LILY: The Results for the Ontology Alignment Contest OAEI 2007

Peng Wang, Baowen Xu

School of Computer Science and Engineering, Southeast University, China
{pwangseu@gmail.com, bwxu@seu.edu.cn}

Abstract. This paper presents the results of LILY, which is an ontology mapping system, for OAEI 2007 campaign. To accurately describe what the real meaning of an entity in the original ontology is, LILY extracts a semantic subgraph for each entity. Then it exploits both linguistic and structural information in semantic subgraphs to generate initial alignments. If necessary, using these initial results as input, a subsequent similarity propagation strategy could produce more alignments, which often can not be obtained by the previous process. The preliminary results of the experiments for four tasks (i.e. benchmark, directories, anatomy and conference) are presented. The discussion of the results and future work of LILY are also given.

1 Presentation of the system

Currently more and more ontologies are distributedly used and built by different communities. Many of these ontologies would describe similar domains, but using different terminologies, and others will have overlapping domains. Such ontologies are referred to as heterogeneous ontologies, which is a major obstacle to realize semantic interoperation. Ontology mapping, which captures relations between ontologies, aims to provide a common layer from which heterogeneous ontologies could exchange information in semantically sound manners.

LILY is a system for solving the issues related to heterogeneous ontologies. One important function of LILY is to match heterogeneous ontologies. LILY uses the *semantic subgraph* to describe the meaning of an entity. Then linguistic and structural similarity algorithm and similarity propagation strategy are exploited to create the alignments between ontologies.

1.1 State, purpose, general statement

When LILY is used to find alignments between heterogeneous ontologies, it tries to utilize all useful information to discover the correct matching results. Currently it does not use any external knowledge such as WordNet. The matching process consists of three main steps: (1) *Extracting semantic subgraph* LILY tries to use a semantic subgraph to represent the real meaning for a given entity in an ontology. A semantic subgraph, which is also a subgraph of the original ontology, is extracted by a variant algorithm based on the connection subgraphs discovery algorithm [1]. (2) *Computing alignment similarity* Through analyzing the literal and structural information in the semantic subgraphs, LILY computes the similarity confidences between entities from different ontologies. (3) *Similarity propagation* In most cases, LILY can find satisfactory alignment results after the second process. If few alignment results are got, a strategy will decide whether to take similarity propagation process. The similarity propagation could produce more alignments that can not be found in the previous processes. The matching process is shown in Fig. 1.

LILY is still being improved and enhanced, and the lasted version is V1.2.

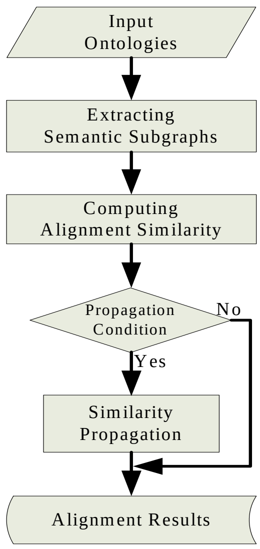

Fig. 1. Matching process

1.2 Specific techniques used

LILY aims to provide high quality alignments between concept/property pairs. The main specific techniques used by LILY are as follows.

Semantic subgraph An entity in a given ontology has its specific meaning. In our ontology mapping view, capturing such meaning is very important to obtain good alignment results. Therefore, before similarity computation, LILY first describes the meaning for each entity accurately. The solution is inspired by the method proposed by Faloutsos et al. for discovering connection subgraphs [1]. It is based on electricity analogues to extract a small subgraph that best captures the connections between two nodes of the graph. Ramakrishnan et al. also exploits such idea to find the informative connection subgraphs in RDF graph. We modify the method for extracting an *n-size* subgraph for a node or edge in an ontology graph. The subgraphs can give the precise descriptions of the meanings of the entities, and we call such subgraphs semantic subgraphs. The details of the semantic subgraph extraction process will be reported elsewhere.

Alignment similarity computation The similarity computation is based on the semantic subgraphs, i.e. all the information used in the similarity computation is come from the semantic subgraphs. LILY uses two kinds of descriptions to interpret the concepts and properties. The first is the basic description, which is a document consisting of the identifier, label and comments. The second is the semantic description. A semantic description of a concept contains the information about class hierarchies, related properties and instances. A semantic description of a property contains the information about hierarchies, domains, ranges, restrictions and related instances. For the descriptions from different entities, we calculate the similarities of the corresponding parts. Finally, all separate similarities are combined with the experiential weights. The descriptions collect the linguistic and structural information of entities. Therefore, for the regular ontologies, LILY can find satisfactory alignments in most cases.

Similarity propagation When the ontologies lack of regular and clear literal descriptions, the above method just returns few alignments. LILY uses similarity propagation strategy to compensate for it. Compared with other similarity propagation methods such as similarity flood [3] and SimRank [4], our similarity propagation defines stronger propagation condition and is based on the semantic subgraphs. The propagation graph is not stable, but is incremental during propagation process. Using similarity propagation can find more alignments that cannot be found in the previous process. 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 it requires a strategy to determine when to use the similarity propagation.

Automatic threshold selection The previous processes will return a similarity matrix, which represents the confidence level between entities from two ontologies. It is need a threshold to filter the low confidence values and keep high confidence ones. However, the threshold is usually set up manually, that cannot adapt to all matching situations. LILY treats the similarity matrix as an image, and then uses the classic image threshold selection algorithm to find a threshold automatically. There are many image thresholding methods [5]. After comparing the effectiveness of a variety of thresholding algorithms, we choose the maximum entropy approach to calculate the threshold [6]. After filtering, final 1-1 alignments are generated using the stable marriage strategy.

1.3 Adaptations made for the evaluation

In the evaluation, the size of semantic subgraph would influence on the alignment results. We set *5-size* semantic subgraphs for most test cases. When the ontologies lack of regular literals, we set *10* to *35-size* semantic subgraphs for capturing more structural information. For large scale ontologies, we just set *2* to *3-size* semantic subgraphs for the purpose of reducing the time of extracting semantic subgraphs.

1.4 Link to the system and the set of provided alignments

A demo version of LILY and the alignment results for OAEI2007 campaign are available at http://ontomappinglab.googlepages.com/oaei2007.

2 Results

In this section, we will analyze the performances and problems during taking the four alignment tracks.

2.1 benchmark

The benchmark tests the performance of matching system during various ultimate situations.

101-104 This test set contains same, irrelevant, ontology language generalization and restriction ontologies. LILY plays well for these test cases. But for the irrelevant ontology 102, LILY returns several alignments because it cannot decide whether the two ontologies are irrelevant, so it tries to find any possible alignments.

201-210 In the test cases, the structure of ontology is preserved, but the labels and identifiers are replaced by random names, misspellings, synonyms and foreign names. The comments have been suppressed in some cases. LILY can produce good results for this test set. Even without right labels and comments information, LILY can find most correct alignments through making use of other information such as instances. Using few alignment results obtained by the basic methods as inputs, the similarity propagation strategy will generate more alignments.

221-247 The test cases can be divided into two subgroups: 221-231 and 232-247. The first subgroup contains 11 kinds of modifications, such as the hierarchy is flattened or expanded, and individuals, restrictions and datatypes are suppressed. Due to the labels and comments are preserved, the modifications have little influence on our system. LILY can find most correct alignments using the labels and comments information. In the second subgroup, the modifications are the combinations of the ones used in 221-231. LILY can obtain good results for 232-247 as well.

248-266 This is the most difficult test set. All labels and identifiers are replaced by random names, and the comments are also suppressed. LILY uses the information from the semantic subgraphs to look for alignments. However, no enough information is provided in the ontologies, and the similarity computation process can only find few alignments. Subsequently, using these initial results as input, LILY exploits the similarity propagation algorithm to discover more alignments. In our experiments, too smaller and too bigger size semantic subgraph can not produce good alignments. *10-35* is a suitable size range in our experience. In 254 and 262, since almost all literal and structure information are suppressed, the similarity propagation can not find more results, so LILY just can produce limit results. When some structure information is preserved, similarity propagation will play a role and can generate more alignment results.

301-304 This test set are the real ontologies. For LILY just can find equivalent alignment relations, the inclusion alignments can not be generated. For 301-302, LILY finds most correct alignments, but it also returns some wrong results. The alignment results for 303 are far from satisfactory. We think the reason might be that 303 is no individuals and with shallow class hierarchy, and there are no direct connections between the classes and properties. Without the external knowledge, LILY can not produce good results for 303. 304 has similar structure and vocabularies to the reference ontology 101, so LILY outperforms other three ontologies.

2.2 anatomy

The anatomy track consists of two real large-scale biological ontologies. Handling such ontologies is a big problem for LILY, because extracting semantic subgraphs would need long time and large memory space. Even though LILY sets up small size semantic subgraphs for this matching task, it needs about 4 days to create the alignment results. For the purpose of producing the alignments in time, the principal technique advantages of LILY are discards in this alignment task.

2.3 directory

The directory track requires to matching two taxonomies describing the web directories. Except the class hierarchy, there is no other information in the ontologies. Therefore, LILY will utilize the hierarchy information to decide the alignments. There are three alignment tasks. The first is matching the 4640 small ontologies pairs. The second task is matching a 10% sample ontology pair. LILY completes the two tasks smoothly. The third task is required to match two large-scale taxonomies. LILY takes 8 days to produce the alignments. Similar to the anatomy track, we just set up the small size semantic subgraphs to assure that the alignment results can be generated in time.

2.5 conference

This track contains 14 real-case ontologies about conference. For a given ontology, we compute the alignments with itself, as well as with other ontologies. For we treat the equivalent alignment is symmetric, we get 105 alignment files totally. The heterogeneous character in this track is various. It is a challenge to generate good results for all ontology pairs in this test set.

3 General comments

3.1 Comments on the results

During the OAEI campaign, we are aware of the strengths and weaknesses of LILY.

Strengths For normal size ontologies, if they have regular literals or similar structures, LILY can achieve satisfactory alignments. The reason lies in two aspects: (1) The semantic subgraphs could represent the real meanings of the concepts or properties, that avoids introducing the unnecessary and noise information to the matching processes. (2) The similarity propagation strategy could compensate for the linguistic matching methods, and it can produce more alignments when ontologies lack of linguistic information.

Weaknesses LILY has two obvious weaknesses. (1) *Processing large scale ontologies* LILY cannot work well for large scale ontologies. Semantic subgraph extraction process and similarity propagation process could take terrible time for large scale ontologies. (2) *Efficiency* LILY needs to extract semantic subgraphs for all concepts and properties. It is a time-consuming process. In similarity propagation, the propagation graph would become large that will also need more time for propagating the similarities.

3.2 Discussions on the way to improve the proposed system

In OAEI07, we find the efficiency is an outstanding problem for LILY. In the matching process, most of time is spent on extracting semantic subgraphs and similarity propagation. The two processes usually account for 80% time in the full matching process. In addition, LILY's time complexity is $O(kn^2)$, where n is the number of entities and k is the average time for calculating an alignment. Therefore, it is very slow when run the large scale ontology matching task. Even we completed two large scale ontology matching tasks (directory and anatomy), we had to use the basic parameters. It causes that some advanced methods in LILY can not be utilized. To sum up, improving the efficiency and finding suitable methods to handle large scale ontologies are the near future work for LILY.

3.3 Comments on the OAEI 2007 test cases

More real ontologies should be added to the test cases. The real ontologies could be better than the ones designed manually for testing the performance of a matching system.

The large scale ontology alignment task is a challenge for some ontology matching systems such as LILY. For the sake of fairness, currently, all reference alignment results for large scale ontologies matching tasks are unknown to all participants. We suggest that it was necessary to provide an open large scale ontology matching task for all researchers. That would be benefit to finding efficient methods for matching large scale ontologies. In addition, different matching systems could compare their results based on such open large scale ontologies.

4 Conclusion

We briefly introduce our ontology matching tool LILY. The matching process and the special techniques used by LILY are presented. The preliminary alignment results are carefully analyzed. Finally, we summarized the strengths and the weaknesses of LILY.

References

1. Faloutsos, C., McCurley, K. S., Tomkins, A.: Fast Discovery of Connection Subgraphs. In the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, Washington (2004).
2. Ramakrishnan, C., Milnor, W. H., Perry, M., Sheth, A. P.: Discovering Informative Connection Subgraphs in Multirelational Graphs. ACM SIGKDD Explorations, Vol. 7(2), (2005)56-63.
3. Melnik, S., Garcia-Molina, H., Rahm, E.: Similarity Flooding: A Versatile Graph Matching Algorithm and its Application to Schema Matching. In the 18th International Conference on Data Engineering (ICDE), San Jose CA (2002).
4. Jeh, G., Widom, J.: SimRank: A Measure of Structural-Context Similarity. In the 8th International Conference on Knowledge Discovery and Data Mining (SIGKDD), Edmonton, Canada, (2002).
5. Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation, Journal of Electronic Imaging, Vol. 13(1): (2004)146–165.
6. Kapur, J. N., Wong, A. K. C.: A new method for gray-level picture thresholding using the entropy of the histogram. Computer Vision, Graphics and Image Processing, Vol. 29(2), (1985)273-285.

Appendix: Raw results

The final results of benchmark task are as follows.

Matrix of results

#	Name	Prec.	Rec.
101	Reference alignment	1.00	1.00
102	Irrelevat ontology	0.00	NaN
103	Language generalization	1.00	1.00
104	Language restriction	1.00	1.00
201	No names	1.00	1.00
202	No names, no comments	1.00	0.80
203	No comments	1.00	1.00
204	Naming conventions	1.00	1.00
205	Synonyms	1.00	0.99
206	Translation	1.00	0.99
207		1.00	0.99
208		1.00	1.00
209		0.92	0.91
210		1.00	0.91
221	No specialisation	1.00	1.00
222	Flatenned hierachy	1.00	1.00
223	Expanded hierarchy	1.00	1.00
224	No instance	1.00	1.00
225	No restrictions	1.00	1.00
228	No properties	1.00	1.00
230	Flatenned classes	0.94	1.00
231	Expanded classes	1.00	1.00
232		1.00	1.00
233		1.00	1.00
236		1.00	1.00
237		1.00	1.00
238		0.98	0.98
239		0.97	1.00
240		0.97	1.00
241		1.00	1.00
246		0.97	1.00
247		0.94	0.97
248		1.00	0.77
249		1.00	0.80
250		0.85	0.67
251		0.96	0.74
252		0.94	0.76
253		0.97	0.75
254		1.00	0.27

257		0.85	0.67
258		0.76	0.74
259		0.94	0.75
260		0.62	0.45
261		0.61	0.42
262		1.00	0.27
265		0.86	0.41
266		0.64	0.42
301	BibTeX/MIT	0.89	0.80
302	BibTeX/UMBC	0.82	0.65
303	Karlsruhe	0.58	0.69
304	INRIA	0.91	0.97