

AgreementMakerLight Results for OAEI 2014

Daniel Faria¹, Catarina Martins^{2,3}, Amruta Nanavaty⁴, Aynaz Taheri⁴,
Catia Pesquita^{1,2}, Emanuel Santos², Isabel F. Cruz⁴, and Francisco M. Couto^{1,2}

¹ LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

² Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, Portugal

³ INESC-ID, Universidade de Lisboa, Portugal

⁴ ADVIS Lab, Department of Computer Science, University of Illinois at Chicago, USA

Abstract. AgreementMakerLight (AML) is an automated ontology matching framework based on element-level matching and the use of external resources as background knowledge. This paper describes the configuration of AML for the OAEI 2014 competition and discusses its results.

Our goal this year was broadening the scope of AML by delving into aspects such as translation and structural matching, while reinforcing the key aspects behind its success last year (i.e., element-level matching, the use of background knowledge, and alignment repair).

AML's participation in the OAEI 2014 was very successful, as it obtained the highest F-measure in 6 of the 8 ontology matching tracks.

1 Presentation of the system

1.1 State, purpose, general statement

AgreementMakerLight (AML) is an automated ontology matching system derived from AgreementMaker [1, 2] and developed to handle large ontology matching problems. It combines the design principles of AgreementMaker (flexibility and extensibility) with a strong focus on efficiency [6]. Furthermore, it draws on the knowledge accumulated in AgreementMaker by reusing and adapting several of its components, but also includes a growing number of novel components.

AML is primarily based on lexical matching techniques, with an emphasis on the use of external resources as background knowledge. It also emphasizes alignment coherence, featuring an improved alignment repair module.

While initially AML was mainly focused on the biomedical domain, we have striven to expand its scope throughout the last year, and it is now a general-purpose ontology matching system. We have also moved towards full automation by employing a general-purpose core matching strategy complemented with an automated background knowledge selection algorithm [5].

1.2 Specific techniques used

The AML workflow for the OAEI 2014 comprises nine different steps, as shown in Figure 1: ontology loading and profiling, translation, baseline matching, background

knowledge matching, word and string matching, structural matching, property matching, selection, and repair. The key differences from last year's workflow are the introduction of the translation and structural matching steps.

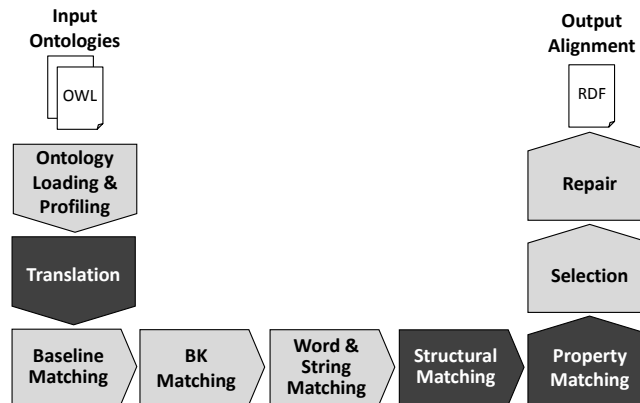


Fig. 1. The AgreementMakerLight matching workflow for the OAEI 2014. Steps in dark gray are conditional.

Ontology Loading & Profiling AML employs the OWL API [7] to read the input ontologies then retrieve the necessary information to populate its own data structures [6]:

- Class localNames, labels and synonym annotations are normalized and stored into the *Lexicon* of the corresponding ontology. AML automatically derives new synonyms for each name by removing leading and trailing stop words [12], and by removing name sections within parenthesis.
- Property names, types, domains, and ranges are stored in the *PropertyList* of the corresponding ontology.
- Relations between classes (including disjointness) and between properties are stored in a global *RelationshipMap*.

Note that AML currently does not store or use comments, definitions, or instances. After loading, the matching problem is profiled taking into account the size of the ontologies, their language(s), and the property/class ratio.

Translation AML features an automatic translation module based on Microsoft® Translator. When there is no significant overlap between the language(s) of the input ontologies, AML employs this module to translate the names of all classes and properties from the language(s) of the first ontology to those of the second and vice-versa. The translation is done by querying Microsoft Translator for the full name (rather than word-by-word). To improve performance, AML stores locally all translation results in dictionary files, and queries the Translator only when no stored translation is found.

Baseline Matching AML employs an efficient weighted string-equivalence algorithm, the *Lexical Matcher* [6], to obtain a baseline class alignment between the input ontologies. The *Lexical Matcher* has been updated to handle multi-language ontologies, by matching only class names in the same language.

Background Knowledge Matching AML has available four sources of background knowledge which can be used as mediators between the input ontologies: the Uber Anatomy Ontology (Uberon) [9], the Human Disease Ontology (DOID) [14], the Medical Subject Headings (MeSH) [10], and the WordNet [8].

The WordNet is only used for small English language ontologies, as it is prone to produce erroneous mappings in other settings. It is used through the JAWS API¹ and with the *Lexical Matcher*. The remaining three background knowledge sources are tested in all non-small single-language problems, by measuring their mapping gain over the baseline alignment [5]. When their mapping gain is very high ($\geq 20\%$), they are used to extend the *Lexicons* of the input ontologies [12]; otherwise, when it is above the minimum threshold (2%) their alignment is merged with the baseline alignment.

Uberon and DOID are both used in OWL format, and each has an additional table of pre-processed cross-references (in a text file). They can be used directly through the cross-references or with the *Lexical Matcher*. MeSH is used as a stored *Lexicon* file, which was produced by parsing the MeSH XML file, and is used only with the *Lexical Matcher*.

Word & String Matching To further extend the alignment, AML employs a word-based similarity algorithm (the *Word Matcher*) and a string similarity algorithm (the *Parametric String Matcher*) [6]. The former is not used for very large ontologies, because it is error prone. The latter is used globally for small ontologies, but only locally for larger ones as it is time-intensive.

For small ontologies, AML also employs the new *Multi-Word Matcher*, which matches closely related multi-word names that have matching words and/or words with common WordNet synonyms or close hypernyms.

Structural Matching For small and medium-sized ontologies, AML also employs a structural matching algorithm, called *Neighbor Similarity Matcher*, that is analogous to AgreementMaker's Descendants Similarity Inheritance algorithm [4]. This algorithm computes similarity between two classes by propagating the similarity of their matched ancestors and descendants, using a weighting factor to account for distance.

Property Matching When the input ontologies have a high property/class ratio, AML also employs the *PropertyMatcher*. This algorithm first ensures that properties have the same type and corresponding/matching domains and ranges. If they do, it compares the properties' names by doing a full-name match and computing word similarity, string similarity, and WordNet similarity.

¹ <http://lyle.smu.edu/tspell/jaws/>

Selection AML employs a greedy selection algorithm, the *Ranked Selector* [6], to reduce the cardinality of the alignment. Depending on the size of the input ontologies, one of three selection strategies is used: strict, permissive, or hybrid. In strict selection, no concurrent mappings (i.e., different mappings for the same class/property) are allowed and a strict 1-to-1 alignment is produced; in permissive selection, concurrent mappings are allowed if their similarity score is exactly the same; in hybrid selection, up to two mappings per class are allowed above 75% similarity, and permissive selection is applied below this threshold.

For very large ontologies, AML employs a selection variant that consists on combining the (lexical) similarity between the classes with their structural similarity, prior to performing ranked selection. This strategy enables AML to select mappings that “fit in” structurally over those that are outliers but have a high lexical similarity.

In the interactive matching track, AML employs an interactive selection algorithm which asks the user for feedback about mappings which are below a high similarity threshold (70%) and have a significant variance (with regard to similarity) between matching algorithms. The algorithm stops when a given threshold of negative answers is reached. This algorithm is based on AgreementMaker’s user feedback module [3].

Repair AML employs a heuristic repair algorithm to ensure that the final alignment is coherent [13].

For the interactive matching track, AML employs an interactive variant of this algorithm, wherein the user is asked for feedback about the mappings selected for removal.

1.3 Adaptations made for the evaluation

The only adaptations made for the evaluation were the preprocessing of cross-references from Uberon and DOID for use in the Anatomy and Large Biomedical Ontologies tracks (due to namespace differences), and the precomputing of translations for the Multifarm track (due to 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>).

1.5 Link to the set of provided alignments

The alignments generated by AML for the OAEI 2014 are available at the SOMER project page (<http://somer.fc.ul.pt/>).

2 Results

2.1 Anatomy

AML had the highest F-measure and recall this year (and all time) in this track, registering a slight improvement over last year’s result (of 0.2%). This improvement was

Table 1. AgreementMakerLight global results on all the OAEI 2014 tracks.

Track	Precision	Recall	F-Measure	Run Time
Anatomy	95.6%	93.2%	94.4%	28 sec
Interactive Matching	91.3%	73.5%	80.1%	19 sec
Benchmark				
biblio	92%	39%	55%	49 sec
cose	46%	46%	46%	140 sec
dog	98%	58%	73%	1506 sec
Conference				
Reference 1	85%	64%	73%	14 sec
Reference 2	80%	58%	67%	
Large Biomedical Ontologies				
Average	90.6%	75.2%	81.9%	307 sec
FMA-NCI small	96.0%	89.9%	92.8%	27 sec
FMA-SNOMED small	92.6%	74.2%	82.4%	126 sec
SNOMED-NCI small	91.7%	72.4%	80.9%	831 sec
FMA-NCI whole	83.2%	85.6%	84.4%	112 sec
FMA-SNOMED whole	89.1%	64.7%	74.9%	251 sec
SNOMED-NCI whole	91.2%	64.5%	75.6%	497 sec
Library				
New OWL	82.4%	77.8%	80.0%	68 sec
Old OWL	71.6%	74.8%	73.1%	71 sec
Multifarm				
Different Ontologies	57%	53%	54%	8 min
Same Ontologies	95%	48%	62%	
Ontology Alignment for Query Answering				
Original Reference	72.2%	69.4%	70.4%	N/A
Repaired Reference	70.1%	69.4%	69.1%	

mainly due to the addition of MeSH as a background knowledge source.

2.2 Benchmark

AML had a good performance in *dog*, ranking third in F-measure, but average performances in the other two test suites. This difference is due to the fact that AML does not handle ontology instances, which are present in the other two test suites, but not in *dog*.

2.3 Conference

AML ranked first in F-measure and recall, and second in precision this year, with a considerable improvement (3% F-measure) over last year's results. This improvement is due to the refinements in the Word and String Matching step.

2.4 Interactive Matching

AML ranked first in F-measure, recall, and precision this year, with a considerable improvement (7.2% F-measure) over last year's result. This improvement is partially due to the (non-interactive) refinements in the Word and String Matching step, but mainly due to the refinement of the interactive selection algorithm. The latter is evidenced by the fact that the difference between AML's interactive and non-interactive performance increased since last year (from 3 to 7.1% in F-measure) while the number of user interactions was approximately the same.

2.5 Large Biomedical Ontologies

AML had the highest F-measure in all six tasks this year, and the highest F-measure of all time in four of them. It improved substantially in all tasks, thanks to the addition of new background knowledge sources (MeSH and DOID) and the refined selection step.

2.6 Library

AML ranked first in F-measure, precision and recall in this track, having the highest F-measure of all time when using the new OWL conversion of the Library thesauri. AML's F-measure using the old OWL conversion was approximately the same as last year, but it had a higher precision and a lower recall due to a more stringent selection step. The improvement when using the new OWL conversion is due to the conversion's differentiation of *skos:altLabel* and *skos:prefLabel*. This effectively enables AML's lexical weighting scheme, which greatly improves its ability to score and select mappings.

2.7 Multifarm

AML had the highest F-measure and recall in both modalities of this track this year (and the highest F-measure of all time) and also the highest precision in matching same ontologies. These results show that, while simple, AML's new translation module is effective. The improvement over last year was dramatic, as last year AML did not perform translation.

2.8 Ontology Alignment for Query Answering

AML ranked second and third in F-measure in this new track, when evaluated on the original and repaired reference alignment respectively. While good, these results do not reflect the fact that AML produced the best set of alignments for the Conference ontologies used in this track, as the number of queries performed was too small to be representative.

3 General comments

3.1 Comments on the results

AML was very successful in this year's evaluation, ranking first in F-measure in 6 of the 8 ontology matching tasks. Furthermore, AML improved substantially over last year's evaluation, overcoming several of its limitations. Throughout the past year, we have striven to make AML a more complete ontology matching system while maintaining a strong emphasis on efficiency, which is accurately depicted in the results.

3.2 Discussions on the way to improve the proposed system

The one key feature still missing from AML is handling (and matching) ontology instances, so this is the aspect where it could improve the most. We are also interested in enabling AML to read SKOS thesauri, to broaden its applicability.

3.3 Comments on the OAEI test cases

This year's Large Biomedical Ontologies reference alignments marks a significant improvement over previous years, as the use of a 'soft' repair (where mappings are flagged rather than removed) makes the evaluation less biased [11].

Currently, our main concern about the evaluation is the (in)completeness of some of the reference alignments, particularly in the Large Biomedical Ontologies and (to a lesser degree) Anatomy tracks. Upon analyzing the alignment produced by AML for these tracks, we observed that many of the false positives were in fact correct mappings that were absent from the reference alignment. Incomplete reference alignments undermine OAEI's evaluation effort, so albeit cumbersome, completing them is paramount. On our part, we will share with the track organizers the false positive mappings found by AML that we deem to be true positives, upon a more extensive analysis.

We would also like to comment on the fact that the Conference track's reference alignment 1 includes many mappings that are apparently erroneous (as they were removed upon creating reference alignment 2). This is evidenced by the considerable drop in precision between references 1 and 2 observed for all systems. While we see the merit in a blind evaluation, we would expect that, if a partial reference alignment is provided to systems, it be fully correct. This is especially relevant given that the reference alignment 1 is used in several other OAEI tracks.

Finally, while we recognize the importance of evaluating ontology alignment applications, as per the new Query Answering track, we hope that in subsequent OAEI editions this evaluation be more representative of the underlying alignments.

4 Conclusion

AML's OAEI 2014 participation was a success, as it ranked first in F-measure in 6 of the 8 ontology matching tracks. This success reflects the effort put into the development of AML throughout the last year, which focused on increasing efficiency and automation, and particularly on expanding AML's scope.

Acknowledgments

DF, CP, ES, CM and FMC were funded by the Portuguese FCT through the SOMER project (PTDC/EIA-EIA/119119/2010) and the LASIGE Strategic Project (PEst-OE/EEI/UI0408/2014). The research of IFC, AN and AT was partially supported by NSF Awards CCF-1331800, IIS-1213013, IIS-1143926, and IIS-0812258 and by a UIC-IPCE Civic Engagement Research Fund Award.

We would like to thank Pedro do Vale, Joana Pinto, and Cláudia Duarte for their collaboration in developing AML.

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. I. F. Cruz, C. Stroe, and M. Palmonari. Interactive user feedback in ontology matching using signature vectors. *2012 IEEE 30th International Conference on Data Engineering*, 0:1321–1324, 2012.
4. I. F. Cruz and W. Sunna. Structural alignment methods with applications to geospatial ontologies. *Transactions in GIS*, 12(6):683–711, 2008.
5. D. Faria, C. Pesquita, E. Santos, I. F. Cruz, and F. M. Couto. Automatic Background Knowledge Selection for Matching Biomedical Ontologies. *PLoS One*, In Press, 2014.
6. 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.
7. M. Horridge and S. Bechhofer. The owl api: A java api for owl ontologies. *Semantic Web*, 2(1):11–21, 2011.
8. G. A. Miller. WordNet: A Lexical Database for English. *Communications of the ACM*, 38(11):39–41, 1995.
9. C. J. Mungall, C. Torniai, G. V. Gkoutos, S. Lewis, and M. A. Haendel. Uberon, an Integrative Multi-species Anatomy Ontology. *Genome Biology*, 13(1):R5, 2012.
10. S. J. Nelson, W. D. Johnston, and B. L. Humphreys. Relationships in medical subject headings (mesh). In *Relationships in the organization of knowledge*, pages 171–184. Springer, 2001.
11. C. Pesquita, D. Faria, E. Santos, and F. M. Couto. To repair or not to repair: reconciling correctness and coherence in ontology reference alignments. In *ISWC International Workshop on Ontology Matching (OM)*, CEUR Workshop Proceedings, 2013.
12. 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.
13. E. Santos, D. Faria, C. Pesquita, and F. M. Couto. Ontology alignment repair through modularization and confidence-based heuristics. arXiv:1307.5322, 2013.
14. L. M. Schriml, C. Arze, S. Nadendla, Y.-W. W. Chang, M. Mazaitis, V. Felix, G. Feng, and W. A. Kibbe. Disease Ontology: a backbone for disease semantic integration. *Nucleic Acids Research*, 40(D1):D940–D946, 2012.