EVOCROS: Results for OAEI 2018

Juliana Medeiros Destro[1], Gabriel Oliveira dos Santos[1], Julio Cesar dos Reis[1], Ricardo da S. Torres[1], Ariadne Maria B. R. Carvalho[1], and Ivan Luiz Marques Ricarte[2]

[1] Institute of Computing, University of Campinas, Campinas-SP, Brazil
{juliana.destro,jreis,rtorres,ariadne}@ic.unicamp.br
gabriel.santos@students.ic.unicamp.br
[2] School of Technology, University of Campinas, Limeira-SP, Brazil
ricarte@unicamp.br

Abstract. This paper describes EVOCROS, a cross-lingual ontology alignment system suited to create mappings between ontologies described in different natural language. Our tool combines semantic and syntactic similarity measures in a weighted average metric. The semantic is computed via NASARI vectors used together with *BabelNet*, which is a domain-neutral semantic network. The tool employs automatic translation to a pivot language to consider the similarity. EVOCROS was tested and obtained high quality alignment in the Multifarm dataset. We discuss the experimented configurations and the achieved results in OAEI 2018. This is our first participation in OAEI.

Keywords: cross-lingual matching · semantic matching · background knowledge

1 Presentation of the system

There is a growing number of ontologies described in different natural languages. The mappings among different ontologies are relevant for the integration of heterogeneous data sources to facilitate the exchange of information between systems. Although automatic monolingual ontology matching has been extensively investigated [7], cross-lingual ontology matching still demands further investigations aiming to automatically identify correspondences between ontologies described in different languages. EVOCROS is our attempt at automatic cross-lingual ontology matching, inspired from experiments on the influence of syntactic and semantic similarity measures in ontology matching algorithms [1]. In this section, we describe the system and the implemented techniques.

1.1 State, purpose, general statement

EVOCROS is a cross-lingual ontology alignment tool based on a composed similarity measure relying on both syntactic and semantic similarity techniques. Syntactic similarity may be understood as a score calculated based on string analysis (extracted from labels of concepts), whereas the semantic similarity is computed taking into account background knowledge. Our approach computes a weighted mean of semantic and syntactic similarities.

1.2 Specific techniques used

The tool is developed in Python 3. It works by comparing the computed similarity between a concept from an ontology (in its automatically translated version) to another concept from a different ontology. The concept terms are translated to a pivot natural language aiming to use available external resources such as thesauri, corpora, dictionaries, *etc.* to overcome the language and alphabet barriers.

Figure 1 presents the workflow of the tool. The first step is the pre-processing of the source and target input ontologies, converting them into owlready2[3] objects. Each concept of the source ontology is compared to all concepts of the target ontology.

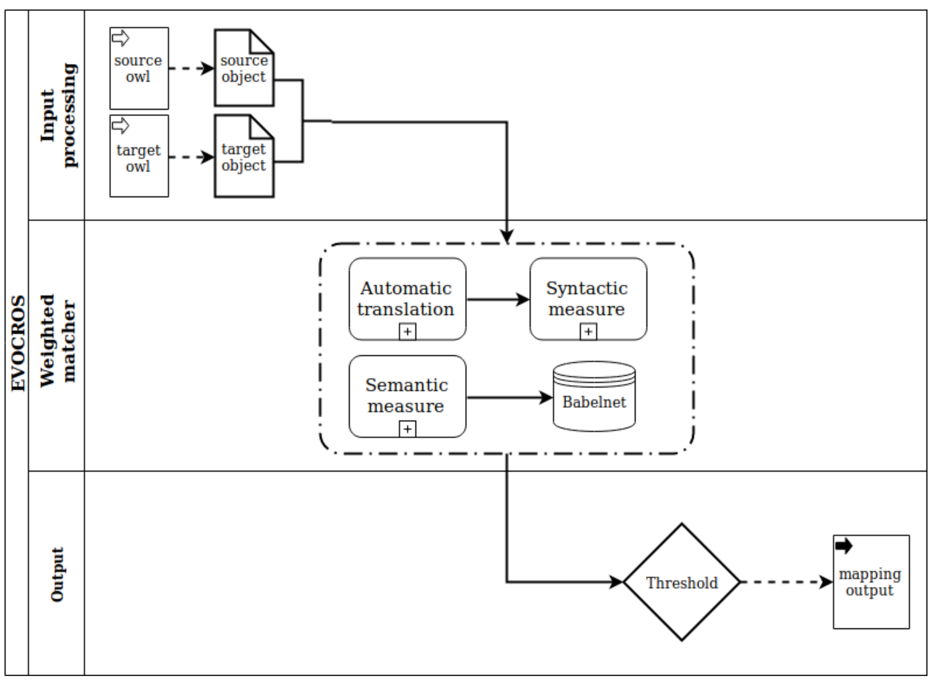

Fig. 1. EVOCROS workflow.

[3] Python 3 library to manipulate ontologies as objects.

Syntactic Similarity Measure. For syntactic similarity measure, the concept labels of both the source and target ontologies are first translated to a pivot language using automatic translation. We are using English as pivot language for OAEI 2018 though the tool accepts any language as pivot. The concepts are then compared by measuring the syntactic similarity via edit distance Levenshtein [3]) as a syntactic similarity measure.

Semantic Similarity Measure. Semantic similarity between terms is a metric to evaluate how similar two given terms are considering their meanings in a certain context. For example, the words “nail” and “hammer” are more similar considering the tool context than “nail” and “finger”. On the other hand, when we consider the anatomy context, “nail” and “finger” are more similar than “nail” and “hammer”.

For semantic similarity, we use the concept label in its original language, without any translation. There are a lot of algorithms to calculate semantic similarity. These algorithms usually explore an external resource such as vocabulary, dictionaries or thesauri to help computing the similarity between two words. EVOCROS explores a *Weighted Overlap* measure [6] relying on the neutral-domain semantic network *BabelNet* [5]. The tool retrieves from *Babelnet* the synsets of the concept labels of both source and target ontologies and compare them to measure the semantic similarity.

Our proposal generates cross-lingual ontology alignments taking into account the combination of semantic and syntactic similarity by computing the weighted average as follows:

Definition 1 (Composed Similarity). *Let $sem(t_1, t_2)$ and $sin(t_1, t_2)$ be the semantic similarity, and the syntactic one between the terms t_1 and t_2, respectively. We assume that the similarities are normalized between 0 and 1. Formally:*

$$simC(t_1, t_2) = \frac{\alpha sin(t_1, t_2) + \beta sem(t_1, t_2)}{\alpha + \beta} \tag{1}$$

where α and β are constants.

If the weighted similarity reaches a threshold, the concept pair is recorded to the output file, generated in RDF format. Otherwise, it is discarded.

1.3 Adaptations made for evaluation

EVOCROS uses a configuration file with the source and target ontologies, and their respective language. In order to participate in OAEI, we modified the tool to receive the source and target ontologies as input parameters and retrieve the ontology language from the *lang* XML tag. The bridge created for SEALS platform is written in Java and executed system calls to run the tool, written in Python 3. Although the tool executed locally using the SEALS client, there were issues during evaluation on SEALS platform and only local results are available in this report.

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

Alignment results are available at https://github.com/jmdestro/evocros-results.

2 Results

In this section, we describe the results obtained from local experiments using a sub-set of Multifarm with the same configuration used in OAEI 2018 evaluation.

2.1 Multifarm

Our experiments were based on ontologies from conference domain from the *MultiFarm dataset 2015* [4]. We used the reference mappings between the ontologies described in English and Spanish mapped into those concepts in the Portuguese Language.

Several weights for similarity measures and different similarity thresholds were evaluated locally. For OAEI 2018, only the following configuration was submitted: **threshold:** 0.66, **syntactic similarity weight:** 0.75, **semantic similarity weight:** 0.25. This was the configuration with the most interesting results. Table 1 presents the used configuration and the results for conference-conference alignment for languages spanish-portuguese (es-pt) and english-portuguese (en-pt).

Table 1. Cross-lingual mapping of conference-conference ontologies from MultiFarm.

Languages	Threshold	Syntactic similarity weight	Semantic similarity weight	Precision	Recall	F-measure
es-pt	0.66	0.75	0.25	0.68	0.33	0.44
en-pt	0.66	0.75	0.25	0.72	0.41	0.52

The choice of weights assigned to each similarity measure played an important role in the results. Tables 2 and 3 present the obtained results for different configurations. Considering the syntactical weights as 0.75 and 0.80 generated the best mappings, that is, they result in alignments with the greatest f-measure. Thus, our technique may be understood as a good alternative to syntactic or semantic only methods, and it might perform even better taking into account the correct parameters.

Table 2. MultiFarm alignment of Conference [ES] - Conference [PT] ontologies, using different threshold and weight.

Threshold	Syntactic weight	Semantic weight	Precision	Recall	F-measure
0.66	0.50	0.50	0.49	0.15	0.23
	0.33	0.67	0.40	0.10	0.16
	0.25	0.75	0.33	0.15	0.21
	0.20	0.80	0.30	0.15	0.20
	0.67	0.33	0.69	0.30	0.42
	0.75	0.25	0.68	0.33	0.44
	0.80	0.20	0.59	0.31	0.40
0.75	0.50	0.50	0.58	0.16	0.25
	0.33	0.67	0.48	0.16	0.24
	0.25	0.75	0.45	0.18	0.25
	0.20	0.80	0.40	0.17	0.24
	0.67	0.33	0.65	0.16	0.26
	0.75	0.25	0.75	0.31	0.44
	0.80	0.20	0.72	0.33	0.45
0.80	0.50	0.50	0.65	0.16	0.26
	0.33	0.67	0.58	0.16	0.25
	0.25	0.75	0.50	0.17	0.26
	0.20	0.80	0.45	0.18	0.25
	0.67	0.33	0.65	0.16	0.26
	0.75	0.25	0.65	0.16	0.26
	0.80	0.20	0.75	0.31	0.44
0.95	0.50	0.50	0.64	0.11	0.18
	0.33	0.67	0.67	0.15	0.24
	0.25	0.75	0.69	0.16	0.26
	0.20	0.80	0.65	0.16	0.26
	0.67	0.33	0.64	0.11	0.18
	0.75	0.25	0.64	0.11	0.18
	0.80	0.20	0.64	0.11	0.18

3 General comments

In this section, we discuss our results and the ways to improve the system.

3.1 Comments on the results (strength and weaknesses)

The tool had satisfactory results but the execution time was exceedingly long due to constant RestAPI calls to *Babelnet*. The results showed an influence of threshold: as the threshold rises, the precision also increases. It may be explained by considering equivalence of only those concepts with a high level of similarity. However, f-measure declines as the threshold increases because large values assigned to threshold make the algorithm disregards concepts that are equivalent, but somehow was assigned a lower level of similarity than expected by the threshold. As a result, the recall drops substantially, because many correct correspondences are ignored, and thus f-measure decreases. Empirically, we concluded

Table 3. MultiFarm alignment of Conference [EN] - Conference [PT] ontologies, using different threshold and weight.

Threshold	Syntactic weight	Semantic weight	Precision	Recall	F-measure
0.66	0.50	0.50	0.57	0.18	0.27
	0.33	0.67	0.42	0.21	0.28
	0.25	0.75	0.32	0.18	0.23
	0.20	0.80	0.28	0.17	0.21
	0.67	0.33	0.69	0.34	0.45
	0.75	0.25	0.72	0.41	0.52
	0.80	0.20	0.68	0.21	0.32
0.75	0.50	0.50	0.60	0.17	0.26
	0.33	0.67	0.52	0.23	0.32
	0.25	0.75	0.50	0.22	0.31
	0.20	0.80	0.43	0.21	0.28
	0.67	0.33	0.58	0.21	0.31
	0.75	0.25	0.70	0.15	0.25
	0.80	0.20	0.75	0.17	0.27
0.80	0.50	0.50	0.58	0.16	0.25
	0.33	0.67	0.57	0.23	0.32
	0.25	0.75	0.52	0.23	0.32
	0.20	0.80	0.50	0.22	0.31
	0.67	0.33	0.61	0.21	0.32
	0.75	0.25	0.61	0.09	0.15
	0.80	0.20	0.73	0.15	0.25
0.95	0.50	0.50	0.64	0.19	0.29
	0.33	0.67	0.61	0.21	0.32
	0.25	0.75	0.61	0.21	0.32
	0.20	0.80	0.61	0.21	0.32
	0.67	0.33	0.64	0.19	0.29
	0.75	0.25	0.64	0.07	0.13
	0.80	0.20	0.64	0.07	0.13

that the thresholds that generate the more accurate mappings were $\lambda = 0.66$ and $\lambda = 0.75$.

3.2 Discussions on the way to improve the proposed system

This was the first evaluation of the system and although there was issues during the evaluation phase of OAEI, preventing the system to be executed in SEALS platform, the local results are encouraging. Our main goals for future work are:

Reduce execution time: the tool has a long execution time due to constant RestAPI calls to *Babelnet* and needs to be optimized with local caches.

Bag of graphs: ontologies can be represented as graphs, thus allowing for partitioning [2] and comparison of sub-graphs. Bag-of-graphs [8] is a graph matching approach, similar to bag-of-words. It represents graphs as feature vectors, highly simplifying the computation of graph similarity and reducing execution time. We propose as future investigation to use a simple vector-based representation for graphs and investigate it for cross-lingual ontology matching.

3.3 Comments on OAEI

There were issues during the evaluation phase, preventing the system to participate in Multifarm track. For future editions of OAEI, we plan to participate submitting EVOCROS on the newly available HOBBIT platform, using a docker image, to ensure system compatibility during evaluation.

4 Conclusion

EVOCROS proposed an approach to cross-lingual ontology matching by combining semantic and syntactic similarity measures. This is the first participation of the system in OAEI. The evaluation with the Multifarm dataset confirmed the quality of mappings generated by our technique. For future work, we plan to improve our cross-lingual ontology alignment proposal considering different combinations of background knowledge, such as specific-domain thesauri to evaluate the semantic similarity. We also plan to further evaluate runtime optimization aspects to fix issues found during the evaluation phase.

Acknowledgements

This work has been supported by São Paulo Research Foundation (FAPESP): grants #2017/23522-3 and #2017/02325-5.

References

1. Destro, J.M., Dos Reis, J.C., Brito, A.M.C.R., Ricarte, I.L.M.: Influence of semantic similarity measures on ontology cross-language mappings. In: Proceedings of the Symposium on Applied Computing (SAC 2017). pp. 323–329. ACM (2017)
2. Hamdi, F., Safar, B., Reynaud, C., Zargayouna, H.: Alignment-based partitioning of large-scale ontologies. In: Advances in knowledge discovery and management, pp. 251–269. Springer (2010)
3. Levenshtein, V.I.: Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics Doklady **10**, 707710 (1966)
4. Meilicke, C., Garcia-Castro, R., Freitas, F., van Hage, W.R., Montiel-Ponsoda, E., de Azevedo, R.R., Stuckenschmidt, H., Sváb-Zamazal, O., Svátek, V., Tamilin, A., dos Santos, C.T., Wang, S.: Multifarm: A benchmark for multilingual ontology matching. Web Semantics: Science, Services and Agents on the World Wide Web **15**, 62–68 (2012)
5. Navigli, R., Ponzetto, S.P.: Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. Artificial Intelligence **193**, 217–250 (2012)

6. Pilehvar, M.T., Jurgens, D., Navigli, R.: Align, disambiguate and walk: A unified approach for measuring semantic similarity. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics. pp. 1341–1351 (2013)
7. Shvaiko, P., Euzenat, J.: Ontology matching: State of the art and future challenges. IEEE Transactions on Knowledge and Data Engineering **25**(1), 158–176 (Jan 2013). https://doi.org/10.1109/TKDE.2011.253
8. Silva, F.B., Tabbone, S., Torres, R.d.S.: Bog: A new approach for graph matching. In: Pattern Recognition (ICPR), 2014 22nd International Conference on. pp. 82–87. IEEE (2014)