Ontologies in Support of Knowledge Exchange in Air Traffic Control Applications

Paulo C. G. Costa, Ph.D.
Associate Professor, SEOR Dept.
Agenda

❖ Knowledge Engineering Letter Soup
❖ Ontologies
❖ Beyond Ontologies
❖ FIXM, AIXM, and WXXM
The K. Eng. Letter Soup

Taxonomy

Extended ER

Relational Model

Ontology

Conceptual Model

Thesaurus

Description Logic

First Order Logic

Conceptual Model

Web Agents

XML

XML Schema

UML

Topic Maps

Semantic Web

XTM

DAML+OIL

Model Logic

OWL

RDF

OWL

Relational Model

Extended ER

Schema

Schema

PGRA

<owl:sameAs rdf:resource="#Poor GMU Research Assistant"/>
Extensible Markup Language (a.k.a. XML)

- XML is:
  - A serialization format. It provides a way for encoding information so it can be parsed when passed between machines.
  - A markup language that defines a set of rules for encoding documents
  - Extensible. It has no predefined tags.
  - THE interchange format for passing data over the Internet

- Usually confused with:
  - HyperText Markup Language (HTML), which is designed to create structured documents.
  - XHTML an XML-based HTML. It serves the same function as HTML, but with the same rules as XML documents. These rules deal with the structure of the markup.
  - XML Schema, an XML written language that allows language grammar definitions with precise structural form in addition to specific data typing conventions for elements, structures, and attributes.

An XML Purchase Order

```
<?xml version="1.0"?>
<purchaseOrder orderDate="1999-10-20">
  <shipTo country="US">
    <name>Alice Smith</name>
    <street>123 Maple Street</street>
    <city>Mill Valley</city>
    <state>CA</state>
    <zip>90952</zip>
  </shipTo>
  <billTo country="US">
    <name>Robert Smith</name>
    <street>8 Oak Avenue</street>
    <city>Old Town</city>
    <state>PA</state>
    <zip>95819</zip>
  </billTo>
  <comment>Hurry, my lawn is going wild!</comment>
  <items>
    <item partNum="872-AA">
      <productName>Lawnmower</productName>
      <quantity>1</quantity>
      <USPrice>148.95</USPrice>
      <comment>Confirm this is electric</comment>
    </item>
    <item partNum="926-AA">
      <productName>Baby Monitor</productName>
      <quantity>1</quantity>
      <USPrice>39.98</USPrice>
      <shipDate>1999-05-21</shipDate>
    </item>
  </items>
</purchaseOrder>
```

http://jmvidal.cse.sc.edu/talks/xmlrdfdaml/xml2.html?style=White
Resource Description Framework (a.k.a. RDF)

- RDF is an **abstract data model**, a way to break down knowledge into discrete pieces.
- It provides structure (a framework) for representing data
- A fact is expressed as a triple of the form (Subject, Predicate, Object)
- While it is most popularly known for its RDF/XML syntax, RDF can be stored in a variety of formats.

### XML/RDF Notation

```xml
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix ex: <http://www.example.org/>.

ex:vincent_donofrio    ex:starred_in
ex:law_and_order_ci .

ex:law_and_order_ci    rdf:type ex:tv_show .

ex:the_thirteenth_floor    ex:similar_plot_as ex:the_matrix .
```

### N3 Notation

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix ex: <http://www.example.org/> .

ex:law_and_order_ci    rdf:type    ex:tv_show .
ex:the_thirteenth_floor    ex:similar_plot_as    ex:the_matrix .
```

Ontology Web Language (a.k.a. OWL)

- OWL is:
  - …built on top of RDF
  - …written in XML, designed to be interpreted by computers
  - …*not* designed for being read by people
  - …a W3C standard composed by 3 sublanguages

- OWL adds vocabulary, a richer set of properties and restrictions to RDF, which enables expressing logical statements (…and therefore applying logical reasoning)

- OWL sublanguages:
  - OWL lite: simpler, decidable
  - OWL DL: based on Description Logics, restricted capabilities
  - OWL Full: unrestricted FOL expressiveness, not decidable
OWL vs. RDF

- Additions to RDF include…
  - Relationship between classes (ex: disjointWith, complementOf, inverseOf)
  - Equality (ex: sameAs, equivalentClass)
  - Richer properties (ex: symmetric, transitive, functional)
  - Class property restrictions (ex: allValuesFrom, someValuesFrom)

- In short, OWL…
  - …provides a much richer vocabulary, including annotations and metadata
  - …is more rigid, specifying what and how to express things
  - …has support for logical reasoning in various flavors

- Caveat: With power comes responsibility
  - It is easy to do bad models in OWL
Summary So Far…

- XML: a syntax
  - Used for creating, structuring, and encoding documents
  - Often used to store data and to allow for communication between applications

- RDF: a data model
  - It can be represented in various syntaxes/formats. RDF/XML is the most commonly used
  - Shines when distributed data needs to be represented

- OWL: an ontology language
  - Enhances RDF syntax so it can express logical statements
  - Used for building ontologies
Agenda

❖ Knowledge Engineering Letter Soup
❖ Ontologies
❖ Beyond Ontologies
❖ FIXM, AIXM, and WXXM
Ontologies

- In Philosophy: the study of nature of being and knowing
- In Information Systems: many definitions

An ontology formally defines a common set of terms that are used to describe and represent a domain. Ontologies can be used by automated tools to power advanced services such as more accurate Web search, intelligent software agents and knowledge management. (Owl Use Cases)

An explicit formal specification on how to represent the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships among them. (dictionary.com)

In information science, an ontology is the product of an attempt to formulate an exhaustive and rigorous conceptual schema about a domain. An ontology is typically a hierarchical data structure containing all the relevant entities and their relationships and rules within that domain (Wikipedia.org).

What is really Important?

A partial specification of a conceptual vocabulary to be used for formulating knowledge-level theories about a domain of discourse. The fundamental role of an ontology is to support knowledge sharing and reuse. (The Internet Reasoning Services project - IRS)
Definition: An ontology is an explicit, formal representation of knowledge about a domain of application. This includes:

- a) Types of entities that exist in the domain;
- b) Properties of those entities;
- c) Relationships among entities;
- d) Processes and events that happen with those entities;

where the term entity refers to any concept (real or fictitious, concrete or abstract) that can be described and reasoned about within the domain of application.

Why Ontologies?

• Ontologies allow structural and semantic definitions of documents, providing:
  › Intelligent search, instead of keyword matching
  › Query answering, instead of information retrieval
  › Integration of data from different sources
  › Improved maintainability, modularity, and extensibility of systems

• In short: ontologies convey the meaning of the resources they represent (not only web resources), allowing software agents to select and combine it in an appropriate way.
The Ontology Spectrum

From:
(DaConta et al., 2003, page 157)

Logical Theory
Description Logic
OWL
UML
First Order Logic
Modal Logic
Is Disjoint Subclass Of, with transitivity property

Conceptual Model
RDF/S
XTM
Extended ER
Thesaurus
ER
DB Schemas, XML Schema

Taxonomy
Relational Model, XML

Weak Semantics

Strong Semantics

From: (DaConta et al., 2003, page 157)

Advancing Along the Interpretation Continuum

**Interpretation Continuum**

*Human interpreted* → *Computer interpreted*

**DATA**
- Relatively unstructured
- Random

**KNOWLEDGE**
- Very structured
- Logical

**Simple Metadata:**
- XML

**Richer Metadata:**
- RDF/S

**Very Rich Metadata:**
- OWL

Display raw documents; All interpretation done by humans

Find and correlate patterns in raw docs; display matches only

Store and connect patterns via conceptual model (i.e., an ontology); link to docs to aid retrieval

Automatically acquire concepts; evolve ontologies into domain theories; link to institution repositories (e.g., MII)

Automatically span domain theories and institution repositories; inter-operate with fully interpreting computer

Moving to the right depends on increasing automated semantic interpretation

*This slide was presented at STIDS 2013 and is being used here with permission from Dr. Leo Obrst*
Syntax: Propositional Logic

• PL is a Language having a Syntax & a Semantics
  – A set of symbols:
    • Logical Constants: True, False (or T, F)
    • Logical Variables (or propositional symbols): p, q, r, …
    • Logical Operators (or connectives): ¬, ∨, ∧, →, ↔, (, )
  – Formulas (Well-formed Formulas or WFFs) of PL (we will call these *propositions*)
    • Any propositional symbol is a WFF of PL
    • If α and β are WFFs, then so are (α ∨ β), (α ∧ β), (α → β), (α ↔ β), and (¬α) [and note that we will dispense with parentheses where we can]
    • Nothing else is a WFF.
  – So the following are WFFs: p, ¬p, p ∨ q, p ∧ q, (p ∧ q) → r
  – Propositions are things that are true or false

Propositions in English:
If John is a management employee, then
John manages an organization.
John is a management employee.
John manages an organization (MP)

Propositions in PL:
p → q
p
q (MP: Modus Ponens)

Still Need Semantics!
Predicate Logic: 
Add Predicates, Individuals, Quantifiers

Propositions & Predicates in English:
If John is a management employee, then John manages an organization.

John is a management employee. 
\[ \text{John manages an organization} \] (MP)

Propositions & Predicates in First Order
Predicate Logic:
\[ p(x) \rightarrow q(x) \]
\[ p(john) \]
\[ q(john) \] (MP: Modus Ponens)

Propositions & Predicates in English:
Everyone who is a management employee manages some organization.

Or:
For everyone who is a management employee, there is some organization that that person manages.

John is a management employee.
\[ \text{There is some organization that John manages.} \]

Propositions & Predicates in First Order
Predicate Logic:
\[ \forall x. [p(x) \rightarrow \exists y. [q(y) \land r(x,y)]] \]

“For all x, if x is a p, then there is some y such that y is a q, and x is in the r relation to y”

p(john)
\[ \exists y. [q(y) \land r(john,y)] \] (MP: Modus Ponens)

Still Need Semantics!
Propositional & Predicate Logic

• Propositional Logic
  – Limitation: cannot speak about individuals (instances)
  – Granularity not fine enough
  – Propositions: truth-functions
    
    \[
    \text{If Plato is human, then Plato is mortal} \quad p \rightarrow q \\
    \text{Plato is human} \quad p \\
    \]  

    \[
    \text{Plato is mortal} \quad q \quad \text{Modus Ponens}
    \]

• Predicate Logic
  – Finer distinctions: can talk about individuals (instances)
    
    \[
    \text{If Plato is human, then Plato is mortal} \quad \forall x: p(x) \rightarrow q(x) \\
    \text{Plato is human} \quad p(plato) \\
    \]  

    \[
    \text{Plato is mortal} \quad q(plato) \quad \text{Modus Ponens}
    \]

  – An instantiated predicate is a proposition, e.g., human(plato) = true
Proof Using Inference Rule of Modus Ponens

**Given:** If motherOf is a subProperty of parentOf, and Mary is the mother of Bill, then Mary is the parentOf Bill

- motherOf is a subProperty of parentOf
- Mary is the motherOf Bill

**Infer:**

Mary is the parentOf Bill

Deduction: A method of reasoning by which one infers a conclusion from a set of sentences by employing the axioms and rules of inference for a given logical system.

The term **ontology** has been used to describe models with different degrees of structure (Ontology Spectrum)

- **Less structure:** Taxonomies (Semio/Convera taxonomies, Yahoo hierarchy, biological taxonomy, UNSPSC), Database Schemas (many) and metadata schemes (ICML, ebXML, WSDL)
- **More Structure:** Thesauri (WordNet, CALL, DTIC), Conceptual Models (OO models, UML)
- **Most Structure:** Logical Theories (Ontolingua, TOVE, CYC, Semantic Web)

Ontologies are usually expressed in a logic-based language

- Enabling detailed, sound, meaningful distinctions to be made among the classes, properties, & relations
- More expressive meaning but maintain “computability”

Using ontologies, tomorrow’s applications can be “intelligent”

- Work at the human conceptual level

Ontologies are usually developed using special tools that can model rich semantics
A Military Example of Ontology

This slide was presented at STIDS 2013 and is being used here with permission from Dr. Leo Obrst
Defined vs. Inferred

This defined class has conditions that are part of the definition: ie any Pizza that has the country of origin, Italy is a RealItalianPizza. It also has conditions that merely describe the members – that all RealItalianPizzas must only have ThinAndCrispy bases.

Pizza
hasCountryOfOrigin has Italy
hasBase only ThinAndCrispyBase
hasBase some PizzaBase
Ontologies vs. OO

- Databases / OO
  - Rigidly defined classes that govern the system behavior
  - All instances are created as members of some class.
  - Changing a class affects all of its instances
  - *Closed World Assumption* / Well suited for top down governance

- Ontologies
  - Rely on logical reasoners to infer class relationships and instance membership
  - Flexible format that adapts its class structure as new information is learned
  - *Open World Assumption* / Well suited for open systems
Agenda

❖ Knowledge Engineering Letter Soup
❖ Ontologies
❖ Beyond Ontologies
❖ FIXM, AIXM, and WXXM
Deterministic Reasoning

(... is not always suitable to the problem)

• All Birds lay eggs
• Many aquatic birds have Duck-like bills
• Many aquatic birds have webbed feet (like ducks)
• All aquatic birds swim very well and can hold breath for a long period
• Joe:
  ‣ Is an egg-laying animal
  ‣ Has a Duck-like bill
  ‣ Has webbed feet
  ‣ Swin very well and can hold breath for a long period
• Therefore: Joe is a…
Joe is a Duck-Billed Platypus (a mammal)

- It lays eggs like a bird or a reptile (this makes it a monotreme mammal)
- The males have poison like a snake in spurs on their hind legs. The poison can kill a dog and cause extreme pain in people.
- They have a bill like a duck.
- They have a tail like a beaver.
- They have webbed feet like a duck.
- The mother's milk comes out through glands on her skin and the babies lick it off of her fur.
Ontologies

- An ontology is an explicit, formal representation of knowledge about a domain of application. This includes:
  - Types of entities that exist in the domain: Fighter, APV, Missile, ...
  - Properties of those entities: hasMaxSpeed, hasNetWeight, hasMaxGRate, ...
  - Relationships among entities: isCommissionedAt, supports, hasLaunchBase, ...
  - Processes and events that happen with those entities: participate in mission x, ...
  - 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;

where the term entity refers to any concept (real or fictitious, concrete or abstract) that can be described and reasoned about within the domain of application.
Knowledge Workflow

Typical Web Service/Agent's Knowledge Flow
PR-OWL Probabilistic Ontology Language

- Upper ontology written in W3C-recommended OWL ontology language.
- Represents probabilistic knowledge in XML-compliant format.
- Based on MEBN, a probabilistic logic with first-order expressive power.
- Open-source, freely available solution for representing knowledge and associated uncertainty.
- Reasoner under development in collaboration with University of Brasilia.
Bayesian Networks

- A formal language for representing knowledge about uncertain quantities
  - nodes represent hypotheses
  - arcs represent direct dependency relationships among hypotheses
  - conditional probabilities encode strengths of dependencies

- A computational architecture for computing impact of evidence on beliefs
  - updates beliefs when new evidence is observed
  - exploits independence assumptions to make computation more efficient

- Representation language for probabilistic knowledge base
  - Domain ontology
  - Fragments of Bayesian networks
  - Composition operators for “gluing together” fragments

- Inference engine for probabilistic knowledge based system
  - Retrieve appropriate fragments from knowledge base and “glue together”
  - Apply inference algorithm to draw appropriate inferences
Related Work: Bayesian SW

- Annotating Web Languages with Bayesian Network Concepts
  - Bayesian Extensions to RDF: Yoshio Fukushige, Matsushita, Japan

- Extending OWL using Bayesian Networks
  - Extending OWL with Bayesian Network concepts: Ding & Peng, UMBC
  - Dealing with uncertainty contexts using BNs: Gu et al., National University of Singapore
Why not BNs?

The Star Trek Problem: Discriminating Starships and making decisions with incomplete and uncertain knowledge.

Bayesian Reasoning
Update prior beliefs as evidence accrues.
All new data can be considered.

The Star Trek Problem: Discriminating Starships and making decisions with incomplete and uncertain knowledge.
Why Not BNs?

How about multiple starships showing up at the same time? One BN for each situation?
Multi-Entity Bayesian Networks
MEBN Fragments

- Building blocks that collectively form a model (MTheory)
- Each one stores a specific "Chunk of knowledge"

The danger to self MFragment
PR-OWL General Architecture

Main Classes / Elements
- Finding
- Simple Argument relationship
- Argument relationship
- Context
- Input
- Resident
- Declarative distribution
- Probability Distribution
- PR-OWL table

SubClasses
- Finding MFragment
- MFragment
- Ordinary Variable
- Exemplar
- Built-In RV
- Boolean RV states
- Categorical RV states
- Entity
- Meta-Entity
- Object Entity

Reified Relationships

Support / Built-in Elements
- Domain MFragment
- Domain Resident
- Finding Resident
- Probability Assignment
- Conditional relationship
PR-OWL vs. OWL
Resources - PR-OWL Website

PR-OWL: A Bayesian extension to the OWL Ontology Language

A Bayesian Framework for Probabilistic Ontologies

What is PR-OWL?
PR-OWL is an open research work aimed to extend the OWL2 ontology Web language so it can represent probabilistic ontologies. In other words, it is a probabilistic extension to OWL that provides a framework for authoring probabilistic ontologies and is based on the Bayesian first-order logic called Multi-Entity Bayesian Networks (MEBN).

A More Detailed Explanation
Uncertainty is ubiquitous. Any representation scheme intended to model real-world actions and processes must be able to cope with the effects of uncertain phenomena.

A major shortcoming of existing Semantic Web technologies is their inability to represent and reason about uncertainty in a sound and principled manner. This not only hinders the realization of the original vision for the Semantic Web, but also raises an unnecessary barrier to the development of new, powerful features for general knowledge applications.

The overall goal of our research is to establish a Bayesian framework for probabilistic ontologies, providing a basis for plausible reasoning services in the Semantic Web. As an initial effort towards this broad objective, this dissertation introduces a probabilistic extension to the Web ontology language OWL2, thereby creating a crucial enabling technology for the development of probabilistic ontologies.

The extended language, PR-OWL (pronounced as "prowl"), adds new definitions to current OWL while retaining backward compatibility with its base language. Thus, OWL-built legacy ontologies will be able to interoperate with newly developed probabilistic ontologies. PR-OWL moves beyond deterministic classical logic (Frege, 1879; Peirce, 1885), having its formal semantics based on MEBN probabilistic
Agenda

❖ Knowledge Engineering Letter Soup
❖ Ontologies
❖ Beyond Ontologies
❖ FIXM, AIXM, and WXXM
The Data Exchange Model Family

**Information Domain**

<table>
<thead>
<tr>
<th>Aeronautical</th>
<th>Weather</th>
</tr>
</thead>
<tbody>
<tr>
<td>An information domain that identifies and describes the infrastructure (e.g., facility, system, airspace), aviation services, and operational rules that ensure stakeholders can operate safely and efficiently within the national airspace system. It describes the definitions, schedules for, configuration and state of the infrastructure, aviation services, and operational rules of the NAS operational environment. This information is produced from aeronautical data provided by other functions and external entities, as well as that derived internally. Aeronautical Information is a dynamic, shared information resource supporting most of the pre-operational, operational, and post-operational processes used in air traffic management.</td>
<td>An information domain that identifies and describes the observation, processing, interpretation, forecasting, distribution and storage of aviation weather information and associated products and services to support all phases of flight. WXXM shall support and facilitate system-wide interoperability and will assure the quality and integrity of the delivered information.</td>
</tr>
</tbody>
</table>

**Exchange Model**

| (AIXM) | (WXXM) |
| Establishes the aeronautical information standard and enables its dissemination, using XML structures, to support enterprise-wide information exchange standards. | Establishes the weather information standard and enables its dissemination, using XML structures, to support enterprise-wide information exchange standards. |

| Flight |
| An information domain that identifies and describes an extensible and dynamic collection of flight-specific data elements describing an individual flight from planning through operation, including preferences and constraints; where appropriate, aeronautical information is leveraged. |
| Establishes the flight information standard and enables its dissemination, using XML structures, to support enterprise-wide information exchange standards. |

**Supporting Domains**

- Security
- Environment
- Traffic
- Etc.

Flight Information Exchange Model (a.k.a. FIXM)

- FIXM is a data interchange format for sharing information about flights throughout their lifecycle.
- It will standardize current and future data elements to increase interoperability and data exchange between systems that require Flight Data.
- FIXM will include formatting attributes for data such as aircraft identifiers and parameters; current flight plan information; operator preferences, constraints and SOPs; flight capabilities, preferences and constraints; security information; etc.

http://www.fixm.aero/
Flight Information Exchange Descriptions

- **FIXM**: the *data interchange format*
- **FO**: the Flight Object will be an extensible and dynamic *collection of data elements* that describes an individual flight
- **FOXS**: the Flight Object eXchange Service will be the technical or *application infrastructure* required to exchange standardized Flight Data
  - It is envisioned to be used for overall management of the Flight Object, potentially including areas such as Flight Object services, read/write privileges, security and encryption, performance, etc
  - FOXS is not intended to necessarily infer creation of a single centralized Flight Object “system”, but rather FOXS may be a collection of functions distributed amongst multiple NAS and international systems

http://www.fixmap.aero/
Flight Information Exchange Tomorrow

Aeronautical Information Exchange Model (a.k.a. AIXM)

- AIXM is a data interchange format designed to enable the management and distribution of Aeronautical Information Services (AIS) data in digital format.
- It has two main components, the AIXM Conceptual Model and the AIXM XML Schema.
- The AIXM Conceptual Model is the component of the AIXM data standard that provides a conceptual model of aeronautical data.
  - It models the important features, properties (attributes and associations) and business rules that make up aeronautical information. As such, it can be used as the basis for the design of an AIM database.
  - The model is designed using the Unified Modeling Language (UML).
- The Schema derives from the AIXM Conceptual Model by mapping its features, attributes and business rules into XML.

http://www.aixm.aero/
Weather Information Exchange Model (a.k.a. WXXM)

- The Weather Information Exchange Models are designed to enable a platform independent, harmonized and interoperable meteorological information exchange covering all the needs of the air transport industry. It has three main components:
  - The conceptual Information Model (WXCM)
  - The Logical Data Model (WXXM)
  - The Exchange Schema (WXXS)
- The conceptual Information Model (WXCM) is a high level conceptual model of the MET domain.
  - It provides a view of the concepts and packages (e.g., ICAO Annex 3, ISO 19100, WMO BUFR tables), that make up the data model.
  - The model is designed using the Unified Modeling Language (UML).
- The logical data model that combines the concepts from the high-level WXCM packages into a coherent model that takes into account concerns related to data exchange.
- The WXXM is the component of the Meteorological Information Exchange used for system-to-system exchange of meteorological information.
- The WXXS derives from the WXXM by mapping its features, attributes and business rules into XML.

http://www.wxxm.aero/
Wait a minute! UML what??

- We talked about XML, RDF, OWL, ontologies, etc., but FIXM, AIXM, and WXXM use UML (although also use XML). What is this?

- These are tools for different purposes
  - UML is meant to model object-oriented software artifacts
  - RDF is meant to model (distributed) data
  - OWL is meant to model knowledge about a domain

- Yet, they are all modeling languages with overlapping features and different foci. There are initiatives to map those languages
  - UML is closer to RDF schema than to RDF or OWL. See a mapping at http://www.w3.org/TR/NOTE-rdf-uml/
  - There are converters from UML to OWL, but OWL is more expressive so there are no converters from OWL to UML
Discussion