Flexible Bootstrapping-Based Ontology Alignment

Prateek Jain, Pascal Hitzler, Amit P. Sheth

Kno.e.sis Center, Wright State University, Dayton, OH

Abstract. BLOOMS (Jain et al, ISWC2010, to appear) is an ontology alignment system which, in its core, utilizes the Wikipedia category hierarchy for establishing alignments. In this paper, we present a Plug-and-Play extension to BLOOMS, which allows to flexibly replace or complement the use of Wikipedia by other online or offline resources, including domain-specific ontologies or taxonomies. By making use of automated translation services and of Wikipedia in languages other than English, it makes it possible to apply BLOOMS to alignment tasks where the input ontologies are written in different languages.

1 Introduction

In this work, we present the extension of BLOOMS called Plug-n-Play BLOOMS (PnP BLOOMS). PnP BLOOMS allows users to plug in a datasource of choice to assist in the task of ontology matching while utilizing the core BLOOMS approach. Thus, PnP BLOOMS can be customized by plugging in other information sources best suited to the needs of the domain and the user. At the same time it has been developed in a modular fashion to allow for de-coupling of the main matching technique with the auxiliary source. The system is available from http://wiki.knoesis.org/index.php/BLOOMS.

The paper is organized as follows. In Section 2, we present our proposed solution, Section 3 contains a minimalistic evaluation which shows the feasibility of our approach and finally we conclude in Section 4.

2 Our Solution

We present the BLOOMS approach following [1], modified as appropriate to cater for our plug-n-play extension. BLOOMS constructs a forest (i.e., a set of trees) T_C (which we call the *BLOOMS forest* for C) for each matching candidate class name C – in the original approach, this roughly corresponds to a selection of supercategories of the class name in the sense of the Wikipedia clas shierarchy. Comparison of the forests T_C and T_B for matching candidate classes C and B then yields a decision whether or not (and with which of the candidate relations) C and B should be aligned.

We have implemented an extension of the BLOOMS framework from [1] to support 'plug-n-play' methodology. Plug-n-Play methodology provides users the flexibility to customize the framework to use the best auxiliary datasource based on their needs and application scenario. The framework allows users to plug-in the auxiliary data source which is best suited to the target domain. This is in part possible due to the nature of the BLOOMS approach, which relies on the comparison of nodes in the categorization tree.

Table 1. Comparison of BLOOMS and PnP BLOOMS on French ontologies of the Benchmark track of OAEI 2009.

Ontology	f-measure BLOOMS	f-measure PnP BLOOMS
206	0.53	0.76
207	0.58	0.77
210	0.50	0.72
Avg.	0.54	0.75

3   Evaluation

In order to evaluate our plug-n-play approach for ontology matching, we utilized the French language ontologies from the Benchmark track of OAEI 2010 initiative.[1] We utilized these ontologies since It allows us to compare the performance of PnP BLOOMS with the old version of BLOOMS. Further, it allows us to check the plug-n-play version of BLOOMS with respect to ease of use and functionality. Table 1 illustrates the comparative performance of PnP BLOOMS with BLOOMS. The table clearly illustrates the advantage obtained by using the PnP approach. The effectiveness of the BLOOMS PnP approach is demonstrated by an increment of 42% over the previous version of BLOOMS.

4   Conclusion and Future Work

In this paper we have presented an extension of our recent work on BLOOMS [1] which allows users to plug-in auxiliary sources of their choice as oracles in the task of ontology matching. It differs from the previous version of BLOOMS which was restricted to using English language Wikipedia to help in this task. The effectiveness of the approach is demonstrated by the increment of 42% over the previous version of BLOOMS. The flexible plug-n-play based approach allows users to customize the system based on their needs and the domain and language of the ontologies. Using PnP BLOOMS, we are also planning on investigating the use of 'cocktails' of auxiliary data sources for the task of ontology matching.

References

1. Jain, P., Hitzler, P., Sheth, A.P., Verma, K., Yeh, P.Z.: Ontology Alignment for Linked Open Data. In: Proceedings of the 9th International Semantic Web Conference, ISWC 2010, Shanghai, China, November 7-11, 2010 (To Appear), Springer-Verlag (2010) 402–417 Available from http://knoesis.org/library/resource.php?id=844.

---

[1] http://oaei.ontologymatching.org/2009/