$\to^{\mathcal{S}}_{\mathcal{G}_{sp}}$	$(l(t, f))[0.3]$
$r_3 : (l(c, f))[1]$	$\to^{\mathcal{S}}_{\mathcal{G}_{sp}}$	$(l(c, f))[1]$
$r_4 : (l(d, Y))[0.8]$	$\to^{\mathcal{S}}_{\mathcal{G}_{sp}}$	$(l(d, h))[0.8]$

Table 3: Specificity grounding

The specificity GOP takes a legal partitioning of $\mathcal{R}^{\times}$ and starts by considering all instances of all most specific conditionals. Any instance of any less specific conditional is added only if it does not directly conflict with an instance already included.

Example 9 shows that there may be more than one legal partitioning for a given knowledge base $\mathcal{R}$ and Def. 11 suggests that the actual legal partitioning used in the join process may influence the outcome of the grounding. This is not the case as the following theorem shows.

Theorem 3. *The $\mathcal{R}$ be a knowledge base. Then for any two legal partitionings A, A' of a $\mathcal{R}^{\times}$ it is $\mathcal{G}^A_{sp}(\mathcal{R}) = \mathcal{G}^{A'}_{sp}(\mathcal{R})$.*

Due to Th. 3 we can omit the actual partitioning used for defining the specificity grounding. Hence, in what follows we will write $\mathcal{G}_{sp}$ instead of $\mathcal{G}^A_{sp}$ with a partitioning A.

Example 10. We consider the knowledge base $\mathcal{S}$ from Ex. 9. When grounding $\mathcal{S}$ with the specificity GOP this yields $\mathcal{G}_{sp}(\mathcal{S})$ shown in the right column of Tab. 3. Notice, that the instance $(l(d, f))[0.9]$ of r_1 is preferred over $(l(d, f))[0.8]$ and $(l(d, f))[0.4]$ because the last two came from two incomparable conditionals r_2 and r_4 and therefore left out according to the conservative conflict solution.

Theorem 4. *The specificity GOP $\mathcal{G}_{sp}$ satisfies (Compatibility), (Stability), (Structural Preservation), (Upper Bound), (Rationality), and (Conflict Freeness).*

In general, the specificity GOP does not satisfy (Maximality). But this is not very surprising considering the following example.

Example 11. Let $\mathcal{R} = \{(p(X))[\alpha], (p(X))[\beta]\}$ be a knowledge base with $\alpha \neq \beta$ and let $U = \{a\}$ a singleton set of constants. As the conditionals in $\mathcal{R}$ are incomparable with respect to specificity and both instances are in direct conflict it follows $\mathcal{G}_{sp}(\mathcal{R}) = \emptyset$. Obviously, $\mathcal{G}_{sp}(\mathcal{R})$ is not maximal as any one of the two instances of the conditionals might be added to $\mathcal{G}_{sp}(\mathcal{R})$ without violating (Conflict Freeness).

Without any external means of preference ordering besides specificity knowledge bases of the type as in Ex. 11 cannot be maximally grounded. However, excluding these classes of knowledge bases shows that the specificity GOP satisfies (Maximality).

Theorem 5. *If $\mathcal{R}$ contains no two incomparable conditionals then $\mathcal{G}_{sp}(\mathcal{R})$ is maximal, i. e., it holds for all $r \in \mathcal{G}_U(\mathcal{R}) \setminus \mathcal{G}_{sp}(\mathcal{R})$ the set $\mathcal{G}_{sp}(\mathcal{R}) \cup \{r\}$ has a direct conflict.*

Related Work

Many proposals like Bayesian Logic Programs and Markov Logic Networks (Getoor and Taskar, 2007) have been developed that extend propositional probabilistic models like Bayes Nets and Markov Networks (Pearl, 1998) to the relational case. However, up to now only few approaches on extending probabilistic conditional logic and ME-inference to first-order logics are available. The approach developed here bases on notions introduced in (Fisseler, 2009) where a first-order probabilistic conditional logic is proposed and also inference based on maximum entropy is employed. Contrary to our approach here, however, in that paper meta-constraints are used to avoid conflicts between grounded rules; so the process of strategic grounding is of no concern in (Fisseler, 2009). The approach presented in (Thimm, 2009) combines relational knowledge representation and ME-inference within a novel semantics for first-order conditional logic. In that paper, open conditionals are interpreted not as a schema for their instances but as a statement that should "in average" hold for all its instantiations. In this way, (Thimm, 2009) circumvents the problem of inconsistent knowledge bases with respect to universal instantiations.

The paper (Kern-Isberner and Lukasiewicz, 2004) combines logic programming with probabilistic reasoning and ME-inference. It also takes a more subjective view on probabilities, but applies a closed world assumption and assigns interval-valued probabilities to conditionals, hereby aiming at covering all probabilities that arise during instantiations. One of the most important papers on the role of the principle of maximum entropy in first-order probabilistic logic is (Bacchus et al., 1996); however, probabilities are assigned a statistical reading there, reflecting properties of populations whereas in this paper, we make use of a possible world semantics and focus on information on individuals within a population.

Please note that by making use of the maximum entropy principle—in our approach as well as in the approaches above—there is no need to define external strategies for combining rules, as in e. g. Bayesian Logic Programs (Getoor and Taskar, 2007; Ch. 10). Once a consistent probabilistic knowledge base is provided, the maximum entropy principle internally combines information in an optimal way.

Summary and Conclusion

This paper presents an approach to first-order probabilistic reasoning by exploring techniques from the propositional context. So, first-order knowledge bases are seen as schemas encoding information on (typical) individuals, and are used in a ground form to which the principle of maximum entropy can be applied for inference (RME reasoning). However, conflicts may arise when different conditionals suggest different probabilities for the same individual. We proposed solutions to solve these conflicts by applying different strategies for grounding. We presented three GOPs that we evaluated with the help of a set of properties. The specificity operator, being the most advanced of these operators, ensures that most specific information on an individual is used for reasoning.

An implementation of the RME framework and the GOPs presented in this paper is integrated into the KREATOR workbench for relational knowledge representation (Finthammer, Loh, and Thimm, 2009).

This paper deals with appropriate grounding strategies for probabilistic conditionals, i. e., from a knowledge base, specific instances of formulas are derived. In some way, this process can be seen as being inverse to generalization processes in concept learning (Mitchell, 1997) or inductive logic programming (Muggleton, 1991). As part of our future work, we will combine the ideas presented in this paper with the work of (Fisseler, 2009) by deriving meta-constraints from the grounding strategies so that the generic constraints equipped with meta-constraints can be seen as least general generalizations of conditionals in the grounded knowledge base.

Acknowledgements. The research reported here was partially supported by the Deutsche Forschungsgemeinschaft (grant KE 1413/2-1). We also thank the anonymous reviewers for their helpful suggestions to improve the original version of this paper.

References

Bacchus, F.; Grove, A.; Halpern, J.; and Koller, D. 1996. From statistical knowledge bases to degrees of belief. *Artificial Intelligence* 87(1-2):75–143.

Delgrande, J. P., and Schaub, T. 1997. Compiling specificity into approaches to nonmonotonic reasoning. *Artificial Intelligence* 90(1-2):301–348.

Delgrande, J. P. 1998. On First-Order Conditional Logics. *Artificial Intelligence* 105(1–2):105–137.

Finthammer, M.; Loh, S.; and Thimm, M. 2009. Towards a Toolbox for Relational Probabilistic Knowledge Representation, Reasoning, and Learning. In *Workshop RelKR at KI-2009, Proceedings*.

Fisseler, J. 2009. First-Order Probabilistic Conditional Logic: Introduction and Representation. In *Workshop RelKR at KI-2009, Proceedings*.

Getoor, L., and Taskar, B., eds. 2007. *Introduction to Statistical Relational Learning*. MIT Press.

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

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

Mitchell, T. M. 1997. *Machine Learning*. McGraw-Hill.

Muggleton, S. H. 1991. Inductive logic programming. *New Generation Computing* 8(4):295–318.

Pearl, J. 1998. *Probabilistic Reasoning in intelligent Systems: Networks of plausible inference*. Morgan Kaufmann.

Thimm, M. 2009. Representing Statistical Information and Degrees of Belief in First-Order Probabilistic Conditional Logic. In *Workshop RelKR at KI-2009, Proceedings*.