Semantic Web Technologien

Einführung in Semantic Web
Visionen und Technologien

Gong Chen

Betreuer: Steffen Koch
Contents

- An Overview of the Semantic Web
  Introduction
  How will the Semantic Web benefit us?

- Model of the Semantic Web
  Unicode / URI
  XML / XML Schema
  RDF / RDF Schema
  OWL
  Logic
  Trust

- Current state of the Semantic Web and our jobs
- Resources
An Overview of the Semantic Web

- Introduction

  - The current Web represents information using
    - natural language (English, Chinese, German,…)
    - graphics, multimedia, flash, …
An Overview of the Semantic Web

- Introduction

- Humans can understand this easily
  - can create mental associations
  - can deduce facts from partial information
  - are used to various sensory information

- But machines are not!
  - difficult to make sense from, e.g., an image
  - drawing analogies automatically is difficult
  - difficult to combine information automatically
Introduction

If the requirement is like:

- Find me the most suitable hotel for my journey
- Tell me if it’s possible to make contact with doctor Jones?

Surely we human can do it, but the computer??

At least not now, maybe in the future when the Semantic Web has been implemented.
An Overview of the Semantic Web

- How will the Semantic Web benefit us?

  Example 1: Reserve a hotel automatically

  Computer knows about your preferences
  - builds up knowledge base using your past
  - combines the local knowledge with remote services:
    ✓ dietary requirements
    ✓ calendaring
An Overview of the Semantic Web

- How will the Semantic Web benefit us?

Example 2: Make an appointment with professor

- Computer looks for the professor’s information on the web
- Communicate with the professor’s personal website to check if there’s still any chance of visiting
An Overview of the Semantic Web

- How will the Semantic Web benefit us?
  - With Semantic Web, the above mentioned will come true and even more…
An Overview of the Semantic Web

- How will the Semantic Web benefit us?
  - The Semantic Web is an extension of the current Web in which information is given well-defined meaning
  - It better enables computers and people to work in cooperation, providing a universally accessible platform that allows data to be shared and processed by automated tools as well as by people.
Model of the Semantic Web

- Semantic Web
  
  A Web that is meaningful to computers

Diagram: Semantic Web layers, from bottom to top:
- Unicode
- URI
- XML + NS + xmlschema
- RDF + rdfschema
- Ontology vocabulary
- Logic
- Data
- Rules
- Trust
- Proof
Model of the Semantic Web

- Unicode / URI

**Unicode**

is an industry standard designed to allow text and symbols from all of the writing systems of the world to be consistently represented and manipulated by computers.

**URI**

A *Uniform Resource Identifier* (URI), is a compact string of characters used to identify or name a resource. The main purpose of this identification is to enable interaction with representations of the resource over a network, typically the World Wide Web, using specific protocols.
Model of the Semantic Web

- XML / XML Schema

We are here
Model of the Semantic Web

- **XML / XML Schema**
  - A text-based meta-language format for data exchange
  - Provides a pathway to transfer data easily between various applications
  - Markup or Tags – identifies structures in the document (<name> </name>)
  - XML Schema – provides a schema to XML files
Model of the Semantic Web

- XML / XML Schema

```xml
<artist>
    <name>...</name>
    <homepage>...</homepage>
    <album>...</album>
</artist>
```

**Structure or Syntax**

- artist
- name
- homepage
- album
XML / XML Schema

Limitations of XML
- Makes no commitment towards domain-specific vocabulary
- Interoperability (of meaning) feasible only for closed collaboration
  - agents in a small & stable community
  - pages on a small & stable intranet
- Not suitable for sharing information in WWW
- XML can’t provide Semantics
Model of the Semantic Web

- RDF / RDF Schema

We are here
Model of the Semantic Web

- RDF / RDF Schema

What is Resource Description Framework?

- Defines a framework for structuring & describing resources (e.g., documents) in the Semantic Web
- Enables the definition of vocabularies for the description of the resources
- Goals:
  - Improved support for interpretation of data by machines
  - Extensibility, interoperability, and reuse of vocabularies
Model of the Semantic Web

- RDF / RDF Schema

The RDF Data Model

- Simple but powerful model for creation of metadata
- Can be expressed in XML
- Consists of three concepts:
  - Resource: an element, a URI, a literal..
  - Properties: directed relations between two resources
  - Statements: triples of two resources bound by a property
    - Usual terminology: (s, p, o) subject, predicate, object
Model of the Semantic Web

- RDF / RDF Schema

RDF Statement & Graph

- Each triple (s, p, o) represents a RDF statement

```
<table>
<thead>
<tr>
<th>subject</th>
<th>predicate</th>
<th>object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael Jackson</td>
<td>is the artist of</td>
<td>Bad</td>
</tr>
</tbody>
</table>
```
Model of the Semantic Web

RDF / RDF Schema

RDF Resource

- The Resource forms the central concept in RDF
- Anything can be described as a resource (E.g., website, book, picture, persons..)
- Resources are identified by URI’s

Model of the Semantic Web

- **RDF / RDF Schema**

  **RDF Property**
  - Represents the predicate of an RDF statement
  - Is labeled with a URI referencing to a RDF property
  - Is directed pointing from the subject of a statement to the object of a statement

  - **http://www.michaeljackson.com**
  - **http://www.music.org/songs/mj/Bad**
  - **music:Artist**
  - **Artist**
Model of the Semantic Web

- RDF / RDF Schema

How does RDF help?

- Vast majority of data processed by machines can be represented in the form of triples

- Subject, Predicate, Object are identified by URI’s
  - Allows to *uniquely* identify them
  - Concepts are *not* just words in a document, but are *tied* to a unique definition found in the Web

- Uniqueness is vital to make a *consistent* statement
  - Michael Jackson denoted by http://www.michaeljackson.com means the same to *everyone*!
Model of the Semantic Web

- RDF / RDF Schema

Why is RDF not enough?

- RDF properties can be regarded as attributes of resources
- RDF properties also represent relationships between resources
- But, RDF does not provide mechanisms for describing:
  - The properties (in terms of their range and domain)
  - The relationships between the properties and other resources
Model of the Semantic Web

- RDF / RDF Schema

RDFS – RDF Schema

- The RDF Vocabulary Description Language
- Enables us to:
  - Define classes of resources
  - Define relationships between the classes
  - Define the kinds of properties that instances of that classes have
  - Define relationships between properties
Model of the Semantic Web

- RDF / RDF Schema

  - `<rdf:Description ID="ModernMusic">`
    - `<rdf:type resource="http://www.w3.org/...#Class"/>`
    - `<rdfs:subClassOf rdf:resource="http://www.w3.org/...#Resource"/>`
  </rdf:Description>

  - `<rdf:Description ID="PopMusic">`
    - `<rdf:type resource="http://www.w3.org/...#Class"/>`
    - `<rdfs:subClassOf rdf:resource="#ModernMusic"/>`
  </rdf:Description>

  - `<rdf:Description ID="Artist">`
    - `<rdf:type resource="http://www.w3.org/...#Property"/>`
    - `<rdfs:domain rdf:resource="#PopMusic"/>`
    - `<rdfs:range rdf:resource="#Person"/>`
  </rdf:Description>

  - `<rdf:Description ID="hasHomepage">`
    - `<rdf:type resource="http://www.w3.org/...#Property"/>`
    - `<rdfs:subPropertyOf rdf:resource="#Artist"/>`
  </rdf:Description>
Model of the Semantic Web

- Web Ontology Language (OWL)

Problems with RDFS

- RDFS too weak to describe resources in sufficient detail
  - No localised range and domain constraints
  - No existence/cardinality constraints
  - No transitive, inverse or symmetrical properties

- Difficult to provide reasoning support
  - No “native” reasoners for non-standard semantics
  - May be possible to reason via FO axiomatisation
Model of the Semantic Web

- Web Ontology Language (OWL)

We are here
Model of the Semantic Web

- Web Ontology Language (OWL)

- An ontology is an engineering artifact:
  - It is constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary.

- Thus, an ontology describes a formal specification of a certain domain:
  - Shared understanding of a domain of interest
  - Formal and machine manipulable model of a domain of interest
Model of the Semantic Web

- Web Ontology Language (OWL)

Web Ontology Language Requirements

Desirable features identified for Web Ontology Language:

- Extends existing Web standards
  - Such as XML, RDF, RDFS
- Easy to understand and use
  - Should be based on familiar KR idioms
- Formally specified
- Of “adequate” expressive power
- Possible to provide automated reasoning support
Model of the Semantic Web

- Web Ontology Language (OWL)
  - OIL (Ontology Interface Layer)
    - Outcome from “On-To-Knowledge” project sponsored by European IST (Information Society Technologies) project
  - DAML (DARPA Agent Markup Language)
    - Began as a DARPA research program
    - DAML-ONT
  - DAML+OIL
    - DAML combines OIL components
    - DAML-S, DAML-L
  - OWL (Web Ontology Language)
    - W3C standard
Model of the Semantic Web

- **Web Ontology Language (OWL)**

**Evolution of Web Ontology Languages**

<table>
<thead>
<tr>
<th>Year</th>
<th>XML</th>
<th>RDF</th>
<th>RDFS</th>
<th>OIL</th>
<th>DAML-ONT</th>
<th>DAML (DAML+OIL)</th>
<th>OWL</th>
<th>OWL-S</th>
<th>DAML-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Define vocabularies
- Extend vocabularies
- Combine vocabularies
- Revision
- Extend HTML tags for semantic description
- For Web services
Model of the Semantic Web

- Logic

We are here

- Data
- Logic
- Ontology vocabulary
- RDF + rdfschema
- XML + NS + xmschema
- Unicode
- URI

Trust
- Proof
- Digital Signature
- Rules
- Data

Self-desc. doc.
Model of the Semantic Web

- Logic
  - Inference, Inference, Inference
  - OWL can be used for inference
    - MJ is a RockMusic artist, so he is also a PopMusic artist.
  - But how?
    - With a formal representation that equivalent to OWL, but easier to do inference.
  - It’s logic
Model of the Semantic Web

- Web Ontology Language (OWL)
Model of the Semantic Web

- Trust
  - If one person says that x is blue, and another says that x is not blue, doesn't the whole Semantic Web fall apart?
  - No!
    - Applications on the Semantic Web will depend on context,
    - Nonmonotonic reasoning can be applied.
    - We will have proof checking mechanisms, and digital signatures.
  - Proof Languages: a language that let's us prove whether or not a statement is true
    - Still missing at this moment
Current state of the Semantic Web and our jobs

The two fundamental technologies, which the Semantic Web bases on: XML and RDF are already developed

Logic remains developing and ontologies still need mapping

Proof and Trust are required to be established

There have been a number of small scale Semantic Web applications written up, but still a long way to go
What can we do?

- There are many ways in which one can contribute to creating the Semantic Web. Here's a few of them:-
  - Publish some globally useful data in RDF.
  - Write an inference engine in the language of your choice.
  - Spread the word: do some education and outreach.
  - Help in the development of RDF Schema.
  - Contribute in representing state in RDF, a rather neglected field of research.
  - Apply your own development backgrounds to the Semantic Web.
  - Instead of using some proprietary system for your next application, consider making it a Semantic Web project instead.
- When you don’t know what to do, just ask in the community for more details.
Resources

- Resources - Projects

- US
  - DAML: http://www.daml.org/
  - Mindswap : http://www.mindswap.org/
  - Protege: http://protege.stanford.edu/
  - HayStack: http://haystack.lcs.mit.edu/

- International
  - RACER http://www.sts.tu-harburg.de/~r.f.moeller/racer/

- Portal:
  - SemanticWeb: http://www.semanticweb.org
  - SemWebCentral: http://projects.semwebcentral.org/
  - SIGSEMIS: http://www.sigsemis.org/
Resources

- Resources – Projects

EU
- On-To-Knowledge http://www.ontoknowledge.org/ 99-02
- OntoWeb: http://www.ontoweb.org/ 01-04
- WonderWeb: http://wonderweb.semanticweb.org 02-04
- Knowledge Web http://knowledgeweb.semanticweb.org/ 04-08
- SWAD-Europe: http://www.w3.org/2001/sw/Europe/
- SEKT: http://www.sekt-project.com/ 04-07
- Sesame: http://www.openrdf.org/
- KAON: http://kaon.semanticweb.org/
- FaCT: http://www.cs.man.ac.uk/~horrocks/FaCT/
- SWAP: http://swap.semanticweb.org/
- SWWS: http://swws.semanticweb.org/
Resources

- Resources – Standards

  - XML: http://www.w3.org/XML/
  - RDF: http://www.w3.org/RDF/
  - OWL: http://www.w3.org/2001/sw/WebOnt/
  - DAML+OIL: http://www.daml.org/2001/03/daml+oil-index.html
  - RDQL: http://www.w3.org/Submission/RDQL/
  - Jena: http://jena.sourceforge.net/
Resources

- Resources – Conferences
  - WWW: World Wide Web
  - ESWC: European Semantic Web Conference
  - DL: Description Logics
  - ISWC: International Semantic Web Conference
  - SWDB: Semantic Web and Databases
  - KR: Knowledge Representation and Reasoning
  - SWEB: Semantic Web Technologies in Electronic Business
  - Protege: Protege conference
  - DC: Dublin Core
  - WI: IEEE/WIC/ACM Web Intelligence
  - Semantic Technology Conference
Question?

- Question?