Instance-based Probabilistic Reasoning in the Semantic Web

Pedro Oliveira and Paulo Gomes
CISUC-DEI-FCTUC
University of Coimbra
Pólo II, 3030 Coimbra, Portugal
pcoliv@student.dei.uc.pt, pgomes@dei.uc.pt

Motivation

- Current approaches for dealing with uncertainty in the Semantic Web rely on the principle that the uncertainty is already asserted
  - Users are responsible for the uncertainty annotation of ontologies (tedious and error-prone task)

Objective

- Perform probabilistic reasoning on OWL-DL ontologies without any kind of uncertainty annotation
- Using Markov Logic
  - Combination of first-order logic and probabilistic graphical models
  - A world that violates a formula is less probable, but not impossible
  - Learn first-order formulas weights and build a Markov Network

Procedure

1—Interpret OWL as FOL

SubClassOf(A,B)

PropertyDomain(P,A)

PropertyRange(P,B)

ClassAssertion(A,a)

ClassAssertion(B,b)

PropertyAssertion(P,a,b)

A(x) => B(x)

P(x,y) => A(x)

P(x,y) => B(y)

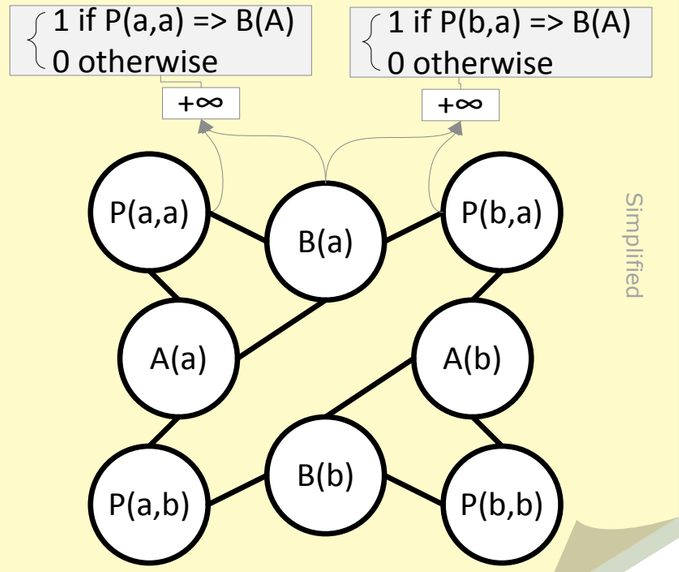
A(a)

B(b)

P(a,b)

2—Build Markov Network

- A variable for each ground atom
- An edge if two atoms appear in the same formula
- A feature for each grounded formula



3—Learn Features Weights

Maximize the pseudo-log-likelihood given the individuals.

The pseudo-log-likelihood of world *x* given weight *w* is

$$\log P_w^*(X = x) = \sum_i \log P_w(X_i = xi \mid N_x(X_i))$$

where *xi* is the truth value of variable *i*, and *Nx(Xi)* is the truth values of its neighbors.

4—Inference

The probability of a world is given by the number of formulas that are true in that world, i.e.,

$$P(X = x) = \frac{1}{Z} \exp(\sum_{i=1}^{F} w_i n_i(x))$$

where *F* is the number of formulas, *ni(x)* the number of true grounding of *Fi* in world *x*, *wi* the weight of *Fi*, and *Z* a normalizing constant.

Applications

- Risk Assessment
  - Using a financial ontology, determine the probability of a loan being problematic
- Social Network Analysis
  - Apply link-mining techniques (e.g., link prediction, link-based classification, link-based cluster analysis) to FOAF networks
- Ontology Mapping
- Ontology Learning
- Ontology individual classification/clustering

Future Work

- Find other ways of gather the uncertainty of ontologies
  - Learn Individuals (i.e., ontology population)
  - Learn the uncertainties directly from textual corpus
  - Use the structure of the ontology (e.g., network analysis)
  - Collective learning of weights (e.g., relational reinforcement learning)

More Information

- **INCERTO** — A Probabilistic Reasoner for the Semantic Web based on Markov Logic

  http://code.google.com/p/incerto

References

- P. Domingos, S. Kok, D. Lowd, H. Poon, M. Richardson, and P. Singla, "Markov Logic," *Probabilistic Inductive Logic Programming*, 2008, pp. 92-117.
- T. Lukasiewicz and U. Straccia, "Managing Uncertainty and Vagueness in Description Logics for the Semantic Web," *Web Semantics Sci Serv Agents World Wide Web*, 2008.
- F. Baader, D. Calvanese, D.L. McGuinness, D. Nardi, and P.F. Patel-Schneider, The Description Logic Handbook: Theory, Implementation, and Applications, Cambridge University Press, 2007.