

AOT / AOTL Results for OAEI 2014

Abderrahmane Khiat¹, Moussa Benaïssa¹

¹ LITIO Lab, University of Oran, BP 1524 El-Mnaouar Oran, Algeria
abderrahmane_khiat@yahoo.com
moussabenaïssa@yahoo.fr

Abstract. In this paper, we describe our two ontology alignment systems AOT and AOTL respectively. The AOT system uses different terminological matchers with a local filter and the AOTL system consists in combining the different terminological with linguistic matchers. The AOT and AOTL systems are designed for the ontology matching tracks in order to discover new semantic correspondences between entities of different ontologies to be aligned. This is the first participation of AOT and AOTL at OAEI 2014, we present the results obtained by running the first version of our systems in different tracks of OAEI 2014 evaluation campaign.

1 Presentation of the system

1.1 State, purpose, general statement

AOT (Ontology Alignment at Terminological level) and AOTL (Ontology Alignment at Terminological and Linguistic level) are automatic ontology alignment systems realized for the purpose to solve the problem of ontology Matching. The AOT system uses various terminological matchers with a local filter in order to find correspondences between ontologies to be aligned. Contrary to AOT system, AOTL combines the similarities calculated by the various string matching algorithms at terminological level without a local filter, then these similarities are combined with similarities calculated using an external resource WordNet i.e. at linguistic level. The next step (for AOTL system) consists in combining the similarities by gives the priority to linguistic matcher; otherwise we have used an average aggregation method. Finally AOT and AOTL applied a filter in order to identify the alignment

For AOT system, we have proposed a local filter (section 1.2.1.3) in order to select better correspondences and we envision to use AOT in order to study the system behavior using different aggregation and filter methods and proposing in the future more efficient filters and aggregation methods and add other matchers.

For AOTL system, we have used an external resource WordNet in order to select semantic correspondences and we plan to use AOTL in order to discover new

semantic correspondences more than select the better one i.e. we are interested in recall more than precision, of course with good the f-measure (balanced).

The details of each step of our systems are described in the following section.

1.2 Specific techniques used

The process of AOT and AOTL systems consists in the following two successive steps: 1) Calculation of Similarities and 2) Combination and Extraction of Alignment.

A. AOT system

1.2.1 Step 1: Calculation of Similarities

1.2.1.1 Phase 1: Extraction of Entities of the Ontologies

In this phase, our system takes as input the two ontologies to be aligned and extract their entities: names, labels, properties (data property and object property) and without forgetting the instances.

1.2.1.2 Phase 2: The Applied Matchers

In this phase, our system calculates the similarities between entities extracted in previous phase, using various string-based matching algorithms. More precisely the different string-based matching algorithms used are: levenshtein-distance, block-distance, Jaro, SLIM-Winkler, Jaro-Winkler, Smith-Waterman and Needleman-Wunsch. The calculations of similarities by each string matching algorithm are represented in matrix.

1.2.1.3 Phase 3: Local Filter

In this phase, our system applies a local filter on each matrix i.e. we choose for each string-based matching algorithm a threshold to realize a filter. We consider that: the similarities which are less than the threshold are set to 0. Our intuition behind this local filter is that the similarities which are less than the threshold can influence the strategy of the average aggregation.

1.2.2 Step 2: Combination and Extraction of Alignment

1.2.2.1 Phase 1: Aggregation of Similarities

In this phase, our system combines the similarities of each matrix (after we have applied a local filter) using the average aggregation method and the result of the aggregation is represented in a matrix.

1.2.2.2 Phase 2: Global Filter and Identification of Alignment

In this final phase, our system applies a second filter on the matrix combined (result of the previous step) in order to select the correspondences found using the maximum strategy with a threshold.

B. AOTL system

We mention in this section the difference between AOT and AOTL system.

First, we have added another matcher at linguistic level for AOTL system in second phase “The applied Matchers”, we have used an external dictionary WordNet.

AOTL does not use phase “Local Filter”, the similarities calculated by each matcher are represented in matrix without a local filter.

In the phase “Aggregation of Similarities”, AOTL system gives priority to WordNet i.e. if the similarity value calculated using WordNet is greater than the similarity value calculated using string matching algorithms, the similarity value of the matrix combined is equal to the similarity calculated using WordNet, else we use the average aggregation method. The result of the aggregation is represented in a matrix.

1.3 Adaptations made for the evaluation

We do not have made any specific adaptation for the first version of AOT and AOTL, for OAEI 2014 evaluation campaign. All parameters are the same for different tracks of OAEI 2014.

1.4 Link to the system and parameters file

The first version of AOT and AOTL systems submitted to OAEI 2014 can be downloaded from seal-project at <http://www.seals-project.eu/>.

1.5 Link to the set of provided alignments (in align format)

The results of AOT and AOTL systems can be downloaded from seal-project at <http://www.seals-project.eu/>.

2 Results

In this section, we present the results obtained by running AOT and AOTL on different tracks of OAEI 2014 evaluation campaign i.e. on the tracks: Benchmark, Conference, Multifarm and Anatomy.

2.1 Benchmark

The Benchmark track contains different series which contain reference ontologies of different sizes and from different domains. The AOT system uses various string-based matching algorithms in order to find correspondences between entities of the two ontologies to be aligned and the AOTL system use another matcher at linguistic level in order to select semantic correspondences. However when these ontologies do not contain terminological information (limited information or random strings) our systems fails to identify the alignment.

The table 1 below presents the results obtained by running AOT and AOTL on the Benchmark tracks of OAEI 2014 evaluation campaign i.e. H-mean of our systems on tracks: biblio and finance.

System	Test group	H-mean Prec.	H-mean Rec	H-mean f-Measure
AOT	Biblio	0.96	0.50	0.68
	Finance	0.77	0.65	0.70
AOTL	Biblio	0.85	0.67	0.75
	Finance	0.75	0.63	0.68

Table 1. The results of AOT and OATL on the Benchmark track of OAEI 2014.

2.2 Anatomy

The Anatomy track contains two large ontologies that describe the biomedical domain of human and mouse anatomy. The Table 2 shows the results obtained by running AOT and AOTL on Anatomy track of OAEI campaign 2014.

System	Test	H-mean Prec.	H-mean Rec	H-mean f-Measure
AOT	Anatomy	0.436	0.775	0.558
AOTL	Anatomy	0.707	0.078	0.14

Table 2. The results of AOT and OATL on the Anatomy track of OAEI 2014.

2.3 Conference

The conference track contains about 16 ontologies that describe the same domain (conference organization). The Table 3 presents the results obtained by running AOT and AOTL on Conference track of OAEI campaign 2014.

System	Test	H-mean Prec.	H-mean Rec	H-mean f-Measure
AOT	Conference	0.8	0.48	0.59
AOTL	Conference	0,78	0,42	0,55

Table 3. The results of AOT and OATL on Conference track of OAEI 2014.

2.4 Multifarm

The Multifarm track contains different ontologies translated into different languages. Our systems AOT and AOTL do not deal efficiently (for now) with the Multifarm track. The Table 4 presents the results obtained by running AOT and AOTL on Multifarm track of OAEI campaign 2014.

System	Test	H-mean Prec.	H-mean Rec	H-mean f-Measure
AOT	Diff-ontologies	0,02	0,17	0,04
	Same-ontologies	0,11	0,12	0,12
AOTL	Diff-ontologies	0,10	0,2	0,3
	Same-ontologies	0,11	0,12	0,12

Table 4. The results of AOT and AOTL on Multifarm track of OAEI 2014.

3 General comments

This is the first time that our systems participate in different tracks of the OAEI 2014 evaluation campaign, and AOT and AOTL are new on the SEALS Platform. However we can conclude with this first participation that AOT provides globally good results in terms of F-measure. Contrary to AOT, the AOTL system provides good results in terms of F-measure on benchmark track but in other tracks the results are not so good.

3.1 Comments on the results

The AOT and AOTL systems are an automatic ontology matching system designed in order to find the correspondence between different entities of ontologies to be aligned. The results obtained by running our systems on different tracks of OAEI 2014 evaluation campaign are slightly good on some tracks but not satisfactory in others.

3.2 Discussions on the way to improve the proposed system

The objective behind the implementation of AOT system is to find the best strategy of aggregation and filter as we have proposed in section 1.2.1.3 (a local filter). Contrary to AOT, the objective behind the implementation of AOTL system is to discover new semantic correspondences by adding other matchers. For now, we have used matchers at terminological and linguistic level.

As we have mentioned before AOT and AOTL systems use terminological information and when these ontologies do not contain this information our two systems fails. Our both systems does not deal with ontologies written in different

languages, and we hope in the future add a module to translate them in the same language.

Another point to be discussed is how to make our systems flexible i.e. the choice of thresholds for the various matchers and ontologies. It is obvious that we cannot set the threshold for all ontologies, in order to find automatically the correspondences between entities of ontologies to be aligned; because each ontology possesses its own specific characteristic?

4 Conclusion

This is the first time the AOT and AOTL systems have participated in OAEI campaign. In this year, our systems have participated in different tracks of OAEI 2014 evaluation campaign.

The AOT system combines the various string-based matching algorithms with average aggregation method. Then we have applied a filter on the combined matrix for the selection of semantic correspondences between ontologies to be aligned. The use of these algorithms is justified by the fact that in the ontologies the terminological information is very important.

Contrary to AOT system, AOTL add at linguistic level an external resource dictionary WordNet for better selection of semantic correspondences.

Finally the results show that our systems can provide some good results. We have used a local filter (in section 1.2.1.3) for AOT system and we envision to study the AOT behavior using different aggregation and filter methods in order to propose in the future new other metrics of filter and aggregation. For the AOTL system, we are interested in the discovery of semantic correspondences by matchers rather than the combination of similarities. We envision using other matchers such as structure-based and reasoning-based matchers in AOTL system.

References

1. J. Euzenat and P. Valtchev, "Similarity-based ontology alignment in owlite," in Proceedings of ECAI, (2004).
2. J. Euzenat and P. Shvaiko. *OntologyMatching*. Springer (2007).
3. M. Ehrig. *Ontology Alignment Bridging the Semantic Gap*. Springer (2007).
4. M. Jaro. Advances in record-linkage methodology as applied to matching the 1985 census of tampa, florida. *Journal of America Statistical Association*, 84(406):414-420, (1989).
5. A. Khiat et M. Benaissa: "Nouvelle Approche d'Alignement d'Ontologies à base d'Instances : trasferet des instances par l'inférence", In The Proceeding of International Conference On Artificial Intelligence and Information Technology, ICA2IT 2014, Ouargla, Algeria, (2014).
6. V. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10:707-710, (1966).
7. W. Winkler. The state of record linkage and current research problems. *Statistics of Income Division, Internal Revenue Service, Publication R99/04*, (1999).

8. M. Ehrig and Y. Sure, "Ontology mapping - an integrated approach," in Proceedings of the European Semantic Web Symposium ESWS, (2004).
9. G. A. Miller, R. Beckwith, C. D. Fellbaum, D. Gross, K. Miller. WordNet: An online lexical database. *Int. J. Lexicograph.* 3, 4, pp. 235–244, (1990).