ASMOV: Results for OAEI 2008

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

[1] INFOTECH Soft, 9200 South Dadeland Blvd, Suite 620, Miami, Florida, USA 33156
[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 Ontology Alignment Evaluation Initiative (OAEI). In this paper, we present a brief overview of the algorithm and its improvements, followed by an analysis of its results on the 2008 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 while optionally accepting feedback from a user, uses a weighted average of measurements of similarity along four different features of ontologies, and performs semantic validation of resulting alignments. 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 systems, using their data source ontologies. The current ASMOV implementation produces mappings between concepts, properties, and individuals, including mappings from object properties to datatype properties and vice versa.

1.2 Specific Techniques Used

The ASMOV algorithm iteratively calculates the similarity between entities for a pair of ontologies by analyzing four features: lexical description (id, label, and comment), external structure (parents and children), internal structure (property restrictions for concepts; types, domains, and ranges for properties; data values for individuals), and individual similarity. 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.

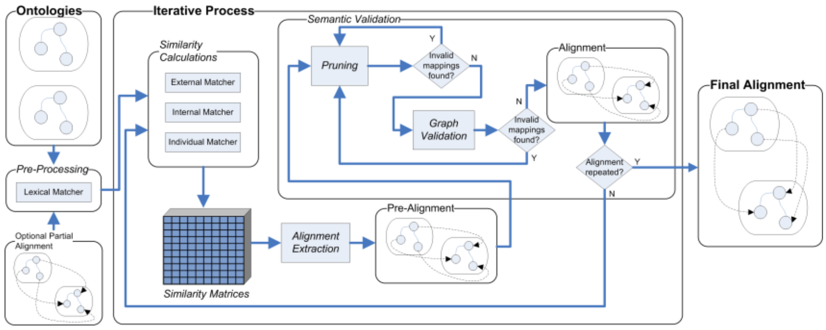

Fig. 1. The ASMOV Mapping Process

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 external structure, internal structure, and individual 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 validation, which detects semantically invalid mappings and their causes. These invalid 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 invalid log when the underlying cause disappears. In the semantic validation process, the pre-alignment is first passed through a pruning process, which detects invalid mappings by analyzing the hierarchical relationships between mapped concepts. This pruning process is performed iteratively until no invalid mappings can be found.

After the pruning process is completed, a graph validation performs a structural analysis using graphs built from the alignment and information from the ontologies, while exploring inconsistencies in equivalence, subsumption, and disjointness relationships. The validation is performed in three phases: class validation, property validation, and concept-property validation. If any invalid mappings are found, the algorithm re-enters the pruning process; otherwise, an alignment is obtained, and the percentage of mappings repeated from the previous alignment is calculated. If this percentage is less than a threshold function, and if the alignment was not previously obtained, the process returns to recalculate the similarity matrices, otherwise the ASMOV system process stops.

Since OAEI 2007, ASMOV has been improved in several important respects. A new, streamlined ontology model has been created, eliminating the use of the Jena ontology model, in order to improve the performance of the system. The lexical similarity calculation has been modified to eliminate the use of Levenshtein distance as an alternative when words are not found in the thesaurus; this calculation, while helping to find some mappings, was also introducing errors, since its value is not comparable to the similarity values obtained using dictionaries. The iterative process has been modified to perform comprehensive pruning and validation in each iteration; this modification has reduced the number of iterations required to find a solution. The ability to use a partial alignment as input to the algorithm has been implemented. A relation classifier has been added to determine whether a relation between two entities mapped to each other is an equality, or whether one is subsumed by the other. And finally, some bugs have been fixed and the overall software code has been improved.

1.3 Adaptations Made for the Evaluation

No special adaptations have been made to the ASMOV system in order to run the 2008 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-2008.

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-2008.

2 Results

In this section, we present our comments on the results obtained from the participation of ASMOV in the four tracks of the 2008 Ontology Alignment Evaluation Initiative campaign. All tests were carried out on a PC running SUSE Linux Enterprise Server 10 with two quad-core Intel Xeon processor (1.86 GHz), 8 GB of memory, and 2x4MB cache.

2.1 Benchmark

The OAEI 2008 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. The benchmarks for 2008 have varied with respect to 2007 such that the results from both benchmarks are not directly comparable. We have run the OAEI 2008 tests using the current ASMOV implementation and ASMOV from OAEI 2007 [7], which was found to be one of the top three performing systems [8]. The results of these benchmark tests for both versions of ASMOV, as well as the time elapsed for each set of tests, are presented in Table 1.

The precision and recall for the entire suite of tests shows the current implementation of ASMOV achieves 95% precision and 86% recall. This represents a 2% improvement in both precision and recall over the previous version for the entire suite of tests. Moreover, Table 1 shows the significant improvement, of an order of magnitude, in execution time achieved in the 2008 version of ASMOV.

Table 1. Benchmark test results for ASMOV version 2008 and version 2007

Level	ASMOV 2008			ASMOV 2007		
	Precision	Recall	Time (sec)	Precision	Recall	Time (sec)
0	1.00	1.00	8.60	1.00	1.00	103.55
1	1.00	1.00	4.91	1.00	1.00	67.06
2	1.00	0.99	6.06	1.00	1.00	70.11
3	0.98	0.97	9.96	0.99	0.98	143.65
4	0.99	0.98	10.07	1.00	0.96	197.09
5	0.96	0.93	8.14	0.98	0.89	222.43
6	0.94	0.88	7.22	0.92	0.82	203.65
7	0.93	0.83	7.60	0.89	0.77	194.56
8	0.90	0.71	6.65	0.84	0.72	183.82
9	0.78	0.46	2.61	0.70	0.44	79.38
10	0.40	0.04	0.54	0.38	0.05	17.96
3xx	0.81	0.77	3.42	0.82	0.82	130.72
All	**0.95**	**0.86**	**75.78**	**0.93**	**0.84**	**1,613.97**

2.1.1 Levels 0 to 4

ASMOV performs very well in this set of tests, producing an overall precision and recall of close to 100%. In level 3, there is a slight decrease in accuracy, due to test 210, which uses French words for identifiers. We should note that, even if ASMOV

2008 does not use a foreign-language dictionary, it still finds most mappings for test 210, by finding similarities over the hierarchy, property structure, and individual membership of the ontologies. In level 4, the lower precision is skewed due to test 240, where an analysis of the ontologies shows that the two "erroneous" mappings found, `Journal` to `Periodical` and `lastName` to `lastName`, should be considered correct mappings and should be present in the reference alignment.

2.1.2 Levels 5 to 8

In these levels, it can be seen that both the precision and recall diminish as the difficulty level increases, as is expected. It is also clear that there is a significant improvement between our 2007 and 2008 versions in both precision and recall, especially for the higher levels of difficulty. We attribute this improvement to the correction of some bugs in the 2007 version. In general, the tests at these levels have been stripped of labels and/or comments, and have had their ids scrambled, so that lexical similarities are not relevant; ASMOV relies on other ontology features to find a substantial number of correct mappings.

2.1.3 Levels 9 and 10

In levels 9 and 10, the most difficult, there is a pronounced decrease in the precision and recall obtained by ASMOV 2008. The results obtained are nevertheless better than those obtained using the 2007 version. In these tests, the information available in the ontologies useful to make a decision on an alignment is increasingly sparse. Level 10 shows low precision and very low recall results; these are the most difficult tests, where almost no information is available to align the ontologies. In test 262, no mappings were found. In this test, any class could be arbitrarily assigned to any other class, and ASMOV deems that the preferred alignment is the one with no mappings. The other two tests, 265 and 266, have slightly more information in terms of a hierarchy, which permits ASMOV to find some correct mappings.

2.1.4 Test 301-304

As indicated by the organizing committee, these tests represent four real-world ontologies of bibliographic references that contain some imperfections and are included for compatibility with previous years. The overall precision and recall for ASMOV 2008 were respectively 81% and 77%, slightly lower than our 2007 version.

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. 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 took 3 hours and 53 minutes in order to generate an alignment.
2. *Optimal precision*: The alignment with optimal precision is obtained by changing the threshold for valid mappings from 1% to 50%. This means that only mappings with confidences greater or equal to 0.5 make it to the alignment. The time cost for the generation of this alignment was 3 hours and 50 minutes.
3. *Optimal recall*: ASMOV uses a threshold for confidence values of 1%, to avoid negligible non-zero confidences. The alignment with optimal recall is generated by changing this threshold to 0%. Under this setup, it took 5 hours and 54 minutes in order to produce the final alignment.
4. *Extended solution*: The alignment was obtained in 51 minutes. Although one would expect that all the mappings within the partial alignment would make it to the final alignment, ASMOV's semantic validation process rejected two of them. Our analysis of the ontologies justifies the rejection performed by ASMOV.

2.3 Directory

For the 2008 version of ASMOV, we believe that a number of improvements and bug fixes in the semantic validation mechanisms have resulted in a more coherent alignment. A noticeable improvement of ASMOV is in the execution time. It took the 2007 version close to 12 minutes to complete the matching tasks while the current version finished in less than 2 minutes. ASMOV was not used to process the mdirectory tests since it does not yet use a multilingual thesaurus. It also could not run the library and vldr tests due to our inability to run the SKOS-to-OWL converter.

2.4 FAO

ASMOV was able to identify a few mappings in this series of tests. This track helped us refine the ontology modeling component of ASMOV with support for ontology extension through the *owl:imports* construct. The total processing time for the FAO tests was 4 hours and 39 minutes.

2.5 Conference

This collection of tests dealing with conference organization contains 15 ontologies. ASMOV is able to generate 75 generic correspondences from those ontologies. The overall time required to process all 75 correspondences was less than 33 seconds. Manual analysis of a small sample of the alignments produced by ASMOV indicates that the overall output of the classification component is promising.

Some issues were encountered with two of this track's ontologies: *paperdyne.owl* and *OpenConf.owl*. Specifically, in *paperdyne.owl* the property `hasAcronym` is declared both as a datatype property and as an inverse functional property; in *OpenConf.owl*, an anonymous class is declared as an enumeration of a mixture of classes and individuals. Neither of these constructs is valid in OWL-DL, according to the OWL specification [9]; ASMOV supports only OWL-DL. Additionally, ASMOV had trouble aligning *Conference.owl* and *MICRO.owl*, possibly due to an inability to compare `oneOf` with `Union` concept declarations.

3 General Comments

3.1 Comments on the Results

Although the current version of ASMOV performed well in the 2008 OAEI benchmark series of tests, its accuracy decreased for a subset of the tests compared to the accuracy obtained with last year's version. However the overall precision and recall of the 2008 version of ASMOV performs better than its 2007 counterpart; an improvement of 2% in both precision and recall was attained. Moreover, ASMOV shows a large improvement in its performance and its ability to process larger ontologies, having reduced processing times by one order of magnitude. Nevertheless, further enhancements to its scalability are still needed.

3.2 Discussions on the Way to Improve ASMOV

As in the 2007 version of ASMOV, the mapping validation in the current implementation is still source dependent, making the alignment process a directional one. As our future work, we intend to improve the mapping validation process so that it does not favor the source ontology. Although ASMOV will always converge, the amount of time needed for execution may be too great when dealing with large ontologies. To address this issue a threshold step function was added to the current version of ASMOV. It is necessary to further study different alternatives for a threshold function, in terms of tradeoff between accuracy and scalability.

3.3 Comments on the OAEI 2008 Test Cases

With the new tests added to the benchmark track we were able to do a proper behavior analysis of ASMOV depending on the semantics within ontologies, which guided the correction of coding errors. In the anatomy series of tests, the newly added test, which includes the previously referenced partial alignment, was useful in identifying issues within our semantic validation process; multiple inheritances was not addressed properly and thus led to the rejection of accurate mappings. The directory tests challenged the taxonomy validation of ASMOV while the conference track tested our relation classifier. The FAO tests made sure that ASMOV is able to properly load ontologies that include the *owl:imports* construct.

An ambiguity exists in the instruction of the execution phase of the OAEI 2008 campaign. Participants are told only to use one set of parameters for all tests in all tracks; however, the anatomy track instructs participants to disregard the names (ids) of the concepts and to rely on their labels and the annotation property values in order to perform the lexical comparison. Since the lexical matcher of ASMOV does leverage the id in its computation, a parameter was added to indicate whether or not to use ids. Therefore, the set of parameters for this track was different than for the other ones ASMOV participated in this year. Furthermore, ASMOV uses one of two lexical databases in order to compute the distance between lexical terms. For the anatomy track, the UMLS Metathesaurus was used while WordNet was used for all other tracks.

4 Conclusion

We have provided a brief description of an automated alignment tool named ASMOV, analyzed its performance at the 2008 Ontology Alignment Evaluation Initiative campaign, and compared it with its 2007 version. The test results show that ASMOV is effective in the ontology alignment realm, and because of its flexibility, 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 with improvement in execution time, ASMOV is now a practical tool for real-world applications that require on-the-fly alignments of small 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 Mr. Patrick Shironoshita 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., Kabuka, M. ASMOV: Ontology Alignment with Semantic Validation. Joint SWDB-ODBIS Workshop, September 2007, Vienna, Austria, 15-20
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 2007. http://www.dit.unitn.it/~p2p/OM-2007/3-o-ASMOV_OAEI_2007.pdf. Accessed 24 Sept 2008.
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/