OPEN DATA EXAM

20th of June 2016. *The exam will take 2 hours. Answer each question in the provided space.* Answers out of such space will not be considered.

Name: …….

Question 1. [7p]

a) Model in RDF(S) the following statements:

"For each conference we want to know the related areas (i.e., research areas from where research papers are welcome) and its name. Each edition (i.e., the yearly event) publishes proceedings which contains the accepted papers in the conference. For each edition we want to keep track of the year and venue (i.e., the city where it was held). For each paper we want to track the title, the primary research area it belongs to (it can be only one), the secondary research areas (they can be many) and list of authors."

Draw a **TBOX RDF graph modelling as many statements as you can**. Identify the statements modelled in your graph by underlining them in the text above. If any, clearly identify in the graph the RDFS constructs. Also clearly distinguish concepts (in rectangles) from literals (in circles). Define your own namespace prefix for the URIs you need to create **[2p]**

b) Now, **assert the following instances** to your RDF graph (draw them below the RDF graph sketched above and clearly **separate them from the TBOX by a dashed line**): *The European Database Technology (EDBT) conference was held in Bordeaux in 2016. EDBT welcomes papers in databases, semantic web and information retrieval. In its proceedings, we can find the paper entitled "DBExplorer: Exploratory Search in Databases" by Manish Singh, Michael Cafarella and Hosagrahar Jagadish* **[1p]**

c) Name at least **one reason for representing and one reason for not representing** a triple object with a literal instead of a URI **[1p]**

Literals are simple values, making them more readable and easier to use for straightforward data representation. For example, representing a person's age as a literal integer (e.g., "25") is straightforward and avoids the complexity of linking to another resource

Using URIs instead of literals enables interlinking between different datasets and enhances reusability. URIs can represent resources that might have additional properties or be referenced elsewhere, facilitating richer and more connected data structures.

d) Write the **triples** corresponding to your **TBOX RDF graph in Turtle format**. Any prefix you created must be properly defined (note that you do not need to define the rdf and rdfs prefixes) **[1p]**

e) What is the meaning of a blank node in SPARQL? **[0,5p]**

f) According to the blank nodes semantics in SPARQL, what are the elements from your RDF graph that can be represented as blank nodes. **Briefly justify your answer [0,5p]**

g) Write a SPARQL query to retrieve *"the EDBT edition that published the most papers"* **[1p]**

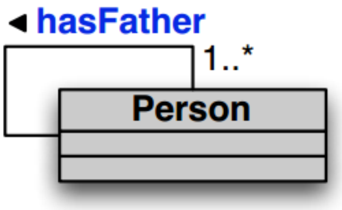
```
SELECT ?edition (COUNT(?paper) AS ?paperCount)
WHERE {
  ?paper exam:isAcceptedIn ?conf .
  ?conf exam:isPublishedIn ?edition ;
            exam:confName "EDBT" .
}
GROUP BY ?edition
ORDER BY DESC(?paperCount)
LIMIT 1
```

Question 2. [3p]

a) For each basic $\mathcal{AL}$ construct below explain its semantics. Provide **both** an intuition of its meaning and its formal semantics in the form of interpretations **[1p]**

Syntax
A
P
$\neg A$
$C \sqcap D$
$\exists R$
$\forall R.C$
$\bot$

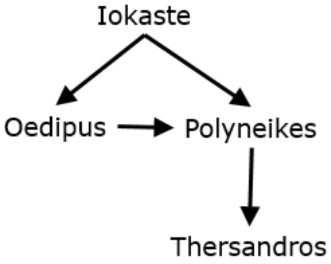
b) Represent the following UML schema in Description Logics (DL). You can use any construct from the $\mathcal{AL}\cdot$ family. **Briefly explain the meaning of your DL assertions [1p]**



c) Considering the *Open-World Assumption*, what would be the result of `Query` given the assertions below? **Justify your answer [1p]**

Note: The figure on the right sketches the hasSon assertions.

```
hasSon(Iokaste,Oedipus)
hasSon(Iokaste,Polyneikes)
hasSon(Oedipus,Polyneikes)
hasSon(Polyneikes,Thersandros)
patricide(Oedipus)
¬patricide(Thersandros)
```



```
Query ≡ ∃hasSon.(patricide ⊓ ∃hasSon.¬patricide)
ABox ⊨ Query(Iokaste)?
```