An Introduction to the Semantic Web

Jeff Heflin
Lehigh University
The Semantic Web

Definition

– The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation. (Berners-Lee et al., Scientific American, May 2001)

Applications

– managing corporate web sites (intranets)
– more automatic generation of web portals
– better indexing of multimedia resources
– web agents and web services
– ubiquitous computing
Ontology

Definition
- a logical theory that accounts for the intended meaning of a formal vocabulary (Guarino 98)
- has a formal syntax and unambiguous semantics
- inference algorithms can compute what logically follows

Relevance to the Semantic Web:
- ontologies define the semantics of the terms used in semi-structured web pages
- identify context
- provide shared definitions
- ease the integration of distinct resources
A Web of Ontologies

The answer to a user’s query might require the combination of data from $S_1$, $S_3$, and $S_4$. 
Semantic Web Standards

World Wide Web Consortium (W3C) Recommendations

- **RDF(S) (1999, revised 2004)**
  - essentially semantic networks with URIs
  - XML serialization syntax

- **OWL (2004)**
  - extends RDF with more semantic primitives
  - based on description logics (DLs)
  - has a model theoretic semantics

```xml
<owl:Class rdf:ID="Band">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasMember" /> <owl:allValuesFrom rdf:resource="#Musician" />
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

A Band is a subset of the groups which only have Musicians as members
```
URIs and Namespaces

- **URI**
  - Uniform Resource Identifier
  - includes URLs
  - but also anything that you can design an identification scheme for
  - helps to prevent collision of names
  - all the “symbols” in RDF are either URIs or Literals

- **Namespace**
  - a mechanism for abbreviating URIs
  - by assigning a prefix for a URI fragment
Description Logic (DL)

- form of knowledge representation
  - useful for formally defining classes
  - studied extensively in 1990’s
  - mature reasoning software
    » e.g., FaCT, RACER, Pellet

- benefits
  - optimized computation of subsumption
    » calculate implicit subClassOf relations
  - ontology integration
    » if two ontologies use class expressions to define their vocabularies in terms of a third ontology, then subsumption can be used to compute an integrated ontology
# OWL Class Constructors

<table>
<thead>
<tr>
<th>Constructor</th>
<th>DL Syntax</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>intersectionOf</td>
<td>$C_1 \sqcap \ldots \sqcap C_n$</td>
<td>Human $\sqcap$ Male</td>
</tr>
<tr>
<td>unionOf</td>
<td>$C_1 \sqcup \ldots \sqcup C_n$</td>
<td>Doctor $\sqcup$ Lawyer</td>
</tr>
<tr>
<td>complementOf</td>
<td>$\neg C$</td>
<td>$\neg$Male</td>
</tr>
<tr>
<td>oneOf</td>
<td>${x_1 \ldots x_n}$</td>
<td>${\text{john, mary}}$</td>
</tr>
<tr>
<td>allValuesFrom</td>
<td>$\forall P.C$</td>
<td>$\forall \text{hasChild}.\text{Doctor}$</td>
</tr>
<tr>
<td>someValuesFrom</td>
<td>$\exists P.C$</td>
<td>$\exists \text{hasChild}.\text{Lawyer}$</td>
</tr>
<tr>
<td>maxCardinality</td>
<td>$\leq nP$</td>
<td>$\leq 1 \text{hasChild}$</td>
</tr>
<tr>
<td>minCardinality</td>
<td>$\geq nP$</td>
<td>$\geq 2 \text{hasChild}$</td>
</tr>
</tbody>
</table>

*borrowed from Ian Horrocks*
# OWL Axioms

<table>
<thead>
<tr>
<th>Axiom</th>
<th>DL Syntax</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>subClassOf</td>
<td>$C_1 \sqsubseteq C_2$</td>
<td>Human $\sqsubseteq$ Animal $\sqcap$ Biped</td>
</tr>
<tr>
<td>equivalentClass</td>
<td>$C_1 \equiv C_2$</td>
<td>Man $\equiv$ Human $\sqcap$ Male</td>
</tr>
<tr>
<td>disjointWith</td>
<td>$C_1 \sqsubseteq \lnot C_2$</td>
<td>Male $\sqsubseteq \lnot$ Female</td>
</tr>
<tr>
<td>sameIndividualAs</td>
<td>${x_1} \equiv {x_2}$</td>
<td>${\text{President_Bush}} \equiv {\text{G_W_Bush}}$</td>
</tr>
<tr>
<td>differentFrom</td>
<td>${x_1} \sqsubseteq \lnot{x_2}$</td>
<td>${\text{john}} \sqsubseteq \lnot{\text{peter}}$</td>
</tr>
<tr>
<td>subPropertyOf</td>
<td>$P_1 \sqsubseteq P_2$</td>
<td>hasDaughter $\sqsubseteq$ hasChild</td>
</tr>
<tr>
<td>equivalentProperty</td>
<td>$P_1 \equiv P_2$</td>
<td>cost $\equiv$ price</td>
</tr>
<tr>
<td>inverseOf</td>
<td>$P_1 \equiv P_2^-$</td>
<td>hasChild $\equiv$ hasParent$^-$</td>
</tr>
<tr>
<td>transitiveProperty</td>
<td>$P^+ \sqsubseteq P$</td>
<td>ancestor$^+$ $\sqsubseteq$ ancestor</td>
</tr>
<tr>
<td>functionalProperty</td>
<td>$\top \sqsubseteq \sqsubseteq 1P$</td>
<td>$\top \sqsubseteq \sqsubseteq 1\text{hasMother}$</td>
</tr>
<tr>
<td>inverseFunctionalProperty</td>
<td>$\top \sqsubseteq \sqsubseteq 1P^-$</td>
<td>$\top \sqsubseteq \sqsubseteq 1\text{hasSSN}^-$</td>
</tr>
</tbody>
</table>

*borrowed from Ian Horrocks*
OWL Inference

- The head of an organization is also a member of it
- A member of a terror organization is a terrorist
- Therefore, the head of a terror organization is a terrorist

```xml
<owl:Property rdf:about="&s;head">
  <rdfs:subPropertyOf rdfs:resource="&s;member" />
</owl:Property>

<owl:Class rdf:about="&t;Terrorist">
  <owl:sameClassAs>
    <owl:Restriction>
      <owl:onProperty rdf:resource="&s;member" />
      <owl:someValuesFrom rdf:resource="&s;TerroristOrg" />
    </owl:Restriction>
  </owl:sameClassAs>
</owl:Class>
```
Is the Semantic Web a Fad?

- Analysts have estimated that 35-65% of system integration costs are due to **semantic** issues.
- Companies that have invested in semantic solutions:
  - Time Inc., BellSouth, Raytheon, Walt Disney Company, General Motors, Cisco Systems, Met Life, etc.
- Growth of the Semantic Web:
  - 2005 → 350,000 RDF/OWL documents
  - Feb. 2006 → 1 million RDF/OWL documents
  - Nov. 2007 → 2.3 million RDF/OWL documents
For more information...

- For more on the Semantic Web
  - http://www.cse.lehigh.edu/~heflin/
  - http://www.semwebcentral.org/
  - http://www.w3.org/2001/sw/
  - http://www.daml.org/
  - http://www.semanticweb.org/