Mix'n'Match: Iteratively Combining Ontology Matchers in an Anytime Fashion

Simon Steyskal[1,2] and Axel Polleres[3,1]

[1] Siemens AG, Siemensstrasse 90, 1210 Vienna, Austria
[2] Vienna University of Technology, 1040 Vienna, Austria
[3] Vienna University of Economics & Business, 1020 Vienna, Austria

1 The Mix'n'Match Framework

Mix'n'Match is a framework to combine different ontology matchers in an iterative fashion for improved combined results: starting from an empty set of alignments, we aim at iteratively supporting in each round, matchers with the combined results of other matchers found in previous rounds, aggregating the results of a heterogeneous set of ontology matchers, cf. Fig.1.

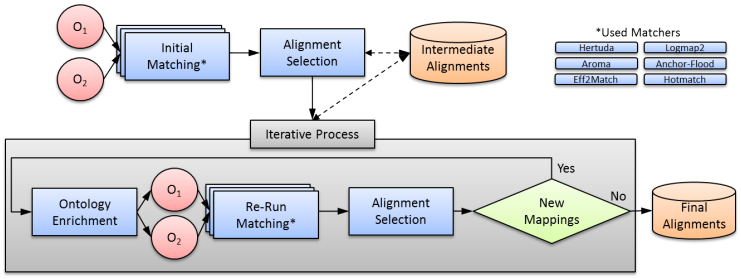

Fig. 1. Framework of Mix'n'Match

Alignment Combination: The combination of the alignments, especially the choice of those which are used for the enrichment step is based on majority votes. By only accepting alignments which were found by a majority of heterogeneous matching tools we aim to ensure a high precision of the found alignments and therefore try to emulate reference alignments as e.g. provided by iterative approval through a human domain expert. Although Mix'n'Match would support the definition of an alignment confidence threshold as additional parameter (i.e, only allowing alignments over a specific threshold to pass) we set this threshold per default to 0 in our experiments: since the calculation of confidence values is not standardized across matchers and some matchers only produce boolean confidence values, e.g. [3]). Other result aggregation methods may be conceivable here, like taking the individual performance of off-the-shelf matchers on specific matching tasks into account [2, 4], but since this approach would lead to a more inflexible alignment process this issue needs more detailed investigations in future versions of Mix'n'Match.

Ontology Enrichment: After mixing of the alignments, enrichment of the ontologies takes place; since most ontology matchers do not support refernce alignments (as specified by the OAEI alignment format[4]), we implement enrichment by simple URI replacement to emulate such reference alignments found in each matching round: for every pair of matched entities in the set of aggregated alignments, a merged entity URI is created and will replace every occurrence of the matched entities in both ontologies. This approach is motivated by the assumption that if two entities were stated as equal by the majority of ontology matchers, their URI can be replaced by an unified URI, stating them as equal in the sense of URIs as global identifiers. Note that, despite the fact that most matchers seem to ignore URIs as unique identifiers of entities, our experiments showed that URI replacement was effective in boosting the confidence value of such asserted alignments in almost all considered matchers.

Intermediate Results and Anytime Behavior: We collect the intermediate results of every finished off-the-shelf matcher in every iteration. Furthermore we keep track of every alignment found so far together with the number of individual matchers which have found this alignment in any previous matching round. This offers the possibility to interrupt the matching process at any time, retrieving only those alignments which have been found by the majority of the ontology matchers at the time the interruption has taken place. In contrast to other ontology matchers which offer this anytime behavior like MapPSO [1], we are not only restricted to gather alignment results of the last finished matching iteration, but also use the alignment results of already finished off-the-shelf matchers in the current matching round.

Evaluation Results To test our approach, we based our evaluations on OAEI evaluation tracks (*Benchmark, Conference, Anatomy*) and retrieved very promising results, typically outperforming the single matchers combined within the Mix'n'Match framework in terms of F-measure. For detailed evaluation results we refer our readers to an extended report accompanying this poster, available at `http://www.steyskal.info/om2013/extendedversion.pdf`.

References

1. J. Bock, J. Hettenhausen. Discrete particle swarm optimisation for ontology alignment. *Information Sciences*, 192:152–173, 2012.
2. I.F. Cruz, F. Palandri Antonelli, C. Stroe. Efficient selection of mappings and automatic quality-driven combination of matching methods. In *Int'l Workshop on Ontology Matching (OM)*, CEUR volume 551, pages 49–60. Citeseer, 2009.
3. M. Seddiqui Hanif, M. Aono. An efficient and scalable algorithm for segmented alignment of ontologies of arbitrary size. *J. Web Sem.*, 7(4):344–356, 2009.
4. A. Nikolov, M. d'Aquin, E. Motta. Unsupervised learning of link discovery configuration. In *ESWC2012*, pages 119–133. Springer, 2012.

[4] `http://alignapi.gforge.inria.fr/format.html`