Abstract: This paper describes an ontology matching system ONTMAT1, and presents the results obtained for the Ontology Alignment Evaluation Initiative (OAEI) 2019. ONTMAT1 compares entities of ontologies to align by structural and terminological methods which use reasoner with wordnet dictionary. Then, based on similarities of individual, datatype properties and the semantic of property restriction, the weight that estimates the performance of structural and linguistic similarities is calculated.

Keywords: Ontology, Alignment, OWL.

1 Presentation of the system

ONTMAT1 (ONTology MATching) is an ontology alignment tool, aiming to align OWL entities (classes, object properties), participating for the first time in OAEI (Conference track).

1.1 State, purpose, general statement

ONTMAT1 uses terminological methods based on n-gram measure and WordNet dictionary [1] that is exploited as background knowledge with pellet reasoner [2], to provide synonyms of names of individuals, concepts, and properties, of ontologies source (O1) and target(O2). Then, the results obtained are saved in: individual matrix ($M_{ind}$), concepts matrix ($M_{con}$), and properties matrix ($M_{p}$), for individuals, concepts and properties, respectively.

Further, a new weight that evaluates the impact of restriction property (object properties [3] and datatype properties) on the structural similarity of concept is calculated. Thus, the impact of terminological similarity is 1 minus this weight. Then, the final result of concepts alignment is the sum of these similarities.

1.2 Approach description

The suggested algorithm is composed from 3 levels as explain in the following:
1. In level 1, the normalization techniques are applied on each entities name of matrices \((\text{M}_{\text{opt}}, \text{M}_{\text{ind}}, \text{M}_p)\), such as lemmatization\(^4\). Then, the n-gram measure is used to assess the similarity among these entities. This measure is opted because it permits the control of the lexicon size and keeping at the same time a reasonable threshold for every composed term (names). The obtained value is assigned to the intersection between entities into every matrix. Since, the metric measures used to align entities may suffer from several drawbacks, such as: the existence of synonyms that expresses the same entity using different words. Entities names are also compared to WordNet synsets using n-gram and the relation among synsets are inferred by reasoner pellet. Then, the relations among these entities are deduced from relations inferred by reasoner.

If synonym relation is inferred, then the value of intersection among these entities in their matrix becomes the average between 1.0 and the value calculated by the n-gram measure, else the existent value is preserved.

2. In level 2, every property restriction defines the class is allocated by a weight \(w_i\) that evaluates the influence of its semantic on this class.

The sublanguage OWL-DL of OWL (Web Ontology Language) certified by the World Wide Web Consortium (W3C)\(^3\) is adopted in this paper to define the offered ontology matching algorithm. This language distinguishes two types of property restrictions: value constraints and cardinality constraints, which give a semantic sense to the assessed weight. A value constraint applies constraints on the range of the property. These constraints on the class \(C\) or an object \(o\) can be:

- **allValuesFrom\((C)\)**, is same to the universal (for-all: \(\forall\)) quantifier of Predicate logic that for each instance of \(C\), every value for Property must satisfies the constraint, so the algorithm can assert that this property has a robust impact on the class. Consequently, from its semantic, the influence of this restriction on the class is considered “strong” and suggested 1.0 as weights \(w_{1i}, w_{2j}\) in \(O_1, O_2\), respectively, affected by ONTMAT1 to allValuesFrom.

- **someValuesFrom\((C)\)**, is like the existential quantifier of Predicate logic that for each instance of \(C\), there exists at least one value for Property that satisfies the constraint. So, the influence of this constraint on a given class can be valued as average and the value 0.75 is affected to \(w_{1i}\) in \(O_1\) and \(w_{2j}\) in \(O_2\).

- **hasValue\((o)\)**, joins a restriction class to a value \(o\), which may be an individual or a data value. This restriction designates a class of all individuals for which the concerned property has at least one value semantically equivalent to \(o\) (it can, also, have supplementary values). So, the effect of this restriction can be considered as weak and the weights \((w_{1i}, w_{2j}\) in \(O_1, O_2\), respectively) assigned are evaluated to 0.25.
• A cardinality constraint is defined by \( \text{maxCardinality}(n) \) and \( \text{minCardinality}(n) \), where \( n \) is the number of values that a property can take. \textit{Owl: maxCardinality}(n) describes a class of all individuals that have \textit{at most} \( n \) diverse values (individuals or data values) for the concerned property. The influence of this constraint is only on \( n \) value, for this reason, it is estimated as a weak constraint and ONTMAT1 affects 0.25 to weights \( w_{ik}, w_{kj} \) in \( O_1, O_2 \), respectively. The same for \( \text{minCardinality}(n) \) that describes a class of all individuals that have \textit{at least} \( n \) various values for the concerned property.

3. Level 3 assesses structural similarity between concepts that established on properties restrictions. Property restrictions can be either \textit{datatype properties} (data literal is the value of properties), or \textit{object properties} (individual is the value of properties)\(^2\). Firstly, restriction names of concepts \( (C_{i1}, C_{2j}) \) to be matched are compared using terminological methods. Secondly, same terminological methods are used to measure similarities among \textit{datatype properties} names of both concepts to align, as well as the average of these similarities is calculated to determine \textit{data similarities}. Finally, similarities among individuals of concepts to match are extracted from \( M_{ind} \) to compute their average \textit{data similarities}.

Afterwards, weights \( w_i \) and \( w_j \) evaluated the influence of property on concepts are multiplied by \textit{data similarities} and \textit{data similarities}. Further, values affected to \( M_p \) will replace by those deduced in this level.

4. The last level consists on aggregation of above similarities of concepts. Consequently, the final similarity is the sum of structural similarity and 1 minus the average of structural weights multiplied by terminological similarity.

1.3 Adaptations made for the evaluation

We have adapted the format of the alignment result of concepts to the reference alignments restricted to name classes, using the “=“ sign for equivalence relation with confidence of 1. Although our system provides other relation called fuzzy relation.

2 Results

In this version we wish to test the techniques used by ONTMAT1, such as, the inferences mechanisms applied on WordNet, and the deduction of the matching among entities using weight based on restriction properties. The track used to do these tests is the conference track. Conference track comprises 16 ontologies from the domain of conference organization.

The results of the evaluation based on crisp reference alignments that contains only classes \( (M1-rar2; M1-ra1; M1-ra2) \) are considered in this study because the objective of this version is to show the influence of the weight and the reasoner on the classes alignment and the properties will treat in the next version
As depicted in Table 1, ONTOMAT1 provides fairly stable alignments when matching conference ontologies. Table 2 illustrates that ONTOMAT1’s performance in discrete and continuous cases increases 16 percent in terms of F-measure over the sharp reference alignment from 0.55 to 0.64, which it is principally driven by increased recall.

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>F-Measure 1</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1-ra1</td>
<td>0.82</td>
<td>0.61</td>
<td>0.49</td>
</tr>
<tr>
<td>M1-ra2</td>
<td>0.77</td>
<td>0.56</td>
<td>0.44</td>
</tr>
<tr>
<td>M1-rar2</td>
<td>0.77</td>
<td>0.58</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Table 2. Results based on the uncertain version of the reference alignment.

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>F-measure 1</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncertain reference alignments (Sharp)</td>
<td>0.82</td>
<td>0.55</td>
<td>0.41</td>
</tr>
<tr>
<td>Uncertain reference alignments (Discrete)</td>
<td>0.82</td>
<td>0.64</td>
<td>0.52</td>
</tr>
<tr>
<td>Uncertain reference alignments (Continuous)</td>
<td>0.82</td>
<td>0.64</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Finally, ONTOMAT1 have generated only one incoherent alignment in the evaluation based on logical reasoning.

2.1 Discussions on the way to improve the proposed system

To improve our application, we will also align the properties of ontologies (O1, O2). Then, adapt it to read all files type, and integrate the translator to test our tool under other tracks as: Instance Matching, MultiFarm.

2.2 Comments on the OAEI test cases

The application seals-ont-client from seal, only test files where the alignment relation between concepts is itself the equivalence relation. However ONTOMAT1, offers other possibilities in terms alignment relations between entities such as; & : Fuzzy. We hope that OAEI takes into consideration those types of relations in the reference alignment file.
3 Conclusion and future work

We have briefly described the mechanisms exploited by our proposition ONTMAT1, and presented the results obtained under the conference track of OAEI 2019.

This is our first participation in OAEI with ONTMAT1, the results are satisfying, and the system presents some limitations in term of recall. In the future, we will make great efforts to improve ONTMAT1 results, and participate in more tracks.

References


---

1 https://www.w3.org/
2 https://www.w3+.org/TR/owl-ref/