Semantic Web

Assignment 7

Dr. Jandson S Ribeiro
jandson@uni-koblenz.de

Isabelle Kuhlmann
iskuhlmann@uni-koblenz.de

Institute of Web Science and Technologies
Department of Computer Science
University of Koblenz-Landau

Submission by: June 20, 2021
Tutorial on: June 24, 2021

Please submit your solutions to your group's OLAT folder.
Always list all group members contributing to the solution!
Do not plagiarize from others!

For all the assignment questions that require you to code, make sure to include the code in the answer sheet, along with a separate Python file.

Team Name: gamma

1. Aman Bisht (amanvista@uni-koblenz.de)
2. Muralikrishna Naripeddi (mnaripeddi@uni-koblenz.de)
3. Ndang Hesley Fonane (fonaneh@uni-koblenz.de)

1 Ontology Matching 10 points

Consider two ontologies O1 and O2 as depicted below:

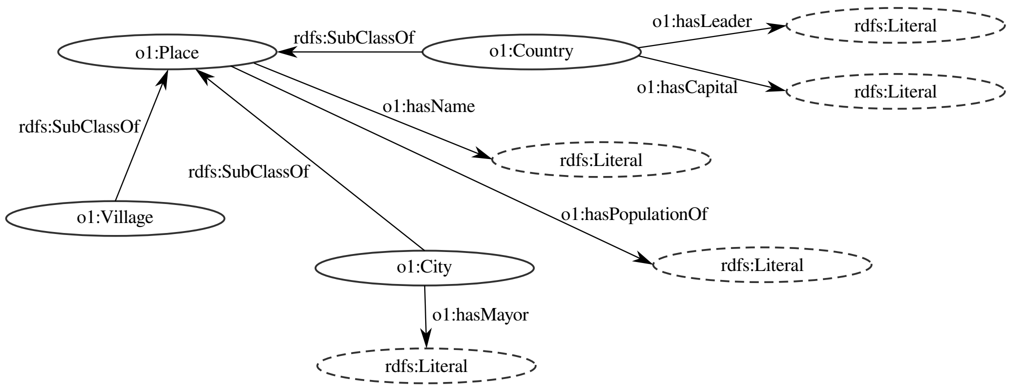

Figure 1: Ontology O1

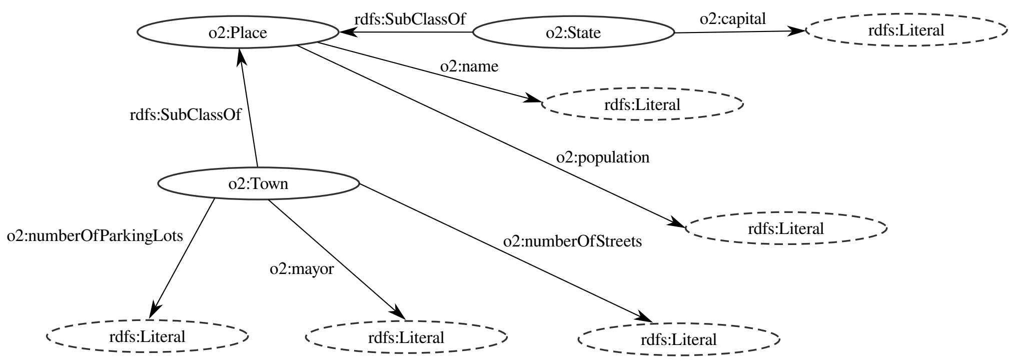

Figure 2: Ontology O2

In this task you shall identify similarities between ontologies O1 and O2. In your analysis, disregard the namespace prefixes. **Give detailed explanations** when answering the following questions:.

1.1 3 points

What are the Levenshtein distances between all possible pairs of `o1:hasLeader`, `o1:hasMayor`, and `o1:hasCapital`?

lev = Levenshtein distance

lev(o1:hasLeader, o1:hasMayor) = lev(o1:hasMayor, o1:hasLeader)

lev(o1:hasLeader, o1:hasMayor)

1. removing L from o1:hasLeader would take 1 steps (o1:haseader)
2. replacing e with M takes 1 step (o1:hasMader)
3. replacing de with yo takes 2 steps (o1:hasMayor)

lev(o1:hasLeader, o1:hasMayor) = 4

lev(o1:hasMayor, o1:hasCapital) = lev(o1:hasCapital, o1:hasMayor)

lev(o1:hasMayor, o1:hasCapital)

1. replacing M with C takes 1 step (o1:hasCayor)
2. replacing yor with pit takes 3 steps (o1:hasCapit)
3. adding al takes 2 steps (o1:hasCapital)

lev(o1:hasMayor, o1:hasCapital) = 6

lev(o1:hasCapital, o1:hasLeader) = lev(o1:hasLeader, o1:hasCapital)

lev(o1:hasCapital, o1:hasLeader)

1. replacing C with L takes 1 step (o1:hasLapital)
2. adding e takes 1 step (o1:hasLeapital)
3. replacing pit with der takes 3 steps (o1:hasLeaderal)
4. removing al takes 2 steps (o1:hasLeader)

lev(o1:hasCapital, o1:hasLeader) = 7

1.2 **4 points**

Calculate the suffix similarity for both `o2:mayor` and `o2:population` to elements in O1. Indicate the pair(s) with the confidence value higher than zero.

Considering suffix similarity as case sensitive

o2:mayor

- o2:mayor with o1:hasMayor has a suffix similarity of $\frac{4}{8} = 0.5$
- o2:mayor with o1:hasLeader has a suffix similarity of $\frac{1}{9} = 0.11$
- o2:mayor with all the other elements in o1 has a suffix similarity of zero

o2:population

- o2:population has zero suffix similarity to all the elements in o1

1.3 3 points

Using graph-based techniques, identify two pairs of entities between non-leaf elements of O1 and O2 that are similar. For each pair $(e1, e2)$ that you provide, explain why they are similar.

1. o1:Place is similar to o2:Place because
   a) o1:Village is equivalent to o2:Town.
   b) o1:City is equivalent to o2:Town.
   c) o1:hasPopulationOf relation is equivalent to o2:population relation.
   d) o1:hasName is equivalent to o2:name.
   e) o1:Country is similar to o2:State.
2. o1:Country is similar to o2:State because
   a) o1:Place is equivalent to o2:Place.
   b) o1:hasCapital is equivalent to o2:capital.
3. o1:City is similar to o2:Town because
   a) o1:Place is equivalent to o2:Place.
   b) o1:hasMayor is equivalent to o2:mayor.

2 Alignment 5 points

Specify an alignment of the following instances to ones found in the Linked Open Data cloud (e.g. dbpedia.org is a good starting point for this kind of search). Provide 2 equality-correspondences per instance (a total of 10). You must provide the level of confidence for each correspondence. For this, you can use the following normalized version of the Levenshtein distance. Given two strings $s1$ and $s2$, let N be the size of the longest string, and L be the Levenshtein distance between $s1$ and $s2$. The normalized Levenshtein distance is $(N - L)/N$. You can use the Levenshtein calculator at `http://www.unit-conversion.info/texttools/levenshtein-distance/` to assist with the task.

- http://example.org/places/Berlin
  1. Instance 1920s_Berlin has normalized Levenshtein distance of $\frac{12-6}{12} = 0.5$
  2. Instance 2016_Berlin_state_election has normalized Levenshtein distance of $\frac{26-20}{26} = 0.23$
- http://example.org/institutions/Reichstag
  1. Instance Reichsrat has normalized Levenshtein distance of $\frac{9-2}{9} = 0.78$
  2. Instance Reichstag_dome has normalized Levenshtein distance of $\frac{14-5}{14} = 0.64$
- http://example.org/people/JimRakete
  1. Instance gunterrakete has normalized Levenshtein distance of $\frac{12-7}{12} = 0.42$
  2. Instance Jim_Carrey has normalized Levenshtein distance of $\frac{10-6}{10} = 0.4$
- http://example.org/artwork/LaTrahisonDesImages
  1. Instance The_Betrayal_of_Images has normalized Levenshtein distance of $\frac{22-15}{22} = 0.32$
  2. Instance Treachery_of_Images has normalized Levenshtein distance of $\frac{22-11}{22} = 0.5$
- http://example.org/places/Myanmar
  1. Instance Religion_in_Myanmar has normalized Levenshtein distance of $\frac{19-12}{19} = 0.37$
  2. Instance Buddhism_in_Myanmar has normalized Levenshtein distance of $\frac{19-12}{19} = 0.37$

Important Notes

Submission

- Solutions have to be submitted to your group's OLAT folder.
- The name of the group and the names of all participating students must be listed on each submission.
- Solution format: all solutions as *one* PDF document. Programming code has to be submitted as Python code to the OLAT folder. Upload *all* `.py` files of your program! Use `UTF-8` as the file encoding. *Other encodings will not be taken into account!*
- Check that your code compiles without errors.
- Make sure your code is formatted to be easy to read.
  - Make sure you code has consistent indentation.
  - Make sure you comment and document your code adequately in English.
  - Choose consistent and intuitive names for your identifiers.
- Do *not* use any accents, spaces or special characters in your filenames.

Acknowledgment

This pdfLaTeX template was adapted by Isabelle Kuhlmann based on the LuaLaTeX version by Lukas Schmelzeisen.

LaTeX

Use `pdflatex assignment_X.tex` to build your PDF.