AgreementMakerLight: A Scalable Automated Ontology Matching System

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

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Dept. Informática, Faculdade de Ciências, Universidade de Lisboa, Portugal
[3] ADVIS Lab, Dept. of Computer Science, University of Illinois at Chicago, USA

Abstract. Ontology matching is a critical task to enable interoperability between the numerous life sciences ontologies with overlapping domains. However, it is a task made difficult by the size of many of these ontologies.

AgreementMakerLight (AML) is a scalable automated ontology matching system developed primarily for the life sciences domain. It can handle large ontologies efficiently, specializes in the use of background knowledge, includes an innovative alignment repair algorithm, and features a graphical user interface which makes it easy to use.

AML has obtained top results in matching life sciences ontologies in the Ontology Alignment Evaluation Initiative, and is being used in several other applications.

1 Background

Ontology matching is the task of finding correspondences (or mappings) between semantically related concepts of two ontologies, so as to generate an alignment that enables integration and interoperability between those ontologies [2]. This task is particularly relevant in the life sciences, given the boom in ontology development which gave rise to hundreds of life sciences ontologies with partially overlapping domains.

At its base, ontology matching is a problem of quadratic complexity as it entails comparing all concepts of one ontology with all concepts of the other. Early ontology matching systems were not overly concerned with scalability, as the matching problems they tackled were relatively small. But with the increasing interest in matching large biomedical ontologies, scalability became a critical aspect, and the traditional all-versus-all ontology matching strategy became unfeasible.

AgreementMakerLight (AML) is a scalable automated ontology matching system developed to tackle large ontology matching problems, and focused in particular on the biomedical domain. It is derived from AgreementMaker, one of the leading first generation ontology matching systems [1].
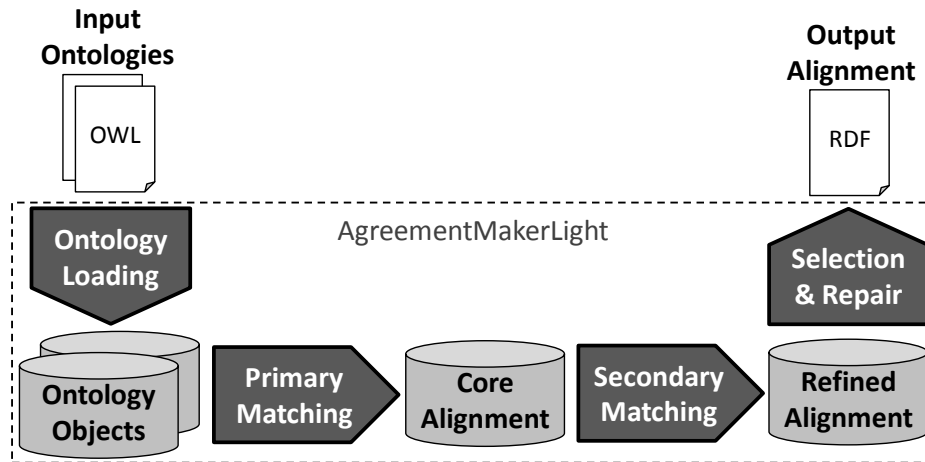
Fig. 1. AgreementMakerLight ontology matching framework.

2 The AgreementMakerLight System

AML is an open source system that is available through github [1] both as a runnable Jar and as an Eclipse project.

2.1 Ontology Matching Framework

The AML ontology matching framework, which is represented in Figure 1, was designed with scalability in mind. It includes several innovative features to maximize the effectiveness of the matching process while maintaining a reduced complexity, and is divided in three main modules: ontology loading, ontology matching, and alignment selection and repair.

The ontology loading module is responsible for reading ontologies and parsing their information into the AML ontology data structures, which were conceived to enable linear-complexity matching [4]. The most important structure for matching is the *Lexicon*, a table of class names and synonyms in an ontology, which uses a ranking system to weight them and score their matches [7].

The ontology matching module contains AML's ontology matching algorithms, or matchers. These are divided into two primary and secondary matchers, with the former being linear-complexity matchers that be employed globally in all matching problems and the latter being polynomial-complexity matchers that can only be applied locally on large problems. The use of background knowledge in primary matchers is a key feature in AML, and it includes a novel automated background knowledge selection algorithm.

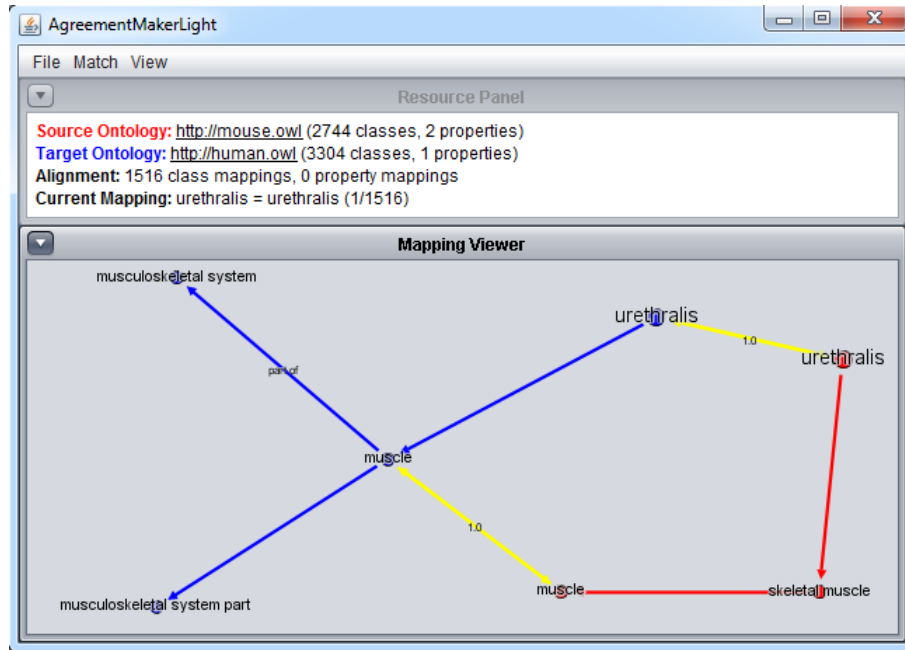The alignment selection and repair module ensures that the final alignment has

---

[1]https://github.com/AgreementMakerLight

Fig. 2. AgreementMakerLight graphical user interface.

the desired cardinality and that it is coherent (i.e., does not lead to the violation of restrictions of the ontologies) which is important for several applications. AML's approximate alignment repair algorithm features a novel modularization step which identifies the minimal set of classes that need to be analyzed for coherence, thus greatly reducing the scale of the repair problem [8].

2.2 User Interface

The GUI was a recent addition to AML, as we sought to make our system available to a wider range of users. The main challenge in designing the GUI was finding a way to visualize an alignment between ontologies that was both scalable and useful for the user. Our solution was to visualize only the neighborhood of one mapping at a time, while providing several options for navigating through the alignment [6]. The result is a simple and easy to use GUI which is shown in Figure 2.

3 AgreementMakerLight in Use

Despite only being in development since 2013, AML already achieved top results in that years' edition of the Ontology Alignment Evaluation Initiative (OAEI) [3]. Namely, it ranked first in F-measure in the anatomy track, and second in the

large biomedical ontologies track and also in the conference track. In addition to its effectiveness in matching life sciences ontologies, AML was characterized by a high F-measure/run time ratio, which attests to its efficiency.

AML has also been used to analyze ontology reference alignments [5], and is currently being used to match chemical and pharmaceutical ontologies, and to analyze the coherence of BioPortal mappings.

AML is easy to use, thanks to its GUI, and also very efficient. In a current personal computer [1] AML's matching procedure takes from under 1 minute for medium-sized problems (up to 10,000 classes per ontology) to at most 20 minutes for very large matching problems (up to 100,000 classes per ontology).

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. J. Euzenat and P. Shvaiko. *Ontology Matching*. Springer-Verlag New York Inc, 2007.
3. D. Faria, C. Pesquita, E. Santos, I. F. Cruz, and F. M. Couto. AgreementMakerLight Results for OAEI 2013. In *ISWC International Workshop on Ontology Matching (OM)*, 2013.
4. D. Faria, C. Pesquita, E. Santos, M. Palmonari, I. F. Cruz, and F. M. Couto. The AgreementMakerLight Ontology Matching System. In *OTM Conferences*, volume 8185 of *LNCS*, pages 527–541, 2013.
5. 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)*, 2013.
6. C. Pesquita, D. Faria, E. Santos, and F. M. Couto. Towards visualizing the alignment of large biomedical ontologies. In *10th International Conference on Data Integration in the Life Sciences*, 2014.
7. 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 *The Semantic Web - ISWC 2013*, volume 8218 of *Lecture Notes in Computer Science*, pages 526–541. Springer Berlin Heidelberg, 2013.
8. E. Santos, D. Faria, C. Pesquita, and F. M. Couto. Ontology alignment repair through modularization and confidence-based heuristics. *CoRR*, arXiv:1307.5322, 2013.

---

[1]With Pentium® Core™ i5 or i7 processor or equivalent, and at least 8GB RAM