ASMOV: Results for OAEI 2009

Yves R. Jean-Mary[1], E. Patrick Shironoshita[1], Mansur R. Kabuka [1,2]

[1] INFOTECH Soft, 1201 Brickell Ave., Suite 220, Miami, Florida, USA 33131
[2] University of Miami, Coral Gables, Florida, USA 33124
{reggie, kabuka}@infotechsoft.com

Abstract. The Automated Semantic Mapping of Ontologies with Validation (ASMOV) algorithm for ontology alignment was one of the top performing algorithms in the 2007 and 2008 Ontology Alignment Evaluation Initiative (OAEI) contests. In this paper, we present a brief overview of the algorithm and its improvements, followed by an analysis of its results on the 2009 OAEI tests.

1 Presentation of the System

In recent years, ontology alignment has become popular in solving interoperability issues across heterogonous systems in the semantic web. Though many techniques have emerged from the literature [1], the distinction between them is accentuated by the manner in which they exploit the ontology features. ASMOV, an algorithm that automates the ontology alignment process, uses a weighted average of measurements of similarity along four different features of ontologies, obtains a pre-alignment based on these measurements, and then semantically verifies this alignment to ensure that it does not contain semantic inconsistencies. A more complete description of ASMOV is presented in [3].

1.1 State, Purpose, General Statement

ASMOV is an automatic ontology matching tool which has been designed in order to facilitate the integration of heterogeneous data sources modeled as ontologies. The current ASMOV implementation produces mappings between concepts, properties, and individuals, including mappings between object and datatype properties.

1.2 Specific Techniques Used

The ASMOV algorithm iteratively calculates the similarity between entities for a pair of ontologies by analyzing four features: lexical elements (id, label, and comments), relational structure (ancestor-descendant hierarchy), internal structure (property restrictions for concepts; types, domains, and ranges for properties; data values for individuals), and extension (instances of classes and property values). The measures obtained by comparing these four features are combined into a single value using a weighted sum in a similar manner to [2]. These weights have been optimized based on the OAEI 2008 benchmark test results.

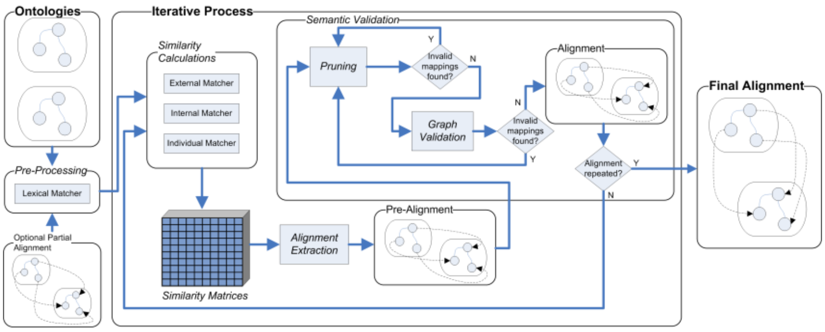

Fig. 1. The ASMOV Mapping Process

Fig. 1 illustrates the fully automated ASMOV mapping process, which has been implemented in Java. In the pre-processing phase, the ontologies are loaded into memory using the Jena ARP parser [4] and ASMOV's ontology modeling component. A thesaurus is then used in order to calculate the lexical similarities between each pair of concepts, properties and individuals. ASMOV can be configured to use either the UMLS Metathesaurus [5] or WordNet [6] in order to derive the similarity measures. A user can also opt to not use a thesaurus; in that case, a text matching algorithm is used to compute the lexical distance.

Following this, the similarities between pairs of entities along the relational structure, internal structure, and extensional dimensions are calculated, and an overall similarity measure (or confidence value) is stored in three two-dimensional matrices, one each for concepts, properties, and individuals. From these similarity matrices, a pre-alignment is obtained by selecting the entity from one ontology with the highest confidence value for a corresponding entity in the other ontology.

This pre-alignment then goes through semantic verification, which detects semantically inconsistent mappings and their causes. These inconsistent mappings are removed from the pre-alignment and logged so that the algorithm does not attempt to map the same entities in a subsequent iteration; mappings are removed from the log of inconsistencies when the underlying cause disappears. Five specific types of inconsistencies are detected by ASMOV:

- Multiple entity correspondences, where the same entity on one ontology is mapped with multiple entities in the other ontology; unless these multiple entities are asserted to be equivalent, this type of mapping is unverified.
- Crisscross correspondences, where if a class c_1 in one ontology is mapped to some other class c_1' in the second ontology, a child of c_1 cannot be mapped to a parent of c_1'.
- Disjointness-subsumption contradiction, where if two classes c_1 and c_2 are disjoint in one ontology, they cannot be mapped to two other classes c_1' and c_2' in the second ontology where one is subsumed by the other. This also

applies to the special cases where c_1‘ and c_2‘ are asserted equivalent, or where they are identical.

- Subsumption incompleteness, if two classes c_1 and c_2 are mapped to two other classes c_1‘ and c_2‘ respectively in the second ontology, and if c_2 is subsumed by c_1, then c_2‘ must be subsumed by c_1‘, otherwise the correspondences are unverified. Similar incompleteness can be verified for the special case of equivalence.
- Domain and range incompleteness: if a class c_1 in one ontology is mapped to some class c_1‘ in the second ontology, and a property p_1 in the first ontology is mapped to some property p_1‘ in the second ontology, and if c_1 belongs to the domain (or range) of p_1 , then c_1‘ must belong to the domain (or, equivalently, range) of p_1‘,

Since OAEI 2008, ASMOV has been improved in three important respects. In particular, instance matching, which had been initially implemented in previous versions, has been thoroughly redesigned, due to the availability of high-quality reference alignments for testing. As can be seen in the results section, this has resulted in high accuracy for the matching of instances, and has also had an effect in the improvement of the accuracy of class and property matching. In addition, the code base for the entire implementation of ASMOV has been thoroughly debugged and tested, particularly to ensure faithful derivation of the entity-set similarity calculations and the semantic verification process as described in [3]. Further, for the anatomy tests in particular, we have worked to improve the performance of the UMLS Metathesaurus adapter, resulting in a significant improvement in execution time.

1.3 Adaptations Made for the Evaluation

No special adaptations have been made to the ASMOV system in order to run the 2009 OAEI tests; however, five Java executable classes have been added in order to respectively run the benchmark series of tests, the anatomy tests, the directory tests, the FAO tests, and the conference tests, and output the results in the OAEI alignment format. The threshold function used to determine the stop criteria for ASMOV was established as a step function, 95% for alignments where both ontologies have more than 500 concepts, and 100% otherwise. Although the rules of the contests stated that all alignments should be run from the same set of parameters, it was necessary to change two parameters for the anatomy tests. These parameters relate to the thesaurus being used (UMLS instead of WordNet) and to the flag indicating whether or not to use ids of entities in the lexical similarity calculations.

1.4 Link to the ASMOV System

The ASMOV system (including the parameters file) can be downloaded from http://support.infotechsoft.com/integration/ASMOV/OAEI-2009.

1.5 Link to the Set of Alignments Produced by ASMOV

The results of the 2008 OAEI campaign for the ASMOV system can be found at http://support.infotechsoft.com/integration/ASMOV/OAEI-2009.

2 Results

In this section, we present our comments on the results obtained from the participation of ASMOV in the five tracks of the 2009 Ontology Alignment Evaluation Initiative campaign. All tests were carried out on a PC running FreeBSD over VMware with two quad-core Intel Xeon processor (1.86 GHz), 8 GB of memory, and 2x4MB cache. Since the tests in the 2008 version were run in a similar machine, but running SUSE Linux Enterprise Server 10 directly on the processors, the execution times are not directly comparable, and should only be used as guidelines.

2.1 Benchmark

The OAEI 2009 benchmark tests have been divided by the organizing committee in eleven levels of difficulty; we have added one more level to include the set of 3xx tests, which have been included in the benchmark for compatibility with previous years. In Table 1, we present the results of running our current implementation of ASMOV against the OAEI 2009 tests, in comparison with the results obtained in the tests in 2008 [7], where ASMOV was found to be one of the top three performing systems [8]. As can be seen, the precision, recall, and F1 measure for the entire suite of tests shows the current implementation of ASMOV achieves 95% precision and 87% recall, and an F1 measure of 91%, which represents a 1% improvement over the 2008 version. The total execution time for all tests was 161 sec..

The accuracy of ASMOV in the benchmark tests is very high, especially for the

Table 1. Benchmark test results for ASMOV version 2009 and version 2008

Level	ASMOV 2009			ASMOV 2008		
	Precision	**Recall**	**F1**	**Precision**	**Recall**	**F1**
0	1.00	1.00	1.00	1.00	1.00	1.00
1	1.00	1.00	1.00	1.00	1.00	1.00
2	1.00	0.99	0.99	1.00	0.99	0.99
3	0.99	0.98	0.98	0.98	0.97	0.97
4	0.99	0.98	0.98	0.99	0.98	0.98
5	0.97	0.93	0.95	0.96	0.93	0.94
6	0.95	0.88	0.91	0.94	0.88	0.91
7	0.94	0.83	0.88	0.93	0.83	0.88
8	0.91	0.71	0.80	0.90	0.71	0.79
9	0.83	0.48	0.61	0.78	0.46	0.58
10	0.40	0.04	0.07	0.40	0.04	0.07
3xx	0.81	0.82	0.81	0.81	0.77	0.79
All	**0.95**	**0.87**	**0.91**	**0.95**	**0.86**	**0.90**

lowest levels of difficulty. It is particularly noteworthy that improvements in both precision and recall were obtained especially at higher levels, with the largest improvement within level 9, the second most difficult. We attribute these improvements mostly to the standardization of the entity-set similarity calculation, as well as to the correction of some coding errors. There is also significant improvement, especially in recall, in testing with the real-world ontologies.

2.2 Anatomy

For the anatomy track, ASMOV uses the UMLS Metathesaurus [5] instead of WordNet in order to more accurately compute the lexical distance between medical concepts. Importantly, improvement in execution time of more than one order of magnitude, for all tests, was achieved by pre-indexing the UMLS Metathesaurus. In addition, the lexical similarity calculation between concept names (ids) is ignored as instructed by the track organizers. ASMOV produces an alignment for all four subtasks of this track:

1. *Optimal solution*: The optimal solution alignment is obtained by using the default parameter settings of ASMOV. It finds 1235 correct and 49 incorrect mappings from the partial alignment, Its total execution time was 4.1 minutes, an order of magnitude improvement over 2008, when it took almost 4 hours.
2. *Optimal precision*: The alignment with optimal precision is obtained by changing the threshold for valid mappings from 0% to 30%. This means that only mappings with confidences greater or equal to 0.3 make it to the alignment. This finds 1,187 correct and only 30 incorrect mappings.from the partial alignment The time cost for the generation of this alignment was 2.7 minutes, compared to 3 hours and 50 minutes in 2008.
3. *Optimal recall*: The alignment with optimal recall is generated by using a threshold of 0% and turning off the semantic verification process, to allow more mappings to form part of the final alignment. Under this setup, 1278 correct mappings and 55 incorrect mappings from the partial alignment are found. it took 4.4 minutes to execute, as opposed the 2008 execution time of 5 hours and 54 minutes.
4. *Extended solution*: The alignment, using the same setup as the optimal solution but with the partial alignment given as input, was obtained in 2.51 min.

2.3 Conference

This collection of tests dealing with conference organization contains 15 ontologies. ASMOV is able to generate all 105 potential alignments from those ontologies, as opposed to 2008 when only 75 alignments were processed. The overall time required to process all 105 correspondences was 187 seconds.

2.4 Directory

The directory tests were completed in 2.75 minutes for all 4639 tasks involved; this is in the same order of magnitude as the 2008 version. A manual analysis of a small sample of the correspondences found in these tests shows that a number of possibly erroneous correspondences found by ASMOV have a very low, but non-zero, confidence value. We therefore expect that the reported accuracy of ASMOV will suffer as a result. Note that the tests were run without setting a confidence value threshold, in compliance with the indication that all tracks be run with the same set of parameters; the use of a threshold would eliminate many of these potentially erroneous correspondences.

2.5 Oriented matching

We have performed gold-standard based evaluation on the Artificial Ontologies Corpus, derived from the benchmark series corpus of 2006 (a subset of the current benchmark series); due to time constraints, it was not possible to obtain results for the Real World Ontologies Corpus, In the Artificial Ontologies Corpus, ASMOV obtains an overall accuracy of 90% precision and 89% recall; in several of the simpler tests, ASMOV finds 100% precision and recall; some reduction in accuracy is observed for the more difficult tests. The execution time for these tests was 75.7 sec.

2.6 Instance Matching

Previous versions of ASMOV contained mechanisms for instance matching based on the same principles as for class and property matching, as outlined in [3]. However, the lack of a gold standard had precluded us from performing rigorous testing on these procedures. With the availability of the instance matching track in OAEI 2009, we have been able to test and improve our algorithms.

The performance of ASMOV in the set of IIMB instance matching tests is quite good, achieving an overall precision very close to 100% and overall recall of 98%. Perfect results are obtained for all the value transformation tests (002-010), as well as for the logical transformation tests (020-029). For the structural transformation tests, slight reductions in accuracy, especially in recall, are found for tests 015, 016, 017, and 019. This slight decrease results from conditions where the best match for an instance in one ontology cannot be chosen from among two or more alternatives in the other ontology; in these cases, ASMOV prefers to not make a selection. The same condition affects the result of test 031.The execution time for all 37 tests was 28 min. 15 sec.; the comparatively longer time is due to the large number of entities in each Abox.

Of the remainder of the instance matching tests, memory consumption issues prevented us from running most of the tests. The only test that could be run was to align the ePrints and Rexa ontologies; this test took 5.7 secs., to execute. The results from this test, and a manual analysis of some results, show that it is necessary to improve some aspects of instance matching, such as the matching of names where either the first or last name is inserted first in the label of an instance. This also shows that it is necessary to improve the scalability of ASMOV when very large ontologies are being aligned.

3 General Comments

3.1 Comments on the Results

The current version of ASMOV has shown improvement overall in recall and F1 measure with respect to the results obtained last year. This is significant since the results in 2008 were already very high. Particularly important is the fact that the improvements have been obtained within some of the most difficult tests, showing the versatility of ASMOV in finding alignments under various conditions. The high accuracy results from the IIMB instance matching tests also show the capability of ASMOV in these tasks. In the anatomy tests, an improvement in execution time of more than one order of magnitude was obtained, and we also expect that the accuracy of the results should have increased with respect to 2008.

3.2 Discussions on the Way to Improve ASMOV

ASMOV still needs to improve its ability to work with very large ontologies and resources. The current implementation of the algorithm utilizes a memory-based approach, where the entire contents of the ontologies are loaded in memory prior to performing the matching process. This process needs to be modified to use permanent storage in order to enable the alignment of very large ontologies. Also, we have started to work on parallelization of the algorithm, by creating separate threads of execution; however, this was still not fully implemented by the time of participation in this contest. In addition, we are also working in the improvement of the general scalability of the ASMOV algorithm for the processing of ontologies with a large number of entities.

3.3 Comments on the OAEI 2009 Test Cases

The new tests added to the OAEI 2009 contests provide important and welcome tools for the improvement of ontology matching systems. Most importantly, with the availability of the instance matching tests and gold standards, and particularly the IIMB benchmarks, we have been able to redesign and thoroughly test the procedures and algorithms coded for the matching of individuals. This has resulted in a much improved version of instance matching for ASMOV. Similarly, the availability of subsumption benchmarks have also allowed us to test the corresponding algorithms. On the other hand, the continuity in the benchmark, anatomy, and conference tracks allows us to evaluate the improvement of our algorithm and implementation as we proceed through its development.

We think it would be advantageous to count with additional gold standards for other alignments, so that the algorithms may be tested, debugged, and improved for a wider variety of conditions. Particularly, we would suggest that subsets of the mouse anatomy and NCI Thesaurus ontologies could be derived and a reference alignment provided for these subsets.

4 Conclusion

We have presented a brief description of an automated alignment tool named ASMOV, analyzed its performance at the 2009 Ontology Alignment Evaluation Initiative campaign, and compared it with its 2008 version. The test results show that ASMOV is effective in the ontology alignment realm, and because of its versatility, it performs well in multiple ontology domains such as bibliographic references (benchmark tests) and the biomedical domain (anatomy test). The tests results also showed that ASMOV is a practical tool for real-world applications that require on-the-fly alignments of ontologies.

Acknowledgments. This work is funded by the National Institutes of Health (NIH) under grant R43RR018667. The authors also wish to acknowledge the contribution of Michael Ryan, Ray Bradley, and Thomas Taylor of INFOTECH Soft, Inc.

References

1. Euzenat J and Shvaiko P. Ontology Matching. Springer-Verlag, Berlin Heidelberg, 2007.
2. Euzenat J. and Valtchev P. Similarity-based ontology alignment in OWL-lite. *In Proc. 15th ECAI*, Valencia (ES), 2004, 333-337.
3. Jean-Mary Y., Shironoshita E.P., Kabuka, M. Ontology Matching with Semantic Verification. Web Semantics: Science, Services and Agents on the World Wide Web. http://dx.doi.org/10.1016/j.websem.2009.04.001.
4. Jena from HP Labs http://www.hpl.hp.com/semweb/
5. Unified Medical Language System (UMLS) http://umlsks.nlm.nih.gov/
6. WordNet http://wordnet.princeton.edu/
7. Jean-Mary Y, Kabuka M. ASMOV: Results for OAEI 2008. http://www.dit.unitn.it/~p2p/OM-2008/oaei08_paper3.pdf. Accessed 28 Sept 2009.
8. Euzenat J, et.al. Results of the Ontology Alignment Evaluation Initiative 2007. http://www.dit.unitn.it/~p2p/OM-2007/0-o-oaei2007.pdf. Accessed 24 Sept 2008.
9. Mike Dean and Guus Schreiber, Editors, W3C Recommendation, 10 February 2004, http://www.w3.org/TR/owl-ref/