Semantic web: from XML to OWL

Pierre Genevès
CNRS
pierre.geneves@inria.fr

Jérôme Euzenat
INRIA
jerome.euzenat@inria.fr

Nabil Layaïda

Marie-Christine Rousset

University of Grenoble
2011–2012

Semantic web: from XML to OWL

Development of the future web

- Expressing information
- Manipulating it
- in the most correct, efficient

  and meaningful way

Semantic web: from XML to OWL

Development of the future web

- Expressing information → Languages
- Manipulating it → Algorithms
- in the most correct, efficient and meaningful way → Logic → Semantics

Foundations of XML

We will talk about languages, algorithms, and programming techniques for efficiently and safely manipulating XML data.

You will learn about:

- Tree structured data (XML)

→ Tree grammars & validation

- XML programming (XPath, XSLT...)

→ Queries & transformations

- Foundational theory & tools

→ *Regular expressions* can be used to *validate* XML data

→ *Finite tree automata* lie at the heart of XML validators

→ *Modal logic* allows implementing XPath query analyzers

Foundations of XML

We will talk about languages, algorithms, and programming techniques for efficiently and safely manipulating XML data.

You will learn about:

- Tree structured data (XML)

→ Tree grammars & validation

- XML programming (XPath, XSLT...)

→ Queries & transformations

- Foundational theory & tools

→ *Regular expressions* can be used to *validate* XML data

→ *Finite tree automata* lie at the heart of XML validators

→ *Modal logic* allows implementing XPath query analyzers

You will **not** learn about:

- Hacking CGI scripts
- HTML
- SQL...

Semantic web

- We will talk about languages, algorithms, and semantics for efficiently and meaningfully manipulating formalised knowledge.

You will learn about:

- Expressing formalised knowledge on the semantic web (RDF)

→ Syntax and semantics

- Expressing ontologies on the semantic web (RDFS, OWL, DL-Lite)

→ Syntax and semantics

→ Reasoning

- Querying formalised knowledge

→ Querying RDF graphs (SPARQL)

→ Querying modulo ontologies (RDFS, OWL)

- Network of ontologies

→ Semantic of alignments

→ Using several independent ontologies at once

Semantic web

- We will talk about languages, algorithms, and semantics for efficiently and meaningfully manipulating formalised knowledge.

You will learn about:

- Expressing formalised knowledge on the semantic web (RDF)

→ Syntax and semantics

- Expressing ontologies on the semantic web (RDFS, OWL, DL-Lite)

→ Syntax and semantics

→ Reasoning

- Querying formalised knowledge

→ Querying RDF graphs (SPARQL)

→ Querying modulo ontologies (RDFS, OWL)

- Network of ontologies

→ Semantic of alignments

→ Using several independent ontologies at once

You will **not** learn about:

- Tagging pictures
- Sharing MP3
- Creating facebook pages...

we assume you already know this ;-)

Organization

- Course website: http://exmo.inrialpes.fr/teaching/sw/
- Slides and papers online
- Specific information (room, changes. . .)
- Send email to Jerome.Euzenat@inria.fr, Pierre.Geneves@inria.fr with "Mosig" in subject
- Do not hesitate to send feedback

Evaluation

- Two 90mn written exams
- Each exam takes place at the end of each part (dates: 23/11, 25/01?)
- Each exam counts for one half of your final mark
- All documents allowed

→ NB: that is a **research** course...
Discover , **Understand** , **Think** , **Justify** ...

Project

- There will be a meeting on October 6th, at 16pm for presenting the teams of AIW!
- There are many project proposals available on the web site.
- It is time to hunt for a good one for you.
- Do not hesitate to contact and visit teams... and to look around on your own.

Foundations of XML

This part will introduce you to the world of XML and to the challenges of dealing with XML data.

Some of these challenges are:

- Existing PL technology offers severely limited (or no) support for XML
  - → how to manipulate XML data efficiently?
  - → how to implement search and extraction from XML data?
  - → how to ensure correct and safe XML processing?
- Existing DB technology can**not** be applied to XML data
  - → XML data are **tree-structured**
- Need for new algorithms and programming techniques
  - → XML motivates **theory** that is still being constructed

About XML

- XML is the World Wide Web Consortium's (W3C: www.w3c.org) **Ex**tensible **M**arkup **L**anguage
- We hope to convince you that XML is not yet another hyped acronym, but is useful technology.
- You will become best friends with one of the most important data structures in Computing Science: the **tree**.
- You will learn to apply a number of closely related XML standards:
  - Representing data: XML itself, DTD, XML Schema, XML dialects
  - Interfaces to connect generic PLs to XML: DOM, SAX
  - Languages to query/transform XML: XPath, XSLT, XQuery
- In the end, you should be able to digest the thick pile of related W3C X__ standards (XQuery, XPointer, XHTML, XInclude, XML Schema...)

References

- Excellent further readings:
  - W3C homepage (for current standards)
    http://www.w3.org
  - course slides of S. Maneth (DB oriented)
    http://www.cse.unsw.edu.au/∼cs4317/
  - book draft of H. Hosoya (PL theory oriented)
    http://arbre.is.s.u-tokyo.ac.jp/∼hahosoya/xmlbook/xmlbook.pdf