CS 586 Assignment-1
REPORT

Submitted by:
Sabita Acharya
676636765

Methods used:

1.

a. For each node, we check if it is a "class" or a "property". If it is a "property", we check if the domain and range of both source and target nodes are same or not. If they are same, we assign a value of 1 to similarity, else the similarity is zero.

b. We also check if the local names of the source and target are similar or not in a number of steps. We compare the actual name, compare the normalized forms of the names, compare the stemmed forms of the names, etc. to determine how similar the two nodes are.

Finally, we combine the results from a and b in different proportions and get a final value.

2. We get the ancestors and children of both the source and target nodes and try to see if there is a match between their local names. We also give certain weight to the local name, the label and comment and check if those values for particular nodes are relevant or irrelevant. If relevant, we find a weighted average and add a certain proportion of it to the result obtained by comparing ancestors.

3 .We stem the local name and count the number of tokens from the source and target that match each other and assign a certain value to them.

4. We take the values obtained from 1, 2 and 3 in certain proportions and determine a final value. This value is then passed to the Mapping function.

Result obtained with the initial baseline code:

Precision = Correct/Discovered: 82.3%
Recall = Correct/Reference: 46.9%
Fmeasure = 2(precision*recall)/(precision+recall): 58.1%

Result obtained after making modifications:

Precision = Correct/Discovered: 79.4%
Recall = Correct/Reference: 56.7%

Fmeasure = 2(precision*recall)/(precision+recall): 65.1%

Some examples of new mappings discovered:

(from cmt-conference):

```
<map>
<Cell>
      <entity1 rdf:resource="http://cmt#Chairman"/>
      <entity2 rdf:resource="http://conference#Chair"/>
      <measure rdf:datatype="http://www.w3.org/2001/XMLSchema#float">0.9</measure>
      <relation>=</relation>
</Cell>
</map>

<map>
    <Cell>
      <entity1 rdf:resource="http://cmt#ProgramCommittee"/>
      <entity2 rdf:resource="http://conference#Program_committee"/>
      <measure rdf:datatype="http://www.w3.org/2001/XMLSchema#float">1.0</measure>
      <relation>=</relation>
    </Cell>
  </map>

<map>
    <Cell>
      <entity1 rdf:resource="http://cmt#email"/>
      <entity2 rdf:resource="http://conference#has_an_email"/>
      <measure rdf:datatype="http://www.w3.org/2001/XMLSchema#float">1.0</measure>
      <relation>=</relation>
    </Cell>
  </map>
```

(from edas-sigkdd):

```
<map>
    <Cell>
      <entity1 rdf:resource="http://edas#hasName"/>
      <entity2 rdf:resource="http://sigkdd#Name_of_conference"/>
      <measure rdf:datatype="http://www.w3.org/2001/XMLSchema#float">0.9</measure>
```

```
    <relation>=</relation>
  </Cell>
</map>
```