Knowledge Engineering and Semantic Web

Exercise Sheet: 1
Will be discussed on: April 23,2024



TUTORS:

Yaser Jaradeh, and some other ORKG team members

QUESTIONS: Please don't hesitate to ask any questions. Questions help you and your peers.

PRINT: Please consider the environment before printing the exercise.

1 Review questions

- 1. Which ones could be considered as main reasons for development of semantic data web?
 - (a) Creation of huge amounts of information and data
 - X Huge amount of information is already existing on the web.
 - (b) Standardization of data for transferring
 - X Transferring of data is already standardized using protocols.
 - (c) Machine readability of data
 - ✓ Data in the web is always represented in documents which are not machine readable.
 - (d) Meaningfulness of data
 - ✓ Semantic Web makes the meaning for the data by indicating relation between resources and defining schemas and ontologies to make data as much meaningful as needed in the domain.
- 2. Which ones could be considered as motivations for development of semantic data web in future?
 - (a) Human readability of data
 - X Documents in the Web are (typically) already human readable.
 - (b) Integration of data from heterogeneous resources
 - ✓ One important issue for web of data and semantic web is collecting and so integrating data which is gathered from different heterogeneous sources.
 - (c) Intelligent information retrieval
 - ✓ When you make your data set with RDF model, it is always interesting to retrieve new facts (information) from it.
 - (d) Syntactical standardization of information
 - × Syntactical standardization is necessary in the progress of semantic web, but it's not a motivation for development of semantic web.

- 3. Choose the correct statements with respect to web standardization:
 - (a) Acronym URI, identifies the unique location of each resource.
 - × What is a location of a resource?
 - (b) The identity of resource could be defined uniquely by URL.

 ✓ The identity of resources are defined by URI not URL.

However, URLs are a subset of URIs.

- (c) Acronym RDF stands for Resource Description Framework, which is standardized by W3C × RDF (Resource Description Framework) is standardized by W3C.
- (d) Data is accessible only by protocol HTTPS in web of data.
 - × What about FTP?
- (e) URI is more general than IRI.
 - × IRI=Internationalized Resource Identifier is more general than URI.
- 4. What is correct about RDF representation of information?
 - (a) Facts are represented by triples
 Value Subject-Predicate-Object
 - (b) Blank nodes can be used just as objects. X No, also as Subject
 - (c) Predicates could be literal or URI X Literals can only be Object
 - (d) Literals can stands in the position of subject. X Literals can only be Object
 - (e) The RDFa makes it possible have meaningful RDF triples.
 - × There is no meaningful or not meaningful RDF triplet, RDFa is a serialization. For meaningfulness of triples we need schema and ontology.

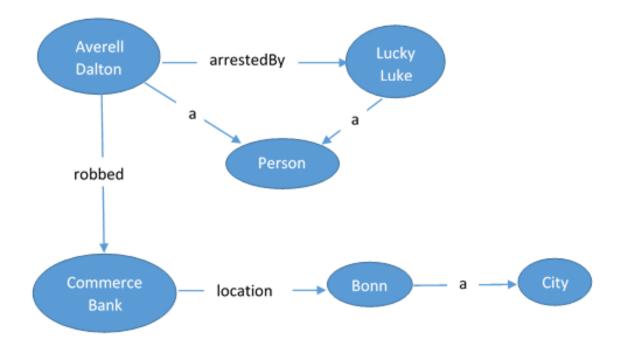
2 Consider the following XML snippet

```
<person>
   <name>Averell Dalton</name>
   <robbed> Commerce Bank </robbed>
   <arrestedBy>
      <person>
         <name>Lucky Luke </name>
      </person>
    </arrestedBy>
</person>
<bank>
   <name > Commerce Bank </name >
   <location>
      <city >
         <name>Bonn</name>
      </city>
   </location>
</bank>
```

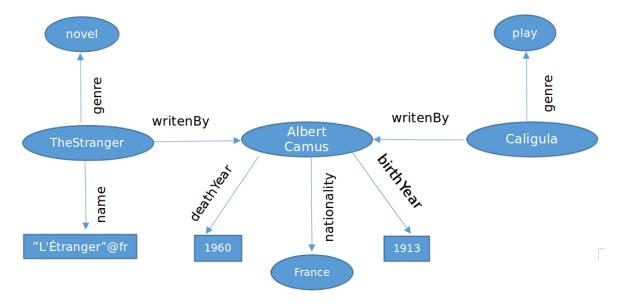
- 1. Try to explain it in your own words.
 - <u>Solution:</u> A Person with the name Averell Dalton has robbed the Commerce Bank. He was arrested by a person, who's name is Lucky Luke. The Bank with the name "Commerce Bank" has a location, and this location is a City with the name "Bonn".
- 2. Transform the XML description into a graph (nodes and links). Use circles or ovals for resource nodes. Use rectangle for literals or datatype values.

Solution:

This XML snippet is amended a little to make it in an uniform definition.



3 Consider the following knowledge graph



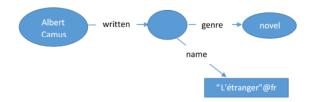
- 1. Count the number of triples and indicate the URIs and literals.
- 2. Write each triple as a simple fact in one sentence of natural language.
- 3. (optional) Rewrite the extracted facts in the last section in XML representation.
- 4. (optional) Consider that we don't have the URI of book "The Stranger", how can we show the respective information?

Solution:

- 1. 8 triples
- 2. (a) The Stranger is written by Albert Camus.
 - (b) The genre of The Stranger is novel.
 - (c) The name of The Stranger in French is "L'étranger".

- (d) Albert Camus died in 1960.
- (e) Albert Camus born in 1913.
- (f) Albert Camus in a French.
- (g) Caligula is written by Albert Camus.
- (h) The genre of Caligula is play.

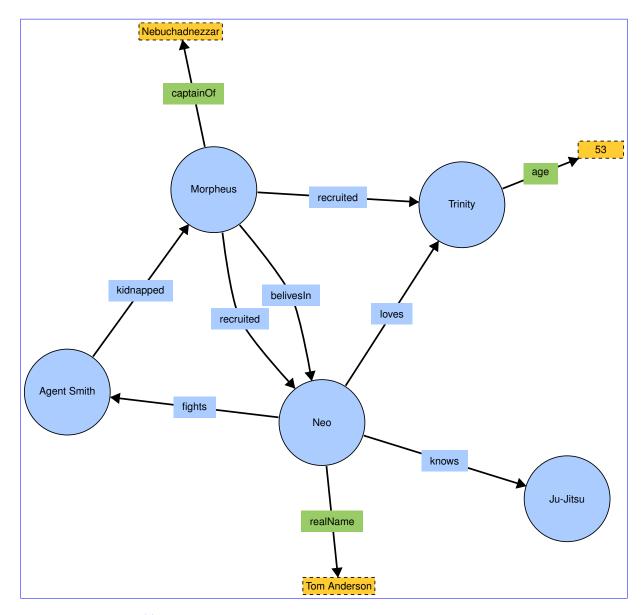
4. With blank node:



4 Assume a familiar subject, such as your family or LUH

- 1. Create a knowledge graph with at least 10 triples about it.
- 2. (optional) Write the triples of your knowledge graph in turtle representation.

Solution:



Oprefix ex:<http://example.org#>.

```
ex:Neo
                ex:realName
                                 "Tom Anderson";
                ex:knows
                                 ex:Ju-Jitsu;
                ex:loves
                                 ex:Trinity;
                ex:fights
                                 ex:AgentSmith.
                ex:kidnapped
                                 ex:Morpheus.
ex:AgentSmith
ex:Morpheus
                ex:captainOf
                                 "Nebuchadnezzar";
                 ex:recruited
                                 ex:Trinity;
                 ex:believesIn
                                 ex:Neo;
                ex:recruited
                                 ex:Neo, ex:Trinity.
                                 "53".
ex:Trinity
                ex:Age
```