MEDLEY Results for OAEI 2012

Walid Hassen

University of Tunis El Manar - Faculty of Sciences of Tunis
Computer Science Department - LIPAH
Campus Universitaire, 1060 Tunis, Tunisia
walidhassenlipah@gmail.com

Abstract. MEDLEY is an alignment method based on lexical and structural treatments. This method includes a specific technique to deal with multilingual ontologies. This paper introduces MEDLEY and summarizes the results for OAEI 2012.

1 Presentation of the system

MEDLEY can be presented as an OWL ontology alignment method that relies on simple similarity metrics. Each ontology pair, can be transformed into graphs structures. This means that links are OWL primitives and nodes are classes, properties, and individuals. The algorithm includes a lexical, structural treatment. Each node can be matched with few ones, then MEDLEY select pairs that maximize the global similarity value.

1.1 State, purpose, general statement

MEDLEY generates alignments between OWL-DL ontologies based on simple lexical metrics and structures matching between links of each node (class, property, instance). Specific treatment is applied for multilinguality issue, using a dictionnary to find equivalence between concepts labelled in diffrent natural languages.

1.2 Specific techniques used

Each entity in the first ontology is aligned each entity in the second, in a primary step, in lexical metrics, then in structural treatment. The algorithm reiterate this process for all ontologies's concepts.

- Lexical treatment : q-gram [1] and levenshtein [2] measures were used to calculate the similarity measures between nodes. In addition, treatments and tokenization stemmatisation were conducted.
- Structural treatment :If an entity belongs to a given ontology has a neighbor that is already part of the alignment set, then the node that neighbor is aligned to must be a neighbor of any prospective match for this entity.

1.3 Adaptations made for the evaluation

The MEDLEY method deals with three test suites used in the Ontology Alignment Evaluation Initiative (OAEI 2012). The method was wrapped in a certain folder structure to be evaluated locally after being integrated in the SEALS platform. The package contains all the libs files required by the method and a zipped .jar file that acts as a bridge.

1.4 Link to the system and parameters file

The release of the MEDLEY method and the parameter file used for OAEI 2012 are located at https://github.com/medley.

2 Results

In this section, we present the results obtained by MEDLEY in the OAEI 2012.

2.1 Benchmark

The benchmark tests sets can be divided into eight groups: 101, 20x, 22x, 23x, 24x, 25x, 26x and 30x. For each group the mean values of precision and recall are computed. Table 1 shows the values of the evaluation metrics. Tables 1, 2 and 3 recapitulate the obtained values for this track.

Table 1. Results on Biblio

Test group	Precision	Recall	F-Measure
101	0.72	1.0	0.84
20x	0.43	0.4	0.408
22x	0.716	1.0	.988
23x	0.781	1.0	0.853
24x	0.633	0.572	0.571
25x	0.51	0.4	0.421
26x	0.322	0.357	0.31

2.2 Conference

In scenario 1, MEDLEY have 0.54 of precision and 0.50, with 0.52 as recall an f-measure about 0.52. In scenario 2, MEDLEY performs 0.59 of precision, 0.42 recall and 0.49 of f-measure.

Table 2. Results on Benchmark 2

Test group	Precision	Recall	F-Measure
101	1.00	1.00	1.00
20x	0.697	0.4	0.493
22x	0.998	1.0	1.0
23x	0.995	1.0	1.0
24x	0.787	0.57	0.63
25x	0.757	0.439	035
26x	0.611	0.354	0.435

Table 3. Results on Benchmark 3

Test group	Precision	Recall	F-Measure
101	0.79	1.00	0.88
20x	0.568	0.4	0.454
22x	0.805	1.0	0.888
23x	0.88	1.0	0.93
24x	0.715	0.571	0.609
25x	0.642	0.4	0.463
26x	0.695	0.352	0.398

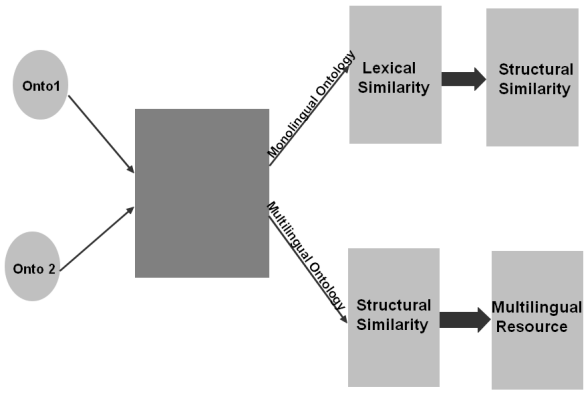

Fig. 1. MEDLEY components

2.3 Multifarm

For treating multilingual ontologies, our method uses an external resource as sketched by figure 1 for the translation stage[1]. Tables 4, 5, 6, 7, 8, 9 and 10 summarize the results.

[1] http://www.freelang.com/dictionnaire/index.php

Table 4. Group **(cz)** as source ontology

Test group	Precision	Recall	F-Measure
cz-de	0.51	0.07	0.13
cz-en	0.33	0.09	0.14
cz-es	0.43	0.07	0.12
cz-fr	0.33	0.05	0.09
cz-nl	0.33	0.06	0.10
cz-pt	0.46	0.08	0.13
cz-ru	0.00	0.00	NaN

Table 5. Group **(de)** as source ontology

Test group	Precision	Recall	F-Measure
de-en	0.40	0.10	0.15
de-es	0.43	0.09	0.15
de-fr	0.40	0.09	0.14
de-nl	0.38	0.09	0.15
de-pt	0.43	0.09	0.15
de-ru	0.00	0.00	NaN

Table 6. Group **(en)** as source ontology

Test group	Precision	Recall	F-Measure
en-es	0.54	0.48	0.51
en-fr	0.62	0.61	0.61
en-nl	0.56	0.42	0.48
en-pt	0.57	0.51	0.54
en-ru	0.05	0.00	0.00

Table 7. Group **(es)** as source ontology

Test group	Precision	Recall	F-Measure
es-fr	0.31	0.04	0.08
es-nl	0.21	0.03	0.05
es-pt	0.50	0.11	0.18
es-ru	0.02	0.00	0.00

3 General comments

We participate this year for the first time in OAEI and see the result obtained by our method. The evaluation and comparison of ontology alignment and schema matching components as OAEI is very useful for the development of such

Table 8. Group **(fr)** as source ontology

Test group	Precision	Recall	F-Measure
fr-nl	0.45	0.10	0.16
fr-pt	0.35	0.08	0.14
fr-ru	0.00	0.00	NaN

Table 9. Group **(nl)** as source ontology

Test group	Precision	Recall	F-Measure
nl-pt	0.31	0.07	0.11
nl-ru	0.03	0.00	0.00

Table 10. Group **(pt)** as source ontology

Test group	Precision	Recall	F-Measure
pt-ru	0.03	0.00	0.00

3.1 Discussions on the way to improve the proposed system

MEDLEY is still a primary work that needs to be adressed on few levels, notably, to deal with greater ontologies.

4 Conclusion

In this paper, we presented MEDLEY as an alignment method. The new proposed method MEDLEY, shows a special focus on multilinguality. The alignment process is based on examining the structures and the informative wealth on each ontlogy pair to align.

References

1. Ukkonen, E.: Approximate string-matching with Q-GRAMS and maximal matchs. Theoretical Computer Science **92**(1) (1992) 191–211
2. Levenshtein, I.V.: Binary codes capables of corrections, deletions, insertions and reversals. Soviet Physics-Doklady **10**(8) (1966) 707–710