UNIVERSITÄT MANNHEIM

Reasoning under Uncertainty with Log-Linear Description Logics

Mathias Niepert

October 2011

Probabilistic Description Logics

1. The system should be usable by individuals knowledgeable only in Semantic Web languages and tools (Protégé, …)
2. It must be possible to express uncertainty with degrees of confidence (real-valued weights) and not necessarily with precise probabilities
3. The user should not have to worry about *inconsistent* and *incoherent* input to the probabilistic reasoner
4. Two types of queries should be supported under uncertainty:
   - The most probable ontology" query and
   - the probability of (conjunctions) of axioms query
5. The worst-case complexity should not exceed that of probabilistic graphical models such as Markov and Bayesian networks

Semi-structured data

Full name	Plato (Πλάτων)
Born	c. 428–427 BC[1] Athens
Died	c. 348–347 BC (age approx 80) Athens
Era	Ancient philosophy

Unstructured data

The exact place and time of Plato's birth are not know
and influential family. Based on ancient sources, mos
BC.[a] His father was Ariston. According to a disputed
Athens, Codrus, and the king of Messenia, Melanthus
Athenian lawmaker and lyric poet Solon.[6] Perictione
the brief oligarchic regime, which followed on the colla
himself, Ariston and Perictione had three other childre
Speusippus (the nephew and successor of Plato as h
were older than Plato.[8] Nevertheless, in his Memorat

amazon mechanical turk
Artificial Artificial Intelligence

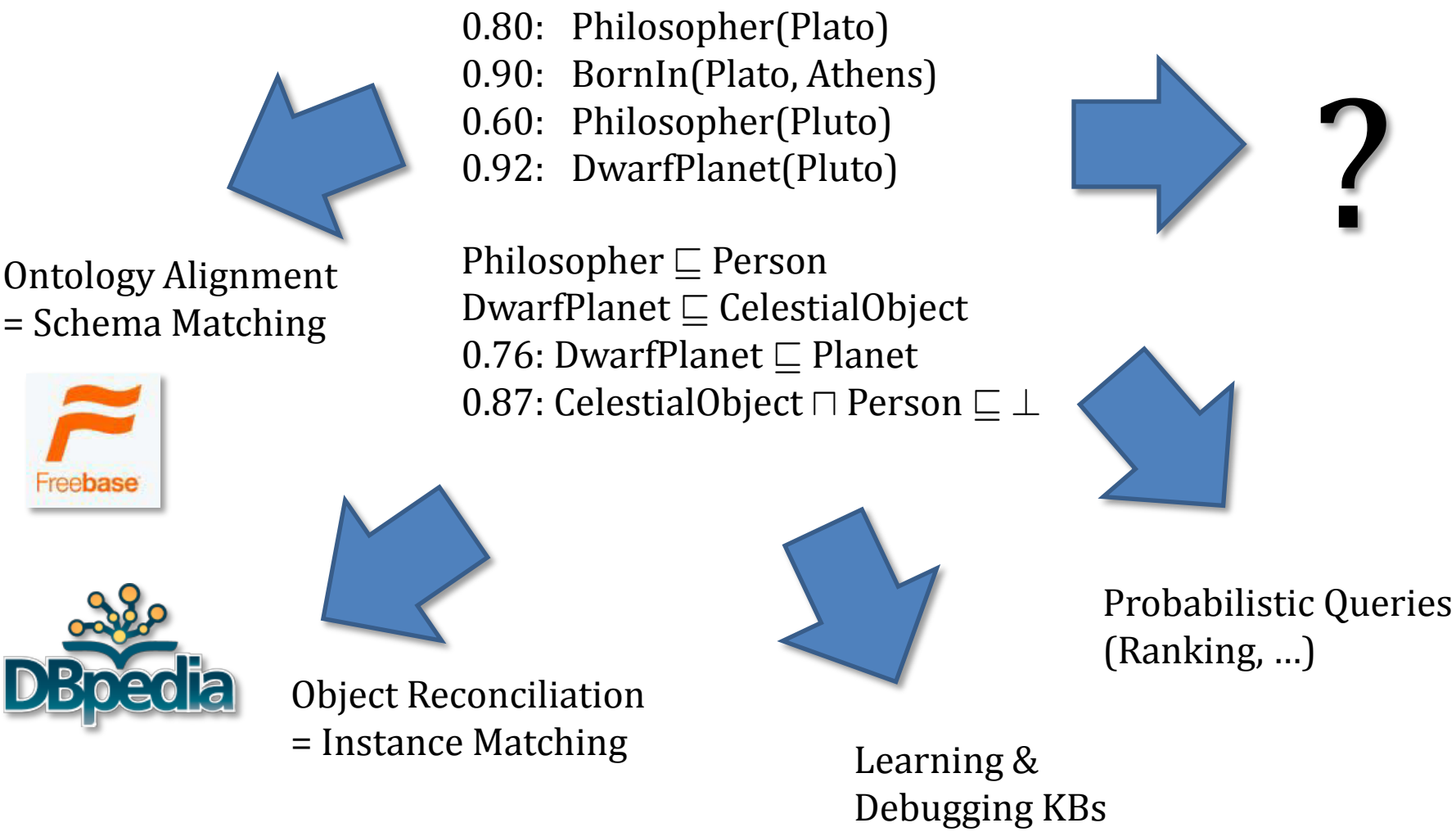
http://inpho.cogs.indiana.edu/
InPhO
the Indiana Philosophy Ontology project

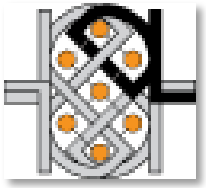
0.8: Philosopher(Plato)
0.9: BornIn(Plato, Athens)
0.6: Philosopher(Pluto)
0.92: DwarfPlanet(Pluto)

Philosopher $\sqsubseteq$ Person
DwarfPlanet $\sqsubseteq$ CelestialObject
0.76: DwarfPlanet $\sqsubseteq$ Planet
0.87: CelestialObject $\sqcap$ Person $\sqsubseteq \bot$

?

Probabilistic Knowledge Bases





Log-Linear Description Logics

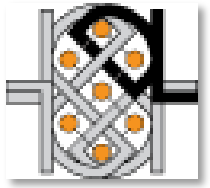
- Probabilistic reasoning for DLs with sound and complete set of inference rules ($\mathcal{EL}^{++}$, …)
- Ontology consists of an **uncertain** $\mathcal{C}^{\mathsf{U}}$ and a **deterministic** $\mathcal{C}^{\mathsf{D}}$ component
- Coherent = no logical contradictions

3

$$P(\mathcal{C}') = \begin{cases} \frac{1}{Z} \exp\left(\sum_{\{(c,w_c)\in\mathcal{C}^{\mathsf{U}}:\mathcal{C}'\models c\}} w_c\right) & \text{if } \mathcal{C}' \text{ is } \textit{coherent} \\ & \text{and } \mathcal{C}' \models \mathcal{C}^{\mathsf{D}}; \\ 0 & \text{otherwise} \end{cases}$$

Normalization constant

Degree of confidence (weights)



Log-Linear Description Logics

2

Two types of probabilistic queries:

- Maximum a-posteriori inference (MAP):
  "Most probable coherent ontology" **{C ⊑ D}**

4

- Conditional (marginal) probability inference:
  "Probability of (conjunction of) axioms"
  P(C ⊑ D | Ev) = 0.47

Application: Ontology Induction

"Is very **A** also a **B**?"

"Can there be anything that is both an **A** and a **B**?"

0.6: $A \sqsubseteq B$
0.9: $D \sqsubseteq A$
...
0.7: $A \sqcap B \sqsubseteq \bot$
0.9: $A \sqcap D \sqsubseteq \bot$
...
0.8: $\exists r.\top \sqsubseteq A$
...

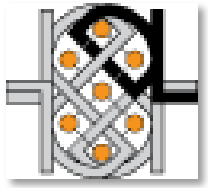
Marginal Inference

0.45: $A \sqsubseteq B$
0.91: $D \sqsubseteq A$
...
0.37: $A \sqcap B \sqsubseteq \bot$
0.29: $A \sqcap D \sqsubseteq \bot$
...

MAP Inference

$A \sqsubseteq B$
$D \sqsubseteq A$
$A \sqcap B \sqsubseteq \bot$
$\exists r.\top \sqsubseteq A$

Axiom type	Algorithm	Precision	Recall	F_1 score
Subsumption	Greedy	0.620	0.541	0.578
	$\mathcal{EL}^{++}$-LL MAP	0.784	0.514	**0.620**
Disjointness	Greedy	0.942	0.980	0.961
	$\mathcal{EL}^{++}$-LL MAP	0.935	0.990	0.961



ELOG in Practice

```
SubClassOf(
   Annotation(<http://URI/ontology#confidence> “0.5"^^xsd:double)
  <http://zoo/Penguin>
  <http://zoo/Bird>
)

DisjointClasses(
 <http://zoo/Bird>
 <http://zoo/Mammal>
)
```

http://code.google.com/p/elog-reasoner/

Probabilistic Description Logics

1. The system should be usable by individuals knowledgeable only in Semantic Web languages and tools (Protégé, …)
2. It must be possible to express uncertainty with degrees of confidence (real-valued weights) and not necessarily with precise probabilities
3. The user should not have to worry about *inconsistent* and *incoherent* input to the probabilistic reasoner
4. Two types of queries should be supported under uncertainty:
   - The most probable ontology" query and
   - the probability of (conjunctions) of axioms query
5. The worst-case complexity should not exceed that of probabilistic graphical models such as Markov and Bayesian networks

Thank you!

Questions?
Criticism?