Results of AML participation in OAEI 2018

Daniel Faria[1], Catia Pesquita[2], Booma Sowkarthiga Balasubramani[3],
Teemu Tervo[2], David Carriço[2], Rodrigo Garrilha[2],
Francisco M. Couto[2], and Isabel F. Cruz[3]

[1] Instituto Gulbenkian de Ciência, 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 a system for automated ontology matching that is characterized by its efficiency, extensibility, and ability to incorporate external knowledge. In OAEI 2018, AML leveraged these features to expand its capabilities to tackle the new tracks. Particular effort was put into extending AML to produce complex mappings, and into improving instance matching approaches. AML was the only system to participate in all OAEI tracks this year, and was the top performing system, or among the top performing systems, in most tracks.

1 Presentation of the System

1.1 State, Purpose, General Statement

AgreementMakerLight (AML) is an ontology matching system based on the design principles of AgreementMaker [1, 2] with an added focus on efficiency, to be able to tackle large-scale ontology matching problems [7]. Its initial focus was the biomedical domain, but it has been continually expanded to address a broad range of ontology and instance matching problems, and it is now a general purpose ontology matching system. AML relies primarily on lexical matching algorithms [8], 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 [6].

This year, our development of AML was mainly focused on tackling complex matching problems from the new Complex Matching track. Alas, just extending AML to handle the complex EDOAL alignment format took up most of our development time. When we were finally able to start developing matching algorithms, it became clear that each of the numerous types of EDOAL mappings would require its own specialized algorithm, and were only able to develop algorithms for some of the simplest cases, found in the Conference dataset.

We were also unable to fully integrate the code for complex matching with the main AML code-base before the OAEI deadline, and thus participated in the Complex Matching track using a different version of AML, AMLC. In addition to this version and the main AML SEALS version, we 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, 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 2018. For further information on AML's matching strategy, we direct the reader to AML's original paper [7] as well as to the OAEI results publications of the last three editions [4, 5, 3].

1.2.1 Complex AML

For the complex matching track, we focused on the challenge based on the conference ontologies. We developed strategies to identify Attribute Occurrence Restrictions and Attribute Domain Restrictions based on patterns similar to [9]. Attribute Occurrence Restrictions were detected by (1) computing the lexical similarities between the source class and the domains/ranges (or superclasses of domains/ranges) of target properties; (2) selecting target properties with domain/range similarity above a given threshold; (3) building a complex mapping with a comparator and a non-negative integer for the properties with similar domain, adding an inverse property restriction for those with similar range.

Attribute Domain Restrictions were discovered by (1) measuring the lexical similarity between the source class and target classes and selecting target classes above a threshold; (2) removing the matched words from source labels; (3) matching the remaining source strings to target properties and selecting target properties above a threshold; (4) composing a complex mapping which is given a score weighted by the two partial similarities (class and property); (5) selecting complex mappings with scores above a threshold.

1.2.2 Main AML

We made only a few minor changes to the main AML code-base for this OAEI edition.

Instance Matching

In previous OAEI editions, AML's matching strategy for instance matching relied only on Data Property values of individuals and on the relations between individuals. This year, due to the new Knowledge Graph track in which individual matching is expected to be mainly based on their annotations, AML added to its instance matching arsenal the same lexical-based strategy it was already using for class and property matching. However, due to problems in parsing the datasets with the OWL API before the OAEI deadline, we were unable to properly configure this matching strategy and ensure its efficiency.

Interactive Matching

We fixed a bug in AML's interaction manager that was causing it to forget user feedback

between the selection and repair steps and thus repeat some questions.

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 virtually identical to last year's, with 95% precision, 93.6% recall, 94.3% F-measure, and 83.2% recall++. It was the best ranking system in this track by F-measure.

2.2 Conference

AML's result was exactly the same as last year's, 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 was 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.

2.3 Multifarm

AML's result was the same as last year's, with 46% F-measure when matching different ontologies and 27% when matching same ontologies. AML was the best ranking system in this track by F-measure in the different ontologies modality.

It is noteworthy that the performance in the same ontologies modality is worse than in the different ontologies, given that the opposite is expected, and indeed was the case for AML prior to 2016. We are unsure as to what led to this relative drop in performance and will have to investigate the matter further.

2.4 Complex Matching

AMLC was configured only for the Conference dataset, in which it obtained 54% precision, 25% recall, and 34% F-measure. It was the only system to produce complex (EDOAL) alignments in this track.

2.5 Interactive Matching

AML had a similar performance to last year's, except that fixing the bug in its interaction manager has prevented repeated queries. We did not yet improve our interactive manager to make feedback requests with sets of conflicting mappings, which would enable us to reduce the total number of user requests AML makes. Thus, the increase in F-measure per user request is relatively low for AML, even if the increase per individual mapping asked is not. It was the system least affected by user errors for the Anatomy dataset, but was substantially more affected than LogMap in the case of the Conference dataset.

2.6 Large Biomedical Ontologies

AML's results were virtually the same as last year's in this track, with an F-measure of 93.3% in FMA-NCI small, 85.5% in FMA-NCI whole, 83.5% in FMA-SNOMED small, 77.2% in FMA-SNOMED whole, 80.1% in SNOMED-NCI small and 76.8% in SNOMED-NCI whole. It was the highest ranked system by F-measure in the FMA-NCI and SNOMED-NCI problems, and the second-highest in the FMA-SNOMED problems.

2.7 Disease and Phenotype

AML generated 2010 mappings in the HP-MP task, 279 of which were unique. It ranked second by F-measure according to the 3-vote silver standard, with 85.6%. In the DOID-ORDO task, it generated by far the most mappings (4749) and the most unique mappings (1886), and as a result had a relatively low F-measure according to the 3-vote silver standard (63.6%).

2.8 Biodiversity and Ecology

AML obtained 86% F-measure in the FLOPO-PTO task and 84.4% F-measure in the ENVO-SWEET task. It ranked first by F-measure in both tasks of this new track.

2.9 SPIMBENCH

AML obtained an F-measure of 86%, ranking second by F-measure. This performance was significantly lower than last year's (92.2%), which was unexpected. We are unsure of whether this is due to a difference in the dataset.

2.10 IIMB

AML obtained a global F-measure of 82.8% across the 20 tasks of this new track, ranking second in F-measure behind LogMap (the only other participating system). It outperformed LogMap in 4 of the tasks, but had a mediocre performance in 6 others.

2.11 Link Discovery

Like in 2017, AML produced a perfect result (100% F-measure) in the Linking and all the Spatial tasks.

2.12 Knowledge Graph

Due to our inability to configure AML's new instance matching strategy prior to the OAEI deadline due to the issues with parsing the datasets for this track, AML took a substantial amount of time to run these datasets, and was unable to finish all of them before the deadline for this manuscript. Nevertheless, for the tasks in which it did complete, it had a high performance in class matching (87% F-measure) but a relatively poor performance in instance matching (28% F-measure).

3 General comments

3.1 Comments on the results

This year, AML was the only system to rise to the challenge of tackling complex ontology matchings, and was the only system to participate in all the tracks. It remained among the highest ranked systems in most of the tracks in which it participated and among the most efficient. The few exceptions to AML's superiority were caused by our inability to test the datasets before the OAEI deadline. We expect to address the remaining challenges in the near future.

3.2 Comments on the OAEI test cases

As always, we welcome the addition of new tracks to the OAEI, and laud the efforts of their organizers, as the effort involved in organizing said tracks cannot be overstated. Nevertheless, we must comment on some of the issues encountered during this OAEI edition, and suggest improvements for the future.

In the new Complex Matching track, we found that the tasks were indeed extremely complex, and in many cases virtually impossible to tackle automatically, as there was insufficient information in the ontologies to derive the type of mappings that were expected. We will work with the organizers to make the tasks more realistic for future OAEI editions, namely by including instance data in the datasets when available.

In the new Knowledge Graph track, the fact that the datasets were not valid OWL (and thus not parsable with the OWL API) before the OAEI deadline was a substantial issue

which prevented us (and undoubtedly other participants) from adequately developing our matching system. For future OAEI editions, we suggest that track organizers test their datasets using a few of the recurring OAEI participating systems.

In the Link Discovery track, we stress once more the need to incorporate TBox information into the datasets so as to enable them to be interpreted automatically without the need for a dedicated parser.

Last but not least, we remain critic of the evaluation in the Disease and Phenotype track by means of silver standards generated from the alignments produced by the participating systems via voting. While we understand that the effort behind building a manually curated reference alignment can be daunting, the current evaluation strategy is unreliable and biased, penalizing systems that are able to find unique mappings that may well be correct. We would welcome an effort to produce a manually curated reference alignments using the silver standards as a starting point.

4 Conclusion

Like in 2017, this year AML was the only matching system to participate in all OAEI tracks, and was among the top performing systems in most of them. AML's performance did not improve in any of the recurring OAEI tracks, as most of our development effort went into tackling new challenges and extending the range of AML. This year, our effort to tackle complex matchings was not well rewarded, as, despite being the only system to generate complex mappings, AML was only able to cover a few of the simplest types of complex mappings. It has become evident that generating such mappings automatically is an extremely difficult task, which requires more effort than that we could devote at this time. Thus, we will continue to address this aspect of ontology matching in the near future.

Acknowledgments

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

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

8. 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.

9. D. Ritze, C. Meilicke, O. Šváb Zamazal, and H. Stuckenschmidt. A pattern-based ontology matching approach for detecting complex correspondences. In *Proceedings of the 4th International Conference on Ontology Matching*, volume 551 of *OM'09*, pages 25–36, Aachen, Germany, Germany, 2009. CEUR-WS.org.

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.