

# Experiments Varying Semantic Similarity Measures and Reference Ontologies for Ontology Alignment

Valerie Cross, Prमित Silwal, and Xi Chen

Computer Science and Software Engineering Department,  
Miami University, Oxford, OH 45056  
crossv@miamioh.edu

**Abstract.** Semantic similarity measures within a reference ontology have been used in a few ontology alignment (OA) systems. Most use a single reference ontology, typically WordNet, and a single similarity measure within it. The mediating matcher with semantic similarity (MMSS) was added to AgreementMaker to incorporate the selection of a semantic similarity measure and the combination of multiple reference ontologies in an adaptable fashion. The results of experiments using the MMSS on the anatomy track of the Ontology Alignment Evaluation Initiative (OAEI) are reported. A variety of semantic similarity measures are applied within multiple reference ontologies. Using multiple reference ontologies with the MMSS improved alignment results. All information-content based semantic similarity measures produced better alignment results than a path-based semantic similarity measure.

The AgreementMaker OA system [1] has a variety of matchers configured in a hierarchical linear weighted combination (LWC) with weights derived from quality measures on their produced mappings. In OAEI 2011, AgreementMaker introduced the mediating matcher (MM) with Uberon as a reference ontology but did not use semantic similarity measures within Uberon. In [2] the MMSS replaces the MM in its OAEI 2011 configuration. Their performances are compared on the anatomy track running both on the same cluster computer. The MM result in Table 1 differs some from that reported in OAEI 2011. With a threshold of 0.90 for semantic similarity between concepts, the same number of correct mappings was produced for both MM and MMSS as shown in Table 1. Investigating the alignment results showed although the same number is correct, they were different ones. More experiments revealed the parametric string matcher (PSM) of AgreementMaker had several of its correct

**Table 1.** Experimental Results on the OAEI Anatomy Track

OAEI 2011	Produced	Correct	Precision	Recall	F-measure
<i>MM, Uberon</i>	1439	1348	93.7	88.9	91.2
<i>MMSS, 0.90, Uberon</i>	1441	1348	93.5	88.9	91.1
<i>MMSS, 0.95 PSM kept, Uberon</i>	1443	1353	93.8	89.2	91.4
<i>MMSS, 0.95 PSM kept, Uberon+FMA</i>	1453	1364	93.9	90.0	91.9

mappings overridden by the MMSS matcher. A new LWC of matchers kept PSM mapping and the similarity threshold was raised [3]. Row 3 shows these results.

AgreementMaker did not participate in 2012 OAEI. GOMMA-bk (Generic Ontology Matching and Mapping Management with background knowledge) [4] uses a composition-based approach by mapping to three reference ontologies (UMLS, Uberon, and FMA). It was the best performer for the anatomy track with respect to f-measure with a value of 0.923. Its precision and recall were 0.917 and 0.928.

The next experiment combined Uberon and the FMA as reference ontologies to see if MMSS performance could improve. The results in the last row of Table 1 show the f-measure of 91.9, almost as good as GOMMA's which required the complete UMLS. The precision of 93.9 was better than GOMMA's but the recall was not as good. All previous experiments with the MMSS used the Lin semantic similarity measure. The two major measure categories are path-based and information-content (IC) based, which have been more prevalent in bioinformatics research. A detailed overview of current semantic similarity measures used in bioinformatics research along with their formulas and references can be found in [5]. One path based measure, the Wu and Palmer (WP) and three IC-based semantic similarity measures, the Resnik, the Lin, and the Jiang-Conrath (JC) are compared. IC measures how specific a concept is in a given ontology. The more specific the higher its IC is. Corpus-based or an ontology-based IC measures can be used. Here the ontology-based method in [6] is used. To compare the performance of these measures without the effects of the other matchers, only the MMSS is used. Table 2 shows the results with both Uberon and FMA.

**Table 2.** Semantic Similarity Measures within reference ontologies

MMSS only, 0.95	Produced	Correct	Precision	Recall	F-measure
WP	1300	1227	94.4	80.9	87.1
Resnik	1283	1226	95.6	80.9	87.6
Lin	1288	1227	95.3	80.9	87.5
JC	1295	1228	94.8	81.0	87.4

The WP measure produced the most mappings with the worst f-measure although all produced nearly identical number of correct mappings. The JC measure had highest recall yet lowest precision. The Resnik measure had the highest f-measure because of it producing less incorrect mappings. More work needs be done to determine if differences exist in the correct mappings.

## References

1. Cruz, I.F., Stroe, C., Caimi, F., Fabiani, A., Pesquita, C., Couto, F.M., Palmonari, M.: Using AgreementMaker to Align Ontologies for OAEI 2011. In: International Semantic Web Conference on Ontology Matching Workshop (2011)
2. Cross, V., Silwal, P., Morell, D.: Using a Reference Ontology with Semantic Similarity in Ontology Alignment. In: International Conference on Biomedical Ontologies (ICBO), Graz, Austria, July 22-25 (2012)

3. Cross, V., Silwal, P.: Measuring Semantic Similarity within Reference Ontologies to Improve Ontology Alignment. In: *Ontology Matching Workshop, 11th Int. Semantic Web Conference (ISWC 2012)*, Boston, USA (2012)
4. Groß, A., Hartung, M., Kirsten, T., Rahm, E.: GOMMA Results for OAEI 2012. In: *Seventh International Workshop on Ontology Matching, ISWC 2012* (2012)
5. Cross, V.: Ontological Similarity. In: *Data Mining in Biomedicine Using Ontologies*, pp. 23–43. Artech House, Norwood (2009) ISBN-13:978-1-59693-370-5
6. Seco, N., Veale, T., Hayes, J.: An Intrinsic Information Content Metric for Semantic Similarity in WordNet. In: *ECAI*, pp. 1089–1090 (2004)