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

| Full name | Plato (Πλάτων) |
| Died     | c. 348–347 BC (age approx 80) Athens |
| Era      | Ancient philosophy |

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

Philosopher ⊑ Person
DwarfPlanet ⊑ CelestialObject
0.76: DwarfPlanet ⊑ Planet
0.87: CelestialObject ⊑ Person ⊑ ⊥
Probabilistic Knowledge Bases

0.80: Philosopher(Plato)
0.90: BornIn(Plato, Athens)
0.60: Philosopher(Pluto)
0.92: DwarfPlanet(Pluto)

Ontology Alignment = Schema Matching

Philosopher ⊆ Person
DwarfPlanet ⊆ CelestialObject
0.76: DwarfPlanet ⊆ Planet
0.87: CelestialObject ⊆ Person ⊆ ⊥

Object Reconciliation = Instance Matching

Probabilistic Queries (Ranking, ...)

Learning & Debugging KBs
Log-Linear Description Logics

• Probabilistic reasoning for DLs with sound and complete set of inference rules ($\mathcal{EL}^{++}$, ...)

• Ontology consists of an uncertain $C^u$ and a deterministic $C^d$ component

• Coherent = no logical contradictions

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

Normalization constant

Degree of confidence (weights)
Two types of probabilistic queries:

- Maximum a-posteriori inference (MAP):
  “Most probable coherent ontology” \(\{C \sqsubseteq D\}\)

- Conditional (marginal) probability inference:
  “Probability of (conjunction of) axioms”
  \[P(C \sqsubseteq D | Ev) = 0.47\]
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: $\exists r. r \sqsubseteq A$

<table>
<thead>
<tr>
<th>Axiom type</th>
<th>Algorithm</th>
<th>Precision</th>
<th>Recall</th>
<th>$F_1$ score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subsumption</td>
<td>Greedy</td>
<td>0.620</td>
<td>0.541</td>
<td>0.578</td>
</tr>
<tr>
<td></td>
<td>$\mathcal{EL}^{++}$-LL MAP</td>
<td>0.784</td>
<td>0.514</td>
<td><strong>0.620</strong></td>
</tr>
<tr>
<td>Disjointness</td>
<td>Greedy</td>
<td>0.942</td>
<td>0.980</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>$\mathcal{EL}^{++}$-LL MAP</td>
<td>0.935</td>
<td>0.990</td>
<td>0.961</td>
</tr>
</tbody>
</table>

0.45: $A \sqsubseteq B$
0.91: $D \sqsubseteq A$

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

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