Re-engineering OntoSem Ontology Towards OWL DL Compliance

Guntis Barzdins, Normunds Gruzitis and Renars Kudins
Institute of Mathematics and Computer Science
University of Latvia

JCKBSE, Tallinn, August 2006
SemTi-Kamols Project

• Integration of Latvian language and semantic web technologies
  – Part of Semantic Latvia initiative*

• Natural language is a challenge and a good measure for advanced semantic web development

• Ontology based natural language processing

• Inspired from the success story of OntoSem framework

• Modified towards latest semantic web approaches

OntoSem Framework

- Based on theory of ontological semantics*

- **Full-fledged** ontology
  - Descendant of Mikrokosmos
    - [http://crl.nmsu.edu/Research/Projects/mikro/index.html](http://crl.nmsu.edu/Research/Projects/mikro/index.html)
  - Disambiguate word meanings
  - Semantic parsing

- Lexical application
  - **Text meaning representation** (TMR)
    - [http://semnews.umbc.edu](http://semnews.umbc.edu)

OntoSem Framework

The Problem

- There are “a priori” defined senses of words, but a sense of a word can be defined by its use-case
Open and Closed Worlds

- **Closed world assumption:**
  - if statement cannot be proved it is assumed to be false

- **Open world assumption:**
  - if statement cannot be proved lack of knowledge is assumed

- Natural language is closer to the **OWA**

<table>
<thead>
<tr>
<th></th>
<th>OWA</th>
<th>CWA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monotonic</strong></td>
<td>Description Logic (OWL-DL)</td>
<td>Data bases</td>
</tr>
<tr>
<td><strong>Non-monotonic</strong></td>
<td>DBs and Frames with defaults (OntoSem)</td>
<td></td>
</tr>
</tbody>
</table>
OntoSem Ontology

- Written in LISP like syntax
- **Poorly documented** formal semantics
- Frame KR schema
  - **Non-monotonic** reasoning (frames with defaults)
  - **Closed world assumption**
OWL-DL Ontologies

- OWL-DL classes overlap

A red artificial flower

A library
Main Idea

• The usual metaphor of building a class with its attributes (UML) is not directly applicable in OWL DL

• Rather, we can use OWL DL to define classes by their logical characteristics and getting much more powerful reasoning support

• Determining types – word-senses using properties and use-cases
Ontosem to OWL-DL

- **Classes**
  - “all” is translated to “all”, instance of owl:Class
  - “objects” and “events” - instances of owl:Class

- **Properties**
  - Properties - instances of owl:ObjectProperty
  - “ontology-slot” is not translated
    - “is-a”, “domain”, etc., are already part of OWL and RDF(S)

- **Facets**
  - “defaults” – non-monotonic logic (CWA)
  - “inverse”, “sem” facets were translated
Soccer frame

(make-frame soccer
  (agent (inv (common striker)))
  (is-a (value (common sports-discipline)))
  (location (sem (common playing-field sports-arena))))

• Universal quantification should be used, otherway we get ontology which is equivalent to DB with mandatory fields which means non-monotonic reasoning (CWA)

• By means of OntoSem semantics, location of “soccer” cannot be both “playing-field” and “sports-arena”
Ontology debugging

- Large ontologies
  - Cyc
  - OntoSem
  - Wordnet, etc.

- Hard to keep consistent
  - Many developers
  - Changing knowledge

- Debug/test ontologies
Ontology debugging

- **Disambiguate** concepts
  - Add information on **disjoint classes - mandatory**
- **Run reasoner**
  - Pellet (open-source)
  - RacerPro (trial), etc.
- **Inconsistencies**
- **Redundancies**
Testing ontology

• Currently **txt2owl** is used for ontology testing
• **Create test-cases**
  – Explanatory dictionary
  – Hand made
• Check if created instances **belongs** to ontology
  – Reasoner
  – Specific application
• **Results**
  – Incomplete data
  – Inconsistencies with real world
Test “produce” event

- Honey: “a sweet sticky fluid made by bees”
Application of OntoSem OWL

• Adapted to
  – OWL-DL
  – Latvian language

• Application txt2owl
  – SWI-Prolog
  – Ontology driven
  – Text to OWL objects – TMR
Text analysis

- Verb - **event** is a main word in sentence
- **Thematic roles** are directly associated with verb*
- **Similarity to RDF triples**

Gudra māte māca meitu.

<Teach rdf:ID="teach_726">
  <GramTense>
    <Present rdf:ID="Present_967"/>
  </GramTense>
  <Agent>
    <Parent rdf:ID="parent_155">
      <GramGender>
        <Female rdf:ID="female_718"/>
      </GramGender>
      <GramCount>
        <Singular rdf:ID="Singular_263"/>
      </GramCount>
      <Literal-object-attribute>
        <Functional rdf:ID="Functional_432"/>
      </Literal-object-attribute>
    </Parent>
  </Agent>
  <Experiencer>
    <Offspring rdf:ID="offspring_918"/>
    ...
  </Experiencer>
</Teach>
Future work

- **Improve** lexical application
- **Understand metaphoric** relations between things, words and senses
  - Implement using SW technologies
- Develop methodology for **ontology testing** and debugging
Thank you!

Questions?

www.semti-kamols.lv