DKP-AOM: results for OAEI 2016

Muhammad Fahad
Centre Scientifique et Technique du Bâtiment *(CSTB),*
290 Route des Lucioles, Sophia-Antipolis, *FRANCE*
firstname.lastname@cstb.fr

Abstract
In this paper, we present the results obtained by our DKP-AOM system within the OAEI 2016 campaign. DKP-AOM is an ontology merging tool designed to merge heterogeneous ontologies. In OAEI, we have participated with its ontology mapping component which serves as a basic module capable of matching large scale ontologies before their merging. This is our second successful participation in the OAEI 2016 campaign and first in the Process Model Matching track of OAEI. DKP-AOM is participating with two versions (DKP-AOM and DKP-AOM_lite). The reference alignments contain correspondences between instances of the class task as well as some correspondences between events. In the lite version of DKP, it does not match classes with the events, as it is of natural semantics that events should not be mapped on classes and vice versa. Therefore, we designed our system with two variants. But, our DKP-AOM system identifies cases where tasks are matched on events (where it makes sense). This is the only difference between two variant, hence for other tracks these two variants produce the same results. In this track, we can see its competitive results in the evaluation initiative among other reputed systems. Finally, we discuss some future work towards the development of DKP-AOM.

Keywords: Ontology matching, Ontology merging, disjoint knowledge, inconsistency, incompleteness, inconciseness, validation of mappings, verification of merged ontology

1 Presentation of the System

Ontology merging is a process of building a new ontology from two or more existing ontologies with overlapping parts. The merged ontology can be either virtual or physical, but must be consistent, coherent and include all the information from the source ontologies [1]. Ontology merging is based on two primary steps. Firstly, the source ontologies are looked-up for correspondences between them. Secondly, duplicate-free and conflict-free union of source ontologies is achieved based on the established correspondences [2]. The first part mainly comes under the ontology matching, whereas the second part targets to achieve the merged ontology based on the results of the first part, i.e., mappings between source ontologies. To produce accurate merged ontology, there should be some mechanism to avoid erroneous intermediate mappings and also to merge them in such a way that produces consistent, complete and coherent merged ontology. There are many hurdles that come across in the generation of desired merged output. Firstly, ontological errors and design anomalies that can occur in the source ontologies detract from reasoning and inference mechanisms, and create bottleneck in their integration tasks [3]. In addition, conceptualization of domain, explication and modeling of knowledge over ontologies and semantic heterogeneities make their integration more difficult [4]. Secondly, even if the individual ontologies are free from errors, some of the identified mappings lead towards the erroneous situations producing several types of errors in the merged ontology [5]. For building an effective ontology merging algorithm, it is essential to incorporate ontological error checking during the validation of ontology mapping process and the verification of merged ontology to attain the accuracy of resultant output.

In order to meet the above mentioned challenges for the ontology merging research, we proposed semi-automatic DKP-OM system implemented in Jena framework for the merging of heterogeneous ontologies with the human user expert [6]. Later, we released a fully Automatic Ontology Merging (AOM) system named DKP-AOM implemented in OWLAPI 3 [7]. The name DKP comes from the concept of performing *Disjoint Knowledge Analysis (DKA)* and *Disjoint Knowledge Preservation (DKP)* during the merging process. Disjoint Knowledge Analysis plays a vital role in controlling the search space for finding similarities between source ontologies. Look-up within disjoint partitions of source ontologies significantly reduces the time complexity of the mapping phase. Disjoint Knowledge Preservation in the merged ontology helps to preserve disjoint axioms in the sub-hierarchies of merged ontology to avoid incompleteness in the resultant merged ontology. In this way, it also pin-points different conflicts between source ontologies based on disjoint axioms in the source ontologies and detects inconsistent mappings. Computed mappings that lead in many cases to a large number of unsatisfiable classes are eliminated so the resultant merged ontology should not suffer from inconsistencies. The next sub-sections provide more details about DKP-AOM and then discuss our results of OAEI participation.

1.1 Adaptations made for the evaluation

As you read above, DKP is an automatically merging system. Therefore it was developed based on user GUIs such as source ontology trees for display, visual alignments between ontologies, merged ontology tree, etc. The original version of DKP has changed and these visual components are removed so that it can participate under the seals platform. However, still it needs proper clean-up to improve its runtime for the future OAEI participations.

1.2 Link to the system

Various versions of my system can be found at my personal site: http://sites.google.com/site/mhdfahad under *plugins* tab. The mapping system is separated from the merging system, and can be downloaded according to needs. For the merging of ontologies, use the same command of seals platform with –o following three paths, two for source ontologies and one for the output merged ontology. As a result of this command, a list of ontology mappings and a resultant merged ontology are produced.

2 Results

In order to show the efficiency and effectiveness of our system, this year we participated in Process Modeling track. The results are very encouraging provided by the OAEI 2016 campaign as our system is acceptable and comparable with other participants, and are discussed in the following subsections.

2.1 Process Model Matching

This track concerns with the task of matching process models, originally represented in BPML. These models have been converted to an ontological representation. The resulting matching task is a special case of an interesting instance matching problem. Organizers have converted the BPMN representation of the process models to a set of assertions (ABox) using the vocabulary defined in the BPMN 2.0 ontology (TBox). For that reason the resulting matching task is an instance matching task where each ABox is described by the same TBox. By offering this track, OAEI hope to gain insights in how far ontology matching systems are capable of solving the more specific problem of matching process models. The collection consists of 9 models ("Cologne", "Frankfurt", "FU_Berlin", "Hohenheim", "IIS_Erlangen", "Muenster", "Potsdam", "TU_Munich", "Wuerzburg"), for each pair exists an alignment in the gold standard. However, there is only an alignment named "Cologne-Frankfurt.rdf" and no alignment "Frankfurt-Cologne.rdf". This is the first time DKP-AOM is participating in this track.

We have participated with two versions of DKP with some differences. The reference alignments contain correspondences between instances of the class task as well as some correspondences between events. In the lite version, we have not matched classes with the events, as it is of natural semantics that events should not be mapped on classes and vice versa. Therefore, we separated our system with two variants. Our DKP-AOM system identifies some cases where tasks are matched on events (where it makes sense). But in its lite version, we did not add this functionality. For an example, consider a scenario where:

BPMN1: Task (Receive Rejection)
BPMN2: Event (Rejected)

Although in real world for someone, it has the impression that the "Rejected-Event" has within the workflow the same semantics as the "Receive rejection Task". In these cases, its about getting informed, receiving a message. That is why in this case an event and a task are used to model the same real world event/task. Indeed its even hard to say, if this is an event or a task. This leads to have two variant of DKP in the participation. The following table 1 shows the comparative analysis of DKP-AOM with other systems participated in the process matching track.

Table 1. Comparative analysis of DKP-AOM with other systems [10]

Participants			Standard				Probabilistic			
Name	**OAEI/PMMC**	**Size**	**P**	**R**	**FM**	**Ranking**	**ProP**	**ProR**	**ProFM**	**Ranking**
AML	OAEI-16	221	0.719	0.685	0.702	1	0.742	0.283	0.410	2
AML-PM	PMMC-15	579	0.269	0.672	0.385	14	0.377	0.398	0.387	4
BPLangMatch	PMMC-15	277	0.368	0.440	0.401	12	0.532	0.272	0.360	8
DKP	OAEI-16	177	0.621	0.474	0.538	8	0.686	0.219	0.333	9
DKP*	OAEI-16	150	0.680	0.440	0.534	9	0.772	0.211	0.331	10
KnoMa-Proc	PMMC-15	326	0.337	0.474	0.394	13	0.506	0.302	0.378	5
Know-Match-SSS	PMMC-15	261	0.513	0.578	0.544	6	0.563	0.274	0.368	7
LogMap	OAEI-16	267	0.449	0.517	0.481	11	0.594	0.291	0.390	3
Match-SSS	PMMC-15	140	0.807	0.487	0.608	4	0.761	0.192	0.307	12
OPBOT	PMMC-15	234	0.603	0.608	0.605	5	0.648	0.258	0.369	6
pPalm-DS	PMMC-15	828	0.162	0.578	0.253	16	0.210	0.335	0.258	16
RMM-NHCM	PMMC-15	220	0.691	0.655	0.673	2	0.783	0.297	0.431	1
RMM-NLM	PMMC-15	164	0.768	0.543	0.636	3	0.681	0.197	0.306	13
RMM-SMSL	PMMC-15	262	0.511	0.578	0.543	7	0.516	0.242	0.329	11
RMM-VM2	PMMC-15	505	0.216	0.470	0.296	15	0.309	0.294	0.301	14
TripleS	PMMC-15	230	0.487	0.483	0.485	10	0.486	0.210	0.293	15

Participants of the Process Model Matching Contest are depicted in grey font, while OAEI participants are shown in black font [for details see ref 10]. The OAEI participants are ranked on position 1, 8/9 and 11 with an overall number of 16 systems listed in the table. In the probabilistic evaluation, however, the OAEI participants (AML, LogMap, DKP, DKP*) gain position 2, 3, 9 and 10, respectively. Our system DKP generates mediocre results, this indicates that the progress made in ontology matching has also a positive impact on other related matching problems, like it is the case for process model matching. While it might require to reconfigure, adapt, and extend some parts of the ontology matching systems, such a system seems to offer a good starting point which can be turned with a reasonable amount of work into a good process matching tool.

Table 2 presents the results obtained by DKP-AOM on the PM track of OAEI campaign 2016.

Test Case ID	Precision	Recall	F-measure	Test Case ID	Precision	Recall	F-measure
Cologne-FU_Berlin	1	1	1	Frankfurt-Potsdam	0.4	1	0.571
Cologne-Frankfurt	0.889	1	0.941	Frankfurt-TU_Munich	0.857	1	0.923
Cologne-Hohenheim	0	0	0	Frankfurt-Wuerzburg	0.5	0.333	0.4
Cologne-IIS_Erlangen	0.5	1	0.667	Hohenheim-IIS_Erlangen	0.5	0.2	0.286
Cologne-Muenster	0.5	1	0.667	Hohenheim-Muenster	1	0.375	0.545
Cologne-Potsdam	0.5	1	0.667	Hohenheim-Potsdam	0	0	0
Cologne-TU_Munich	0.692	1	0.818	Hohenheim-TU_Munich	0	0	0
Cologne-Wuerzburg	0.5	0.333	0.4	Hohenheim-Wuerzburg	1	0.25	0.4
FU_Berlin-Hohenheim	0	0	0	IIS_Erlangen-Muenster	0.714	0.385	0.5
FU_Berlin-IIS_Erlangen	1	0.857	0.923	IIS_Erlangen-Potsdam	0.857	0.857	0.857
FU_Berlin-Muenster	1	0.5	0.667	IIS_Erlangen-TU_Munich	0.5	0.222	0.307
FU_Berlin-Potsdam	1	0.929	0.963	IIS_Erlangen-Wuerzburg	1	0.333	0.5
FU_Berlin-TU_Munich	0.5	0.333	0.4	Muenster-Potsdam	0.714	0.455	0.556
FU_Berlin-Wuerzburg	0.667	0.333	0.444	Muenster-TU_Munich	0.5	0.222	0.307
Frankfurt-FU_Berlin	0.4	1	0.571	Muenster-Wuerzburg	1	0.333	0.5
Frankfurt-Hohenheim	0	0	0	Potsdam-TU_Munich	0.5	0.333	0.4
Frankfurt-IIS_Erlangen	0.4	1	0.571	Potsdam-Wuerzburg	0.667	0.333	0.444
Frankfurt-Muenster	0.2	1	0.333	TU_Munich-Wuerzburg	0	0	0
Global	**0.718**	**0.547**	**0.621**				

Table 2. presents the results obtained by running DKP-AOM

2.2 Conference

The goal of conference track is to find alignments among 16 ontologies relatively smaller in size (between 14 and 140 entities) but rich in semantic heterogeneities about the conference organization domain. As a result, Alignments are evaluated automatically against reference alignments. Therefore, it is very interesting to measure the Precision, Recall and F-measure of our system and also does a comparison between existing systems to see their performance on real world datasets. Table 2 presents the results obtained by running DKP-AOM on the Conference track of OAEI campaign 2016. Our system DKP-AOM has produced very competitive results among top ranked systems. Our precision measure is significantly high, recall is good giving comparable F-measure value to depict a real effort towards detecting heterogeneities for the goal of ontology matching.

Matcher	Runtime	Precision	F-Measure	Recall
DKP-AOM	9913	0.844	0.626	0.498

Table 2. DKP-AOM results on conference track ontologies

3 Conclusion and Future Directions

The participation of DKP-AOM in OAEI 2016 is a success in the Process Model Matching track. Our aim was to implement BPMN model matching; therefore, we have only implemented processing model strategy in our last version of DKP-AOM that participated in 2015. Therefore, it produces (more or less) the same output in the evaluation tracks as OAEI 2015, hence we haven't discuss output on other tracks. We can see DKP-AOM has produced competitive results in the evaluation Process Model initiative among other reputed systems.

References

1. Bruijn, J.d., Ehrig, M., Feier, C., Martín-Recuerda, F., Scharffe, F., and Weiten., M., Ontology mediation, merging and aligning. In Semantic Web Technologies. Wiley 2006
2. Euzenat, J., and Shvaiko, P., Ontology Matching. Springer, 2007, ISBN 978-3-540-49611-3.
3. Fahad, M., Qadir, M.A., Noshairwan, M.W., Ontological Errors - Inconsistency, Incompleteness and Redundancy. In Proceedings of 10th Intl Conference on Enterprise Information Systems, pp. 253-285, 2008, Spain, Springer,
4. Klein, M., (2001): Combining and relating ontologies: an analysis of problems and solution. In Proc. of Workshop on Ontologies and Information Sharing (IJCAI), pp. 53-62. Seattle, USA (2001)
5. Fahad, M., and Qadir, M.A., A Framework for Ontology Evaluation, 16th ICCS Supplement Proceeding, vol. 354, 2008, France, pp.149-158.
6. Fahad, M., Qadir, M.A., Noshairwan, W., Iftakhir, N., DKP-OM: A Semantic based Ontology Merger, Proceedings of 3rd International Conference on Semantic Technologies (I-Semantics 07) Graz, Austria, 2007, Pages 313-322
7. Fahad, M., Moalla, N., Bouras, A., Detection and Resolution of Semantic Inconsistency and Redundancy in an Automatic Ontology Merging System, Journal of Intelligent Information System (JIIS), Vol. 39(2) pp. 535-557, 29/4/2012, DOI 10.1007/s10844-012-0202-y
8. Fahad, M., Moalla, N., Bouras, A., Qadir, M.A., Farukh, M., Disjoint Knowledge Analysis and Preservation in Ontology Merging Process, proceedings of 5th International Conference on Software Engineering Advances (ICSEA'10), IEEE CS, August 22-27, 2010 - Nice, France.
9. Fahad, M., Moalla, N., Bouras, A., Towards ensuring Satisfiability of Merged Ontology, International conference on computational science, ICCS 2011, Procedia Computer Science 4 (2011), pp. 2216–222, 1-3 june, 2011
10. Process Matching Results: http://web.informatik.uni-mannheim.de/oaei/pm16/results.html

11. Fahad, M., Merging of axiomatic definitions of concepts in the complex OWL ontologies. Artificial Intelligence Review, (2016) doi:10.1007/s10462-016-9479-5 pp 1–35