Lily Results for OAEI 2018

Yezhou Tang, Peng Wang*, Zhe Pan, Huan Liu

School of Computer Science and Engineering, Southeast University, China
pwang@seu.edu.cn

Abstract. This paper presents the results of Lily in the ontology alignment contest OAEI 2018. As a comprehensive ontology matching system, Lily is intended to participate in six tracks of the contest: conference, anatomy, largebio, phenotype, biodiv and spimbench. The specific techniques used by Lily will be introduced briefly. The strengths and weaknesses of Lily will also be discussed.

1 Presentation of the system

With the use of hybrid matching strategies, Lily, as an ontology matching system, is capable of solving some issues related to heterogeneous ontologies. It can process normal ontologies, weak informative ontologies [5], ontology mapping debugging [7], and ontology matching tunning [9], in both normal and large scales. In previous OAEI contests [1–3], Lily has achieved preferable performances in some tasks, which indicated its effectiveness and wideness of availability.

1.1 State, purpose, general statement

The core principle of matching strategies of Lily is utilizing the useful information correctly and effectively. Lily combines several effective and efficient matching techniques to facilitate alignments. There are five main matching strategies: (1) Generic Ontology Matching (GOM) is used for common matching tasks with normal size ontologies. (2) Large scale Ontology Matching (LOM) is used for the matching tasks with large size ontologies. (3) Instance Ontology Matching (IOM) is used for instance matching tasks. (4) Ontology mapping debugging is used to verify and improve the alignment results. (5) Ontology matching tuning is used to enhance overall performance.

The matching process mainly contains three steps: (1) Pre-processing, when Lily parses ontologies and prepares the necessary information for subsequent steps. Meanwhile, the ontologies will be generally analyzed, whose characteristics, along with studied datasets, will be utilized to determine parameters and strategies. (2) Similarity computing, when Lily uses special methods to calculate the similarities between elements from different ontologies. (3) Post-processing, when alignments are extracted and refined by mapping debugging.

In this year, some algorithms and matching strategies of Lily have been modified for higher efficiency, and adjusted for brand-new matching tasks like Author Recognition and Author Disambiguation in the Instance Matching track.

1.2 Specific techniques used

Lily aims to provide high quality 1:1 concept pair or property pair alignments. The main specific techniques used by Lily are as follows.

Semantic subgraph An element may have heterogeneous semantic interpretations in different ontologies. Therefore, understanding the real local meanings of elements is very useful for similarity computation, which are the foundations for many applications including ontology matching. Therefore, before similarity computation, Lily first describes the meaning for each entity accurately. However, since different ontologies have different preferences to describe their elements, obtaining the semantic context of an element is an open problem. The semantic subgraph was proposed to capture the real meanings of ontology elements [4]. To extract the semantic subgraphs, a hybrid ontology graph is used to represent the semantic relations between elements. An extracting algorithm based on an electrical circuit model is then used with new conductivity calculation rules to improve the quality of the semantic subgraphs. It has been shown that the semantic subgraphs can properly capture the local meanings of elements [4].

Based on the extracted semantic subgraphs, more credible matching clues can be discovered, which help reduce the negative effects of the matching uncertainty.

Generic ontology matching method The similarity computation is based on the semantic subgraphs, which means all the information used in the similarity computation comes from the semantic subgraphs. Lily combines the text matching and structure matching techniques.

Semantic Description Document (SDD) matcher measures the literal similarity between ontologies. A semantic description document of a concept contains the information about class hierarchies, related properties and instances. A semantic description document of a property contains the information about hierarchies, domains, ranges, restrictions and related instances. For the descriptions from different entities, the similarities of the corresponding parts will be calculated. Finally, all separated similarities will be combined with the experiential weights.

Matching weak informative ontologies Most existing ontology matching methods are based on the linguistic information. However, some ontologies may lack in regular linguistic information such as natural words and comments. Consequently the linguistic-based methods will not work. Structure-based methods are more practical for such situations. Similarity propagation is a feasible idea to realize the structure-based matching. But traditional propagation strategies do not take into consideration the ontology features and will be faced with effectiveness and performance problems. Having analyzed the classical similarity propagation algorithm, *Similarity Flood*, we proposed a new structure-based ontology matching method [5]. This method has two features: (1) It has more strict but reasonable propagation conditions which lead to more efficient matching processes and better alignments. (2) A series of propagation strategies are used to improve the matching quality. We have demonstrated that this method performs well on the OAEI benchmark dataset [5].

However, the similarity propagation is not always perfect. When more alignments are discovered, more incorrect alignments would also be introduced by the similarity propagation. So Lily also uses a strategy to determine when to use the similarity propagation.

Large scale ontology matching Matching large ontologies is a challenge due to its significant time complexity. We proposed a new matching method for large ontologies based on reduction anchors [6]. This method has a distinct advantage over the divide-and-conquer methods because it does not need to partition large ontologies. In particular, two kinds of reduction anchors, positive and negative reduction anchors, are proposed to reduce the time complexity in matching. Positive reduction anchors use the concept hierarchy to predict the ignorable similarity calculations. Negative reduction anchors use the locality of matching to predict the ignorable similarity calculations. Our experimental results on the real world datasets show that the proposed methods are efficient in matching large ontologies [6].

Ontology mapping debugging Lily utilizes a technique named *ontology mapping debugging* to improve the alignment results [7]. Different from existing methods that focus on finding efficient and effective solutions for the ontology mapping problems, mapping debugging emphasizes on analyzing the mapping results to detect or diagnose the mapping defects. During debugging, some types of mapping errors, such as redundant and inconsistent mappings, can be detected. Some warnings, including imprecise mappings or abnormal mappings, are also locked by analyzing the features of mapping result. More importantly, some errors and warnings can be repaired automatically or can be presented to users with revising suggestions.

Ontology matching tuning Lily adopted ontology matching tuning this year. By performing parameter optimization on training datasets [9], Lily is able to determine the best parameters for similar tasks. Those data will be stored. When it comes to real matching tasks, Lily will perform statistical calculations on the new ontologies to acquire their features that help it find the most suitable configurations, based on previous training data. In this way, the overall performance can be improved.

Currently, ontology matching tuning is not totally automatic. It is difficult to find out typical statistical parameters that distinguish each task from others.

Background Knowledge Matching Lily used matching strategy based on background knowledge this year. Lily has two sources of background knowledge:

the UMLS Metathesaurus, two synonyms files which contain a series of synonyms of many common medical terms and we obtain it via API of bioportal.com in advance. These two background knowledge sources are all specific to the biomedical domain such as largebio and phenotype track. Using background knowledge can greatly improve the matching effectiveness and efficiency to some extent. In the future, Lily will explore more effective background knowledge for other OAEI tracks or other matching tasks in the real world.

Virtual Document This year Lily used virtual document matching technology in some matching tasks[12]. Basically, as a collection of weighted words, the virtual document of a URIref declared in an ontology contains not only the local descriptions but also the neighboring information to reflect the intended meaning of the URIref. Document similarity can be computed by traditional vector space techniques, and then be used in the similarity-based approaches to ontology matching. Different matching tasks may have different neighbour information and weighted parameters to tune.

1.3 Adaptations made for the evaluation

For anatomy and conference tasks, Lily is totally automatic, which means Lily can be invoked directly from the SEALS client. It will also determine which strategy to use and the corresponding parameters. For a specific instance matching task, Lily needs to be configured and started up manually, so only matching results were submitted.

1.4 Link to the system

SEALS wrapped version of Lily for OAEI 2018 is available at `https://drive.google.com/open?id=1irGjC4tZdofpG57kHXpblBJcf75ZwUWf`.

2 Results

2.1 Anatomy track

The anatomy matching task consists of two real large-scale biological ontologies. Table x shows the performance of Lily in the Anatomy track on a server with one 3.46 GHz, 6-core CPU and 8GB RAM allocated. The time unit is second (s).

Table 1. The performance in the anatomy task

Matcher	Precision	Recall	Recall+	F-Measure
Lily	0.872	0.795	0.518	0.832

Compared with the result in OAEI 2016 [11], there is no obvious progress(with 0.83 F-Measure). As can be seen in the overall results, Lily lies in the middle position of the rank, which indicates that it is still possible to make further progress. Inside current Lily for anatomy, we used LOM(Large scale ontology matching) technique as mentioned in PART 1.2. In the future, we will add background knowledge into Lily for better matching result.

2.2 Conference track

Lily's performance in the Conference track was exactly the same as OAEI 2016. Obviously, Lily did not output satisfactory results in this track. The performance of Lily was even worse than StringEquiv in some tasks, which is a strange phenomenon. We will further analyze this task and our system to find out the reason later.

2.3 Disease and Phenotype track

Lily participated in this track for the first time. Lily generated almost the most unique mappings(733 in HP-MP task and 1167 in DOID-ORDO task).

Table 2. The performance in the disease and phenotype task

Matcher	Task	Mappings	Unique	Precision	Recall	F-Measure
Lily	HP-MP	2118	733	0.682	0.647	0.664
Lily	DOID-ORDO	3738	1167	0.589	0.783	0.672

However, Lily obtained a relatively low F-measure according to the 3-vote silver standard(0.664 and 0.672 separately). In our matching algorithm, we used classic virtual document technique and background knowledge matching strategy[12]. For the latter, we used a dictionary of synonyms extracted from BioPortal in advance. The reason why our precision is not high may be that the threshold of our virtual document was set too low, which caused many incorrect mappings. In addition, we think current consensus alignment(reference) using voting strategy is unreasonable to some extent for Lily. Since it may be not exactly the same as the gold matching results. For example, it perhaps missed some true mappings. However, these mappings are possible in unique mappings that Lily output but this voting strategy didn't count this part possibly, which led Lily to a low recall value relatively. Anyway, we will further optimize the algorithm inside Lily to make it cope with biological matching tasks better next year.

2.4 Biodiversity and Ecology track

Table 3. The performance in the biodiversity and ecology task

Matcher	Task	Precision	Recall	F-Measure
Lily	FLOPO-PTO	0.813	0.586	0.681
Lily	ENVO-SWEET	0.866	0.641	0.737

Lily obtained 68% F-measure in the FLOPO-PTO task and 73.7% F-measure in the ENVO-SWEET task. The results are not good because of low recall value relatively. In this task, we only considered simple text information(localName, label) for matching and ignored other potential information(structural information etc.). Consequently, Lily couldn't find more true mappings lacking of those information.

2.5 Spimbench track

This is an instance-mactching track which aims to match instances of creative works between two boxes. And ontology instances are described through 22 classes, 31 DatatypeProperty and 85 ObjectProperty properties.

There are about 380 instances and 10000 triples in sandbox, and about 1800 CWs and 50000 triples in mainbox.

Table 4. The performance in the spimbench task

Matcher	Scale	Precision	Recall	F-Measure
Lily	sandbox	0.8494	1.0000	0.9185
Lily	mainbox	0.8546	1.0000	0.9216

As is shown in Table 4, Lily utilized almost the same startegy to handle these two different size tasks. We found that creative works in this task were rich in text information such as titles, descriptions and so on. Lily could make good use of it and got the highest F-Measure with shortest time. However, garbled texts and messy codes were mixed up with normal texts. And Lily relied too much on text similarity calculation and set a low threshold in this task, which accounted for the low percision.

3 General comments

In this year, a lot of modifications were done to Lily for both effectiveness and efficiency. The performance has been improved as we have expected. The strategies for new tasks have been proved to be useful.

On the whole, Lily is a comprehensive ontology matching system with the ability to handle multiple types of ontology matching tasks, of which the results are generally competitive. However, Lily still lacks in strategies for some newly developed matching tasks. The relatively high time and memory consumption also prevent Lily from finishing some challenging tasks.

4 Conclusion

In this paper, we briefly introduced our ontology matching system Lily. The matching process and the special techniques used in Lily were presented, and the alignment results were carefully analyzed.

There is still so much to do to make further progress. Lily needs more optimization to handle biological ontologies with limited time and better matching results. Thus, more complex and effective matching algorithms will be applied to Lily next year. Meanwhile, we have just tried out ontology matching tuning. With further research on that, Lily will not only produce better alignments for tracks it was intended for, but also be able to participate in the interactive track.

5 Acknowledgments

This work was supported by the National Natural Science Foundation of China (61472076 and 61472077).

References

[1] Peng Wang, Baowen Xu: Lily: ontology alignment results for OAEI 2009. In The 4th International Workshop on Ontology Matching, Washington Dc., USA (2009)
[2] Peng Wang, Baowen Xu: Lily: Ontology Alignment Results for OAEI 2008. In The Third International Workshop on Ontology Matching, Karlsruhe, Germany (2008)
[3] Peng Wang, Baowen Xu: LILY: the results for the ontology alignment contest OAEI 2007. In The Second International Workshop on Ontology Matching (OM2007), Busan, Korea (2007)
[4] Peng Wang, Baowen Xu, Yuming Zhou: Extracting Semantic Subgraphs to Capture the Real Meanings of Ontology Elements. Journal of Tsinghua Science and Technology, vol. 15(6), pp. 724-733 (2010)
[5] Peng Wang, Baowen Xu: An Effective Similarity Propagation Model for Matching Ontologies without Sufficient or Regular Linguistic Information, In The 4th Asian Semantic Web Conference (ASWC2009), Shanghai, China (2009)
[6] Peng Wang, Yuming Zhou, Baowen Xu: Matching Large Ontologies Based on Reduction Anchors. In The Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI 2011), Barcelona, Catalonia, Spain (2011)
[7] Peng Wang, Baowen Xu: Debugging Ontology Mapping: A Static Approach. Computing and Informatics, vol. 27(1), pp. 2136 (2008)
[8] Peng Wang: Lily results on SEALS platform for OAEI 2011. Proc. of 6th OM Workshop, pp. 156-162 (2011)

[9] Yang, Pan, Peng Wang, Li Ji, Xingyu Chen, Kai Huang, Bin Yu: Ontology Matching Tuning Based on Particle Swarm Optimization: Preliminary Results. In The Semantic Web and Web Science, pp. 146-155 (2014)
[10] Peng Wang, Wenyu Wang: Lily results for OAEI 2015. In The 10th International Workshop on Ontology Matching, Bethlehem, PA, USA(2015)
[11] Peng Wang, Wenyu Wang: Lily results for OAEI 2016. In The 11th International Workshop on Ontology Matching, Kobe, Japan(2016)
[12] Qu Y, Hu W, Cheng G: Constructing virtual documents for ontology matching[C]// Proc. International Conference on World Wide Web. 23-31(2006)