Semantic Web (CSE632)
Winter 2020
Assignment 1

Instructions

1. There are no optional questions.
2. Assignment is not a group activity. Each student has to work on it by himself/herself and submit the assignment.
3. Plagiarism check and policies will be strictly enforced. Students can be selected at random for a “viva” on the assignment and if the responses are not satisfactory, they will get a 0 in the assignment.
4. Submit the answers in the form of a pdf. Include the code in the pdf as well as separate java files. Provide comments in the code and document every step of your code (except variable assignment). Provide the self-contained jar files to run the java files. Provide the JSON-LD file for Question 3.
5. Deadline will not be extended.

Questions **Max marks: 50**

1. Convert the following statements into RDF triples **10 marks**
    a. juiceMadeOfFruit is a property having domain as FruitJuice and range as Fruit
    b. Apple is an instance of Fruit
    c. juiceMadeOfFruit is a sub role of juiceMadeOf
    d. MixedFruitJuice is made up of Banana, Orange, Pineapple, and Watermelon
    e. MixedFruitJuice is made of 2 Oranges, 1 Pomegranate and 1 Pineapple
    f. OrangeJuice is made of 3 Oranges, and 1 tablespoon of salt
    g. MixedFruitJuice is a subclass of FruitJuice
    h. Fruit, FruitJuice, and MixedFruitJuice are classes
    i. juiceMadeOfFruit, and juiceMadeOf are properties
    j. Cost of a glass of juice is INR 25

2. RDF containers and collections are used to describe group of things. It can be used, for example, to list the authors of a book or to list the members in a band. Refer to the textbook (Foundations of Semantic Web Technologies), the W3C page on RDF, and W3Schools for their usage.

    Given the following paragraph, write it in Turtle Syntax using appropriate containers. Also explain the reason why you chose that particular container. **15 marks**

    ```
    Mary likes 3 types of fruit juices, (1) orange juice (O), (2)
    apple juice (A) and (3) mixed fruit juice (M) in the following
    ```

preference order: M, O, A but during her fruit meal she prefers to eat only one among the following fruits, (1) orange, (2) apple or (3) pineapple. Mixed fruit juice is made up of only oranges, apple, papaya, and banana.

3. ConceptNet (http://conceptnet.io/) is a freely-available semantic network, designed to help computers understand the meanings of words that people use. As a modern Linked Open Data resource, the data in ConceptNet is available as a JSON-LD API (https://github.com/commonsense/conceptnet5/wiki/API), a format that aims to make linked data easy to understand and easy to work with. **10 marks**

    a. For a given word, use the API and get the JSON-LD output
    b. Write a Java program to convert the JSON-LD output to TTL (Turtle) format using a library such as Apache Jena (https://jena.apache.org/). Provide a self-contained jar file that can be run. We will be running this jar file using your JSON-LD file as input and we expect to see the TTL file as output.
    c. Visualize the triples from the output as a graph using a tool/library such as http://librdf.org/raptor/libraptor2.html + GraphViz. Make sure that all the node and property/edge labels are clearly visible. Feel free to use any library, but clearly specify the name and the URL of the library/tool used.

4. Write a Java program using Apache Jena or RDF4J (https://rdf4j.org/) to convert the entries in the csv file (NetflixList.csv) to triples. Note the following
    a. If a cell has multiple values (for example, separated by commas), they have to be represented separately. This enables you to answer queries such as “which two films have the same director and have at least one common actor?”.
    b. Handle empty cell values and special characters

    The expected output is the following
    a. A TTL (Turtle) file with the triples
    b. A Java program along with a self-contained jar file that can be run. We will be running this jar file with the csv file as input.
    c. List of classes and properties that you used in the conversion. Specify the domain and range of all the properties.

    15 marks