A Distribution Semantics for non-DL-Safe Probabilistic Hybrid Knowledge Bases

Marco Alberti Evelina Lamma Fabrizio Riguzzi Riccardo Zese

Dipartimento di Ingegneria – University of Ferrara

Dipartimento di Matematica e Informatica – University of Ferrara

Introduction

- Logic Programming (LP) or Description Logics (DLs) for modeling complex domains
- Domain closure assumption: closed-world assumption for LP and open-world assumption for DLs
- In many domains, such as in legal reasoning, different closure assumptions
- Combination of LP and DL
- Minimal Knowledge with Negation as Failure (MKNF) [Lifschitz IJCAI91]
- MKNF was applied to define hybrid knowledge bases (HKBs) [Motik, Rosati JACM10]: combination of a logic program and a DL KB.

Uncertainty

- LP: distribution semantics [Sato ICLP95]
- DLs: combination with probability theory using graphical models, Bayesian networks, Markov networks, Nilsson's probabilistic logic, probabilistic databases, DISPONTE ("DIstribution Semantics for Probabilistic ONTologiEs") [Bellodi et al URSW11]
- [Alberti et al AI*IA16]: DL-safe Probabilistic Hybrid KBs (PHKBs) under the distribution semantics combining LPADs with DLs under DISPONTE semantics.

Overview

- An LP rule is DL-safe if each of its variables occurs in at least one positive non-DL-atom in the body; a HKB is DL-safe if all its LP-rules are DL-safe.
- Semantics proposed in [Alberti et al AI*IA16]: not applicable to non-DL-safe PHKBs
- New semantics that coincides with the previous one if the PHKB is DL-safe.

MKNF Hybrid Knowledge Bases

- An MKNF Hybrid Knowledge Base (HKB) is a pair $\mathcal{K} = \langle \mathcal{O}, \mathcal{P} \rangle$ where $\mathcal{O}$ is a DL knowledge base and $\mathcal{P}$ is a set of LP rules of the form

$$h \leftarrow a_1, \ldots, a_n, \sim b_1, \ldots, \sim b_m$$

$\sim$ is default negation; a negative literal is a default negated atom.

Dls

- DLs are a fragment of First Order Logic (FOL) used to model ontologies
- DLs can be directly translated into FOL by exploiting a function π
- A DL knowledge base (KB) is defined using concepts, roles and individuals.
- A KB is a triple: a TBox $\mathcal{T}$ containing *concept inclusion axioms* $C \sqsubseteq D$, an ABox $\mathcal{A}$ containing *concept membership axioms* $a : C$ and *role membership axioms* $(a, b) : R$, and possibly an RBox $\mathcal{R}$ containing transitivity axioms $Trans(R)$ and role inclusion axioms $R \sqsubseteq S$
- A DL KB is assigned a semantics in terms of interpretations $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$, where $\Delta^\mathcal{I}$ is a non-empty *domain* and $\cdot^\mathcal{I}$ is the *interpretation function*.

DL Example

$$\begin{array}{rcl}
\mathcal{O} & = & \exists hasAnimal.pet \sqsubseteq \neg cruelToAnimals \\
& & (kevin, fluffy) : hasAnimal \\
& & (kevin, tom) : hasAnimal \\
& & fluffy : cat \\
& & tom : cat \\
& & cat \sqsubseteq pet \\
\pi(\mathcal{O}) & = & \forall X, Y(hasAnimal(X, Y) \land pet(Y) \rightarrow \neg cruelToAnimals(X)), \\
& & hasAnimal(kevin, fluffy), \\
& & hasAnimal(kevin, tom), \\
& & cat(fluffy) \\
& & cat(tom) \\
& & cat(X) \rightarrow pet(X)
\end{array}$$

$$\mathcal{O} \models kevin : \neg cruelToAnimals$$

LP Example

$$\mathcal{P} = \begin{array}{l} protest \leftarrow activist(X), {\sim} cruelToAnimals(X). \\ activist(kevin). \\ activist(nadia). \end{array}$$

$\mathcal{P} \models protest$

Hybrid Knowledge Bases

- An HKB is positive if no negative literals occur in it.
- Given a HKB $\mathcal{K} = \langle \mathcal{O}, \mathcal{P} \rangle$, an atom in $\mathcal{P}$ is a DL-atom if its predicate occurs in $\mathcal{O}$, a non-DL-atom otherwise.
- An LP rule is DL-safe if each of its variables occurs in at least one positive non-DL-atom in the body; a HKB is DL-safe if all its LP-rules are DL-safe.
- An HKB is given a semantics by transforming it into an MKNF formula.

From HKB to MKNF

- MKNF: syntax of first order logic augmented with modal operators **K** and **not** .
- The transform π defined for DLs is extended as follows to support LP rules:
  - if C is a rule of the form $h \leftarrow a_1, \ldots, a_n, \sim b_1, \ldots, \sim b_m$ and **X** is the vector of all variables in C,
    $\pi(C) = \forall \mathbf{X}(\mathbf{K}\, a_1 \wedge \ldots \wedge \mathbf{K}\, a_n \wedge \mathbf{not}\, b_1 \wedge \ldots \wedge \ldots \mathbf{not}\, b_m \supset \mathbf{K}\, h)$
  - $\pi(\mathcal{P}) = \bigwedge_{C \in \mathcal{P}} \pi(C)$
  - $\pi(\langle \mathcal{O}, \mathcal{P} \rangle) = \mathbf{K}\, \pi(\mathcal{O}) \wedge \pi(\mathcal{P})$

From HKB to MKNF

- Δ is the Herbrand universe of the signature at hand.
- An MKNF structure is a triple (I, M, N) where I as a first-order interpretation over Δ and M and N are sets of first order interpretations over Δ.
- Entailment of a closed formula by an MKNF structure is defined as follows:

$$
\begin{aligned}
(I, M, N) &\models p && \Leftrightarrow p \in I \\
(I, M, N) &\models \neg\varphi && \Leftrightarrow (I, M, N) \not\models \varphi \\
(I, M, N) &\models \varphi_1 \wedge \varphi_2 && \Leftrightarrow (I, M, N) \models \varphi_1 \text{ and } (I, M, N) \models \varphi_2 \\
(I, M, N) &\models \exists x : \varphi && \Leftrightarrow (I, M, N) \models \varphi[\alpha/x] \text{ for some } \alpha \in \Delta \\
(I, M, N) &\models \mathbf{K}\,\varphi && \Leftrightarrow (J, M, N) \models \varphi \text{ for all } J \in M \\
(I, M, N) &\models \mathbf{not}\,\varphi && \Leftrightarrow (J, M, N) \not\models \varphi \text{ for some } J \in N
\end{aligned}
$$

Hybrid Knowledge Bases

- An MKNF interpretation is a set M of interpretations over Δ.
- An interpretation M is an MKNF model of a closed formula φ iff
  - $(I, M, M) \models \varphi$ for all $I \in M$
  - for all $M' \supset M$, for some $I' \in M'(I', M', M) \not\models \varphi$
- A formula φ entails a formula ϕ, written $\varphi \models_{\mathrm{MKNF}} \phi$, iff for all MKNF models M of φ and for all $I \in M$ $(I, M, M) \models \phi$.

Probabilistic Logic Programming

- Distribution Semantics [Sato ICLP95]
- A probabilistic logic program defines a probability distribution over normal logic programs (called instances or possible worlds or simply worlds)
- The distribution is extended to a joint distribution over worlds and interpretations (or queries)
- The probability of a query is obtained from this distribution

PLP under the Distribution Semantics

- A PLP language under the distribution semantics with a general syntax is Logic Programs with Annotated Disjunctions (LPADs)
- Heads of clauses are disjunctions in which each atom is annotated with a probability.
- LPAD $\mathbb{P}$ with n clauses: $\mathbb{P} = \{C_1, \ldots, C_n\}$.
- Each clause C_i takes the form:

$$h_{i1} : \Pi_{i1}; \ldots ; h_{iv_i} : \Pi_{iv_i} \leftarrow b_{i1}, \ldots , b_{iu_i}$$

,

- Each grounding $C_i\theta_j$ of a clause C_i corresponds to a random variable X_{ij} with values $\{1, \ldots , v_i\}$
- The random variables X_{ij} are independent of each other.

Example

$$\mathbb{P} = \begin{array}{l} protest : 0.6 \leftarrow activist(X), \sim cruelToAnimals(X). \\ activist(kevin). \\ activist(nadia). \end{array}$$

Distribution Semantics

- Case of no function symbols: finite Herbrand universe, finite set of groundings of each clause
- Atomic choice: selection of the k-th atom for grounding $C_i\theta_j$ of clause C_i
- Represented with the triple (C_i, θ_j, k)
- Example $C_1 = \mathit{protest} : 0.6 \leftarrow \mathit{activist}(X), \sim\mathit{cruelToAnimals}(X).$, $(C_1, \{X/\mathit{kevin}\}, 1)$
- Composite choice κ: consistent set of atomic choices
- The probability of composite choice κ is

$$P(\kappa) = \prod_{(C_i,\theta_j,k)\in\kappa} \Pi_{ik}$$

Distribution Semantics

- Selection σ: a total composite choice (one atomic choice for every grounding of each clause)
- A selection σ identifies a logic program w_σ called world
- The probability of w_σ is $P(w_\sigma) = P(\sigma) = \prod_{(C_i,\theta_j,k)\in\sigma} \Pi_{ik}$
- Finite set of worlds: $\mathcal{W}_\mathbb{P} = \{w_1, \ldots, w_m\}$
- $P(w)$ distribution over worlds: $\sum_{w\in W_T} P(w) = 1$

Distribution Semantics

- We consider only *sound* LPADs, where each possible world has a total well-founded model, so $w_\sigma \models q$ means that the query q is true in the well-founded model of the program w_σ.
- Ground query q
- $P(q|w) = 1$ if q is true in w and 0 otherwise
- $P(q) = \sum_w P(q, w) = \sum_w P(q|w)P(w) = \sum_{w \models q} P(w)$

Example

$$\mathbb{P} = \begin{array}{l} \mathit{protest} : 0.6 \leftarrow \mathit{activist}(X), {\sim}\mathit{cruelToAnimals}(X). \\ \mathit{activist}(\mathit{kevin}). \\ \mathit{activist}(\mathit{nadia}). \end{array}$$

$$P(\mathit{protest}) = 0.6 \cdot 0.6 + 0.6 \cdot 0.4 + 0.4 \cdot 0.6 = 0.84$$

Distribution Semantics

- A composite choice, or a set of composite choices, determine sets of worlds.
- Given a composite choice κ, the set of worlds determined by κ is the set of worlds identified by total choices that are supersets of κ, i.e., $\omega_\kappa = \{w_\sigma \mid \kappa \subseteq \sigma\}$.
- Given a set K of composite choices, the set of worlds determined by K is $\omega_K = \bigcup_{\kappa \in K} \omega_\kappa$
- Two sets K_1 and K_2 of composite choices are equivalent if $\omega_{K_1} = \omega_{K_2}$.

Distribution Semantics with Function Symbols

- Infinite countable Herbrand base, each world infinite, countable, 0 probability
- Uncountable $\mathcal{W}_{\mathbb{P}}$
- Given an LPAD $\mathbb{P}$, let $\Omega_{\mathbb{P}}$ be the set of sets of worlds determined by countable sets of countable composite choices.
- $\Omega_{\mathbb{P}}$ is a σ-algebra over $\mathcal{W}_{\mathbb{P}}$ [Riguzzi IJAR16]
- A probability measure $\mu : \Omega_{\mathbb{P}} \rightarrow [0, 1]$ can be defined over $\Omega_{\mathbb{P}}$.

Distribution Semantics with Function Symbols

- $\mu(\kappa) = P(\kappa)$
- A set of composite choices is pairwise incompatible if any two choices from the set are incompatible;
- The probability of a pairwise incompatible set of composite choices is the sum of the probabilities of its elements: $\mu(K) = \sum_{\kappa \in K} \mu(\kappa)$
- Given a ground query q, a composite choice κ is an explanation for q if $w \models q$ for all $w \in \omega_\kappa$.
- A set K of composite choices is covering for q if $\{w \mid w \models q\} \subseteq \omega_K$.

Distribution Semantics with Function Symbols

- For each countable set K of countable composite choices, there exists a pairwise incompatible countable set K' of countable composite choices that is equivalent to K.
- For sound LPADs, each query q has a countable covering set K of countable explanations [Riguzzi IJAR16]
- Since there exists a pairwise incompatible set K' that is equivalent to K, we can define the probability of q as $\mu(K')$.

Probabilistic Description Logics

- DISPONTE applies the distribution semantics to probabilistic ontologies
- Probabilistic knowledge bases $\mathbb{O}$ are sets of certain and probabilistic axioms.
- Certain axioms are regular DL axioms
- Probabilistic axioms take the form $\Pi :: a$, where Π is a real number in $[0, 1]$ and a is a DL axiom.
- An atomic choice for an axiom a is a pair (a, i), where i is 1 if a is selected and 2 otherwise.

Example

$$
\begin{aligned}
\mathbb{O} \quad = \quad & & \exists hasAnimal.pet \sqsubseteq \neg cruelToAnimals \\
& & (kevin, fluffy) : hasAnimal \\
0.3 & :: & (kevin, tom) : hasAnimal \\
& & fluffy : cat \\
& & tom : cat \\
0.4 & :: & fluffy : cat \\
& & cat \sqsubseteq pet
\end{aligned}
$$

- Composite choices, set of composite choices and the other concepts from the previous subsection can be defined similarly.
- A world, here, is obtained by including in it all certain axioms and a subset of the uncertain axioms.
- The probability of the world is given by the product of the probability Π for the included axioms and $1 - \Pi$ for the excluded ones.
- The probability of a query is then the sum of the probabilities of the worlds where the query holds

Example

$$
\begin{array}{rcl}
\mathbb{O} = & & \exists hasAnimal.pet \sqsubseteq \neg cruelToAnimals \\
& & (kevin, fluffy) : hasAnimal \\
0.3 & :: & (kevin, tom) : hasAnimal \\
& & fluffy : cat \\
& & tom : cat \\
0.4 & :: & fluffy : cat \\
& & cat \sqsubseteq pet
\end{array}
$$

$$P(kevin : \neg cruelToAnimals) = 0.3 \cdot 0.4 + 0.3 \cdot 0.6 + 0.7 \cdot 0.6 = 0.72$$

Probabilistic Hybrid Knowledge Bases

- A Probabilistic Hybrid Knowledge Base (PHKB) is a pair $\mathbb{K} = \langle \mathbb{O}, \mathbb{P} \rangle$ where $\mathbb{O}$ is a DISPONTE knowledge base and $\mathbb{P}$ is an LPAD without function symbols.
- In [Alberti et al AI*IA16] a PHKB's semantics is given by first grounding it over all the constants in the PHKB.
- A world is the deterministic ground HKB obtained by selecting, for each clause $h_{i1} : \Pi_{i1}; \ldots ; h_{in_i} : \Pi_{in_i} \leftarrow b_{i1}, \ldots , b_{im_i}$, one of the disjuncts in the head and some of the DL axioms.
- The world's probability is the product of the probabilities of the selected head disjuncts and the selected axioms.

Probabilistic Hybrid Knowledge Bases

Definition

Given a world w, the probability of a query q is defined as $P(q|w) = 1$ if $w \models_{\mathrm{MKNF}} \mathbf{K}\, q$ and 0 otherwise.
The probability of the query is its marginal probability:

$$P(q) = \sum_{w} P(w)P(q|w) \tag{1}$$

Example

- KB $\mathbb{K}$ modeling the insurgence of a protest against animal testing:

$$\begin{array}{rcll}
\mathbb{P} & = & (C_1) & protest : 0.6 \leftarrow \\
& & & activist(X), \sim cruelToAnimals(X). \\
& & & activist(kevin). \\
& & (C_2) & activist(nadia) : 0.3. \\
\mathbb{O} & = & & \exists hasAnimal.pet \sqsubseteq \neg cruelToAnimals \\
& & & (kevin, fluffy) : hasAnimal \\
& & (E_1) & 0.4 \ :: \ fluffy : cat \\
& & & cat \sqsubseteq pet
\end{array}$$

- This KB has 16 worlds and the query *protest* is true in four of them, those containing $activist(nadia)$ and $protest \leftarrow activist(nadia), \sim cruelToAnimals(nadia)$, plus other two, those in which $activist(nadia)$ is absent and $fluffy : cat$ and $protest \leftarrow activist(kevin), \sim cruelToAnimals(kevin)$ are present.
- $P(protest) = 0.3 \cdot 0.6 + 0.7 \cdot 0.4 \cdot 0.6 = 0.18 + 0.168 = 0.438.$

DL-safety

- This semantics gives wrong results for non-DL-safe PHKBs
- A non-DL-safe HKB may not have the same MKNF models of its grounding over its constants.

Example

- Let $\mathcal{K} = \langle \mathcal{O}, \mathcal{P} \rangle$, where

$$\begin{aligned}\mathcal{P} =& person(X) \leftarrow \sim dog(X).\\ \mathcal{O} =& guard \sqcap person \sqsubseteq soldier\\ & \exists commands.soldier \sqsubseteq commander\\ & john : \exists commands.guard\end{aligned}$$

- Not DL-safe
- In a model of $\mathcal{K}$'s, no individual is a *dog* in all interpretations, so each individual is a *person*.
- In all interpretations, the *guard* that *john commands* is a *person*, and due to the first axiom, a *soldier*;
- In each interpretation *john commands* a *soldier*, and is a *commander*. Thus, $\mathcal{K} \models \mathbf{K}\, commander(john)$.

Example

- However, the grounding over the known individuals yields the following clause:

$$\mathcal{P} = \mathit{person}(\mathit{john}) \leftarrow \sim \mathit{dog}(\mathit{john}).$$

  so the only individual known to be a *person* is *john*
- The grounding of the HKB does not entail $\mathbf{K}\,\mathit{soldier}(\mathit{john})$ because *john* does not command himself in all models
- The *guard* that *john commands* cannot be inferred to be a *soldier* so $\mathcal{K} \not\models \mathbf{K}\,\mathit{commander}(\mathit{john})$.

Semantics for non-DL-safe PHKBs

- Grounding the PHKB over the countable supply of constants provided by the standard name assumption [Motik, Rosati JACM10].
- Δ: the resulting countable set of constants
- Δ countable as the Herbrand base of LP with function symbols $\Rightarrow$ ground the program with Δ and use the same approach for the semantics
- A possible world is obtained by selecting one annotated disjunct for each ground clause in $\mathbb{P}$, and some of the axioms in $\mathbb{O}$
- We assign probabilities to sets of worlds, rather than to individual worlds.

Semantics for non-DL-safe PHKBs

- A selection σ determines the *world* w_σ, i.e., the HKB composed of:
  - one rule for each grounding substitution θ of each rule C in $\mathbb{P}$, where $(C, \theta, k) \in \sigma$, whose head is the k-th disjunct of $C\theta$ and whose body is $C\theta$'s body;
  - the axioms a for which $(a, 1)$ is in the selection.

Semantics for non-DL-safe PHKBs

- Given a *PHKB* $\mathbb{K}$, $\mathcal{W}_{\mathbb{K}}$ is the set of all $\mathbb{K}$'s possible worlds.
- A composite choice, or a set of composite choices, determine sets of worlds, as for LPADs.
- $\Omega_{\mathbb{K}}$ is the set of sets of worlds determined by finite or countable sets of finite or countable composite choices;
- A probability measure $\mu : \Omega_{\mathbb{K}} \rightarrow [0, 1]$ is defined over $\Omega_{\mathbb{K}}$.

Semantics for non-DL-safe PHKBs

- If a query q has a countable covering set K of countable explanations, then there exists a pairwise incompatible set K' with the same property, and whose probability $\mu(K')$ is defined; that is defined as q's probability given $\mathcal{K}$.

Definition

Let $\mathbb{K}$ be a PHKB and K be a countable covering set of countable explanations for a query q. Then q's probability given $\mathbb{K}$ $P_{\mathbb{K}}(q)$ is the probability of a pairwise incompatible set K' of explanations equivalent to K, which is guaranteed to exist.

Example

$$\mathbb{P} = person(X) : 0.5 \leftarrow \sim dog(X).$$
$$\mathbb{O} = guard \sqcap person \sqsubseteq soldier$$
$$\exists commands.soldier \sqsubseteq commander$$
$$john : \exists commands.guard$$

- In the last axiom there is an (unknown) individual that is a *guard* and that *john commands*. Let us call her u.
- $\mathcal{K} \models \mathbf{K}\, commander(john)$ is entailed by the worlds where he first disjunct is selected for the clause with substitution X/u. So $\{\{(C_1, X/u, 1)\}\}$ is a (finite) covering set of (finite) explanations. Its probability is 0.5.

Property

Proposition

Given a DL-safe PHKB without function symbols, the probability of any query is the same under the semantics in Definition 1 and the one in Definition 2.

Proof.

A DL-safe KB is equivalent to its grounding over the constants that occur in it, and if function symbols are not allowed there are finitely many worlds; each world that entails the query is identified by a selection. The set of such selections is a pairwise incompatible covering set of explanations for the query, and its probability is identical to the one given in Definition 1.

Conclusions and Future Work

- Conclusions
  - Semantics for Probabilistic Hybdrid Knowledge Bases for non-DL-safe PHKB
  - For DL-safe PHKB it coincides with the existing one
- Future work
  - Prove that each query has a countable set of countable explanations
  - Reasoner: use **SLG**($\mathcal{O}$) [Alferes et al TOCL13] for HKBs under the well founded semantics.
  - **SLG**($\mathcal{O}$) integrates a DL reasoner into the **SLG** procedure in the form of an oracle in order to manage the DL part of the HKBs.
  - Similar approach for PHKBs, integrating the TRILL probabilistic DL reasoner [Zese et al AMAI16] with the PITA algorithm [Riguzzi, Swift ICLP11]
  - Detailed comparison with alternative approaches for existential constructs in probabilistic logics

THANKS FOR LISTENING

AND

ANY QUESTIONS ?