PR-OWL: Probabilistic Web Ontology Language

Walter Perez Urcia

Universidade de São Paulo
Instituto de Matemática e Estadística
Departamento de Ciências da Computação

Outubro 2015

Outline

1. Motivation
2. Ontologies
3. Probabilistic Ontologies

Motivation

Semantic Web

"The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation" [5]

- Extension of the web through standards by W3C
- Change focus from data driven to knowledge driven

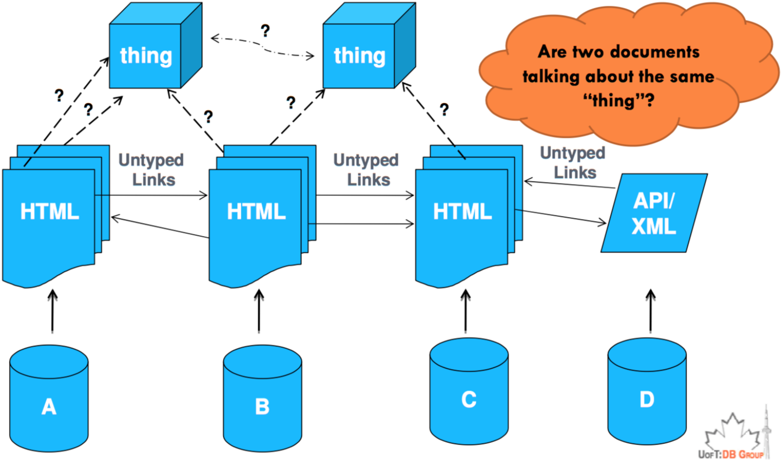

Figure 1: Document-oriented Web

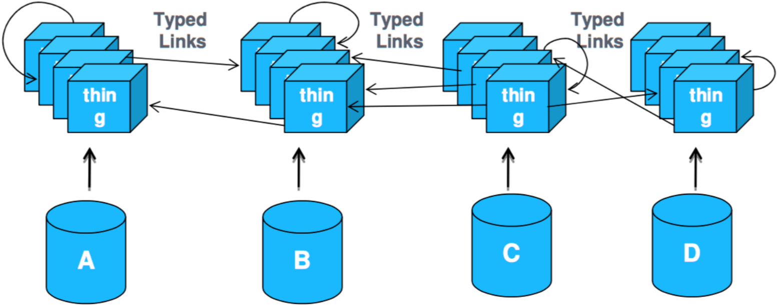

Figure 2: Data-oriented web

The main challenges that automated reasoning systems will have to deal in order to deliver on the promise of the Semantic Web are:

- Vastness
- Vagueness
- Uncertainty
- Inconsistency
- Deceit

The most used solution for Semantic Web are Ontologies (W3C).

Ontologies

An ontology is an explicit, formal knowledge representation that express knowledge about a domain of application. This includes:

- Types of entities that exist in the domain
  - e.g.: Person, Enterprise, ...
- Properties of those entities
  - e.g.: firstName, lastName, ...
- Relationships among entities
  - e.g.: motherOf, ownerOf, ...
- Processes and events that happen with those entities
  - e.g.: choosing best proposal, ...

How to represent ontologies?

How to represent ontologies?

Web Ontology Language (OWL)

- Created at 2004
- Developed by the W3C
- As a language to represent ontologies for the Semantic Web

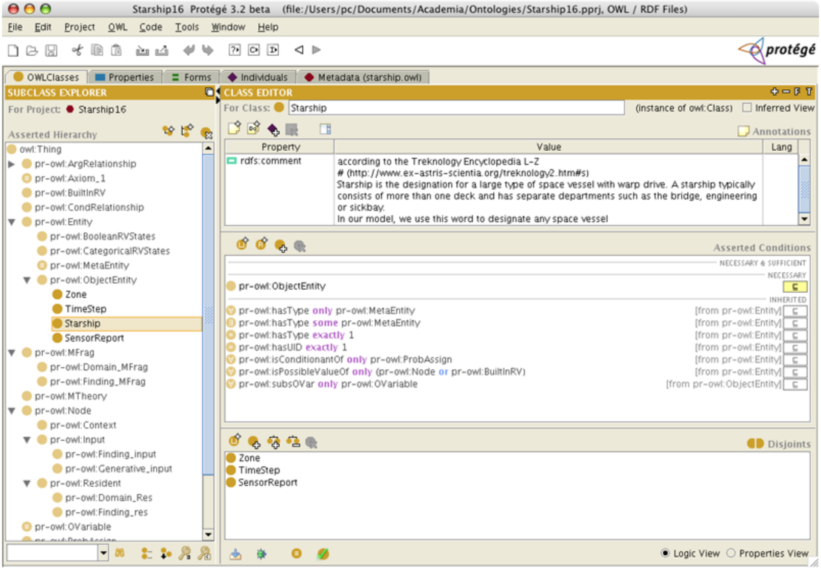

Figure 3: Protegé program

Problems

There is uncertainty on the Web.

Problems

There is uncertainty on the Web.
Suppose we search:

- Python
  - Result entity with types: animal or programming language?

Problems

There is uncertainty on the Web.
Suppose we search:

- Python
  - Result entity with types: animal or programming language?
- Python Cascabel
  - Result entity with types: animal or python library?

How to deal with uncertainty?

Probabilistic Ontologies

An ontology is an explicit, formal knowledge representation that express knowledge about a domain of application. This includes:

- Types of entities that exist in the domain
- Properties of those entities
- Relationships among entities
- Processes and events that happen with those entities

A probabilistic ontology is an explicit, formal knowledge representation that express knowledge about a domain of application. This includes:

- Types of entities that exist in the domain
- Properties of those entities
- Relationships among entities
- Processes and events that happen with those entities
- Statistical regularities that characterize the domain
- Inconclusive, ambiguous, incomplete, unreliable, and dissonant knowledge related to entities of the domain
- Uncertainty about all the above forms of knowledge

How to represent probabilistic ontologies?

How to represent probabilistic ontologies?

Some possible solutions

- Plates
  - Represent fragments of graphical models
  - Very useful with continuous attribute values
  - Can not represent directly uncertainty

How to represent probabilistic ontologies?

Some possible solutions

- Plates
  - Represent fragments of graphical models
  - Very useful with continuous attribute values
  - Can not represent directly uncertainty
- Bayesian Networks (BNs)
  - Represent uncertainty
  - Limited to attribute-value representation

How to represent probabilistic ontologies?

Some possible solutions

- Plates
  - Represent fragments of graphical models
  - Very useful with continuous attribute values
  - Can not represent directly uncertainty
- Bayesian Networks (BNs)
  - Represent uncertainty
  - Limited to attribute-value representation
- Hidden Markov Models
  - They are a special case of Dynamic BNs
  - Capability for recursion

How to represent probabilistic ontologies?

Some possible solutions

- Probabilistic Relational Models
  - Extend BNs to handle multiple entity types
  - Can not express quantified first-order sentences
  - Does not support recursion

How to represent probabilistic ontologies?

Some possible solutions

- Probabilistic Relational Models
  - Extend BNs to handle multiple entity types
  - Can not express quantified first-order sentences
  - Does not support recursion
- DAPER
  - Combine entity-relational model with DAGs models
  - Express probabilistic knowledge
  - Does not support quantifiers

All of these representations can not successfully combine Logic and Uncertainty

Multi-Entity BNs [2]

- Extends Bayesian Networks to provide first-order expressive power
- Extends first-order logic to provide probability distributions
- Support for recursion and quantifiers
- Built up from MEBN fragments or MFrags
- Every first-order sentence can be represented by an MFrag

MEBN Fragment

An MFrag is a 5-tuple $F = (C, I, R, G, D)$ where

- C = finite set of *context* value assignments
- I = finite set of *input* random variable terms
- R = finite set of *resident* random variable terms
- G = a fragment graph
- D = local probability distributions for each $r \in R$

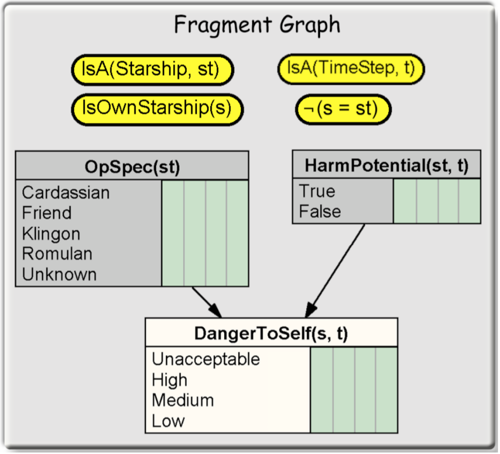

Figure 4: Danger To Self MFrag

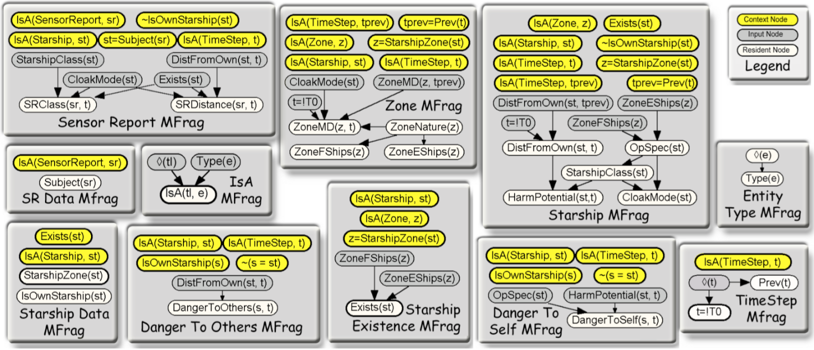

Figure 5: Star Trek MEBN

- Extension of OWL used for Ontologies
- Used to represent probabilistic ontologies
- Based on MEBN theory
- Still an active area for research

Main Researchers

- Marcelo Ladeira, PhD (Universidade de Brasilia)
- Kathryn B. Laskey, PhD (George Mason University)

Applications and studies

- UnBBayes: A probabilistic network framework for Probabilistic Ontologies [4]
- Probabilistic Ontology and Knowledge Fusion for Procurement Fraud Detection in Brazil [1]
- PR-OWL 2 Case Study: A Maritime Domain Probabilistic Ontology [3]

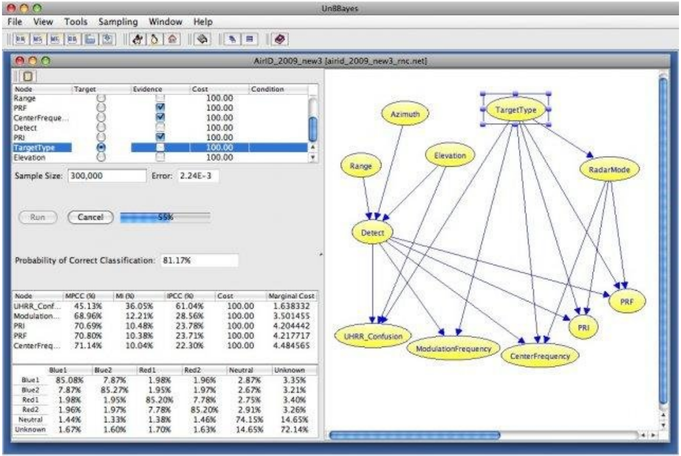

Figure 6: UnBBayes GUI

Referências I

RommelN. Carvalho, Shou Matsumoto, KathrynB. Laskey, PauloC.G. Costa, Marcelo Ladeira, and LaécioL. Santos.
Probabilistic Ontology and Knowledge Fusion for Procurement Fraud Detection in Brazil.
In *Uncertainty Reasoning for the Semantic Web II*, volume 7123 of *Lecture Notes in Computer Science*, pages 19–40. Springer Berlin Heidelberg, 2013.

Kathryn B. Laskey.
MEBN: A Language for First-ORder Bayesian Knowledge Bases.
Artificial Intelligence, 172, 2008.

Referências II

- Kathryn Blackmond Laskey, Richard Haberlin, Paulo Costa, and Rommel Novaes Carvalho.
  PR-OWL 2 Case Study: A Maritime Domain Probabilistic Ontology, 2011.

- Show Matsumoto Marcelo Ladeira Paulo Costa Rommel Carvalho, Laecio Santos.
  UnBBayes-MEBN: Comments on Implementing a Probabilistic Ontology Tool.
  IADIS Applied Computing 2008 conference, 2008.

- James Hendler Tim Berners-Lee and Ora Lasilla.
  The Semantic Web.
  Scientific American, 2001.