AML and AMLC Results for OAEI 2019

Daniel Faria[1], Catia Pesquita[2], Teemu Tervo[2]
Francisco M. Couto[2], and Isabel F. Cruz[3]

[1] BioData.pt & INESC-ID, Lisboa, Portugal
[2] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[3] ADVIS Lab, Department of Computer Science, University of Illinois at Chicago, USA

Abstract. AgreementMakerLight (AML) is an ontology matching system designed with scalability, extensibility and satisfiability as its primary guidelines, as well as an emphasis on the ability to incorporate external knowledge. In OAEI 2019, AML's development focused mainly on expanding its range of complex matching algorithms, but there were also improvements on its instance matching pipeline and ontology parsing algorithm. AML remains the system with the broadest coverage of OAEI tracks, and among the top performing systems overall.

1 Presentation of the System

1.1 State, Purpose, General Statement

AgreementMakerLight (AML) is an ontology matching system inspired on Agreement-Maker [1, 2] and drawing on its design principles, but with an added focus on scalability to tackle large ontology matching problems [8]. While initially focused primarily on the biomedical domain, it is currently a general purpose ontology matching system that is able to successfully tackle a broad range of problems.

AML is primarily based on lexical matching algorithms [9], but also includes structural algorithms for both matching and filtering, as well as its own logical repair algorithm [10]. It makes use of external biomedical ontologies and the WordNet as sources of background knowledge [7].

This year, our development of AML was mainly focused on expanding the arsenal of complex matching algorithms of AML to improve its performance in the new Complex Matching track. The complex matching version of AML, dubbed AMLC, remains separate from the main AML submission, as we have been as of yet unable to integrate the complex code into the main code-base.

In addition to these two versions, we again participated in the SPIMBENCH and Link Discovery tracks via the HOBBIT platform. In the case of SPIMBENCH, we participated with the HOBBIT adaptation of the main AML code-base. In the case of Link Discovery, we participated with two specialized versions of AML (AML-Spatial and AML-Linking for the Spatial and Linking tasks respectively) as had been the case in

OAEI 2017 and 2018, due to the unique characteristics of these matching tasks and to the unavailability of the TBox assertions in the HOBBIT datasets.

1.2 Specific Techniques Used

This section describes only the features of AML that are new for the OAEI 2019. It also describes AMLC, a variant of AML tailored to complex matching. For further information on AML's simple matching strategy, please consult AML's original paper [8] as well as the AML OAEI results publications of the last four editions [4, 5, 3, 6].

1.2.1 AML

Ontology Parsing

We made a few extensions to AML's ontology parser to enable it to infer the types of ontology properties declared only as rdf:property (which the OWL API interprets as annotation properties by default). There were critical to correctly interpret and match the datasets for the Knowledge Graph track.

Instance Matching

We refined AML's instance matching pipeline to more adequately distinguish between cases where lexical matching should be the primary strategy complemented by property-based matching, and cases where property-based matching should be the primary strategy, by using the ratios of labels per instances and property values per instances as deciding factors. These improvements were critical to AML's effectiveness on the Knowledge Graph track.

1.2.2 AMLC

For the complex matching track, we developed algorithms to tackle additional types of EDOAL mappings, namely mappings involving union class constructs. Furthermore, we refined the Attribute Occurrence Restrictions and Attribute Domain Restrictions algorithms developed last year to take into account instance data when available.

These changes allowed AML to match ontologies from the GeoLink dataset, in addition to those from the Conference dataset.

1.3 Adaptations made for the evaluation

As was the case last year, the Link Discovery submissions of AML are adapted to these particular tasks and datasets, as their specificities (namely the absence of a Tbox) demand a dedicated submission. The same is also true to some extent of AML's Complex Matching submission.

As usual, our submission included precomputed dictionaries with translations, to circumvent Microsoft® Translator's query limit.

1.4 Link to the system and parameters file

AML is an open source ontology matching system and is available through GitHub: https://github.com/AgreementMakerLight.

2 Results

2.1 Anatomy

AML's result was the same as in previous years, with 95% precision, 93.6% recall, 94.3% F-measure, and 83.2% recall++. It remains the best ranking system in this track by both F-measure and recall++.

2.2 Conference

AML's result was exactly the same as in recent years, with 74% F-measure according to the full reference alignment 1, 70% F-measure according to the extended reference alignment 2, 78% F-measure according to the discrete uncertain reference alignment, and 77% according to the continuous one. It remains the best ranking system in this track or tied for best by F-measure according to 4 of the 5 sets of reference alignments available. It ranks second by F-measure on the violation free version of reference alignment 2, as enforcing the removal of conservativity violations can produce undesired practical effects that are not aligned with AML's guiding principles, so our repair algorithm does not take them into account.

2.3 Multifarm

AML's results were similar to last year, ranking first with 45% F-measure in the different ontologies modality, but second with only 27% F-measure in the same ontologies modality. We are still unsure as to why AML performs worse in the same ontologies modality.

2.4 Complex Matching

AMLC was configured only for the Conference and Geolink datasets. It also produced results in the Hydrography dataset, but these were expectedly mediocre.
On the conference dataset, AMLC was the only system to participate in the non-populated version (using the simple reference alignment as input). It improved its recall in relation to last year (37% versus 25%) but this came at the expense of precision and so resulted in an identical F-measure of 34%. On the populated version, it had the highest range of coverage (query F-measure) with 46-50%.
On the GeoLink dataset, AMLC obtained a comparably modest F-measure of 32% (the top system had 60%).

2.5 Interactive Matching

AML had an identical performance to last year, as no changes were made to its interactive algorithms. It remains the system with the best F-measure in both the Anatomy and Conference datasets across all error rates (though it also has the best non-interactive F-measure in these datasets).

2.6 Large Biomedical Ontologies

AML had an F-measure of 93.3% in FMA-NCI small, 84.1% in FMA-NCI whole, 83.5% in FMA-SNOMED small, 69.7% in FMA-SNOMED whole, 81.8% in SNOMED-NCI small and 76.5% in SNOMED-NCI whole. In comparison with last year, its performance decreased in all large tasks, due to the erroneous addition of an imprecise matching algorithm in the matching pipeline when testing new configurations. Despite this, it remains the best performing system in five of the six tasks.

2.7 Disease and Phenotype

AML generated 2029 mappings in the HP-MP task, 330 of which were unique. It ranked third by F-measure according to the 3-vote silver standard, but this does not necessarily reflect its actual performance, as the unique mappings were not evaluated. If half of AML's unique mappings were proven correct, which is highly likely given the high precision AML obtains in other biomedical tasks, it would rank first in F-measure. In the DOID-ORDO task, it generated by far the most mappings (4781) and the most unique mappings (2342), and as a result had a relatively low F-measure according to the 3-vote silver standard (65.1%). Again, assessing the correctness of the unique mappings would be essential to gauge AML's true performance.

2.8 Biodiversity and Ecology

AML obtained the highest F-measure in both datasets, with 78.8% in the FLOPO-PTO task and 80.8% in the ENVO-SWEET task. It ranked first in recall and produced both the most mappings and the most unique mappings.

2.9 SPIMBENCH

AML obtained the same results as last year, with an F-measure of 86%, ranking third by F-measure.

2.10 Link Discovery

As in previous years, AML produced a perfect result (100% F-measure) in the Linking and all the Spatial tasks. It was among the most efficient systems in the later, and the only system participating in the former.

2.11 Knowledge Graph

AML was able to complete only four of the five tasks due to an unforeseen timeout in the largest task (which it had been able to carry out in testing). It produced an average F-measure of only 70% if the missing task is counted as zero, but of 88% when it is ignored. In fact, it ranked either first or second in F-measure in all the four tasks it completed.

3 General comments

3.1 Comments on the results

This year, AML was again the system that tackled the most OAEI tracks and datasets, maintaining its status as one of the broadest and best performing matching systems available to the community.

However, unlike AML's performance in traditional (simple) matching tracks, there is clearly room for improvement for AML in complex matching, as it had modest F-measures. We will strive to refine and improve AML's complex matching pipeline and contribute to the development of this branch of ontology matching.

3.2 Comments on the OAEI test cases

We once again laud the efforts of the organizers of both returning and especially new tracks, as the effort involved in organizing them cannot be overstated.

Nevertheless, we must again comment on the unsatisfactory evaluation in the Disease and Phenotype track by means of silver standards generated from the alignments produced by the participating systems via voting. We understand the effort required to build a manually curated reference alignment, but we believe that it is paramount to invest in it, in order to enable a proper evaluation of matching systems.

4 Conclusion

Like in recent years, AML was the matching system that participated in the most OAEI tracks and datasets, and it was among the top performing systems in most of them. AML's performance did not improve in any of the long-standing OAEI tracks, as most of our development effort went into tackling new challenges and extending the range of AML. We improved substantially our results in the knowledge graph track in comparison with last year, thanks to the extensions to AML's ontology parsing algorithm and its instance matching pipeline. We were also able to extend AML's complex matching algorithm portfolio, but despite this, AML complex matching performance requires further improvement. We will continue to invest in addressing this aspect of ontology matching in the near future

References

1. I. F. Cruz, F. Palandri Antonelli, and C. Stroe. AgreementMaker: Efficient Matching for Large Real-World Schemas and Ontologies. *PVLDB*, 2(2):1586–1589, 2009.

2. I. F. Cruz, C. Stroe, F. Caimi, A. Fabiani, C. Pesquita, F. M. Couto, and M. Palmonari. Using AgreementMaker to Align Ontologies for OAEI 2011. In *ISWC International Workshop on Ontology Matching (OM)*, volume 814 of *CEUR Workshop Proceedings*, pages 114–121, 2011.

3. D. Faria, B. S. Balasubramani, V. R. Shivaprabhu, I. Mott, C. Pesquita, F. M. Couto, and I. F. Cruz. Results of AML in OAEI 2017. In *OM-2017: Proceedings of the Twelfth International Workshop on Ontology Matching*, page 122, 2017.

4. D. Faria, C. Martins, A. Nanavaty, D. Oliveira, B. S. Balasubramani, A. Taheri, C. Pesquita, F. M. Couto, and I. F. Cruz. AML results for OAEI 2015. In *Ontology Matching Workshop*. CEUR, 2015.

5. D. Faria, C. Pesquita, B. S. Balasubramani, C. Martins, J. Cardoso, H. Curado, F. M. Couto, and I. F. Cruz. OAEI 2016 results of AML. In *Ontology Matching Workshop*. CEUR, 2016.

6. D. Faria, C. Pesquita, B. S. Balasubramani, T. Tervo, D. Carriço, R. Garrilha, F. M. Couto, and I. F. Cruz. Results of aml participation in oaei 2018. In *OM-2018: Proceedings of the Thirteenth International Workshop on Ontology Matching*, pages 125–131, 2018.

7. D. Faria, C. Pesquita, E. Santos, I. F. Cruz, and F. M. Couto. Automatic Background Knowledge Selection for Matching Biomedical Ontologies. *PLoS One*, 9(11):e111226, 2014.

8. D. Faria, C. Pesquita, E. Santos, M. Palmonari, I. F. Cruz, and F. M. Couto. The AgreementMakerLight Ontology Matching System. In *OTM Conferences - ODBASE*, pages 527–541, 2013.

9. C. Pesquita, D. Faria, C. Stroe, E. Santos, I. F. Cruz, and F. M. Couto. What's in a "nym"? Synonyms in Biomedical Ontology Matching. In *International Semantic Web Conference (ISWC)*, pages 526–541, 2013.

10. E. Santos, D. Faria, C. Pesquita, and F. M. Couto. Ontology alignment repair through modularization and confidence-based heuristics. *PLoS ONE*, 10(12):e0144807, 2015.