Inference in Probabilistic Ontologies with Attributive Concept Descriptions and Nominals
– URSW 2008 –

Rodrigo B. Polastro - Fabio G. Cozman
rodrigopolastro@usp.br - fgcozman@usp.br

Decision Making Lab, Escola Politécnica, Univ. São Paulo, Brazil

Decision Making Lab

- Universidade de São Paulo $\longrightarrow$ Engineering School
- Interest: representation of **uncertainty** in decision making.

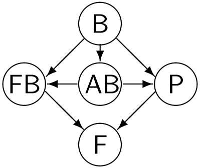
Some Easy Facts

1. $P(A) \geq 0$
2. $P(S) = 1$
3. $P\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ [if $A_i \cap A_j = 0$]
4. Conditional probability: if $P(B) > 0$, $P(A|B) = \frac{P(A,B)}{P(B)}$

Bayesian Networks

- A Bayesian network encodes, using a directed acyclic graph,

$$P(X_1, \ldots, X_n) .$$



- Each node represents a random variable X_i.
  - *Parents* of X_i: $\mathrm{pa}(X_i)$.

- Semantics (Markov condition):

$$p(X_1, \ldots, X_n) = \prod_i p(X_i | \mathrm{pa}(X_i)) .$$

Description Logics

- Description logics offer attractive trade-offs between expressivity and complexity.
- Now used in ontologies, semantic web.

Our goal:

- $\text{Brazilian} \sqsubseteq \text{SouthAmerican}$,
- $P(\text{FootballFan} \mid \text{Brazilian}) \geq 0.85$.
- $\text{Brazilian}(\text{Tweety})$.
- $P(\text{BaseballFan}(\text{Tweety}))$?

Probabilistic Description Logics

- There are many!

No independence:

Heinsohn 94, Jaeger 94, Sebastiani 94, Dürig and Studer 2005, Lukasiewicz 2002.

Independence:

Koller et al 97, Yelland 99, Staker 2002, Ding e Peng 2004, Costa e Laskey 2005, Nottelmann 2006

A Simple Description Logics ($\mathcal{ALC}$)

- Individuals, concepts, roles.
- *Conjunction* ($C \sqcap D$),
  disjunction ($C \sqcup D$),
  negation ($\neg C$),
  existential restriction ($\exists r.C$),
  value restriction ($\forall r.C$).
- Terminologies: $C \sqsubseteq D$ and $C \equiv D$.
- Assertions (Abox):
  $\mathsf{Fruit}(\mathsf{appleFromJohn})$,
  $\mathsf{buyFrom}(\mathsf{houseBob}, \mathsf{John})$.

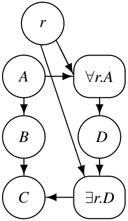
A logic: CR$\mathcal{ALC}$

- Consider $\mathcal{ALC}$.
- Add probabilistic inclusions $P(C|D)$, where C is concept name.
  - Interpret as $\forall x : P(C(x)|D(x)) = \alpha$.
- Add probabilistic assessment $P(r) = \alpha$ for roles.
  - Interpret as $\forall x, y : P(r(x,y)) = \alpha$.
- Require acyclicity (common assumption).
- Assume Markov condition on terminology (independence!).

Example

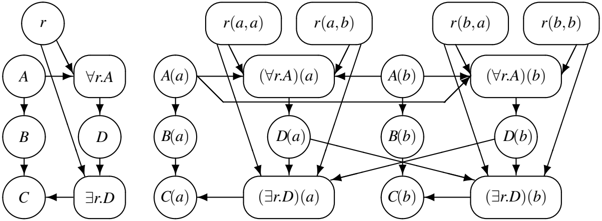
Consider a terminology $\mathcal{T}_2$ with concepts A, B, C, D, where:

- $P(A) = \alpha_1$,
- $B \sqsubseteq A$,
- $P(B|A) = \alpha_2$,
- $D \equiv \forall r.A$,
- $C \equiv B \sqcup \exists r.D$,
- and $P(r) = \alpha_3$.



Example

Consider the same terminology $\mathcal{T}_2$, and a domain $\{a, b\}$.



A Few Results

- Previous example, with inference $P(C(a_0))$ for domain $a_0, \ldots, a_{n-1}$, for several n.

n	1	2	3	5	9	10	20	50
Loopy: $P(C(a_0))$	0.5175	0.5383	0.5291	0.4885	0.4296	0.4223	0.4049	0.4050
Exact: $P(C(a_0))$	0.4350	0.4061	0.4050	0.4050	0.4050	—	—	—

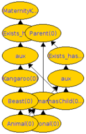
- Exact inferences with SamIam package at reasoning.cs.ucla.edu/samiam.

More Results

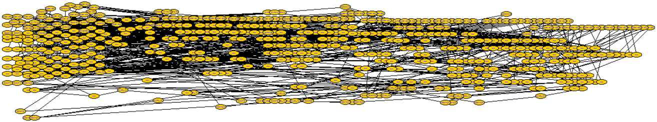
- Same problem, but $P(r) \in [0, 1]$.
- Inference: $P(C(a_0))$ for several n.
- inference with a version of loopy propagation that handles imprecision in probability values (L2U).

n	1	3	5	10	20
L2U: $P(C(a_0))$	[0.405000 0.464500]	[0.405000 0.406783]	[0.405000 0.405030]	[0.405000 0.405000]	[0.405000 0.405000]

Another Test - Kangaroo ontology



(a) $|\mathcal{D}| = 1$



(b) $|\mathcal{D}| = 10$

Inferences $P(Parent(0)|Human(1))$ for domain size n.

n	2	5	10	20	30	40	50
L2U	0.2232	0.3536	0.4630	0.5268	0.5377	0.5396	0.5399

Nominals

- Extensively used in real ontologies to make assertions over individuals or to define a concept with them.
- One-of:
  $\mathsf{WineFlavor} \equiv \{\mathsf{delicate}, \mathsf{moderate}, \mathsf{strong}\}$.
- hasValue:
  $\exists \mathsf{hasColor}.\{\mathsf{red}, \mathsf{white}\}$
- Usually add considerable complexity.
  Nominals are difficult to handle even in standard description logics.

CR$\mathcal{ALC}$ and Nominals

We wish to attach sensible semantics to sentences such as

$$P(\mathsf{Merlot}(\mathsf{a})|\exists\mathsf{Color}.\{\mathsf{red}\}) = \alpha,$$

We allow nominals only as domains of roles in restrictions $r.\{a\}$

We do not allow constructs such as $\mathsf{WineFlavor} \equiv \{\mathsf{delicate}, \mathsf{moderate}, \mathsf{strong}\}$.

The construct $r.\{a\}$ is interpreted directly either as:

in existential restrictions:
$\exists x : r(x,y) \wedge (y = a)$,

in universal restrictions:
$\forall x : r(x,y) \rightarrow (y = a)$.

Example

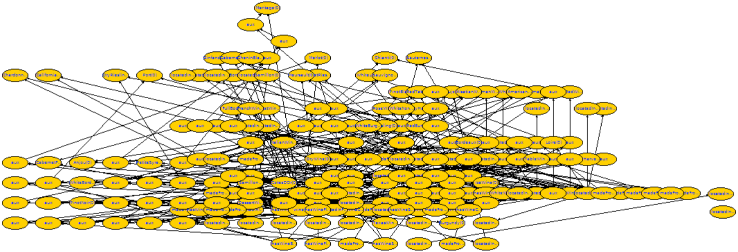
$\exists$hasColor.$\{$red$\}$ is interpreted as:

$$x \in D : \text{hasColor}(x,y) \wedge (y = \text{red}),$$

where $\mathcal{D}$ is the domain with the elements being described and y ranges over all the nominals that “are” colors.

Example - Wine Ontology

- Wine ontology relies extensively in use of nominals.
- Network generated from Wine ontology in $\text{CR}\mathcal{ALC}$ with a domain size 1:



Example - Wine Ontology

Example 1:

The probability of a wine to be Merlot given it has medium body, red color, moderate flavor, dry sugar and is made from merlot grape:

$$P(\mathsf{Merlot}(a) \mid evidences_1(a)) = 1.0$$

Example 2:

The probability of a wine to be Merlot given it has medium body, red color, moderate flavor and dry sugar:

$$P(\mathsf{Merlot}(b) \mid evidences_2(b)) = 0.5$$

Example 3:

The probability of a wine to be Merlot given it has sweet sugar and is made from merlot grape:

$$P(\mathsf{Merlot}(c) \mid evidences_3(c)) = 0.0$$

Domain Size

Note that in the case of the Wine ontology the Markov condition let us break the network in independent smaller networks of domain size 1.

- Thus inference is independent of the size of the domain!

There are cases where this split is not possible, but we have developed a *first-order* loopy propagation algorithm to deal with them.

Conclusion

- Added nominals (in a limited setting) to a simple description logics keeping it viable.
- Some next steps that seems easy to take towards SHOIN logic:
  - inverse roles;
  - cardinality.
- Future work: extend the first order loopy propagation to deal with the nominals.
- Hope to gradually close the remaining gap and a complete probabilistic version of OWL in future work.