ONTMAT: Results for OAEI 2017

Saida Gherbi[1] and M[ed]Tarek Khadir[2]

[1]LabGed, ESTI, Annaba 23000, Algeria
Saida_gharbi23@yahoo.fr
[2]LabGed, University Badji Mokhtar Annaba, 23000, Algeria
Khadir@labged.net

Abstract: This paper describes ONTMAT an ontology matching system, and presents the results obtained for the Ontology Alignment Evaluation Initiative (OAEI) 2017. ONTMAT is an ontology matching process, which compares the instances of ontologies to align in order to deduce the relations between their concepts. Then, based on hierarchical and binary relations between the concepts inside the ontologies it performs entities matching.

Keywords: Ontology, Alignment, OWL.

1 Presentation of the system

ONTMAT (ONTology MATching) is an ontology alignment tool, aiming to align OWL entities (classes, object properties i.e. binary relations), participating for the first time in OAEI (Conference track).

1.1 State, purpose, general statement

ONTMAT uses a terminological methods based on WordNet dictionary [2],which is exploited as background knowledge to provide a set of the relations between individuals names of the ontologies source ($O1$) and target($O2$). Then, if the name does not exist in WordNet the approach handles the n-gram measure instead of the dictionary. Moreover, from this set of individual relations we will deduce the equivalence or subsumption relation among their concepts. The equivalent concepts are recorded in an alignment matrix (A*M*), and the concepts related by subsumptions relations are registered within a temporary alignment matrix (*TAM*) [4].

Furthermore, the *TAM* elements and the concepts neighbors of A*M* are compared by using the inference roles with the terminological techniques cited previously and the retained alignment will be added to *AM*. The concepts neighbors are those related by hierarchical or binary relations with AM concepts. Here, we first align the neighboring concepts because they have more chance to be similar [1], after we will align the other concepts by using the same technics. Next, inference technics are applied on *AM* to align the binary relations.

1.2 Approach description

In our proposition we suppose that Wordnet is hierarchically organized as W(S,≤,A_g,g), where S is a set of synsets {s_1,s_2,...,s_i}(i is a positive integer), and an annotate function A_g associates the gloss g to each synset. Furthermore, the relations ≤ between concepts s_1, s_2 may be presented in the following logical relations [4] as:

1) $s_1 \subseteq s_2$;means thats_1 is a hyponym or meronym of s_2;
2) $s_1 \supseteq s_2$; express that s_1 is a hypernym or holonym of s_2;
3) $s_1 \equiv s_2$; signified that s_1 and s_2 belong to the same synset are similar[1];
4) $s_1 \perp s_2$ when s_1 and s_2 are the siblings in the part of hierarchy they are connected by a relation of antonymy.

The entities aligned can be related by one of the hierarchical relation presented in the set $HR = \{\equiv, \subseteq, \supseteq\}$where (≡: equivalence; ⊆:subclass), fuzzy relation symbolized by "&", or binary relation. Further, the binary ontologies relations (O_1, O_2)are also aligned by an element of the set *HR*. The algorithm will explain in the following items:

1. In level 1 we compare the instances names (I_{O1},I_{O2}) of ontologies($O1, O2$)to deduce the relations among their concepts. To do this, WordNet is exploited because we cannot assume with certainty that two entities are dissimilar if they have different names (synonyms), or they are equivalent if they have the same name(homonyms). If the name does not existed in WordNet we will measure the similarity among names by the n-gram measure. Then the equivalent instances will construct the instances matrix *IM*. The concepts (C_1,C_2) of ($O1, O2$) that have the same sets of instances in *IM* are considered as equivalent concepts as proven in [4], and ($C_1, C_2, \equiv$) can be added to *AM*. Although, if the instances set of C_1 are included in the instances set of C_2then ($C_1, C_2, \subseteq$) will be inserted in *TAM*.
2. The level 2 starts by applying terminological techniques on the concepts names. Next the results obtained will be combined with inferences methods illustrated in [4] to be inserted in *AM*: (C_1, C_2, Rel_1) of *TAM* confirmed or modified, where $Rel_1 \in \{\equiv, \subseteq, \supseteq, \&\}$.
3. The concepts neighbors sorted by hierarchical relations of the *AM* elements (C_1, C_2, Rel_1), are (C_1', C_2')linked to (C_1, C_2)respectively by an element of the set *HR*. The neighbors(C_1', C_2')joined by "≡" with(C_1, C_2)of AM in ($O1, O2$), will be aligned by using the inferences techniques applied on the background knowledge [3]. The background knowledge is the ontology source when the neighbors belong to O_1, and ontology target if we match the neighbors existing in O_2. The other neighbors will be matched using the terminological methods.
4. The fourth level exploits the description logic roles proven in [4] to match the concepts (C_1', C_2') associated to (C_1, C_2) by binary relations (B_1, B_2) in ($O1, O2$), as following:

- If $(C_1, C_2, \equiv) and\ (C_2, C_2', B_1)$ then (C_1, C_2', B_1) will be inserted to A_{BRS}M (Alignment Binary Relation Source Matrix)
- If $(C_1, C_2, \equiv) and\ (C_1, C_1', B_2)$ then (C_2, C_1', B_2) will be added to A_{BRT}M (Alignment Binary Relation Target Matrix)

Thus, binary relations can be aligned because we have: $B_1 h_R B_2$ Iff dom(B_1) h_R dom(B_1) and ran(B_1) h_R ran(B_1); where $h_R \in HR$[4], for instance;
If $(C_1, C_2, \subseteq) and \ (C_1, C_1', B_1)$and $(C_1', C_2', \subseteq) and \ (C_2, C_2', B_1)$ then $(B_1, B_2, \subseteq)$ will be added to A_{BR} (Alignment Binary Relation Matrix).

5. Finally, the concepts not yet aligned, will be matched via the terminological methods.

1.3 Adaptations made for the evaluation

We have adapted the format of the alignment result to the reference alignments restricted to name classes, using the "=" sign for equivalence relation with confidence of 1. Although our system provides other relations as subsumption, and binary relations without measure, as well as the alignment of binary relation by the *HR*.

2 Results

In this version we wish to test the techniques used by ONTMAT, such as, the inferences mechanisms applied onWordNet and the ontologies source and target, and the deduction of the matching among entities based on instances. The most appropriate track to do these tests is the conference track.
Conference track comprises 16 ontologies from the domain of conference organization. Most ontologies of this track were equipped with OWL DL axioms; which is useful to test our inferences approach. Table 1shows the evaluation result obtained by running ONTMAT under the SEALS client with the command:
java -jar F:/temp/seals-omt-client.jar F:/temp/ONTMAT -t

This command tests two predefined ontologies from the Conference. From Table 1 we can write that ONTMAT perform well because these ontologies are the same structure.

Table 1. Results for two predefined ontologies

Precision	Recall	F-Measure
1.0	0.455	0.625

The results obtained by the global test as illustrated in Table 2, are not well as the results of the precedent table in term of precision and F-measure. Although, the global recall is *0.434*.

Table 2. Results for conference track

Test Case ID	Precision	Recall	F-measure
cmt-conference	0.6	0.2	0.3
cmt-confof	0.4	0.25	0.308
cmt-edas	0.444	0.615	0.516

cmt-ekaw	0.217	0.455	0.294
cmt-iasted	0.143	1.0	0.25
cmt-sigkdd	0.176	0.5	0.26
conference-confof	0.052	0.467	0.094
conference-edas	0.052	0.412	0.092
conference-ekaw	0.059	0.32	0.1
conference-iasted	0.03	0.286	0.054
conference-sigkdd	0.059	0.533	0.106
confof-edas	0.04	0.421	0.073
confof-ekaw	0.04	0.4	0.073
confof-iasted	0.02	0.444	0.038
confof-sigkdd	0.02	0.571	0.039
edas-ekaw	0.016	0.217	0.03
edas-sigkdd	0.022	0.467	0.042
ekaw-iasted	0.015	0.6	0.029
ekaw-sigkdd	0.017	0.636	0.033
iasted-sigkdd	0.02	0.733	0.039
Global	0.034	0.434	0.063

2.1 Discussions on the way to improve the proposed system

To improve our application, we will also align the properties of ontologies $(O1, O2)$. Then, adapt it to read all files type, and integrate the translator to test our tool under other tracks as: Instance Matching, MultiFarm.

2.2 Comments on the OAEI test cases

The application seals-omt-client from seal, only test files where the alignment relation between concepts is itself the equivalence relation. However ONTOMAT, offers other possibilities in terms alignment relations between entities such as; & : Fuzzy and binary relations. We hope that OAEI takes into consideration those types of relations in the reference alignment file.

3 Conclusion and future work

We have briefly described the mechanisms exploited by our proposition ONTMAT, and presented the results obtained under the conference track of OAEI 2017.
This is our first participation in OAEI, the results are not satisfying, and the system presents some limitations. In the future, we will make great efforts to improve ONTMAT results, and participate in more tracks.

References

1. Euzenat, J., Shvaiko. P. “Ontology Matching” , pages 37-39; 73-87, 92-93. Springer-Verlag Berlin Heidelberg, 2007.
2. Fellbaum, C. :WordNet: An Electronic Lexical Database, MIT Press, Cambridge, MA (1998).
3. GherbiS., BelleiliH.andKhadirM., 2013. BRMAP : Un outil d’Alignement des ontologies. 7ème édition de la conférence maghrébine sur les avancés des systèmes décisionnels.
4. GherbiS., 535 M. T. Khadir, Inferred ontology concepts alignment using and an external dictionary, Procedia Computer Science 83 (2016) 648- 652, the 7th International Conference on Ambient Systems, Networks and Technologies (ANT 2016) / The 6th International Conference on Sustainable Energy Information Technology (SEIT-2016) / Aliated Workshops.doi:http://dx.doi.org/10.1016/j.procs.2016.04.145.URL http://www.sciencedirect.com/science/article/pii/ S1877050916301752.
5. Giunchiglia, F., Shvaiko, P., Yatskevich, M., 2004. SMatch an algorithm and an implementation of semantic matching. In Proc. 1st EuropeanSemanticWeb Symposium (ESWS), volume 3053 of Lecture notes in computer science, pages 61–75, Hersounisous (GR), 10-12 May 2004.