Ontologies and the Semantic Web

Universidad Politécnica de Madrid
A. Gómez-Pérez, O. Corcho, G. Aguado
{asun, ocorcho, lupe}@fi.upm.es

Masters where this course is taught

Ontologies and the Semantic Web is taught as a four-monthly subject in the following official master programs at the Universidad Politécnica de Madrid.

- Master in Information Technologies
- European Master in Logic Computation (ERASMUS Mundus Program)
- Master in Research in Artificial Intelligence

- It is credited 6 ECTS
- Presential classes: 2 hours /week
- English /Spanish

URL: http://delicias.dia.fi.upm.es/wiki/index.php/Master-Eng-07-08

Pre-requisites

Students that take this course have to hold a degree in Computer Science (5-year degree) or any 3-year Information Technology degree or related area

It is highly recommendable to have attained a certain level in the following subjects and technologies as they will not be explained in the classes.

- Knowledge representation systems: frames, semantic networks and description logics
- Web Technologies: HTML, XML, SOAP, WSDL, UDDI, etc.
- Java and JDBC

Objectives

The general objective is to provide students with a sound grounding of scientific, methodological and technological fundamentals in the Semantic Web domain that will be later used to build applications that can integrate, combine and infer heterogeneous and distributed information.

According to this general objective, students should be able:

1. To identify and resolve real world problems by applying these technologies successfully.
2. To build systems in different domains (for instance, knowledge management, biomedicine, e-commerce, e-learning, etc.) and applications for those areas.
3. To integrate applications developed with semantic web technologies with other *software and hardware* systems
4. To assimilate technological changes

In order to reach these objectives, motivate students and promote a favourable attitude towards the semantic web, emphasis will be put on the following aspects:

1. Students can understand the core of basic concepts and fundamental theories, complemented with a wide range of problems that illustrate the theoretical aspects.
2. Students can master the two most important pillars in the course: ontologies and annotation (systems, languages, etc)
3. Students can get used to handling real world problems extracted from national and European projects, though adapted for the classes, rather than handling "toy" problems.
4. Students can understand the relationship between the different components of a semantic-based system and how they can be integrated with other computer systems (*hardware* and *software*).
5. Students can establish relations between some fundamental aspects in this domain such as: environment interactions, high degree of autonomy and system adaptability.

Course contents

This course is organized in four sections. Each section is composed of several units.

Section 1: Introduction to the Semantic Web

Unit 1. General overview of the semantic web with special emphasis on ontologies and resource annotation (documents, texts, web pages, web services, DBs, etc). Description of the types of problems this technology can be applied to.

Section 2: Computational linguistics

Unit 2. Introduction to some computational linguistics concepts useful in building ontologies (terminological aspects: concepts, terms, relations between them, definitions, etc). Types of terminological resources (lexicons, thesauri, mono-, multilingual dictionaries, controlled-language vocabularies, terminological DBs) that can be used as a starting point in ontology building.

Unit 3. Multilingual representation in ontologies.

Section 3. Ontologies.

Unit 4. Theoretical aspects: definition, scope, types of ontologies, ontology repositories.

Unit 5. Languages used in ontology implementation: (RDF(S) and OWL) as well as query languages: SPARQL.

Unit 6. Tools used in building and storing ontologies (Sesame, Jena, Protégé, NeOn toolkit) as well as in ontology reasoning (RACER).

Unit 7. Life cycles and development methodologies used in building ontologies from scratch. Ontology networks used in building ontologies through collaborative work

Unit 8. Activities to be performed when preparing the ontology specification requirements

Unit 9. Ontology reusing and reengineering and resource learning (ontological and non-ontological resources) as a way of speeding up the process of building ontologies by using knowledge that has been already agreed upon.

Unit 10. Methods, techniques and tools used to carry out alignments and mappings between ontologies and between ontologies and other resources.

Section 4. The Semantic Web

Unit 11. Methods, techniques and tools used in (semi)-automatic annotation of texts and multimedia documentation.

Unit 12. Procedures and methods to turn content information from databases into semantic contents usable in the semantic web. Accessing data bases in terms of semantic contents.
Unit 13. Architectures and languages used in creating semantic web services (WSMO, OWL-S)
Unit 14. Applications using semantic web technologies that have been built in national and European projects in different domains (e-commerce, knowledge management, semantic portals, etc.)

To allow students to consolidate knowledge and skills acquired throughout the course some assignments related to each unit have been designed. Students will work in pairs and all the coursework to be done will be related to a specific domain. That domain will be agreed by the teacher and the students. The aim is to enable students to apply the knowledge acquired in the course in order to face new situations and solve real problems. Thus, students will be well prepared to adapt to the continuous technological evolution in this field.

The units mentioned above will be explained as follows:

- Unit 1. Introduction to the Semantic Web
- Unit 2. Terminological resources used for ontological purposes
- Unit 4. General overview of ontologies
- Unit 5. Languages used in the semantic web
- Unit 6. Editors and reasoners
- Unit 7. Methodologies used in building ontologies
- Unit 8. Ontology specification requirements
- Unit 9. Ontology reusing, reengineering and learning
- Unit 10. Ontology alignments and *mappings*
- Unit 3. Multilingualism in ontologies
- Unit 11. Textual and multimedia annotation
- Unit 12. Semantic access to distributed DBs
- Unit 13. Semantic Web Services
- Unit 14. Applications

Methodology

Table 1 describes the coursework that students will have to prepare for each unit and the technologies that will be used.

Units	Specific assignments
Unit 1. Introducing the Semantic Web	• Define a type of problem where the semantic web technologies can b and identify the application domain. • Reading the TBL paper. • Revision of the contents identified as fundamentals in this subject.
Unit 2. Terminological resources	• Identify a whole set of terminological resources as possible candidat domain application. • Compare them and select the most appropriate subset considering th to solve.
Unit 4. Fundamental of Ontologies	• Define use scenarios in a domain and find available ontologies in tha • Classify the selected ontologies according to the criteria explained. • Compare the overlapping degree of these ontologies with the vocab the selected terminological resources mentioned in Lesson 2.
Unit 5. Semantic Web Languages	• Implement a knowledge model of that domain in SESAME. • Identify and implement at least 10 queries in SPARQL. • Identify the differences between a query in SQL on a DB and the pre SPARQL
Unit 6. Editors and Reasoners	• Use an ontology editor based on description logics (NeOn toolkit, Pr build the knowledge model implemented in Lesson 5. • Perform at least 10 inferences using reasoners. • Comment on the similarities and differencies, advantages and disadv of using a query language against a reasoner. Design and perform so and interoperability tests importing the ontologies to the ontology ed mentioned in lesson 4. • Report on the usability of the employed tools.
Unit 7. Methodologies used to build ontologies	• Identify the life cycle model to be used to build the ontology conside resource reuse. • Design and plan the ontology development, the activities to be carrie resources to be used, etc.

Units	Specific assignments
Unit 8. Ontology Specification activities	• Specify the domain ontology to be built, with great detail. • Use the specifications to improve the ontology built with the editor mentioned in lesson 6.
Unit 9. Reusing, reingeneering and learning	• Reuse non-ontological esources (lexicons, thesauri, etc.) identified in lesson 2 and transform them in ontological resources by re-engineering processes. • Reuse the ontological resources identified in lesson 4 and carry out re-engineering processes, if necessary, in order to use them with the selected ontology editor. • Integrate these resources with the ontology in the ontology editor in order to provide a new version of the ontology.
Unit 10. Alignments and Mappings	• Do mappings between two of the ontologies of the same domain selected in lesson 4. • Do mappings between ontological and non-ontological resources.
Unit 3. Multilingalism in ontologies	• Using the NeOn Toolkit localization tools, translate the ontology into another language, at least.
Unit 11. Annnotation	• Annotate texts, web pages and multimedia documents using the ontology already built and some annotation tools (eg., GATE, MnM, M-OntoMat-Annotizer, Armadillo, AKTive media) • Report on the usability of the employed tools.
Unit 12. Semantic access to distributed DBs	• Use the R20 language to overlay two DBs with the ontology built. • Perform at least 10 queries in each DB. • Identify the advantages of R20 and ODEMapster • Report on the usability of the employed tools.
Unit 13. Semantic Web Services	• No specific assignment has been proposed. • Group work: Students will establish relations among the different ontologies built in the different groups. If there are several ontologies of the same domain, they will propose techniques to evaluate them.
Unit 14. Applications	• No assignments have been proposed.

Con formato: Inglés (Reino Unido)

Evaluation

Students have two ways of passing this subject. They can either take one or the other

1. Continuous evaluation by:
   a. doing the weekly coursework. 12 assignments (50% of the final mark)
   b. presenting a final work that compiles the coursework carried out and incorporates the corrections and suggestions proposed by the teachers during the tutorial sessions. This presentation will take place in an oral session in class, with all students. (50% of the final mark) Date to be fixed.

2. Final exam in February

- For those students that fail, they will have another chance in September. Requirements to pass: Final work compiling the different assignments (100%)

There is no exam in June

Recommended references

- Gómez-Pérez, A. Fernández-López, M. Corcho, O. **Ontological Engineering**. Springer Verlag 2003.

For each unit, a selection of readings will be provided.

Specifications for Semantic Web Languages: RDF(S) and OWL, proposed by W3C

- http://www.w3.org/RDF/
- http://www.w3.org/2004/OWL/

Most part of the teaching materials used in class has been prepared by the teachers of the OEG. Occasionally, some transparencies available at the REASE repository and created under the auspices of the KWeb excellence network (http://rease.semanticweb.org/) are also used.

In the following links, students can find some interesting documentation related to semantic web projects:

- DIP: http://dip.semanticweb.org/
- Knowledge Web: http://knowledgeweb.semanticweb.org/
- NeOn: http://www.neon-project.org/web-content/
- X-Media: http://www.x-media-project.org/