Enriching the Web with Ontology-lexica

Monday 21st May 2012

Presenters:

John McCrae, Brian Davis, Jorge Gracia

Tutorial Programme/Overview

This tutorial covers the usage of the *lemon* model, which has been developed in the Monnet project, to enable the representation of ontologies on the web using RDF and linking them to existing ontology formats, in particular the Web Ontology Language (OWL). In this tutorial we will first provide a background to the *lemon* model, followed by an introduction to core *lemon* model. We will then provide a hands-on session using our lexicon editing tool, *lemon source.* We will then provide a description of some of the modules used in *lemon*, followed by a practical session on using *lemon* GATE tools.

Schedule

Time	Session
09:00-09:45	Background
09:45-10:30	Lemon Core Model
10:30-11:00	Coffee Break
11:00-11:30	Lemon Source (hands-on)
11:30-12:15	Syntax, Structure, Semantics
12:15-13:00	Applications (hands-on)

Important Links

Tutorial web site: http://www.sc.cit-ec.uni-bielefeld.de/content/enriching-web-ontology-lexica
Lemon Cookbook: http://www.lexinfo.net/lemon-cookbook.pdf
Monnet Project home page: http://www.monnet-project.eu/
General Architecture for Text Engineering: http://gate.ac.uk/

Tutorial Description/Outline/Contents

Background (Brian Davis)

The session aims to give a background of the current standards for representing ontologies and lexica. We shall focus on the representation of ontologies and semantic networks on the web, in particular the Web Ontology Language (OWL) and the Resource Description Framework (RDF). We will also describe existing formats for the representation of machine-readable dictionaries and lexica and motivate the need for new formats for sharing such resources on the web as *lexical linked data*. Furthermore, we shall talk about the need for ontology lexicalisation and how the models described in this tutorial enable this vision.

Lemon Core (Jorge Gracia)

This session aims at giving an overview of the *lemon* model, which adapts the ISO Lexical Markup Framework (LMF) standard to Web ontologies, focusing on the *lemon core* components. We will discuss their design criteria and how they benefit NLP applications. We will illustrate, by means of some examples, the abilities lemon has for linking lexical information to online ontologies, as well as the potential use of lemon to establish cross-lingual links across ontologies.

Lemon Source hands-on (John McCrae)

For this course we will be using the Lemon Source editor. This is accessible at http://monnetproject.deri.ie/lemonsource. A local version of the server will be accessible from the Tutorial web site and on USB sticks from the organizers.
The participants should attempt to perform the following tasks

1. Create a new lexicon.
2. Add an entry to this lexicon.
3. Add a property (e.g., part of speech) to this entry.
4. Adding a sense that points to a DBPedia entity.
5. Add a variant form of the entry and properties to describe it.
6. Create a second entry and create a link to the previous entry.
7. Creating a subcategorization frame for an entry.

More detailed instructions will be available on the day

Syntax, structure, semantics (John McCrae)

This session will cover the more advanced aspects of the *lemon* model. First, we will handle the modelling of phrase structure trees within the model and how the results of a parser can be described using the *lemon* vocabulary. Secondly, we shall cover the description of the syntactic usage of terms by means of frames and we will describe how these frames can be linked to predicates in the ontology. Finally, we shall cover the description of relationships between different elements and the usage of data categories in the model.

Applications hands-on (Brian Davis)

This session will cover the exploitation of lemon resources in a widely used Natural Language Processing framework — GATE — General Architecture for Text Engineering. Participants should

have GATE version 6.X or later installed on there machines. Materials for this session will be made available on the tutorial website and on USB sticks from the organizers. A familiarity with GATE is not expected as this session will only introduce tools to work with *lemon* resources. This session will include the hands on tasks:

1. GATE in brief.
2. Ontology lexicalisation and ontology aware gazetteers in GATE.
3. *lemon* custom gazetteer plugin for GATE.
4. Generating a hierarchical gazetteer from *lemon* lexica.
5. Pattern matching over annotations in GATE using the JAPE (Java Annotation Patterns Engine) language.
6. Overview of semantic annotation and ontology population JAPE.
7. Introduction to *lemon*2JAPE compiler
8. Generating simple JAPE lexico-semantic patterns from *lemon* entries.

More detailed instructions will be available on the day.