



Research Studies on Semantic Web

Sabin Corneliu Buraga

busaco@infoiasi.ro

<http://www.infoiasi.ro/~busaco/>

Faculty of Computer Science

“Al.I.Cuza” University of Iasi, Romania

Topics

- Seeds of the Semantic Web
- Extensible Markup Language (XML)
- Resource Description Framework (RDF)
- Semantic Web
- Our Research

Seeds of the Semantic Web

- **Web 1.0 (1990)**
 - Tim Berners-Lee (CERN) develops a minimal hypertext system for the Internet using Vannevar Bush's idea of universal tool for human knowledge storage - MEMEX ('45)
 - **client/server architecture**
 - **hypertext = hypermedia**
(Ted Nelson, 1965)



Seeds of the Semantic Web

- Origin and Goals of the Web
 - Human communication thru shared knowledge
 - Exploitation of computing power in real life
- Web Architecture
 - Resources are identified by URIs (Uniform Resource Identifiers)
 - Internet protocols (e.g. HTTP) allow representation to be looked up
 - Documents parsed as markup
 - Markup contains URIs

Seeds of the Semantic Web

- **Revolutions**
 - **The HTML Revolution**
shared information space
 - **The XML Revolution**
culture of open exchange of information
 - **The RDF Revolution**
foundation of the Semantic Web

Seeds of the Semantic Web

- The XML Revolution
 - “XML is the ASCII of the future”
(Tim Bray, XML Files, oct.'98)
 - Meta-language (subset of SGML)
 - XML adds **type** and **structure** to **information**
 - Simple, flexible, and suitable for Web applications (platform independent)
 - Formal description: Document Type Definition, XML Schema, namespaces
 - Universal industry support

Seeds of the Semantic Web

- The XML Revolution
 - XML is a simple data format that balances the needs of people to read/write data with needs of machine to read/write data
- ```
<elementName attributeName="value">
 <empty />
</elementName>
```

# Seeds of the Semantic Web

- XML family
  - Syntax: XML (1998, 2000)
  - Validation: DTD, namespaces (1999), XML Schema (2000)
  - Hyper linking: XLink (2000), XPointer (2000)
  - Transformations and semantics: XSL - Extensible Stylesheet Language (2001)  
uses XPointer and XSLT languages

# Seeds of the Semantic Web

- Processing XML
  - **Document Object Model (DOM)**
    - XML document as an abstract tree of objects
    - Implementation independent
    - Very complex and feature-rich
    - DOM - level 1 (1998), DOM - level 2 (2000), ...
  - **Simple API for XML (SAX)**
    - XML document as a flow of element-events
    - More suitable for small applications

# Seeds of the Semantic Web

- **The RDF Revolution**
  - Standardized foundation for processing metadata  
**Metadata**=data about data
  - World-Wide Web Consortium's recommendation (Lassila and Swick, 1999)
  - Consists of a model for representing named properties and property values
  - Based on **classes**
  - A collection of classes = RDF **schema**

# Seeds of the Semantic Web

- Resource inspection
- Data investigation
- Intelligent agents
- Intellectual property rights
- Personal/general security (digital signatures)

Intelligent structures

Knowledge representation

Standardization

Semantic Web

# Seeds of the Semantic Web

- RDF Model: unordered set of triples
  - **Resources**  
named by URIs (Uniform Resource Identifiers)
  - **Properties**  
specific aspects/attributes/relations  
used to express resources
  - **Statements**  
{subject, predicate, object}  
subject=resource, predicate=property,  
object=value of property
- Semantics = assertions
- Vocabulary = resources, properties

# Seeds of the Semantic Web

- RDF model provides an abstract framework for defining and using metadata
- Concrete syntax is based on XML
- Example:

"Sabin is the owner of his Web page."

```
<rdf:RDF>
```

```
 <rdf:Description about="http://www.infoiasi.ro/~busaco/">
```

```
 <s:Owner>Sabin Corneliu Buraga</s:Owner>
```

```
 </rdf:Description>
```

```
</rdf:RDF>
```

# Seeds of the Semantic Web

- **The RDF Revolution**
  - Most important goal:  
to denote semantics for data based on **XML** in a homogeneous, platform-independent, and object-oriented way
  - To give support for Tim Berners-Lee's **Semantic Web** (2001)

# Semantic Web

- Ideas
  - The essential property of the WWW is its universality: “anything can link to anything”
  - For the Semantic Web to function, computers must have access to structured collections of information and sets of inference rules that they can use to conduct automated reasoning  
⇒ new models for **knowledge representation**
  - “The Semantic Web will enable machines to **comprehend** semantic documents and data, not human speech and writings.”  
(Tim Berners-Lee, 2001)

# Semantic Web

- Solutions
  - **Ontologies**

“An ontology is a specification of a conceptualization.” (Tom Gruber)

    - **Ontolingua** (1993)
    - **SHOE - Simple HTML Ontology Extension** (1997)
    - **OIL - Ontology Interference Layer** (2000)  
based on description logics

# Semantic Web

- Solutions
  - Agents
    - “Agent” as an Ascription
      - “The concept of agent embodied in humans helping humans is often one where expertise is indeed mixed with knowledge of you.”  
(Nicholas Negroponte, 1997)
    - “Agent” as a Description
      - A software entity which functions continuously and autonomously in a particular environment, often inhabited by other agents and processes.

# Semantic Web

- Solutions
  - Using properly designed Web agents, “the Semantic Web can assist the evolution of human knowledge as a whole.” (Tim Berners-Lee, 2001)
  - The consumer and producer agents can reach a shared understanding by exchanging ontologies, which provide the vocabulary needed for discussion.

## Our Research

- The essential property of the WWW is its universality:  
“anything can link to anything”
- Goal:
  - A formal approach to model spatial and **temporal** relations between (fragments) of Web sites
  - The model will be processed by machines in a platform-independent way

# Our Research

- Past and present
  - Modeling (spatial) relationships between different Web resources - i.e. files stored on a particular (distributed) file system:
    - RDF descriptions of a distributed file system
  - Searching for different structures of Web documents
    - Encoding structural information of a Web page by using XML markups

# Our Research

- Present and future
  - Modeling temporal relationships between (fragments of) Web sites to study the evolution of a distributed (multimedia) Web resources
    - Temporal Logics - Interval Temporal Logic (Allen, 1984)
    - Temporal Logic of Actions (Lamport, 1991)
    - Temporal Petri Nets?

## Our Research

- Future
  - To assign actions for each temporal change of the Web site's structure/state
  - The RDF model based on this approach can emphasize the dynamics of the (local) Web

# References

- T.Berners-Lee - "Weaving the Web", Cambridge Massachusetts, 1999
- S.C.Buraga - "Web Technologies" (in Romanian), Matrix Rom, Bucharest, 2001
- Semantic Web: <http://www.SemanticWeb.org>
- W3C Technical Reports, Boston, 2001: <http://www.w3.org/TR/>
- For other information, see <http://www.infoiasi.ro/~busaco/>



# Research Studies on Semantic Web

**Sabin Corneliu Buraga**

[busaco@infoiasi.ro](mailto:busaco@infoiasi.ro)

<http://www.infoiasi.ro/~busaco/>

Faculty of Computer Science

"Al.I.Cuza" University of Iasi, Romania

**Thank you!... Any questions?**