M2R Exam – Semantic web: from XML to OWL
Semantic web part

Duration : 1h30
Any document allowed – no communication device allowed

January 2011

Note: Please, carefully read all the questions before answering.

RDF and ontologies

Here are the 8 triples of an RDF graph G about writers and their works: (all identifiers correspond in fact to URIs, _:b is a blank node):

$$\langle \texttt{d:Poe}, \texttt{o:wrote}, \texttt{d:TheGoldBug}\rangle \quad \langle \texttt{d:Baudelaire}, \texttt{o:translated}, \texttt{d:TheGoldBug}\rangle$$
$$\langle \texttt{d:Poe}, \texttt{o:wrote}, \texttt{d:TheRaven}\rangle \quad \langle \texttt{d:Mallarmé}, \texttt{o:translated}, \texttt{d:TheRaven}\rangle$$
$$\langle \texttt{d:TheRaven}, \texttt{rdf:type}, \texttt{o:Poem}\rangle \quad \langle \texttt{d:Mallarmé}, \texttt{o:wrote}, \texttt{\_:b}\rangle$$
$$\langle \texttt{\_:b}, \texttt{rdf:type}, \texttt{o:Poem}\rangle \quad \langle \texttt{d:TheGoldBug}, \texttt{rdf:type}, \texttt{o:Novel}\rangle$$

1. Draw an RDF graph corresponding to these statements
2. Express in English the meaning of these statements.

Consider the RDFS ontology o containing, in addition to those of G, the following statements:

$$\langle \texttt{o:Novel}, \texttt{rdfs:subClassOf}, \texttt{o:Literature}\rangle$$
$$\langle \texttt{o:Poem}, \texttt{rdfs:subClassOf}, \texttt{o:Literature}\rangle$$
$$\langle \texttt{o:translated}, \texttt{rdfs:range}, \texttt{o:Literature}\rangle$$
$$\langle \texttt{o:wrote}, \texttt{rdfs:domain}, \texttt{o:Writer}\rangle$$

3. Does this allow to conclude that `d:Poe`, `d:Baudelaire` or `d:Mallarmé` is a `o:Writer`? Explain why.
4. Can you express in OWL the statement that "anyone who write Literature is a Writer"?

SPARQL query containment

Consider the following queries q_1 and q_2 on the RDF graph of the previous exercise:

- $q_1 =$ `SELECT` $?w$ `FROM` G `WHERE` $(\langle ?w\ \texttt{o:wrote}\ ?x\rangle$ `AND` $\langle ?x\ \texttt{rdf:type}\ \texttt{o:Poem}\rangle)$ `UNION` $\langle ?w\ \texttt{o:translated}\ ?x\rangle$;
- $q_2 =$ `SELECT` $?w$ `FROM` G `WHERE` $(\langle ?w\ \texttt{o:wrote}\ ?l\rangle$ `UNION` $\langle ?w\ \texttt{o:translated}\ ?l\rangle)$ `AND` $\langle ?l\ \texttt{rdf:type}\ \texttt{o:Poem}\rangle$.

5. In the course, we defined the distinguished variables $\vec{B}$, the queried graph G and the query pattern P. Identify them in q_1 and q_2.
6. Provide the answers of q_1 and q_2 with respect to the graph G.

Query containment $q \sqsubseteq q'$ between two queries $q = \texttt{SELECT}\ \vec{B}\ \texttt{FROM}\ G\ \texttt{WHERE}\ P$ and $q' = \texttt{SELECT}\ \vec{B}\ \texttt{FROM}\ G\ \texttt{WHERE}\ P'$ is defined by the fact that for any RDF graph, the answers to q are included in those to q' ($\forall G, \mathcal{A}(\vec{B}, G, P) \subseteq \mathcal{A}(\vec{B}, G, P')$).

7. What does the answer to the previous questions tell you about query containment between q_1 and q_2?
8. Do you think that query containment holds in some direction between q_1 and q_2 (either $q_1 \sqsubseteq q_2$ or $q_2 \sqsubseteq q_1$)?
9. Provide a proof for this. This may be done semantically by using the interpretation of query patterns or syntactically by translating queries into logic and showing that the query containment statement is a theorem.

Query modulo ontology

We now consider the ontology o and the following queries:

- $q_3 = \texttt{SELECT}\ ?y\ \texttt{FROM}\ o\ \texttt{WHERE}\ \langle ?x, \texttt{o:translated}, ?y\rangle$;
- $q_4 = \texttt{SELECT}\ ?y\ \texttt{FROM}\ o\ \texttt{WHERE}\ \langle ?y, \texttt{rdf:type}, \texttt{o:Literature}\rangle$.

10. Do you think that query containment holds in some direction between q_3 and q_4 (either $q_3 \sqsubseteq q_4$ or $q_4 \sqsubseteq q_3$)? Tell why.
11. Can you provide a definition for query containment modulo an ontology o ($q \sqsubseteq_o q'$)?
12. Does it return different answers for q_3 and q_4?(either $q_3 \sqsubseteq_o q_4$ or $q_4 \sqsubseteq_o q_3$)? Tell why.

Network of ontologies

We now consider an ontology o' which defines the class `op:Buch` and contains the following statements:

$$\langle \texttt{d:Baudelaire}, \texttt{o:translated}, \texttt{d:Confessions}\rangle\ \langle \texttt{d:DeQuincey}, \texttt{o:wrote}, \texttt{d:Confessions}\rangle$$

and o'' which defines the class `opp:Roman` and contain the following statements:

$$\langle \texttt{d:Confessions}, \texttt{rdf:type}, \texttt{opp:Roman}\rangle\ \langle \texttt{d:Musset}, \texttt{o:translated}, \texttt{d:Confessions}\rangle$$

They are related together by the following three alignments:

- $A_{o,o'} = \{\langle \texttt{o:Literature}, \equiv, \texttt{op:Buch}\rangle\}$
- $A_{o',o''} = \{\langle \texttt{op:Buch}, \sqsubseteq, \texttt{opp:Roman}\rangle\}$
- $A_{o'',o} = \{\langle \texttt{opp:Roman}, \equiv, \texttt{o:Novel}\rangle\}$

So that we have a network of ontology $\langle \{o, o', o''\}, \{A_{o,o'}, A_{o',o''}, A_{o'',o}\}\rangle$.

13. Do you think that this network of ontologies is well designed? Why?
14. Is this network consistent? Provide a model for this network of ontologies.
15. Provide the constraints that the alignments impose on models.
16. What does this entail for the class (`rdf:type`) of `d:Confessions` and `d:TheRaven` at o in this network?